

**NEURO LINGUISTIC PROGRAMMING AUTOMATION
FOR IMPROVEMENT OF ORGANISATIONAL
PERFORMANCE**

A DISSERTATION

SUBMITTED TO THE SCHOOL OF COMPUTING AND DIGITAL MEDIA

AND THE COMMITTEE ON RESEARCH DEGREES

OF LONDON METROPOLITAN UNIVERSITY

IN PARTIAL FULFILMENT OF THE REQUIREMENTS

FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

Mohammad Hossein Amirhosseini

September 2018

© Copyright by Mohammad Hossein Amirhosseini 2018

All Rights Reserved

Abstract

Neuro Linguistic Programming (NLP) is a methodology used for recognition of human behavioural patterns and the modification of the behaviour. A significant part of this process is influenced by the theory of representational systems which based on the five main senses. Meta model is another important technique in this process. This technique can be adopted to allow an individual to gain a better understanding of their own issues as well as those of others. Another vital factor in NLP are Meta programs, which are habitual ways of inputting sorting and filtering the information found in the world around us. The difference in Meta programs results in significant differences in behaviour from one person to another, the type of personality can be recognised through utilising and analysing the Meta programs. There are different methods to predict the personality type based on Meta programs and Myers-Briggs Type Indicator® (MBTI) is currently considered to be one of the most popular and reliable methods. Traditionally, the application of NLP relies on consultation with a profession qualified in implementation of this technique. To circumvent the limitations in reliability of this process, attempts of automation of this technique have been carried out. These attempts aim to eliminate the effect of human error such as lack of skill and experience, inconsistency in judgement, inaccuracy or mistakes as well as the impact of personal opinion. Nonetheless, many shortcomings are integral of the methodologies adopted in these attempts. Primarily, these automations are in the format of computerisation of the NLP practice and no artificial intelligence techniques have been implemented to substitute the role of the human practitioner. Hence, improvement of reliability and accuracy remain a challenge for application of NLP, which this research aims to address using artificial intelligence

techniques such as natural language processing. The second challenge in this field is the opportunity of applying NLP to benefit a group of people in order to make NLP applicable for organisations rather than individuals alone. This research aims to create this prospect in order to extend the application of NLP for improvement of organisational performance.

The focus of this research is on the automation of the three main branches of NLP, which includes (1) identification of the preferred representational system, (2) the Meta model and (3) personality type prediction based on the Meta programs. Hence, it aims to generate an intelligent software for recognising the preferred representational system and personality type of employees as individuals and also as a group. This recognition offers organisations a specific output of information and relevant advice to improve task allocation, communication and teamwork. Moreover, this research also aims to significantly increase the efficiency, accuracy and reliability of using NLP by substituting the dependence on human judgement by an automated software. Limitations of previous computerisations of NLP are also aimed to be responded to by incorporation of artificial intelligence. To achieve these objectives, the means of analysing the behavioural pattern of individuals by software is to be explored. Moreover, the implementation of natural language processing for identifying the preferred representational system, personality type and application of the NLP Meta model during a human-computer conversation will be investigated.

To examine the function of the software and the reliability of its output, three evaluations are to be conducted. Firstly, the results of using the software is to be compared to the use of a questionnaire, which the responses to would be analysed by an experienced NLP practitioner. Both of these methods are to focus on the identification of the preferred

representational system. Secondly, the application of the Meta model in a human-computer conversation is to be compared to an NLP practitioner's analysis of the same conversation. Thirdly, the analysis of personality type is to be evaluated by comparing the use of the intelligent software to the use of a computerised questionnaire.

Natural Language Processing and machine learning techniques were used for the automation process and an intelligent software has been developed. The automation is successful in eliminating human errors, thereby the software is able to perform with a higher level accuracy, reliability and efficiency. The performance of the software has been tested and compared to the performance of humans and existing methods. Regarding the representational system identification, the results of the software are similar to an experienced NLP practitioner. However, in various parts of the process, the software responded more accurately than a human practitioner. The results of the automated Meta model have shown increased accuracy in identification of the language patterns used in conversation. The recovery of information has shown to be more efficient in the software in comparison to an NLP practitioner. Finally, the results of the software regarding the personality type prediction was highly accurate and reliable after comparing with an official MBTI questionnaire. The novel methodology created in this research will assist the NLP practitioners and psychologists to obtain an improved understanding of their clients' behavioural patterns and the associated cognitive and emotional processes. It can also facilitate the organisational performance improvement in organisations.

Acknowledgments

The words ‘thank you’ are not a strong enough statement of my gratitude. The work on this thesis has been an inspiring, often exciting, sometimes challenging, but always a thought-provoking and intriguing experience. It has been made possible by the many people who have supported me.

I wish to express my sincere gratitude to my Ph.D. supervisors. Thank you to my lead supervisor, Professor Hassan Kazemian for his guidance and support of my work. His passion for high quality research, dedication to work closely with his students, enthusiasm to challenge new research fields and methods and his kindness are the reasons he is an inspiration to me and to all of his students. I also wish to thank my second supervisor, Professor Karim Ouazzane, and my third supervisor, Doctor Chris Chandler for their support.

Thank you to all of my family and friends for their unwavering support and encouragement. I am especially thankful and appreciative of my parents, Doctor Khatereh Azimi and Doctor Hamid Amirhosseini, for their emotional and financial support as well as their continuous encouragement throughout this Ph.D. and my time in education. Thank you also to my brother and sister for their endless support and encouragement. Their happiness and their success in education has always motivated me to push and challenge myself further. A special thank you to my wife, Fatemeh Kalabizadeh. I wish to express my deepest gratitude for her constant support, understanding and love. Thank you also to her family for their kindness and unending encouragement during this Ph.D.

Publication

Amirhosseini, M.H., Kazemian, H., Ouazanne, K. and Chandler, C. (2018) 'Natural Language Processing approach to NLP Meta model automation', *IEEE World Congress on Computational Intelligence: Proceedings of the International Joint Conference on Neural Networks*, Rio de Janeiro 8-13 July, pp.186-193.

Amirhosseini, M.H. and Kazemian, H. (2019) 'Automating the process of identifying the preferred representational system in Neuro Linguistic Programming using Natural Language Processing', *Cognitive Processing Journal*, Springer. doi: 10.1007/s10339-019-00912-3

Table of Contents

| | |
|---|----|
| 1. Introduction | 1 |
| 1.1 Aim and objectives | 1 |
| 1.2 Research questions | 2 |
| 1.3 Introduction to Neuro Linguistic Programming | 3 |
| 1.4 History and development of Neuro Linguistic Programming | 4 |
| 1.5 NLP and organisational performance | 6 |
| 1.6 Background of NLP automation | 8 |
| 1.7 Research challenges | 19 |
| 1.8 Contribution of this thesis to the knowledge | 20 |
| 1.9 Overview of the thesis | 22 |
| 2. Background and literature review | 24 |
| 2.1 Introduction | 24 |
| 2.2 Definition of Neuro Linguistic Programming | 24 |
| 2.3 Representational systems..... | 27 |
| 2.4 Meta model..... | 31 |
| 2.5 Meta programs..... | 36 |
| 2.5.1 Basic Meta programs | 38 |
| 2.6 Personality types..... | 41 |

| | | |
|-------|--|----|
| 2.6.1 | Related methodologies for personality type prediction | 53 |
| 2.6.2 | Logistic Regression | 56 |
| 2.6.3 | Naïve Bayesian..... | 57 |
| 2.6.4 | Random Forest..... | 59 |
| 2.6.5 | K Nearest Neighbor (KNN)..... | 60 |
| 2.6.6 | Linear Discriminant Analysis (LDA)..... | 60 |
| 2.6.7 | Support Vector Machine (SVM) | 61 |
| 2.6.8 | Extreme Gradient Boosting | 62 |
| 2.7 | Summary..... | 63 |
| 3. | Neuro Linguistic Programming Automation methodology | 64 |
| 3.1 | Introduction | 64 |
| 3.2 | Automating the preferred representational system identification..... | 64 |
| 3.2.1 | Natural Language Processing as a tool for automation | 64 |
| 3.2.2 | Software structure and development procedure | 66 |
| 3.2.3 | Tokenization process | 70 |
| 3.2.4 | Lexical and Syntactic analysis..... | 70 |
| 3.2.5 | Comparison process..... | 73 |
| 3.3 | Automating the process of using the Meta model | 77 |
| 3.3.1 | Software structure..... | 77 |
| 3.3.2 | Programming language and the relevant library..... | 77 |

| | |
|---|-----|
| 3.3.3 Asking a question | 78 |
| 3.3.4 Defining the key words..... | 81 |
| 3.3.5 Tokenization process | 83 |
| 3.3.6 Lexical and Syntactic analysis..... | 83 |
| 3.3.7 Comparison process..... | 84 |
| 3.3.8 Decision making process | 85 |
| 3.3.9 Informing the person..... | 86 |
| 3.3.10 Repetition process..... | 87 |
| 3.4 Automating the process of personality type prediction | 89 |
| 3.4.1 Related methodologies for personality type prediction | 89 |
| 3.4.2 Dataset for training the model | 90 |
| 3.4.3 Proportionality in dataset..... | 90 |
| 3.4.4 Categorizing the type indicators in four dimensions | 95 |
| 3.4.5 Pre-processing the dataset..... | 98 |
| 3.4.6 Vectorise with count and Term Frequency-Inverse Document Frequency (TF-IDF) | 98 |
| 3.4.7 Classification task..... | 99 |
| 3.4.8 Developing Gradient Boosting model for the dataset | 99 |
| 3.4.9 Employees' MBTI personality type prediction | 106 |
| 3.4.10 Multiple choice questionnaire..... | 106 |

| | |
|--|-----|
| 3.4.11 Designing questions..... | 107 |
| 3.4.12 Comparison process for predicted personality types | 109 |
| 3.4.13 File processing procedure for predicted personality types | 109 |
| 3.5 Participants and data gathering procedure..... | 110 |
| 3.6 Data analysis procedure..... | 114 |
| 3.7 Demonstration of the software | 114 |
| 3.8 Summary..... | 122 |
| 4. Results and Discussion..... | 124 |
| 4.1 Introduction | 124 |
| 4.2 Results of automated representational system identification | 124 |
| 4.3 Results of automated Meta model system..... | 139 |
| 4.4 Results of automated personality type prediction | 148 |
| 4.5 Results of using representational system identifier in the company | 157 |
| 4.6 Results of using Meta model system in the company | 159 |
| 4.7 Summary..... | 163 |
| 5. Conclusion | 165 |
| 5.1 Conclusion of the research | 165 |
| 5.2 Recommendations and future work..... | 169 |
| References..... | 171 |
| Appendix 1 – The preferred representational system questionnaire..... | 185 |

Appendix 2 - MBTI personality type questionnaire189

List of Tables

| | | |
|------|--|-----|
| 2.1 | Color coding system for describing cognitive functions of each MBTI personality type | 44 |
| 2.2 | Description of personality types | 44 |
| 2.3 | Top five careers for each personality type | 46 |
| 2.4 | Career matches by personality type | 47 |
| 2.5 | Researches on personality type prediction and personality model used | 55 |
| 3.1 | Example of visual representational system predicates | 74 |
| 3.2 | Example of auditory representational system predicates | 74 |
| 3.3 | Example of kinaesthetic representational system predicates | 75 |
| 3.4 | Example of auditory digital representational system predicates | 75 |
| 3.5 | Specific key words to be used in different steps of the Meta model | 82 |
| 3.6 | The breakdown of the total number of people of each type in the dataset | 92 |
| 3.7 | Percentage of occurrence of each MBTI personality type in the dataset | 93 |
| 3.8 | Personality type distribution in the general population | 94 |
| 3.9 | Distribution across type indicators | 95 |
| 3.10 | Correlation efficient between personality type indicators | 97 |
| 3.11 | Accuracy of MBTI personality type prediction by software | 100 |
| 3.12 | Prediction accuracy for each MBTI binary class | 101 |

| | | |
|------|---|-----|
| 3.13 | Comparison of prediction before and after configuration | 102 |
| 3.14 | Comparison of accuracy of post classification and user classification methods for recurrent neural network | 104 |
| 3.15 | Comparison of accuracy Extreme Gradient Boosting model and recurrent neural network model | 105 |
| 4.1 | Difference between questionnaire and automated representational system identifier results | 128 |
| 4.2 | Recognised predicates by human | 130 |
| 4.3 | Recognised predicates by software | 132 |
| 4.4 | Comparing the human and software performance | 133 |
| 4.5 | Comparing the performance of the software and human | 143 |
| 4.6 | Comparing the number of recovery questions related to each deletion category asked by the software and human | 145 |
| 4.7 | Comparing the number of recovery questions related to each distortion category asked by the software and human | 146 |
| 4.8 | Comparing the number of recovery questions related to each generalization category asked by the software and human..... | 147 |
| 4.9 | Personality type frequency in organisation based on Gradient Boosting model results | 154 |

List of Figures

| | | |
|-----|--|----|
| 1.1 | VAK test | 8 |
| 1.2 | VAK test result | 9 |
| 1.3 | A sample of using Manteya in outlook | 11 |
| 1.4 | Manteya result | 12 |
| 1.5 | Manteya result | 13 |
| 1.6 | LifeSet™ questionnaire | 14 |
| 1.7 | LifeSet™ questionnaire | 15 |
| 2.1 | NLP elements | 26 |
| 2.2 | Representational systems | 27 |
| 2.3 | Reception, perception and comprehension processes | 32 |
| 2.4 | The Meta model mechanism | 33 |
| 2.5 | Personality types key | 41 |
| 2.6 | Personality types in Myers-Briggs Type Indicator® | 42 |
| 2.7 | The cognitive functions of each personality type | 43 |
| 3.1 | Development procedure of the first phase of implementation | 69 |
| 3.2 | Development procedure of the second phase of implementation | 88 |
| 3.3 | Number of occurrence for each MBTI personality type in the dataset | 91 |
| 3.4 | Percentage of occurrence for each MBTI personality type in the dataset | 92 |

| | | |
|------|---|-----|
| 3.5 | Distribution across type indicators | 96 |
| 3.6 | Pearson correlation coefficient between personality type indicators | 97 |
| 3.7 | Scoring strategy for the computerised MBTI questionnaire | 108 |
| 3.8 | Answer sheet for the preferred representational system questionnaire | 112 |
| 3.9 | Software beginning the conversation | 115 |
| 3.10 | Software detecting generalisation | 116 |
| 3.11 | Scoring strategy for the computerised MBTI questionnaire | 118 |
| 3.12 | Scoring strategy for the computerised MBTI questionnaire | 121 |
| 3.13 | Scoring strategy for the computerised MBTI questionnaire | 121 |
| 4.1 | Questionnaire results for the preferred representational system of participants | 125 |
| 4.2 | Software results for the preferred representational system of participants | 126 |
| 4.3 | Comparing the number of preferred representational systems identified by a questionnaire and software..... | 127 |
| 4.4 | Number of predicates related to each representational system identified by human | 130 |
| 4.5 | Number of predicates related to each representational system identified by software..... | 131 |
| 4.6 | Comparing the number of predicates identified by human and software | 132 |
| 4.7 | Comparing the number of visual predicates, identified by the human and software | |

| | |
|--|-----|
| | 134 |
| 4.8 Comparing the number of auditory predicates, identified by the human and software..... | 135 |
| 4.9 Comparing the number of kinaesthetic predicates, identified by the human and software | 136 |
| 4.10 Comparing the number of auditory digital predicates, identified by the human and software | 137 |
| 4.11 Number of recovery questions about deletion, distortion and generalization, asked by the software | 141 |
| 4.12 Number of recovery questions about deletion, distortion and generalization, asked by human | 141 |
| 4.13 Comparing the number of Deletion, Distortion and Generalization, identified by the human and software | 142 |
| 4.14 Comparing the number of recovery questions related to each category of deletion, distortion and generalization processes, asked by the software and human | 144 |
| 4.15 Number of personality type in the organisation based on the questionnaire results | 149 |
| 4.16 Number of personality type in the organisation based on the software results..... | 150 |
| 4.17 Comparing the number of identified personality types via questionnaire and software | 151 |

4.18 Frequency of each personality type in the company identified by the software152

4.19 Frequency of each personality type in the company identified by the software153

4.20 Most popular personality type in the company based on the system results.155

4.21 Most popular personality type in the company based on the questionnaire results.156

4.22 Number of employees in each category of preferred representational system...158

4.23 percentage of each preferred representational system in the company.158

4.24 percentage of each category of recovery questions asked b the software.160

4.25 number of recovery questions related to each category of the deletion, distortion and generalization.161

4.26 Percentage of recovery questions related to each process of the deletion, distortion and generalization.162

Glossary

| Term | Definition |
|------------------------------------|---|
| NLP | Standing for Neuro Linguistic Programming, which is a psychological approach for personality development. |
| Natural Language Processing | A subfield of artificial intelligence and linguistic, which aims to enable computers to understand the words or sentences written in human languages. |
| Representational system | Different ways that people represent or store information in their mind |
| Predicate | Sensory words in language, which can reveal the use of the related sensory modality to give an indication of an individual's preferred system of use. |
| Visual | Sensory channel for coding and storing information in mind through seeing |
| Auditory | Sensory channel for coding and storing information in mind through hearing |
| Kinesthetic | Sensory channel for coding and storing information in mind through feeling |
| Olfactory | Sensory channel for coding and storing information in mind through smelling |
| Gustatory | Sensory channel for coding and storing information in mind through tasting |
| Auditory Digital | A non-sensory system referring to how people sort experience following its occurrence focusing on self-talk, discrete |

| | |
|-----------------------|---|
| | words, facts, figures and logic. |
| Meta Model | A method for identification of language patterns to detect generalisation, distortion and deletion of information in speech with the aid of specific questions to recover the information not presented through language. |
| Deletion | Process of deleting portions of the presented concept |
| Distortion | Process of distorting portions of the presented concept |
| Generalisation | Process of generalising portions of the presented concept |
| Meta programme | Cognitive strategies and habitual ways of inputting information, sorting them and filtering the world around us. |
| MBTI | An assessment that can measure psychological preferences to understand how a person perceives the world and make decisions. It can determine the personality type of a person and based on that personality type, provides general assumptions about how that personality type is best appropriate for success in terms of careers and communication. |

1. Introduction

This chapter aims to introduce the concepts underpinning the research undertaken as well as the theories and models covered in this thesis. Firstly, the aims and objectives of this research are presented followed by the research questions attempted to be addressed. Following this, an introduction to Neuro Linguistic Programming will be presented before a brief history on the development of Neuro Linguistic Programming. The relationship between Neuro Linguistic Programming and organisational performance will also be discussed. Moreover, a review of the automation of Neuro Linguistic Programming and the previous related works will be presented. After this section, the research challenges in this field will be overviewed. This chapter will also include a synopsis of the contribution of this research to the knowledge in its field and an overview of the thesis.

1.1 Aim and objectives

This research aims to automate the three important aspects of Neuro Linguistic Programming, including (1) identification of the preferred representational system, (2) the Meta model and (3) personality type prediction based on the Meta programs, in order to be used in organisations for the purpose of organisational performance improvement. The overall objective of this research will be a comprehensive methodology for a software suite to be used in organisations in order to analyse the employees' developmental and

behavioral patterns and recommend solutions for improving the performance of the organisation's workforce. Other research objectives can be summarised as follows:

- Automating the recognition of the personality type of each employee and the most popular personality type in the organisation for application in improvement of the task allocation process in organisations
- Automating the recognition of the preferred representational system of each employee and the most popular preferred representational system in the organisation for application in improving communication and teamwork in organisations.
- Increasing the accuracy, reliability and efficiency of the current methods for personality type prediction.
- Providing a tool for employees and managers to use in order to identify any personal, communicational and organisational problems in the organisation.
- Eliminating the contributing human factors and errors such as lack of skill and experience, personal judgment and opinion, inaccuracy or mistakes of NLP practitioners, from the process of applying Neuro Linguistic Programming.

1.2 Research questions

Based on the research challenges discussed in section 1.5 and the aim and objectives discussed in section 1.6, research questions have been defined and listed as follows:

- How the current behavioural patterns of an employee or a group in an organisation can be understood using a software instead of human (a person who works as a consultant, NLP practitioner or psychologist)
- How the process of identifying the preferred representational system of a person and the most popular preferred representational system in an organisation can be completed automatically using a software, using Neuro Linguistic Programming techniques.
- How to apply the NLP Meta model automatically during a conversation between a human and computer using Natural Language Processing techniques
- How to predict the personality type of a person and the most popular personality type in an organisation using an intelligent software with improved performance in comparison to the previous automation attempts.

1.3 Introduction to Neuro Linguistic Programming

Neuro Linguistic Programming (NLP) is a powerful practical approach to personal development (Andreas and Faulkner, 1996) which emphasises on how an individual's brain connects to the surrounding world and the influence of this connection on one's behaviour (Joey and Yazdanifard, 2015). NLP techniques have been used in a variety of fields such as business, education, sales and healthcare. In addition to the influence this technique provides for an NLP practitioner in assisting clients, it can also enable

individuals to reach in and embark on personal development (O'Connor and McDrumott, 1997; Casale, 2012). Application of NLP has been deployed by well-known companies such as NASA, IBM, McDonald's and the U.S. Army (Witkowski 2010). It has also been declared that NLP is widely and often informally applied within the UK educational system (Singer and Lalich 1996). Over time NLP has become a popular technique amongst the majority of psychologists and also university employees (Tosey and Mathison 2003). Thus, success in the application of NLP has been achieved across many different disciplines, thereby increasing confidence in its utility.

1.4 History and development of Neuro Linguistic Programming

NLP was developed in the early 1970s by Bandler and Grinder. In the beginning, it was created as a methodology for modelling communication to understand why some psychotherapists are more successful than others (Janicki 2010). Bandler and Grinder synthesised a model from three different researchers: Milton Erickson, Fritz Perls and Virginia Satir (Lazarus 2010). Their main focus was on three elements. These were (1) mental processes, (2) non-verbal behaviour, and (3) language that the psychotherapists have been using (Janicki 2010). They successfully developed useful language patterns that could improve communication and as such was a starting point for other researchers to create more models for NLP (Lazarus 2010). In the next stage of NLP development, the meta-model was created in 1975. Bandler and Grinder claimed that the map of the world for a person is influenced by three elements: (1) deletion, (2) distortion, and (3)

generalisation (Oberholzer 2003). Bandler and Grinder studied the hypnotherapist Milton Erickson, who worked in the field of family therapy and medical hypnosis. They created a model named 'the Milton model' in order to replicate Erickson's results (NLP centre 2006). Further development was made by scientists such as David Gordon and Leslie Cameron who created further methods in NLP such as reframing, and anchoring. Furthermore, they could demonstrate the importance of using representational systems (Pegasus NLP 2011).

Initially NLP focused on the strategy people used to process information and how this strategy can be recognised and understood. This then developed into a collection of tools, techniques and frameworks to be used in different disciplines (Tosey and Mathison 2006). According to Tan (2003) the evolution and development of NLP can be described as a quick change from a model for therapy into a model for personal excellence and communication. Indeed, NLP was originally defined as a psychological methodology for modelling excellence (Tosey, Mathison and Michelli 2005; Tosey and Mathison 2006; Tosey 2010). More recently, however, NLP has been recognised as a science for improving communication, management and leadership skills (Tosey 2010). Lazarus (2010) also acknowledges NLP as a series of approaches, techniques and communication tools that can help people, companies and organisations to achieve their goals.

1.5 NLP and organisational performance

Competition between companies in a globalising world is difficult. Many successful companies however, acknowledge that their competitive advantage lies in human resources (Singh & Abraham, 2008). In fact, employees are being increasingly recognised as one of the most important assets of a company. There has also been an increase in awareness of the nature of human resource, changing from focus on physical skills to that of soft skills (Joey and Yazdanifard, 2015). In this climate, it is necessary for managers to be able to conceptualise what needs to be done in order to achieve the organisational objectives. It is also significant that they optimise their communication skills with their employees in order to be receptive and understanding towards their concerns. Good communication skills would also allow for managers motivational and effective in their leadership (Singh & Abraham, 2008).

The capacity of NLP proves applicable to many areas of human resources, potentiating improvement of effective communication amongst people (individuals and groups) whilst able to serve as a motivator to employees (Joey and Yazdanifard, 2015). Moreover, Biswal and Prusty (2011) points out that Neuro Linguistic Programming has uses in a much broader range of applications to increase understanding and effectiveness in communication and it can be an advantageous way of thinking about individual or group communication.

Lavan (2002) argues that “what makes the real difference between medium and high performance is the balance in values and personal unseen limiting beliefs, rather than a lack of knowledge or ability”. This suggests a need for a specific tool to address the changing environment through finding effective ways in organisations, allowing

managers to communicate in a way that can struggle some issues like lack of motivation, limited beliefs or ineffective emotional states (Oberholzer, 2013).

Furthermore, Tripathi and Tripathi (2002) believe that organisational success can be achieved through upbringing of employees. This argument proposes that organisations can have the same resources but achieve different results on the basis of the nurture of their employees (Tripathi and Tripathi, 2002). Thus, it can be suggested that the application of NLP would be positively correlated to organisational success, based on the working environment of the organisation. This as a result can have an effect on employees' performance and behaviour and can therefore, impact the organisation's efficiency. With the focus of NLP on personal excellence, it could be said that NLP can effect on organisational success via development of a companies' staff (Tripathi and Tripathi, 2002).

Another important point to consider is that even the traditional concept of organisational success contains elements of focus on the well-being of employees. This is because of the increasingly radical changes to the nature of work in many fields and the psychological demands brought about as a result of this evolution. (Abraham & Singh, 2008). Examples of this would be the increasingly digitalised mode of carrying out many tasks in the workplace or the shifts of focus or responsibilities brought about as a result of retrospective development of an organisation. This would naturally change the demands of an organisation from its employees, necessitating a route of optimising the performance of staff. The capacities of NLP in personal development and enhancement of communication therefore, present as a valuable tool for application in improvement of organisational performance via evaluation and development of employees.

1.6 Background of NLP automation

In this research, three key aspects of NLP have been utilised. These are representational systems, the Meta model and Meta programs which collectively aid in recognition of the different aspects of personality types. There have been previous attempts to automate the identification process of the preferred representational systems and Meta programs. However, they have only been as competent as a simple computerisation of the concept rather than an intelligent automation. One of the few efforts for NLP automation is an online NLP test on <http://www.vaknlp.com>. This website makes an effort to explain the relationship between the human senses and different types of personality. This test focuses on representational systems including the visual, auditory and kinesthetic types. The tool used is a collection of 10 questions (Fig 1.1) that attempt to identify the preferred representational system of the user.

www.vaknlp.com/vak.htm

1. Your main interests:
 You know fine wines.
 You are talented working with your hands.
 You prefer taking souvenir photos.

2. Your sense of humor:
 Sarcastic.
 Subtle.
 Sometimes mocking.

3. The image you want to project:
 Well spoken and trustworthy.
 Sane with balanced interests.
 Wise from life's experiences.

4. When you don't want to do anything:
 You listen to a CD or to the radio.
 You watch television or read magazines.
 You take a long, long bath.

5. In general with pets:
 You watch each other carefully.
 You pet them affectionately.
 You play with them while talking to them.

6. Your negative characteristics:
 To be a pushover.
 To be selfish.
 To be rigid.

7. Your influence:
 You are persuasive.
 You engage people's feelings.
 You present an attractive image.

8. To make you a new friend:
 You are sweet and charming.
 You are funny and make them laugh.
 You compliment them and are solicitous.

9. For you, a good film must:
 Make you dream.
 Be intelligent.
 Be a beautiful spectacle.

10. Your friends like you because:
 You entertain them.
 You help them to relax.
 They have great conversations with you.

Calculate the results Clear all

Figure 1.1: VAK test (<http://www.vaknlp.com>, 2016)

During this research, these 10 questions were tested by different users at different times. However, it was recognised that this survey is very limited in accuracy to identify the profile of the respondents. An example of this is seen below.

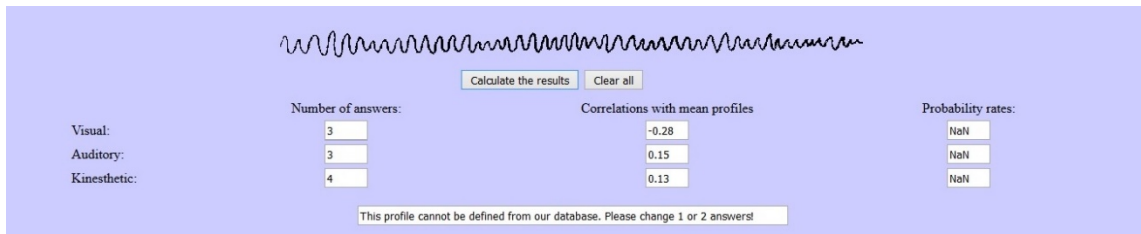


Figure 1.2: VAK test result (<http://www.vaknlp.com>, 2016)

It was recognised that this software is able to identify how many answers were related to each one of visual, auditory and kinesthetic representational systems. The user's preference is then identified based on the highest number of answers correlated with a representational system, based on associations of answers and systems previously defined for the software. Thus, it seems that for each set of numbers, a profile has been defined for the software. Its function would therefore, stem from comparing the acquired set of numbers with the defined set of numbers and followed by the display of the result as the user's preference. However, as shown in figure 1.2, if the pattern of numbers acquired from the user is different to the defined set of numbers, the system is unable to identify the user's preference. Subsequently, the system would ask the user to change some of their answers. This results in a significant impact on the accuracy and reliability of the results obtained.

Another software named ‘Manteya Email’ was also found to utilise the NLP representational systems theory. This software was created by Manteya, the online persuasion experts, and it was available on their website (Manteya, 2016). The software was also available to reach on the ANLP (The Association for Neuro Linguistic Programming) website (ANLP, 2016). This software is capable of analysing the model of writing used by individuals in their emails in order to aid them to communicate more effectively. A powerful psychological database is used for this process to understand how people build rapport through computers when they are communicating via emails (ANLP, 2016). Neil Trigger who is the founder of Manteya, claims that he utilised a unique PhD research for this application where he has found a method for analysing incoming Emails and scoring their content according to psychological methods (Manteya, 2016). According to the ANLP website (2016) “When you reply to an email, it automatically cross-references your email with the score the original sender has accumulated over time. If there is a mismatch identified, the system will inform you of what you need to change to make it more persuasive.” Hence, the software can improve interactions with anyone you communicate with in trying to improve the psychological impact of your email.

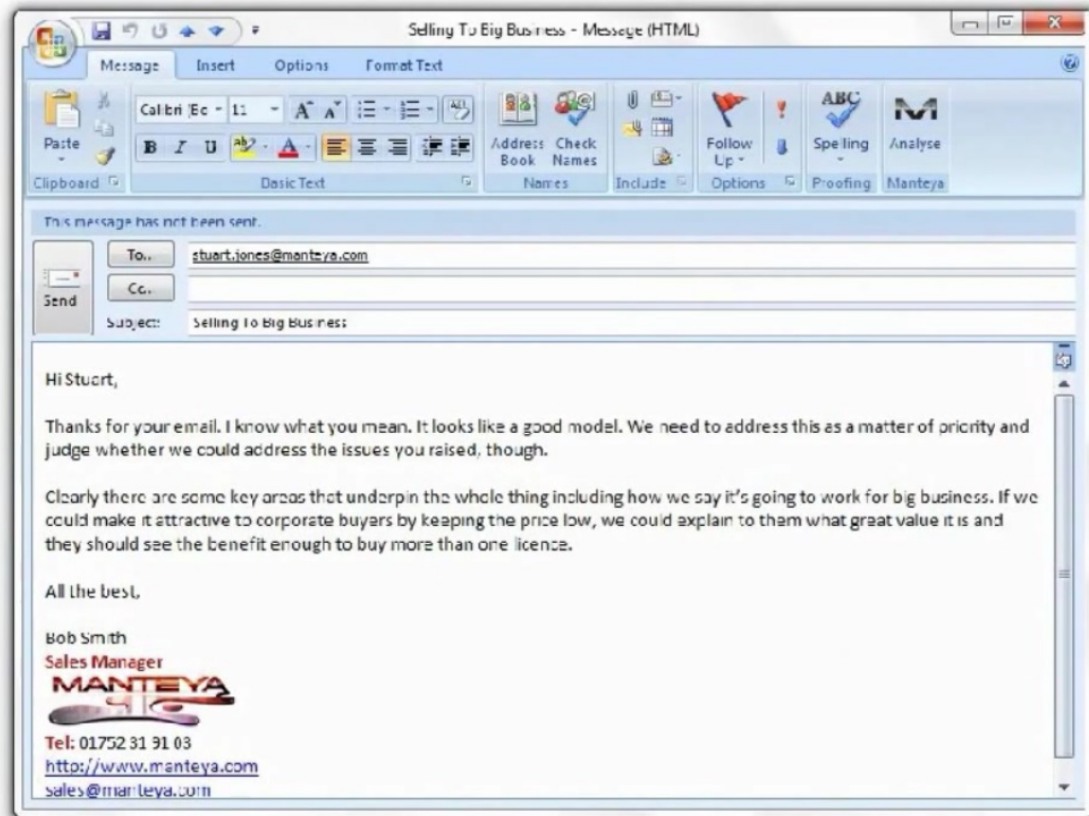


Figure 1.3: A sample of using Manteya in outlook (Youtube, 2016)

Figure 1.3 is a screen shot of a video on YouTube that describes how the Manteya email software works. In this email, the software would recognise specific words such as 'see', 'clearly', 'explain', 'looks', 'say' and etc. in the text. Then, these words will be analysed and upon clicking the Manteya button, a report about the text would be displayed as shown in Figure 1.4.

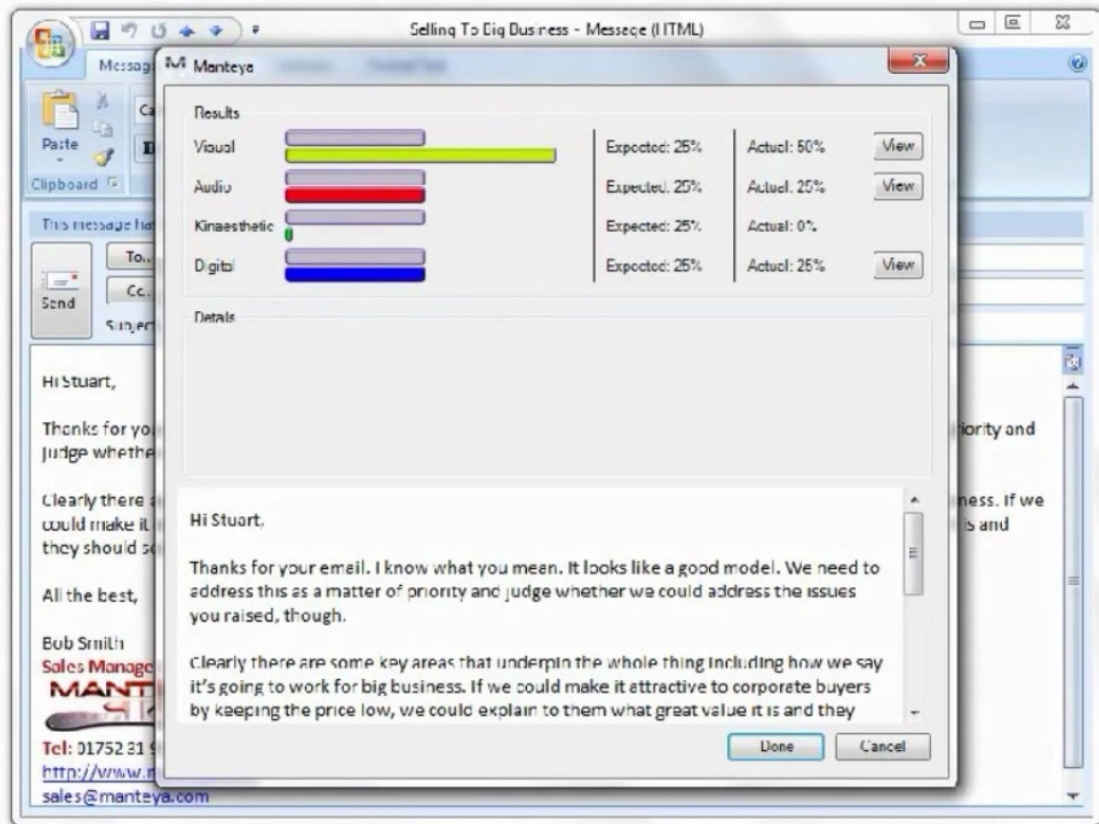


Figure 1.4: Manteya result (Youtube, 2016)

Figure 1.4 shows that the written text has more emphasis on words that are associated to use by visual people. By clicking on the view button, the software would show how you may change the text to make it more psychologically effective. Then by clicking on the update button, the software would change the text automatically replacing words with the suggested words to improve elements of the language. The result is shown in figure 1.5.

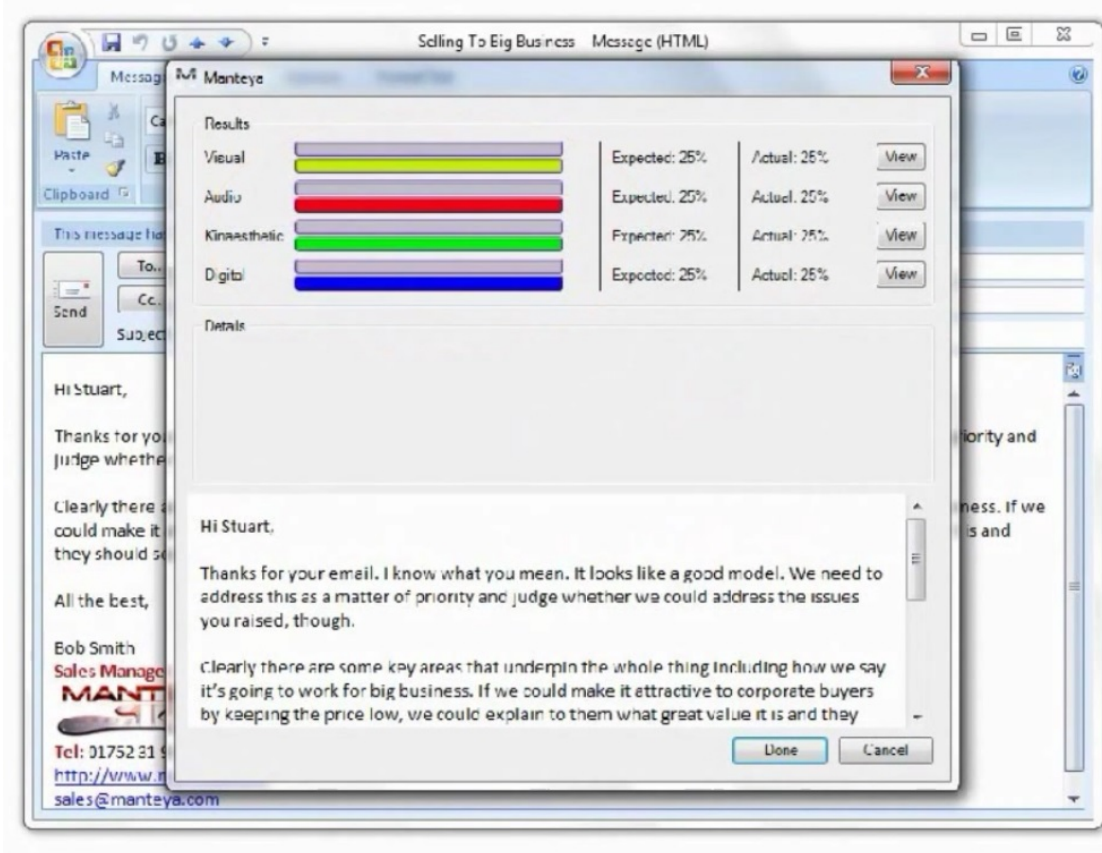


Figure 1.5: Manteya result (YouTube, 2016)

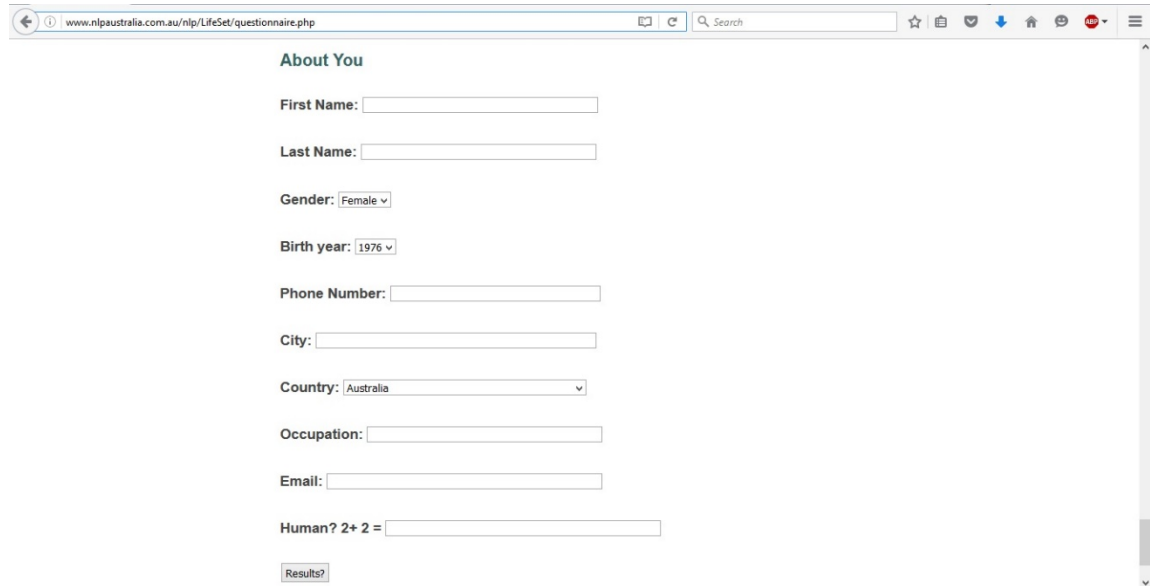
Another attempt for NLP automation was by Australia's Elite NLP Training Company. This company was founded by Terrence McClendon in 1979. After many years of training and coaching people in NLP techniques, he developed the 'LifeSet™ Meta Programs' survey, which is an online survey available on the Australia's Elite NLP Training Company's website.



Figure 1.6: LifeSet™ Questionnaire (McClendon, 2016)

McClendon utilises this tool when he is undertaking corporate training and behavioural modeling. For instance, he uses this survey to identify the characteristics of the best performers in an organisations, for example the best salesperson or the best counselor. Then, he assembles a training session in order to develop the same successful skills in others (Australia`s Elite NLP Training Company website, 2016). This survey includes 60 questions, taking 20 minutes to be completed. This survey is able to gain an insight into the characteristics of the user which may be influential to their performance. In other words, this survey is like a personality test that can identify the dominant orientations in key Meta Programs, focusing on six key Meta Programs. There are 10 questions available for identifying each one of these six key Meta programs. This online survey was completed during this research. After answering all questions and submitting

the survey, a link about the NLP Personality Profile result was sent via email for viewing or printing. Figure 1.7 show how you can submit the survey.



The image shows a web browser window displaying the LifeSet™ Questionnaire. The browser's address bar shows the URL www.nlpaustralia.com.au/nlp/LifeSet/questionnaire.php. The page title is "About You". The form contains the following fields:

- First Name:
- Last Name:
- Gender:
- Birth year:
- Phone Number:
- City:
- Country:
- Occupation:
- Email:
- Human? 2+ 2 =

At the bottom of the form is a button labeled "Results?".

Figure 1.7: LifeSet™ Questionnaire (McClendon, 2016)

The structure of this survey was simplified to the format of a multiple-choice questionnaire. As a result, there was no need for complicated coding or using machine-learning techniques. Another important aspect about this online survey is that it is only able to analyse on an individual level. Hence, its application for a group of people or an organisation, would require collection of data from employees before an NLP practitioner or a psychologist has to compare and analyse the results of individuals to reach a conclusion about the model in place. Thus, the process would not be entirely automatic nor intelligent. It can be said that this tool can only help a NLP practitioner or psychologist as an ordinary assistant.

After analysing the previous related works, it can be concluded that the existing methods are more like an online self-assessment questionnaire and where answers are based on the individual's judgment and opinion of themselves. Many also provide discrete options to be chosen from rather than allowing a more candid expression. Moreover, some services, although classified as 'automatic', do not provide immediate results. They often require answers to be sent to a NLP practitioner for analysis before the results are sent to clients. Another shortcoming of the available online surveys is their simplicity, which results in limited considerations and ultimately, reduced accuracy or error. Additionally, none of the services available uses artificial intelligence in their attempts of automation. During the literature review of this research, no windows application or online application was found with the ability of having a conversation with a human being, or in other words, the ability of computer-human interaction. It has also been understood that the process of using the Meta Model has always been considered as a face-to-face technique during conversation and there have been no attempts to automate this practice or use computers for improvement of this process.

On the other hand, the interest in automated personality prediction from social media has been significantly increased between researchers in both Natural Language Processing and Social Science fields (Nguyen et al., 2016). So far, the application of traditional personality tests has mostly been limited to clinical psychology, counseling and human resource management. However, automated personality prediction from social media has a wider application, such as social media marketing or dating applications and websites (Gjurkovic and Snajder, 2018).

Most researches on personality prediction have focused on the Big Five or MBTI personality models, which are the two most used personality models in the world. The Big Five personality model classifies personality traits in five different categories: (1) extroversion, (2) agreeableness, (3) conscientiousness, (4) neuroticism and (5) openness (Goldberg, 1990). On the other hand, MBTI which stands for the Myers-Briggs personality type indicator, classifies personality types in 16 ways via four dimensions. These are (1) introversion/extroversion, (2) sensing/intuition, (3) thinking/feeling and (4) judging/perceiving (Myers et al., 1990).

Research proposes that considering controversy about reliability and validity of these two models, the MBTI model has more applications, especially in industry and for self-discovery of personality types (Barbuto, 1997). In fact, MBTI is an assessment that can measure psychological preferences to understand how a person perceives the world and makes decisions (Gregory, 2011). It can determine what type of personality the person has and based on that personality type, provide general assumptions about how that personality type can be advantageous and drawn on for success in terms of career development and communication (Gregory, 2011).

Research on personality type prediction from textual data is scarce. However, important steps have been taken in this endeavor through machine learning. For example, there have been success in using machine learning techniques and artificial neural networks for prediction of the MBTI personality types. It has been proven that deep feed forward neural networks are useful in successfully predicting the MBTI personality types from textual datasets. Champa and Anandakumar (2010) have used applied a three layer

feed forward architecture on handwritten textual data. The success of their work proving deep neural architectures to be proficient for MBTI personality type prediction with considerable accuracy. Another successful research by Golbeck and et al (2011) presents a method of using machine learning techniques for predicting a user's personality type through publicly available Twitter profile information. Following this work, Komisin and Guinn (2012) use classical machine learning methods, including Naïve Bayes and SVM, on word choice features, using a bag of words model, were able to accurately predict MBTI personality type as well. Another venture on predicting personality types based on social media was carried out by Wan and et al (2014) where machine learning methods were used to predict the Big-Five personality types of Weibo (a chinese social network) users. In this study they conducted an inventory test with 131 users of Weibo, extracting the texts from their use of this application. In the next step, they extracted the five most relative dimensionalities by studying the relevance between personality results of users and all types of user-generated information. This led to success in predicting the personality types of users via machine learning methodologies. Similarly, a study by Tanderla et al (2017) attempted to implement deep learning architectures to predict an individual's Big-Five personality type based on information on their Facebook account. They conducted a comprehensive analysis of the accuracy of their result, comparing their model to previous research which have used older machine learning algorithms for building their models. The results of their comparison was successful in showing their model to outperform the accuracy of previous similar research. One of the most recent studies in this field was carried out by Li, Wan and Wang (2017) who focused on the use of textual information to predict personality characteristics. This

information was harvested through principal component analysis and correlation analysis. Following this, the gray prediction model, the multiple regression model and the multi-tasking model were used to successfully predict the result.

1.7 Research challenges

One of the most important issues to be mentioned about the previous automation attempts is that their methodologies were more like a simple computerisation and artificial intelligence techniques have not been used in the automation process. As a result, their obtained results were still suffering from limitations and inaccuracy. Thus, this remains as a challenge in regards to improving the accuracy and reliability of the Neuro Linguistic Programming techniques. The second key issue is that the previous attempts are focusing on an individual client and they cannot be used for a group of people or implemented in an organisation. The third challenge is that artificial intelligence techniques have been used for personality type prediction in social media but there have not been any attempt to use AI techniques for automating the identification of the personality types in organisations. This research aims to address these three main challenges.

1.8 Contribution of this research to the knowledge

This research attempts to develop a new methodology for a more competent, comprehensive process of detecting the preferred representational system as well as identifying the Meta programs and predicting the personality type. This research also intends to create a new methodology for implementing the Meta model in order to increase the success rate of this method. This is carried out by attempting to remove the limitations found in the manual and computerised services available and increasing the accuracy, reliability and efficiency of the current methods through intelligent automation. As a result, the contributing human factors and errors such as lack of skill and experience, personal judgment and opinion, inaccuracy or mistakes of psychologists and NLP practitioners are eliminated from the process. Alternatively, an intelligent system with the ability to analyse natural language is developed with the capability of acknowledging the meaning of the words, sentences and context used in order to detect the pattern of language associated with the preferred representational system as a more accurate and reliable NLP method. In addition, the software created in this research aims to provide a more effective alternative for implementation of the Meta model branch of NLP for personal development. The software is also able to recognise the Meta programs and predict the relevant personality type by using a new machine learning methodology with a better accuracy in comparison to other existing methods.

While previous accomplishments in computerised analysis have been limited to multiple-choice questions and analysis of a body of text, this study develops a means of establishing a human-computer interaction where this analysis takes place via a

conversation initiated by the software. Progression of this is entirely adapted to the answers provided by the user. This allows for a highly accurate analysis where multiple questions are asked to ensure a precise prediction.

In this research, it is recognised that classification techniques such as logistic regression, Naïve Bayes, Random forest, K Nearest neighbor (KNN), linear discriminant analysis (LDA) and Support Vector Machine (SVM) have been used for personality type prediction whereas, Extreme Gradient Boosting technique has not been used. The latter was used in this research for the first time to predict personality types based on a MBTI personality type indicator.

The developed software could also be used for analysis of a group of people instead of just one person, identifying the most common preferred representational system and the most common personality types in an organisation. While previous research has focused on social media or costumers' behavior, this research intends to focus on industry and organisational advancement. This is to be accomplished via personal development of employees as well as their analysis as a group for painting a picture of an organisation's body of staff for better targeting of management and leadership strategies. Moreover, solutions for advancement of communication across an organisation and improvement of the task allocation processes would be presented as suggestions towards improvement of organisational performance as a whole.

1.9 Overview of the thesis

This thesis includes five chapters beginning with an introduction that starts with the aims and objectives of this research followed by the research questions attempted to be addressed. This chapter continues with a brief overview of Neuro Linguistic Programming followed by the history and development of Neuro Linguistic Programming. Moreover, the relationship between Neuro Linguistic Programming and its potential for application in improvement of organisational performance is discussed. Background of Neuro Linguistic Programming automation and other previous relevant researches are mentioned in the following section. Furthermore, the research challenges in this field are overviewed. The last part of this chapter presents the contribution of this research to the knowledge in its field before a final overview of the thesis is explained.

The second chapter is a thorough literature review of Neuro Linguistic Programming including the variety of features and techniques associated with this methodology. All definitions and practices considered in this research are explained in this chapter. Beginning with the definition of NLP, this chapter moves on to an overview of the representational systems in the second section of this chapter and identification process of the preferred representational system is discussed. In the third section, another important technique of NLP, the Meta model, which has been implemented in this research, is explained and discussed. This chapter continues to review the Meta programs in the following section, assessing the basic Meta programs and the relationship between Meta programs and personality type recognition. Finally, the related methodologies for personality type prediction are discussed.

The third chapter focuses on the methodologies used in this study. First, the methodology for automating the process of identifying the preferred representational system is explained. Following this section, the methodology for automating the process of using the Meta model is explained. This chapter continues with the methodology for automating the process of personality type prediction in the next section. Finally, the data gathering procedure and the methodology for analysis of results is explained in the last section of this chapter.

Chapter 4 presents the results of this research as well as a discussion on the same basis. In the first section, the results of the automated identifier of the representational system are presented and discussed. Secondly, the results of automated Meta model applier are presented and discussed in the following section. Finally, in the third section the results of the automated personality type predictor are presented and discussed.

The fifth chapter concludes the findings of this research overviewing the implementation process and the achievements of this study. Future work and recommendations are explained at the end of in this chapter. Finally, all the references used in this research followed by appendices are available at the end of this thesis.

2. Background and Literature Review

2.1 Introduction

This chapter starts with the definition of Neuro Linguistic Programming based on the academic sources. Then, representational systems will be explained and the process of identifying the preferred representational system of a person will be discussed. Moreover, the definition and application of Meta model will be explained and different processes of applying the Meta model will be described. Furthermore, different aspects of Meta programs and the relationship between the Meta programs and personality types will be discussed. Following this section, Myers-Briggs Type Indicator® will be explained and discussed. Finally, the related methodologies for personality type prediction will be discussed.

2.2 Definition of Neuro Linguistic Programming

NLP is recognised as a collection of techniques that can help to identify the way in which people think, how they use words and language to communicate and behave. moreover, detecting patterns in people's behaviour (Transform Destiny, 2015). Andreas and Faulkner (1996) explained that 'Neuro' refers to the nervous system and the mental pathways of the five senses of hearing, sight, touch, taste, and smell. In other words, it is referring to the person's neurological system. The idea behind this, is that people use their senses to experience and interpret everything around them and it is possible to interpret

these five different senses to make them understandable for both types of thought processes that are conscious and unconscious. Thus, this thought process can have an effect on the emotions and behaviour of the person (Ready & Burton, 2010).

‘Linguistic’, on the other hand, refers to the use of language and how specific words and phrases mirror the mental scene. This word also refers to the ‘silent language’ of gestures, body language and habits that reveal further information (Andreas and Faulkner, 1996). Ready and Burton (2010) also stated that ‘linguistic’ is referring to the way that people use body language and words and how these factors can effect on the process of experiencing for a person, the conceptualisation of this experience and how it can be used to communicate with other people.

The term ‘Programming’ is borrowed from the field of computer science, to suggest that our thoughts, feelings, and actions are simply habitual programs that can be changed by upgrading the ‘mental software’ (Andreas and Faulkner, 1996). In fact, this word refers to the process of how different experiences are coded first, followed by how a person can use internal strategies and processes or specific thinking patterns to make decisions or solve problems (Ready & Burton, 2010). Figure 2.1 shows different elements of Neuro Linguistic Programming.

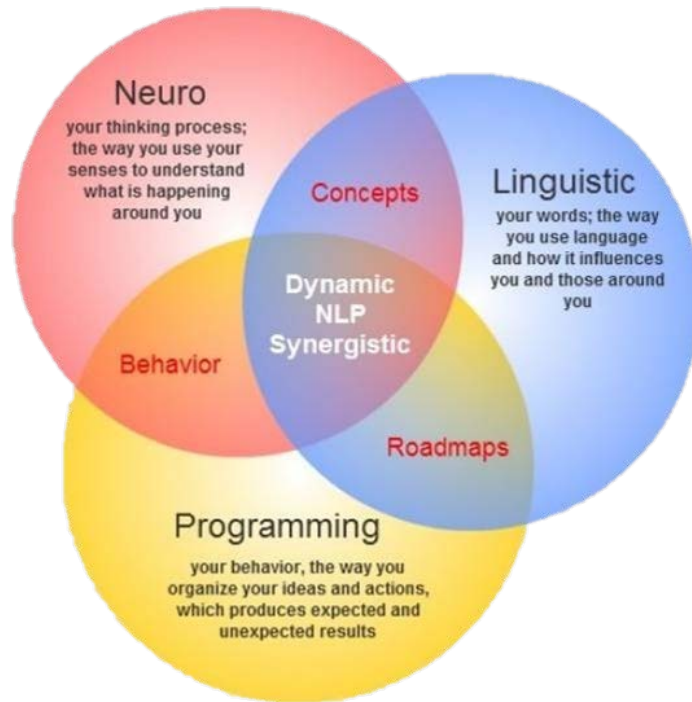


Figure2.1: NLP elements (Bryant, 2016)

It can be said that NLP is a combination of the art and science of personal excellence (O'Connor and Seymour, 1993). It is art because everyone can bring their unique personality and style to what they do, and this achievement can never be captured in words or techniques. In addition, it is Science because there is a method and process for discovering the patterns used by outstanding individuals to achieve outstanding results. In other words, NLP is the study of what makes the difference between the excellent and the average (O'Connor and Seymour, 1993). An earlier definition by Casale (2012) defined NLP as “A school of psychological techniques that effectively communicate with the listener's subconscious or unconscious mind”.

2.3 Representational Systems

There are variety of techniques included in NLP with varying steps in the personal development process. However, one of the most important stages is the identification of the preferred representational system of an individual. In the context of NLP, representational systems are the different ways that we represent or store information in our mind (Ellerton 2007). This occurs via the five main sensory modalities through which people comprehend the world, coding and storing information in their mind through seeing, hearing, feeling, tasting and smelling. Following this, they then filter this information with their beliefs and values in order to re-represent experiences to themselves and finally act on the result (O'Connor and Seymour 1993; Linder-Pelz 2010).

Figure 2.2 shows this process.

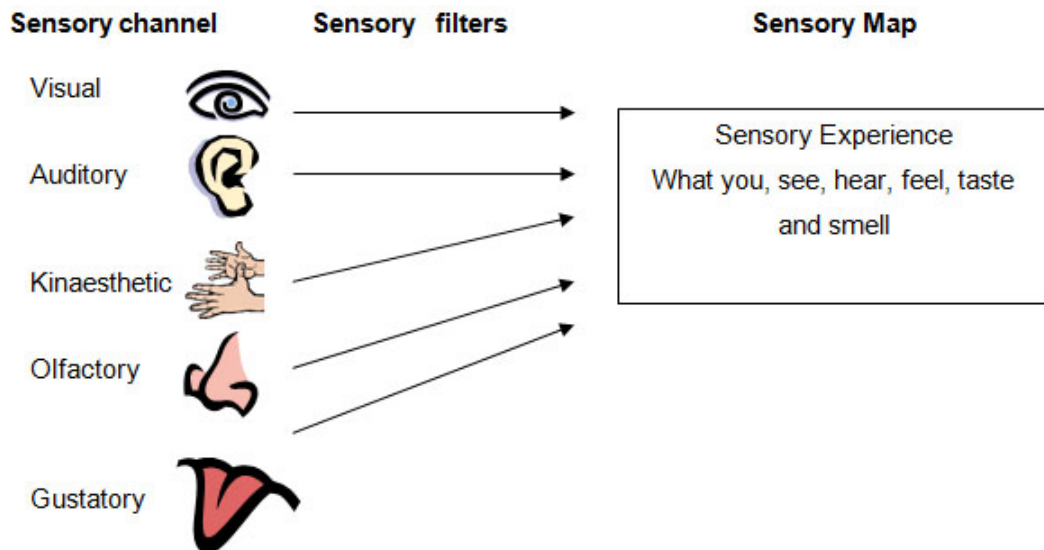


Figure 2.2: Representational systems (Caroll, 2011)

Therefore, through examining the representational systems in NLP, we can assess how the human mind processes information and interprets meanings (McAfee 2014). Palmiero, Di Matteo and Benardinelli (2014) also pointed out that how people represent conceptual knowledge is a long-debated issue and one important approach is to assume that conceptual knowledge is distributed across different attribute domains, such as vision and touch.

While people use all sensory-based representational systems as a means for learning, each person has a dominant preferred system that is used more often than others. This preferred system is conveyed through different ways in an individual's speech, learning methods, and other communicatory pathways (NLP Dynamics Ltd. 2013). There are different generalisations of characteristics, which are based on people's preferred representational systems. Hence, understanding the preferred representational system of an individual reveals a lot about likely characteristics, behavioural patterns and learning processes, which can be key to NLP modelling, and personal development processes (NLP Dynamics Ltd 2013).

Ellerton (2007) suggests that there are six representational systems in total instead of five. Five of them correspond to the main senses which are visual, auditory, kinaesthetic, olfactory and gustatory (VAKOG) and the sixth one is identified as the auditory digital representational system. This is a non-sensory system and refers to how people classify their experience following its occurrence (Monkeypuzzle training and consultancy 2016). This focuses on self-talk, discrete words, facts, figures and logic. Ellerton (2015) also proposes that people often work with three representational systems, the visual, auditory and kinaesthetic (VAK) and the two other representational systems, gustatory and

olfactory, do not play a major role and are often included within kinaesthetic. Through a NLP study, however, Rayner Institute (2015) recognised four primary representational systems in total. Smell and taste were disregarded as they are not normally a primary sense for most people, and the category of auditory digital was added. This study led NLP practitioners to confine their consideration to the VAK and auditory digital representational system when assessing the preferred representational system.

Each representational system is associated with specific tendencies of characteristics. McAfee (2014) explains that visual people usually memorise via observation of imagery and they are interested to see how a concept looks like. They are less distracted by noise and have trouble remembering long verbal instructions. In other words, what they see has priority and is more important than what they understand and experience through hearing or feeling (Monkeypuzzle training and consultancy 2016). On the other hand, Bensted (2014) discusses that auditory people typically can be easily distracted by noise. They can learn and memorise by listening and tone of voice can be very important to them. They like music and can repeat things easily. For this group, what they hear has priority and is more important than what they understand or experience through seeing or feeling (Monkeypuzzle training and consultancy 2016). People with a kinaesthetic preference memorise by doing or walking through something. They are more interested in a program that gives them a gut feeling or in something that feels right (McAfee 2014). They also respond very well to physical rewards and touching (Bensted 2014). This group loves physical activities and they are more interested in trying something out and less interested in theory (Monkeypuzzle training and consultancy 2016). Finally, people with an auditory digital preference, spend a fair amount of time talking to themselves (Bensted. 2014).

They usually memorise by steps, procedures and sequences and it is important for them to know if the program makes sense (McAfee 2014). In fact, for this group Logic is a priority and is more important than how they understand and experience through seeing, hearing or feeling (Monkeypuzzle training and consultancy 2016). In other words, they are more interested in facts and science, therefore before doing something, they need to understand it (Monkeypuzzle training and consultancy 2016).

Language is recognised as a key identifier of dependence to sensory modalities. Recognition of sensory words termed ‘predicates’ in language can reveal the use of the related sensory modality and thus give an indication of an individual’s preferred system of use. Accordingly, adapting the language used to match an individual’s, based on their preferred system, will assist them in understanding what you wish to communicate (Brefi Group Limited 2004). Thus, the preferred representational system can be recognised through analysis of the language used in conversation, considering the sentences and words used by an individual for predicates. There are defined patterns for identification of the preferred representational system by psychologists and NLP practitioners. This method, however, is unguarded against human factors such as lack of experience, personal judgment, mistakes and inaccuracy that may have a direct or indirect impact on the identification of systems.

2.4 Meta model

NLP consists of a variety of techniques and escalating levels of processes to aid personal development in clients and oneself, one of the most significant techniques being the Meta model. The Meta model is the first formal model in NLP, first described by Richard Bandler and John Grinder in the first edition of their book, 'Structure of Magic' published in 1975. They had observed the use of certain language patterns and essential questions by successful therapists that enabled them to correctly and efficiently identify and address the issues of their clients. The Meta model is now established as the identification of language patterns to detect generalisation, distortion and deletion of information in speech with the aid of specific questions to recover the information not presented through language (Bandler and Grinder, 1975; Freeth, 2016). As people speak about a problem or a situation, the words that they choose, may distort, delete and generalise portions of the presented concept. Thus, by considering these language patterns, the information concealed behind the words can be identified and recovered (Bandler and Grinder, 1975).

The discrepancy in the information presented by language was in fact, identified to be rooted in the processing of information acquired through the senses. It has been recognised that the nervous system uses deletion, distortion and generalisation of the raw sensory input in order to process reality more readily and into a more manageable version (Davis, 2015). Fig. 2.3 shows how information input may be developed through this process.

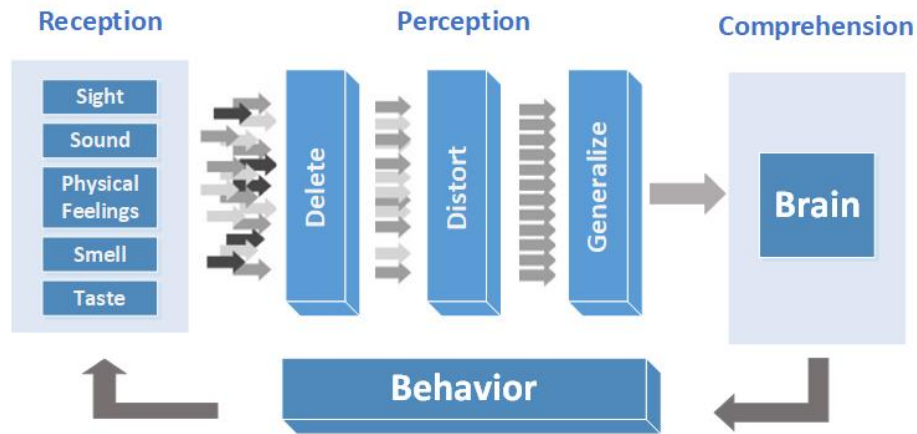


Figure 2.3: Reception, perception and comprehension processes

Deletion refers to the portions of the mental map, which do not appear in the verbal expression due to being eliminated. These gaps of information are recognised by the NLP practitioner and retrieved in conversation (Freeth, 2016; Carroll, 2016). Distortion, on the other hand, is about alteration of the information from its initial form. Upon detection, this is explored in conversation and the original information is recovered (Freeth, 2016). Carroll (2016) defines distortion as “the process of representing parts of the model differently to how they were originally represented.” Finally, generalisation is about simplification of information through which concepts may be merged. The practitioner then retrieves lost information by prompting the client to become progressively more specific throughout the conversation (Freeth, 2016). Fig. 2.4 shows how the Meta model deals with these processes.

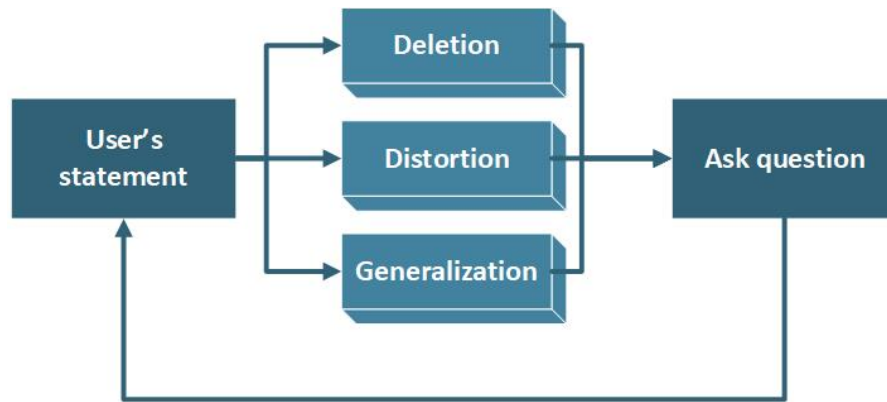


Figure 2.4: The Meta model mechanism

The main focus in identifying the process of deletion has five important elements. These are (1) unspecified nouns, (2) simple deletions, (3) comparative deletions, (4) unspecified verbs and (5) ‘Ly’ adverbs (Davis, 2015). Nouns included in a sentence, which are not specifically referred to in the statement, can be categorised as unspecified nouns. The missing information may be deleted completely or it may be replaced with an unspecified pronoun (Davis, 2015). On the other hand, simple deletions, refer to the missing elements of a sentence, which are key to the statement being made. In the case of unspecified nouns, the sentence has an object which is merely unspecified. Whereas in simple deletion, it would be the case of information missing entirely (O’Brien, 2009). Davis (2015) mentions that “simple deletions are where part of the meaning is left out or lost and you can notice them in a sentence with ‘it’ and ‘that’ and also when referring to missing descriptions (adjectives).” Comparative deletion happens when the person uses hypnotic words to make a comparison, but does not explain what is being compared and hence it is left unstated (O’Brien, 2009). Unspecified verbs are verbs that neither describe the action completely, nor are they fully informative with regards to the statement. In this

case, one may fill in the gap with their own experience. This process is called ‘mind reading’ (Elston, 2017). Finally, ‘Ly’ adverbs are words that end with ‘Ly’ such as ‘slowly’ or ‘creatively’. Stoker (2014) points out that the problem with ‘Ly’ adverbs is that they present a judgment, which tends to be accepted by other people without questioning whether it is true or not. This may cause problems as people may dismiss the validity of the judgment being passed, thus accepting it as the truth and inquiring no further.

The second major process in the Meta model is distortion, which focuses on five important language patterns (1) mind reading, (2) lost performative, (3) cause effect pattern, (4) complex equivalence and (5) linguistic Presuppositions (Carroll, 2016). Mind reading occurs when a person assumes that they know what the other person is thinking or feeling without confirming with the individual. This can lead a person to take an action or withhold from an action, because they think that they know how the other person is likely to react (Freeth, 2016). Lost performatives happen when a person presents a personal belief as a universal truth, which can lead people to accept that belief as the truth without questioning it (Davis, 2015). Cause effect patterns, on the other hand, imply a relationship with time. This suggests that in the case of one event taking place, a second event will automatically follow (Freeth, 2016; Elston, 2017). The fourth important language pattern in distortion is complex equivalence, which takes place when two experiences, ideas, objects or their meanings are interpreted as being synonymous (Carroll, 2016). Finally, the fifth language pattern in distortion is linguistic presupposition, which is one of the most powerful aspects of the Meta model language patterns. It refers to statements where unstated elements are assumed to be true, in order for the statement to make sense (Hoag, 2017). Hence, these statements are simply presuppositions.

Linguistic presuppositions are categorised into four groups. These are (1) linguistic presuppositions of awareness, (2) linguistic presuppositions of time, (3) linguistic presuppositions of order and (4) adverbs and adjectives (Davis, 2015).

The third phase of the Meta model is generalisation. Identification of this process consists of two important elements, (1) modal operators and (2) universal quantifiers (Avery, 2015). Modal operators refer to a person's feelings regarding carrying out a task. Examples of this would be the difference in a person's mood regarding a task they enjoy doing and a task they have to do regardless (Davis, 2015). Modal operators are categorized into two groups; (1) necessity and (2) possibility. Modal operators of necessity define rules that must be followed and there are undefined consequences in case of breaking these rules. On the other hand, modal operators of possibility reduce flexibility by creating limits on what can or cannot be done and thus define arbitrary barriers (Avery, 2015). Universal quantifiers are another type of generalisation which takes a single case or situation and apply it to all other cases at all times (Freeth, 2016).

There are defined outlines for psychologists and NLP practitioners for using the Meta model during counselling or therapy. Many human limiting factors are bound to contribute to this process, such as lack of experience or skill, personal judgment, and inaccuracy, which may have a direct or indirect impact on the outcome of the Meta model.

2.5 Meta Programmes

One of the most important factors in NLP are Meta programmes. Brian (2013) explained that Meta programmes are cognitive strategies that a person runs all the time and they are different ways that a person can sort information. Davis (2015) also stated that Meta programmes are habitual ways of inputting information, sorting them and filtering the world around us. In other words, they are our thinking styles or typical strategies and patterns. According to Ellerton (2004) Meta programmes act like filters that determine how people perceive the world around them. It can also have a major influence on behaviours as well as how people communicate with others. Furthermore, Meta programmes can be considered as deep-rooted mental programs, which are able to filter our experiences automatically and guide our thought processes. As a result, this leads to significant differences in behaviour from one person to another (Ellerton, 2004).

In the early stage of NLP development, Meta programmes emerged when Leslie Cameron Handler and Richard Bandler collaborated together (Hall and Bodenhamer, 1997). Leslie's initial work focused on 'textbook NLP' and in this process discovered that the NLP processes do not always work (James and Woodsmall, 1988). Eventually, she and Richard Bandler discovered that people use different strategies for doing different things (Hoag, 2017). For instance, they use different strategies when they want to make a decision or when becoming convinced about something. As a result, the initial list of NLP Meta programmes was presented by Leslie and Bandler (James and Woodsmall, 1988) and they identified about 60 different patterns (Ellerton, 2004). Many of these Meta programmes have been combined together by subsequent researchers to form a much

smaller and more useful set (Ellerton, 2004). There are different sets of Meta programmes introduced by different researchers and the number of Meta programmes in each set and their descriptions of the patterns are slightly different (Ellerton, 2014). However, Davis (2015) suggests there is no real set or list of Meta programmes as they are always evolving and the names change and their usefulness may also change with the context.

First, Leslie codified the initial list of Meta programmes for therapeutic use (Davis, 2015). Then Roger Bailey and Ross Stewart developed them for use in business (James and Woodsmall, 1988) and Bailey created a profiling instrument named 'LAB profile' which stands for the language and behavior profile. (Charvet, 1997). According to Charvet (1997), "LAB profile is a way of thinking about people and groups that allows us to notice and respond with just the right Influencing Language." It is structured and tailored to each situation to allow us to understand how people process information, how they get motivated and how they make decisions. In fact, it is a set of questions that can be used as a formal survey for a group of people or it can be fed into a casual conversation (Success Strategies Company, 2016). When you ask a question from someone, even if the person answers the question indirectly, he or she reveals a pattern. LAB profile can teach us to pay attention to how people talk when they respond to questions instead of what they talk about (Success Strategies Company, 2016).

Bailey also reduced the number of patterns from 60 to 14 in order to make detecting and using these patterns simpler (Charvet, 1997). Following this, Woodsmall developed Meta programmes for use in business and therapy and integrated them with Myers-Briggs Personality Inventory (Hall and Bodenhamer, 1997). The results were published in a book named 'Time Line Therapy and The Basis of Personality' in 1988.

He reduced the number of patterns again and made a smaller set of Meta programmes, which includes only four basic and key Meta programmes. These four Basic Meta programmes, also known as the MBTI that is Abbreviation for Myers-Briggs Type Indicator[®], describe the preferences of an individual in four dimensions and these basic dimensions combine into one of 16 different Personality Types (Mind academy, 2014).

2.5.1 Basic Meta Programmes

According to Fretwell, Lewis and Hannay (2013) these four dimensions or basic Meta programme are Extroversion-Introversion (E-I), Sensation-Intuition (S-N), Thinking-Feeling (T-F), and Judgment-Perception (J-P). The first Meta programme (Extroversion-Introversion) is regarding external behaviours (NLP World Ltd, 2017) and reflects where people prefer to focus their attention (Fretwell, Lewis and Hannay, 2013). Eagle (2017) explains that extroverts typically prefer to work with other people rather than systems and machines. They usually think aloud and like to be around people and they may be found in the centre of attention. When their energy is low, the best way to boost their mood is to be around other people. Fretwell, Lewis and Hannay (2013) also mentions that extroverts obtain their energy from other people. They work very fast and can act quickly and sometimes even without thinking. They can communicate easily with other people and prefer oral communication. This group of people do not like complicated procedures and also are not patient with slow time-consuming jobs. On the other hand, introverts typically prefer to be quiet and to work with systems and machines rather than people. They try to

think before they speak and they are usually shy. When they do not have enough energy, they will generally spend time by themselves in order to regain their energy (Eagle, 2017). They are more interested in the inner world of experiences including ideas and concept. They are usually very thoughtful and their energy comes from ideas and thoughts. Introverts are generally careful with details and are able to concentrate on one thing or one project for a long period of time. In contrast to extroverts, introverts prefer written communication and work better alone (Fretwell, Lewis and Hannay, 2013).

The second Meta programme (Sensation-Intuition) is about internal processes (NLP World Ltd, 2017). It reflects how individuals obtain information and how they understand the world around them (Fretwell, Lewis and Hannay, 2013). Sensors typically rely on experiences and are good with timing. They are able to recognise patterns and quickly respond to those patterns (Eagle, 2017). They usually prefer concrete details and rely on their five senses to observe what is happening (Myers & McCaulley, 1989; Fretwell, Lewis and Hannay, 2013). On the other hand, intuitors are great with concepts and ideas. They enjoy dreaming and creative processes (Eagle, 2017). They enjoy unusual things, change and novelty. They are practical and realistic (Myers & McCaulley, 1989).

The third Meta programme (Thinking-Feeling) is about internal state (NLP World Ltd, 2017) and reflects the preference of an individual for processing data, making decisions and analysing their understanding of different things (Fretwell, Lewis and Hannay, 2013). Thinkers are logical and sometimes may be seen as cold and spiritless as they are not associated with feelings (Eagle, 2017). People with 'thinking' preference are interested in logic and facts and they feel uncomfortable when dealing with the feelings

of others (Fretwell, Lewis and Hannay, 2013). On the other hand, people with ‘feeling’ preference may be seen as being too emotional. When making a decision, they usually refer to their feelings instead of their mind (Eagle, 2017). They typically use personal or social values when they want to make a decision and they consider the impact of their decision on other people. They also enjoy pleasing other people and offering appreciation or sympathy (Fretwell, Lewis and Hannay, 2013).

Finally, the fourth Meta programme (Judgment-Perception) is about releasing and controlling (NLP World Ltd, 2017) and people with this preference prefer to organise themselves to the outside world (Fretwell, Lewis and Hannay, 2013). Perceivers try to change themselves to fit in, when something is not right. They are flexible, relaxed, and open to new ideas. They are easy-going and go along with what others are doing. People with perception preference prefer to adapt situations rather than control them (Eagle, 2017; Fretwell, Lewis and Hannay, 2013). On the other hand, judges usually try to change others to behave and think like them. They feel uncomfortable if they are unaware of what is going to happen. They are more interested in leading, organised situations and an orderly life. Judges are punctual and prefer to control their life through detailed planning (Eagle, 2017; Fretwell, Lewis and Hannay, 2013).

As discussed above, the preferences of an individual are categorised in four dimensions and each dimension is representing two types of personalities. Figure 2.5 shows 8 personality types key used in Myers-Briggs Type Indicator®.



Figure 2.5: Personality types key (Tieger, Barron and Tieger, 2007)

2.6 Personality types

Personality is derived from the Latin word *persona*, which means describing the behaviour, character, or a private person (Darsana, 2013). It has been discussed that the meaning of personality is reflected in the very nature of the attitude of a person that can be distinguished from other people (Alwi et al, 2003). Personality according to Allport (Hall and Lindzey, 2005) is a dynamic organisation within the individual as a psychophysical system. It determines the unique way in which an individual adapts to an environment. Personality is a description of the individual's self-image that influences

their behaviour uniquely and dynamically, it is because the behaviour may change through the process of learning, experience, education, etc. This opinion clarifies the opinion by Setiadi (2003) that personality is the dynamic organisation of the system that uniquely determines the individual's adjustment to the environment.

As it was discussed above, the preferences of an individual are categorised in four dimensions and different combinations of the personality type's key in these categories represents 16 different personality types based on Myers-Briggs Type Indicator®. Figure 2.6 shows these 16 personality types that result from the interactions among the preferences of an individual.

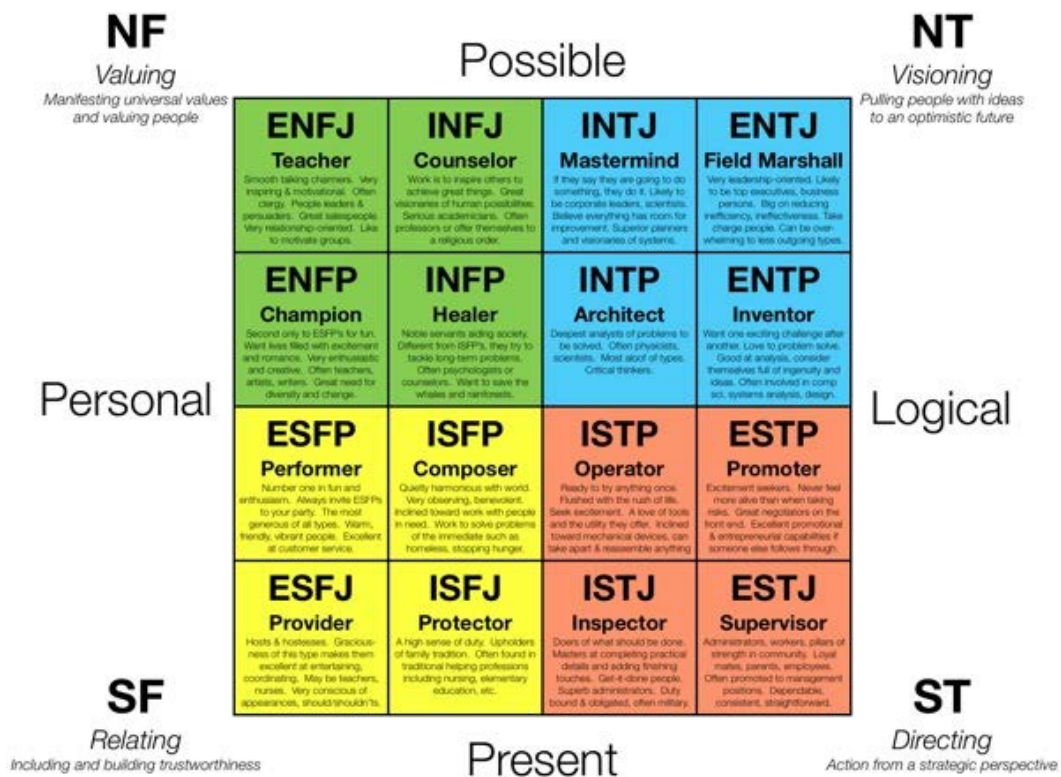


Figure 2.6: Personality types in Myers-Briggs Type Indicator® (Tieger, Barron and Tieger, 2007)

As a result, the Myers-Briggs Type Indicator® (MBTI®) personality inventory has been used in this research in order to understand the type of personality and suggest the most appropriate job position for the identified personality. The most popular Meta programmes and personality types will be identified as well, and the current organisational culture and task allocation can be modified based on this information. Each key word in figure 2.6 represents a specific type of personality and figure 2.7 describes the cognitive functions of each MBTI personality type. The background colour of each type represents its dominant function and the colour of the text represents its auxiliary function.

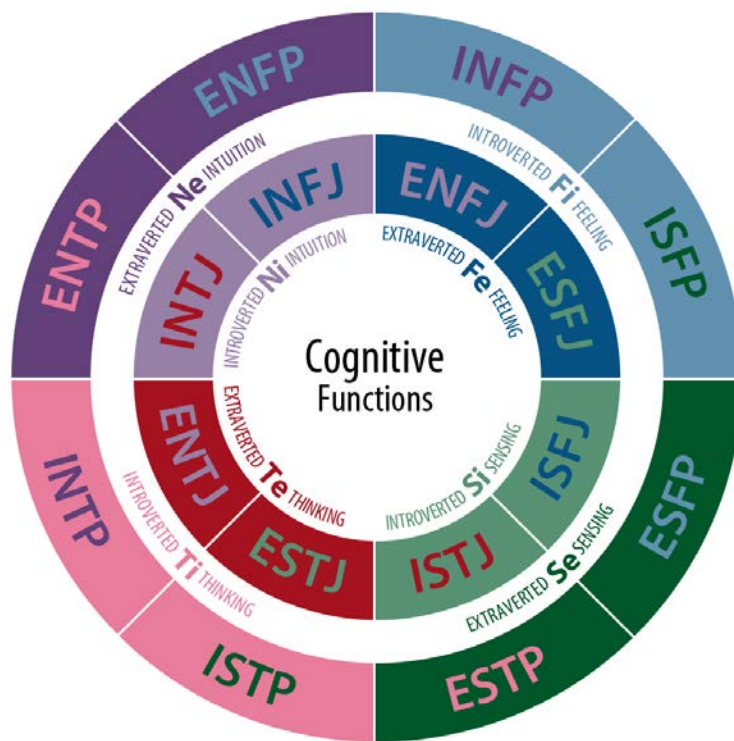


Figure 2.7: The cognitive functions of each personality type (Beech, 2013)

The color coding system in figure 2.7 is explained in table 2.1.

| Color | Meaning |
|--------------|-----------------------|
| Dark purple | Extraverted intuition |
| Light purple | Introverted intuition |
| Dark green | Extraverted sensing |
| Light green | Introverted sensing |
| Dark blue | Extraverted feeling |
| Light blue | Introverted feeling |
| Pink | Introverted thinking |
| Red | Extraverted thinking |

Table 2.1: Color coding system for describing cognitive functions of each MBTI personality type

Furthermore, table 2.2 shows the description of each personality type.

| Personality type | Characteristics |
|------------------|---|
| ENFJ | People lovers who are energetic, articulate and diplomatic. They excel in cooperative roles that require them to be expressive and logical. |
| INFJ | Thoughtful, creative people driven by firm principals and personal integrity. They do well in behind-the-scenes roles that require them to communicate on a personal level. |
| INTJ | Creative perfectionists who prefer to do thing their own way. They perform well in non-social roles that require them to think theoretically. |

| | |
|-------------|---|
| ENTJ | Natural leaders who are logical, analytical and good strategic planners. They gravitate toward authoritarian roles that require them to be organised and efficient. |
| ENFP | Curious and confident creative type who see possibilities everywhere. They perform well in expressive roles that require them to be alert and communicative. |
| INFP | Sensitive idealists motivated by their deeper personal values. They excel in roles that require them to be compassionate and adaptable. |
| INTP | Independent and creative problem-solvers. They gravitate toward roles that require them to be theoretical and precise. |
| ENTP | Enterprising creative people who enjoy new challenges. They excel in risky roles that requires them to be persistent and non-conformist. |
| ESFP | Lively and playful people who value common sense. They gravitate toward roles that require them to be expressive and interact with others. |
| ISFP | Warm and sensitive types who like to help people in tangible ways. They do well in roles that require them to be sympathetic and attentive. |
| ISTP | Straightforward and honest people who prefer action to conversation. Perform well in roles that require them to make use of tools. |
| ESTP | Pragmatists who love excitement and excel in crisis. They excel in high-stakes roles that require them to be resourceful. |
| ESFJ | Gregarious traditionalists motivated to help others. They gravitate toward social roles that require them to care for the well-being of others. |
| ISFJ | Modest and determined workers who enjoy helping others. They do well in roles that require them to provide services to others without being in a position of authority. |
| ISTJ | Hard workers who value their responsibilities and commitments. They excel in behind-the-scenes roles that require them to be reliable. |

| | |
|-------------|---|
| ESTJ | Realist who are quick to make practical decisions. They perform well in social roles that require them to lead. |
|-------------|---|

Table 2.2: Description of personality types (Tieger, Barron and Tieger, 2007)

The Myers-Briggs Type Indicator® (MBTI®) also suggest the best careers, which are suitable for each one of these personality types. Table 2.3 shows the top five careers for each personality type.

| Personality Type | Appropriate Jobs |
|-------------------------|--|
| ESTJ | Insurance sales agent, Pharmacist, Lawyer, Judge, Project manager |
| ISTJ | Auditor, Accountant, Chief financial officer, Web development engineer, Government employee |
| ESFJ | Sales representative, Nurse/Healthcare worker, Social worker, PR account executive, Loan officer |
| ISFJ | Dentist, Elementary school teacher, Librarian, Franchise owner, Customer service representative |
| ESTP | Detective, Banker, Investor, Entertainment agent, Sports coach |
| ISTP | Civil engineer, Economist, Pilot, Data communications analyst, Emergency room physician |
| ESFP | Child welfare counselor, Primary care physician, Actor, Interior designer, Environmental scientist |
| ISFP | Fashion designer, Physical therapist, Massage therapist, Landscape architect, Storekeeper |
| ENTJ | Executive, Lawyer, Market research analyst, Management/Business consultant, Venture capitalist |

| | |
|-------------|---|
| INTJ | Investment banker, Personal financial advisor, Software developer, Economist, Executive |
| ENFJ | Advertising executive, Public relation specialist, Corporate coach/trainer, Sales manager, Employment specialist/HR professional |
| INFJ | Therapist/Mental health counselor, Social worker, HR diversity manager, Organisational development consultant, Customer relations manager |
| ENTP | Entrepreneur, Real estate developer, Advertising creative director, Marketing director, Politician/Political consultant |
| INTP | Computer programmer/Software engineer, Financial analyst, Architect, College professor, Economist |
| ENFP | Journalist, Advertising creative director, Consultant, Restaurateur, Event planner |
| INFP | Graphic designer, Psychologist/Therapist, Writer/Editor, Physical therapist, HR development trainer |

Table 2.3: Top five careers for each personality type (Tieger, Barron and Tieger, 2007)

Table 2.4 shows some appropriate careers for each personality type and in this table the variety of jobs is more than table 2.3.

| Personality Type | Appropriate Jobs | | | |
|-----------------------------|-------------------------|------------------|-----------------|-------------|
| ESTJ (Overseers) | Accountant | Engineer | Judge | Underwriter |
| | Senior Manager | Police Officer | Insurance Agent | Professor |
| | Auditor | Business | Economist | Lecturer |
| | Military Officer | Administrator | Teacher | Manager |
| | Banker | Business Analyst | Editor | Librarian |
| | Financial Officer | Computer | Technical | Researcher |

| | | | | |
|------------------------------------|---|---|--|--|
| | Government Worker | Specialist Detective Scientist | Specialist Nursing Administrator | |
| ISTJ (Examiners) | Accountant Military Officer Police Officer Administrator Auditor Medical Doctor | Dentist Financial Officer Detective Scientist Math Teacher Lawyer/Attorney | Judge Manager Computer Programmer Computer Specialist Librarian Executive | Steelworker Electrician Mechanical Engineer Systems Analyst Technical Specialist Technician |
| ESFJ (Supporters) | Accountant Bookkeeper Child Care Church Worker | Human Resources Counselor Family Doctor Homemaker | Nurse Teacher Social Worker Office Manager | Administrator Speech Pathologist Organization Leader Receptionist |
| ISFJ (Defenders) | Administrator Career Counselor Child Care Police Officer Church Worker Clerical Supervisor Counselor Medical Doctor | Social Worker Actor/Actress Counselor Human Resources Medical Technologist Office Manager Shopkeeper Researcher | Scientist Senior Manager Early Childhood Development Librarian Nurse Auditor Bookkeeper Business Analyst | Health Service Homemaker Writer Military Accountant Administrative Assistant |
| ESTP (Persuaders) | Carpenter Craftsman Paramedic/EMT | Detective Driver Firefighter | Farmer Comedian IT Support | Marketer Project Manager |

| | | | | |
|--------------------------------|---------------------|---------------------|--------------------------|------------------------------------|
| | Police Officer | Military | Entrepreneur | Computer Technician Sales Agent |
| ISTP (Craftsmen) | Athlete | Firefighter | Engineer | Project Manager |
| | Carpenter | Forensic | Steelworker | Sales Agent |
| | Construction Worker | Pathologist | Transportation Operative | Scientist |
| | Mechanic | Pilot | Entrepreneur | Systems Analyst |
| | Military | Police Officer | Human Resources | Technical Specialist |
| | Motorcyclist | Probation Officer | Marketer | Business Analyst |
| | Paramedic/EMT | Detective | Computer Programmer | |
| | Farmer | Driver | Computer Specialist | |
| | | Electrical Engineer | | |
| ESFP (Entertainers) | Athlete | Comedian | Receptionist | Teacher |
| | Artist | Interior Decorator | Supervisor | Journalist |
| | Actor/Actress | Marketer | Broadcaster/News anchor | Shopkeeper |
| | Coach | Musician | | Animal Trainer |
| | Fashion Designer | Painter | Pediatrician | Firefighter |
| | Entrepreneur | Performer | Singer | Dancer |
| | Social Worker | Photographer | Nurse | Public Speaker |
| | Recreation Worker | Public Relations | Manager | Events Coordinator |
| | | | Sales Agent | |
| ISFP (Artists) | Artist | Clerical | Early Childhood | Mechanic |
| | Carpenter | Supervisor | Development | Writer |
| | Teacher | Composer | Park Ranger | Personal Service Worker |
| | Chef | Counselor | Librarian | |
| | Child Care | Social Worker | Editor | X-Ray Technician |

| | | | | |
|---------------------------|------------------|--------------------|--------------------|-----------------|
| | Pediatrician | Designer | Musician | Medical Staff |
| | Church Worker | Physical Therapist | Naturalist | Dental Staff |
| | | Veterinarian | Nurse | Homemaker |
| ENTJ (Chiefs) | Business | Politician | Scientist | Professor |
| | Administrator | Computer | Banker | |
| | Executive | Consultant | Systems Analyst | |
| | Entrepreneur | Lawyer/Attorney | Teacher | |
| | Manager | Judge | | |
| INTJ (Strategists) | Systems Analyst | Corporate | Professor | Judge |
| | Computer | Strategist | Psychologist | Teacher |
| | Programmer | Researcher | Business | Organization |
| | Entrepreneur | Scientist | Administrator | Founder |
| | Computer | Medical Doctor | Military Officer | Dentist |
| | Specialist | Lawyer/Attorney | Manager | |
| | Engineer | | | |
| ENFJ (Mentors) | Teacher | Church Worker | Artist | Designer |
| | Professor | Trainer | Editor | Engineer |
| | Social Worker | Human Resources | Administrator | Occupational |
| | Career Counselor | Librarian | Actor/Actress | Therapist |
| | Counselor | Writer | Events Coordinator | Politician |
| | Therapist | Nurse | Facilitator | Project Manager |
| | Psychologist | Diplomat | Homemaker | Sales Agent |
| | Psychiatrist | Consultant | Musician | Senior Manager |
| INFJ (Confidants) | Counselor | Psychologist | Teacher | Actor/Actress |
| | Clergy | Social Worker | Musician | Medical Doctor |
| | Therapist | Educational | Photographer | Alternative |
| | Writer | Consultant | Professor | Medicine |
| | Missionary | Child Care | Librarian | Designer |

| | | | | |
|-------------------------------------|---------------------|-------------------|------------------------|-------------------------|
| | Psychiatrist | Child Development | Artist | Human Resources Trainer |
| ENTP (Originators) | Engineer | Journalist | Psychologist | Photographer |
| | Computer Analyst | Lawyer/Attorney | Consultant | Sales Agent |
| | Computer Programmer | Marketer | Scientist | Comedian |
| | Entrepreneur | Politician | Systems Analyst | Actor/Actress |
| | Inventor | Psychiatrist | Writer | Musician |
| INTP (Engineers) | Professor | Systems Analyst | Investigator | Legal Mediator |
| | Scientist | Computer | Teacher | Project Manager |
| | Archaeologist | Animator | Logician | Financial Planner |
| | Architect | Computer | Marketer | Banker |
| | Mathematician | Programmer | Translator/Interpreter | Researcher |
| | Artist | Computer | er | Consultant |
| | Technical Writer | Specialist | Musician | Photographer |
| | Philosopher | Economist | Network Specialist | |
| | Strategic Planner | Engineer | Judge | |
| | Inventor | Lawyer/Attorney | | |
| | Psychiatrist | | | |
| | Psychologist | | | |
| ENFP (Advocates) | Journalist | Editor | Psychologist | Entrepreneur |
| | Social Worker | Broadcaster/News | Art Director | Dietitian/Nutritionist |
| | Conference Planner | caster | Project Manager | |
| | Diplomat | Teacher | Artist | Human Resources |
| | Writer | Professor | Painter | Consultant |
| | Nurse | Musician | Actor/Actress | Marketer |
| | Public Relations | Lawyer/Attorney | Merchandise Planner | Occupational Therapist |
| | Church Worker | | | |

| | | | | |
|----------------------------------|--------------------|---------------|------------------------|---|
| | Politician | Counselor | Career Counselor | Social Scientist |
| | Housing Director | Therapist | Speech Pathologist | |
| INFP (Dreamers) | Activist | Teacher | Filmmaker | Web Designer |
| | Church Worker | Musician | Librarian | Photographer |
| | Missionary | Therapist | Editor | Human Resources |
| | Educational | Actor/Actress | Graphic Designer | Video Editor |
| | Consultant | Artist | Writer | Holistic Health |
| | Social Worker | Journalist | Psychologist | Practitioner |
| | Physical Therapist | Professor | Translator/Interpreter | Speech Pathologist |
| | Fashion Designer | Minister | er | Employee |
| | Counsellor | | Social Scientist | Development Specialist Researcher |

Table 2.4: Career matches by personality type (Personalitymax™, 2018)

As a result, if the companies understand the personality type of their employee, and compare the job position with the suggested careers for that type of personality, they will be able to understand if that employee is located in a right position and if suitable tasks are given to that employee. This can have a direct effect on the productivity of the employee and improvement of the organisational performance.

2.6.1 Related methodologies for personality type prediction

Big Five and MBTI personality models are the most used personality models in the world and they have been used in most researches on personality type prediction. It has been claimed that despite some disputes about reliability and validity of Big Five and MBTI personality models, the MBTI model has more applications, especially in industry (Barbuto, 1997). The Big Five model classifies personality types into 5 categories (Goldberg, 1990) but MBTI model classifies personality types into 16 categories via four dimensions (Myers et al., 1990).

Classic machine learning techniques and neural networks have been used successfully for predicting MBTI personality types. One of the earliest studies on personality prediction was by Champa and Anandakumar (2010) who used a three-layer feedforward architecture on handwritten textual data. Their work can be considered as a proof that deep neural architectures are proficient for MBTI personality type prediction with considerable accuracy. A little later, another method using machine learning techniques was presented by Golbeck et al (2011), they could accurately predict a user's personality type based on MBTI personality type indicator and by considering the information presented on their Twitter. In another study, Komisin and Guinn (2012) used Naïve Bayes and Support Vector Machine (SVM) techniques to predict an individual's personality type based on their word-choice. Their database was built based on in-class writing samples that were taken from 40 graduate students along with their MBTI personality type. They compared the performance of these two techniques and discovered that Naïve Bayes technique performs better than SVM on their small dataset. Two years

later, Wan and et al (2014) used a machine learning method to predict the Big Five personality type of users through their texts in Weibo, a Chinese social network, they were able to successfully predict the personality type of the users. Li, Wan and Wang (2017) used the grey prediction model, the multiple regression model and the multi-tasking model to predict the user personality type based on the Big Five model and their text samples. They compared the performance of these three models and found that the grey prediction model performs better than the two other models. Tandra and et al (2017) in another research used Big Five personality model and some deep learning architectures to predict a person's personality based on the user's information in Facebook. They compared the performance of their method with other previous researches that used classical machine learning methods and the results showed that their model successfully outperformed the accuracy of previous similar researches. Furthermore, in another study, Hernandez and Knight (2017) used various types of recurrent neural network (RNN) such as simple RNN, GRU, LSTM, and Bidirectional LSTM to build a classifier capable of predicting people's MBTI personality type based on text samples from their social media posts. Myers-Briggs Personality Type Dataset from Kaggle was used in their research. They compared the results and found that LSTM gave the best results. A more recent research was done by Cui and Qi (2017) who used Baseline, Logistic Regression, Naïve Bayes, SVM to predict an individual's MBTI personality type from one of their social media posts. They compared the results of all these methods and realised that SVM performed better. They used the same database with the previous research which was Myers-Briggs Personality Type Dataset from Kaggle. Table 2.5 shows the researches and the personality model used.

| Research | Personality model |
|-------------------------------|-------------------|
| Champa and Anandakumar (2010) | MBTI |
| Golbeck and et al (2011) | MBTI |
| Komisin and Guinn (2012) | MBTI |
| Wan and et al (2014) | Big Five |
| Li, Wan and Wang (2017) | Big Five |
| Tandera and et al (2017) | Big Five |
| Hernandez and Knight (2017) | MBTI |
| Cui and Qi (2017) | MBTI |

Table 2.5: Researches on personality type prediction and personality model used

In this research, it was found that classification techniques such as logistic regression, Naïve Bayes, Random forest, K Nearest Neighbour (KNN), Linear Discriminant Analysis (LDA) and Support Vector Machine (SVM) have all been used for personality type prediction based on MBTI or Big Five personality type models. However, Extreme Gradient Boosting (XGBoost) technique has not been implemented. The idea and theory behind the mentioned classification techniques will be explained in the following section.

2.6.2 Logistic Regression

Logistic regression classifier is one of the basic linear models for classification which can be used to analyse the relevance between multiple independent variables and a definite dependent variable. This model fits the data to a logistic curve to appraise the probability of occurrence of an event (Hyeoun-Ae, 2013). In other words, logistic regression is a specific category of regression which can be used to predict for binary or categorical dependent variables (Punnoose and Ajit, 2016). The difference between logistic regression and traditional multiple regression is that in logistic regression, maximum probability estimation is used rather than the least squares estimation which is used in traditional multiple regression (Elsalamony, 2014). In other words, logistic regression predicts a logit transformation of the dependent variables by using a mathematical model of a set of descriptive variables, whereas multiple regression uses a mathematical model of a set of descriptive variables in order to predict the mean of a continuous dependent variable.

Logistic regression classifiers use starting values of the predicted parameters and then compute the sample which is derived from a population with those parameters. The values of the estimated parameters are attuned iteratively until the highest probability value is obtained (Kleinbaum, 2010). Logistic regression can be considered as an approach to learning functions of the following form:

$$F: A \rightarrow Y, \text{ or } P(Y|A) \quad (8)$$

In this case, Y is a discrete-value target, and A can be considered as (A_1, A_2, \dots, A_n) any attribute containing discrete or continuous independent attributes (Elsalamony, 2014). When there is a binary variable, the value of 0 and 1 can be assigned to the two outcomes of this variable. Numerical value of 1 can represent a positive response and numerical value of 0 can represent a negative response. The proportion of a positive response will be understood by considering the mean of this variable. If the proportion of observations with an outcome of 1 is considered as P , then probability of a outcome of 0 can be considered as $1-P$ (Domínguez-Almendros, 2011; Kleinbaum, 2010). Equation (9) shows the logit transformation:

$$\ln\left(\frac{P}{1-P}\right) = k_0 + k_1A \quad (9)$$

In this equation, $k_0 + k_1A$ is the familiar equation for the regression line.

2.6.3 Naïve Bayesian

Naïve Bayes is a classification technique that has been popular and attractive between researchers because of its simplicity as well as performance (Mitchell, 1997). Naïve Bayes classifier is a special case of Bayesian classifier. Bayesian classifiers can be used to predict the probability if a sample belongs to a specific class (Elsalamony, 2014). In other words, Bayesian classifiers perform classification based on the idea that the role of a class is to predict the values of features for members of that class (Poole and

Mackworth, 2017). Thus, if an agent knows the class, it can predict the other features and if the agent does not know the class, the class given the feature values can be predicted using Bayes' rule. In fact, the learning agent builds a probabilistic model of the features and this model can be used to predict the classification of a new sample (Poole and Mackworth, 2017). This technique can be useful for large databases because of its high level of accuracy and speed in classification. It is also intuitive and fast to train with simple models (Elsalamony, 2014).

Bayesian classifiers use Bayes theorem which describes the relationship between $P(x)$, $P(x|Y)$, $P(Y)$ and $P(Y|x)$ as shown in equation (10):

$$P(x|Y) = \frac{P(Y|x)P(x)}{P(Y)} \quad (10)$$

$P(x|Y)$ is the probability of instance Y being in class x . $P(Y|x)$ is the probability of generating instance Y given class x . $P(x)$ is the probability of occurrence of class x and $P(Y)$ is the probability of instance Y occurring. The simplest case of Bayesian classifiers is the naive Bayesian classifier which performs classification based on probabilities arrived and works with a base assumption that all variables are conditionally independent of each other (Poole and Mackworth, 2017). Thus, in order to simplify the task, Naïve Bayesian classifiers assume that attributes have independent distribution and thereby estimate (Bishop, 2006). This is shown in equation (11):

$$P(Y|x) = P(Y_1|x) * P(Y_2|x) * ... * P(Y_n|x) \quad (11)$$

$P(Y|x)$ is the probability of class x generating distance Y . $P(Y_1|x)$ is the probability of class x generating the observed value for feature 1 and it will be multiplied by $P(Y_2|x)$ which is the probability of class x generating the observed value for feature 2. This process will continue until the last probability which is $P(Y_n|x)$ and it is for the probability of class x generating the observed value for feature n .

2.6.4 Random Forest

Random forest is a popular ensemble learning technique that uses a set of decision trees that grow in randomly selected subspaces of data, in order to build a predictor ensemble (Biau, 2012). In other words, in this technique a set of decision trees are built and then merged together in order to get a more accurate and stable prediction. Denil and et. al. (2014) also explained that “random forests are a type of ensemble method which makes predictions by averaging over the predictions of several independent base models.” It has been discussed that the main idea behind ensemble methods is that a strong learner can be formed when a set of weak learners come together. In this technique, the process starts with a standard machine learning technique called decision tree. In ensemble terms, decision tree technique corresponds to the weak learner. When the decision tree is built, an input can be entered at the top and then when the input traverses down the tree, the data is bucketed into smaller sets (Benyamin, 2012).

2.6.5 K Nearest Neighbor (KNN)

K Nearest Neighbour (KNN) classifier is a popular method to solve classification problems because of its simplicity and comparatively high convergence speed (Aldayel, 2013). In this technique, the k nearest instances $\{x_1, x_2, \dots, x_k\}$ will be considered from an instance (m) and then KNN decides based on the most frequent class in the set $\{y_1, y_2, \dots, y_k\}$. The most frequent class is assumed to be the class of that instance (m). The KNN technique can determine the neighbours using a distance metric that measures the vicinity of instance m to k of sorted instances (Vivencio and et al. 2007; Shouman and et al. 2012).

2.6.6 Linear Discriminant Analysis (LDA)

In many applications of machine learning, dimensionality reduction is important and linear discriminant analysis (LDA) is one of the most important methods which have been proposed in this regard (Fukunaga, 1990). In this method one or more discriminant functions will be created to maximise the variance between the categories relative to the variance with the categories (Nagadevara et al, 2008). In other words, this method maximises the ratio of the between-class distance to the within-class distance, in order to find the optimal discriminant vectors. As a result, the maximum class discrimination will be achieved (Ye and Xiong, 2006).

Punnoose and Ajit (2016) explained the linear discriminant analysis as “deriving a variate or z-score, which is a linear combination of two or more independent variables that will discriminate best between two (or more) different categories or groups.” The probabilities that a particular member or observation belongs to a class will be estimated using the z-score which is calculated via the discriminant functions.

2.6.7 Support Vector Machine (SVM)

Support Vector Machine is a supervised learning method that can be used to implement the principles of statistical learning theory. Both linear and nonlinear binary classification problems can be solved using this method (Cortes and Vapnik, 1995). A support vector machine performs classification through building an N-dimensional hyperplane for achieving class separation. A good separation can be achieved by a hyperplane that has the biggest distance to the nearest training data points of any class. This means that a larger margin will result in a lower generalisation error of the classifier (Punnoose and Ajit, 2016). In fact, hypothesis space of a linear function in a high dimensional feature space will be used in this technique, and it will be trained with a learning algorithm from optimisation theory that can implement a learning bias derived from statistical learning theory (Jakkula, 2011).

2.6.8 Extreme Gradient Boosting

Boosting is a method based on creating a very accurate prediction rule through combining rough inaccurate rules-of-Thumb (Freund and Schapire, 1997). In this process, a sequence of weak learners is fitted onto modified data and the predictions from all of them are combined through a weighted majority vote. This will help to produce the final prediction. In each step, there might be some samples that were misclassified in the previous iteration. As a result, the data modification is necessary at each step and it includes assigning higher weights to the training samples that were misclassified. Those samples that are difficult to predict during the iterations progress, will receive increasing influence and this will force the weak learner to focus on the samples that are missed by its ancestor.

In gradient boosting technique, new models will be fitted during the learning process, to provide a more accurate estimation of the response variable (Natekin and Knoll, 2013). Extreme Gradient Boosting is a boosted tree algorithm that follows the principle of gradient boosting (Friedman, 2001). It is able to perform better due to using a more regularised model formalisation in order to control over-fitting (Punnoose and Ajit, 2016).

2.7 Summary

Definitions of Neuro Linguistic Programming based on the academic sources were discussed in the beginning of this chapter and the theory behind the three terms ‘Neuro’, ‘Linguistic’ and ‘Programming’ were explained in details. Next, representational systems were discussed and the process of identifying the preferred representational system of a person was explained. In addition, specific tendencies of characteristics, which are associated with each representational system were discussed. Moreover, the definition and application of Meta model was discussed and different processes of applying the Meta model were described. The language patterns that can be identified via Meta model and the strategies for recovering the missing information in a conversation were also discussed in detail. Furthermore, different aspect of Meta programs and the relationship between the Meta programs and personality types were explained in this chapter. Finally, Myers-Briggs Type Indicator[®] was introduced and 16 personality types and their associated characteristics were explained. In addition, appropriate jobs for each personality type and the impact of understanding the personality type on the productivity of the employees in an organisation and improvement of the organisational performance was discussed. Finally, related methodologies for personality type prediction were overviewed.

3 Neuro Linguistic Programming automation methodology

3.1 Introduction

The methodology used in this research will be presented in this chapter. The implementation process in this research was divided into three different phases and each phase will be explained separately. The first phase is regarding the automation of identifying the preferred representational system. The second phase is the Meta model automation. The third phase is the automation of the personality type prediction. This section will be followed by explaining the data gathering procedure and data analysis strategy. Finally, a demonstration of the software is presented in this chapter.

3.2 Automating the preferred representational system identification

3.2.1 Natural Language Processing as a tool for automation

Intelligent systems are created in order to work well in different situations and environments. Their intelligence allows them to reach the maximum probability of success even with insufficient knowledge regarding a situation. As a result, artificial intelligence can be considered as a powerful tool for automating the process where a human interacts directly with a computer (Gudwin 2000). According to Chopra, Prashar

and Sain (2013) Natural Language Processing is a subfield of artificial intelligence and linguistic, it aims to enable computers to understand the words or sentences written in human languages.

Natural language processing is defined as a computerised approach, based on the use of a variety of theories and technologies to analyse the human language. This enables the language input to be processed and understood, whilst the same natural language can be generated by the system in order to communicate with the user (Chopra, Prashar and Sain, 2013; Liddy, 2001). Natural Language Processing is a multidisciplinary field of study, covering computer science, linguistics, psychology and artificial intelligence, focusing on the interaction between computers and the natural language of the user (Chopra, Prashar and Sain, 2013). According to Liddy (2001) there are seven levels in natural language processing; (1) phonology, (2) morphology, (3) lexical, (4) syntactic, (5) semantic, (6) discourse and (7) pragmatic.

The phonology level deals with interpretation of sound in speech to identify words and will be applied only if the text origin is speech (Enayet, 2010). Nugues (2006) states that “morphology is the study of how root words and affixes are composed to form words”. It is therefore an analysis and identification of the structure of words (Chopra, Prashar and Sain, 2013; Nugues, 2006). Lexical analysis, on the other hand, is about understanding the position of words in a sentence, their meaning and their relation to other words in that sentence (Enayet, 2010). Syntactic analysis focuses on analysing the words with regards to the grammatical structure of the sentence. The structural dependency relationships between the words in a sentence will also be recognised in the following stage of processing (Liddy, 2001). In the semantic analysis stage, the focus is on the

interactions among word-level meanings in a sentence and the way the lexical meaning is combined morphologically and syntactically to form the meaning of the statement (Liddy, 2001; Briscoe, 2013). Following this stage, discourse level looks at the connections between sentences in a text and deals with the properties of the whole statement in conveying meaning (Liddy, 2001). This is to take into account the dependence of each sentence on the previous and following sentences for conveying its meaning (Chopra, Prashar and Sain, 2013). Finally, pragmatic analysis focuses on the use of language in context, deriving the purposeful use of the language in different situations (Briscoe, 2013). After considering the stages of analysis through natural language processing, it was determined that this would be an ideal tool for automating the process of identifying the preferred representational system, recognising the Meta programs and personality type and also using the Meta model in the human-computer conversation.

3.2.2 Software structure and development procedure

Natural Language Processing has been perceived as the most relevant and powerful techniques for automating the preferred representational system identification process. A software has been developed using Python, a very flexible programming language, as well as NLTK which is a very powerful Natural Language Toolkit for Python. According to the NLTK official website, NLTK is a leading toolkit for developing Python programmes using human language data. NLTK has a user-friendly interface for over 50 corpora and lexical resources like WordNet. NLTK has a suite of text processing libraries

for classification, tokenisation, stemming, tagging, parsing, and semantic reasoning. In addition to wrappers for industrial-strength NLP libraries, and an active discussion forum (NLTK official website 2016).

Dialogue Systems can be used to enable communication between computer and human in natural language, instead of complex commands or procedures (Yan et al., 2017). They can be divided into two main categories including (1) chat oriented systems and (2) task-oriented systems (Su et al, 2016). Chat oriented systems usually aim to communicate with the user in order to provide reasonable and interesting responses which is contextually relevant (Banchs and Li 2012; Yan et al, 2016). On the other hand, task-oriented systems try to assist users to achieve specific goals. For instance, helping users to find a specific product or flight (Bohus and Rudnicky, 2009). It can be said that the developed system in this research is influenced by the second category, which is related to task-oriented systems, as the software is going to communicate with the user in order to analyse the users' answers, understand the developmental and behavioural patterns, and then provide some recommendations for improvement of performance.

The software is able to have a conversation with human users and communicate through an interactive environment starting with a brief introduction followed by response-based questions. The individual's answers are communicated through typing out of relevant response. The answers will be analysed by the software both throughout and at the end of the conversation. Based on an overall analysis, the software will then recognise the preferred representational system and based on the relevant characteristics of that representational system, it suggests relevant solutions to the user for improvements in communication and learning. Furthermore, the software is able to recognise the most

common preferred representational system between employees in a company and suggest the relevant solutions to the manager for improvements in communication between employees and their learning level. Figure 3.1 shows this part of the software development procedure.

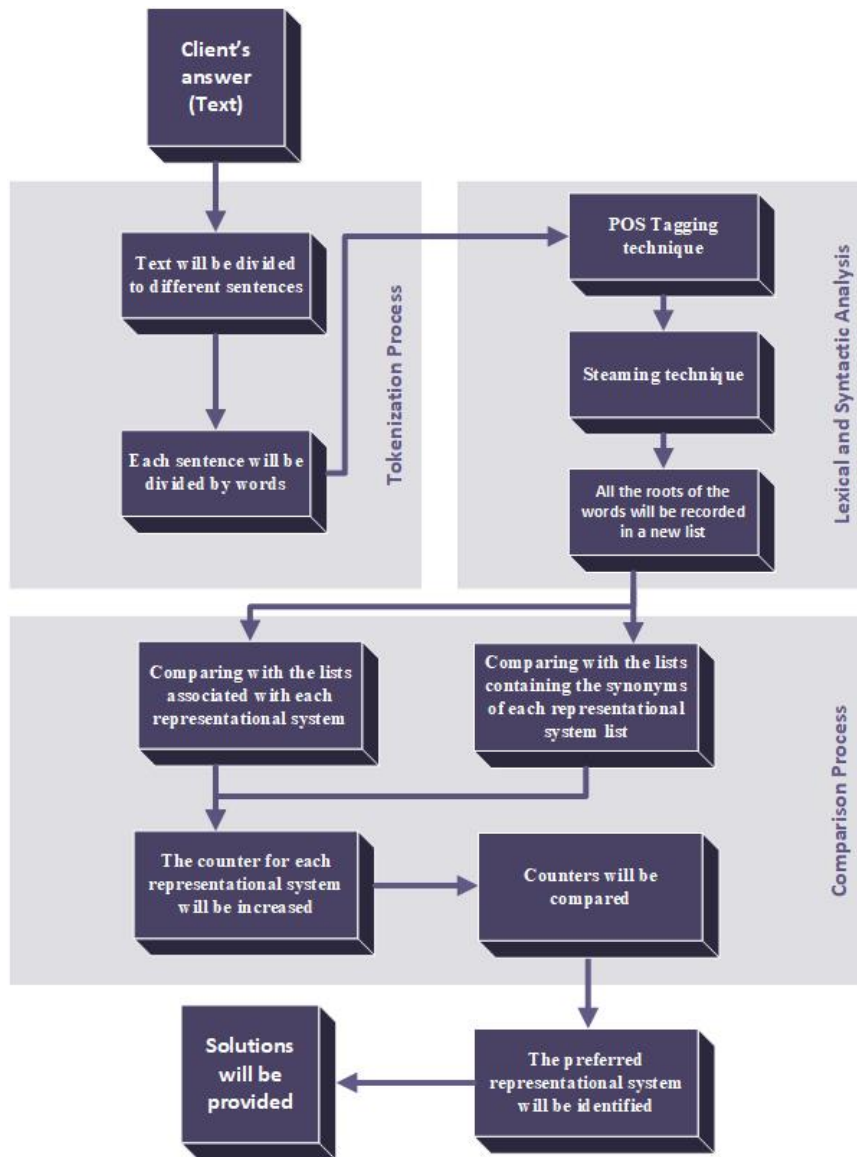


Figure 3.1: Development procedure of the first phase of implementation

3.2.3 Tokenisation process

The first step in analysis of response is via the tokenisation technique. According to Manning, Raghavan and Schütze (2009) “a token is an instance of a sequence of characters in some particular document that are grouped together as a useful semantic unit for processing”. In fact, these basic units must be clearly segregated, otherwise it would be impossible to carry out any analysis or generation (Webster and Kit 1992). As a result, an individual’s answer will be recorded as a string and this string will be divided to different sentences and each sentence will be analysed separately. All of the sentences will be recorded in a list called ‘sentence_list’ and each sentence will be divided by words. Then all of those words will be recorded in a different list again which is called ‘word_list’. These lists will be then used for lexical and syntactic analysis in the next step.

3.2.4 Lexical and Syntactic analysis

A one-dimensional language like a written language is composed of letters and symbols and can be considered as a code describing some reality (Horn, 2008). It needs to have rules in order to describe how its words or sentences are connected to that reality and how to put them together into a language representation (Horn, 2008). As a result, Part-Of-Speech tagging or POS tagging technique will be used in this step. In this process, the software recognises the role of each word in each sentence. Jurafsky and Martin (2014) explained that in this process a part-of-tagging speech marker will be assigned to each

word for an input text. Bird, Klein and Loper (2009) also in a book named ‘Natural Language Processing by Python’ stated that a part-of-speech tagger can process a sequence of words and then attach a part of speech tag to each one of those words. There are different lists of Parts-Of-Speech tag sets and one of the most common is the Brown corpus which has been used for this research. As a result, all verbs, nouns, adjectives, adverbs and other relevant elements in each sentence will be recognised.

Hidden Markov Model (HMM) is a very common and useful tagging technique that has been used to build the POS tagger. Below is the bigram HMM equation:

$$x_m = \arg \max_n P(x_n | x_{m-1}, y_m) \quad (1)$$

The nearby words and tags are checked in order to solve the tagging problem:

$$x_m = \arg \max_n P(x_n | x_{n-1})P(y_m | x_n) \quad (2)$$

In this equation $P(y_m | x_n)$ represents word likelihood and $P(x_n | x_{n-1})$ represents tag co-occurrence. The full model aims to identify the best sequence of tags for the whole sentence:

$$\begin{aligned} \hat{X} &= \arg \max_{X \in \mathcal{X}} P(X | Y) \\ &= \arg \max_{X \in \mathcal{X}} \frac{P(X)P(Y | X)}{P(Y)} \\ &= \arg \max P(X)P(Y | X) \end{aligned} \quad (3)$$

The chain rule is used to expand this equation:

$$\begin{aligned} P(X)P(Y | X) &= \\ \prod_{m=1}^k P(y_m | y_1 x_1 \dots y_{m-1} x_{m-1} x_m) P(x_m | y_1 x_1 \dots y_{m-1} x_{m-1}) \end{aligned} \quad (4)$$

In order to approximate these two factors, the trigram assumption is simplified. Therefore, the probability of a word depends only on its tag:

$$P(x_m | y_1 x_1 \dots x_{m-1} x_m) = P(y_m | x_m) \quad (5)$$

Then the tag history is approximated by the two most recent tags:

$$P(x_m | y_1 x_1 \dots x_{m-1}) = P(x_m | x_{m-2} x_{m-1}) \quad (6)$$

Lastly, the equation is replaced:

$$P(X)P(Y | X) = P(x_1)P(x_2 | x_1) \prod_{m=3}^k P(x_m | x_{m-2} x_{m-1}) \left[\prod_{m=1}^k P(y_m | x_m) \right] \quad (7)$$

Following the process of POS tagging, stemming technique has been used to remove all prefixes and suffixes from the words and thus identifies the root of each word. Manning, Raghavan and Schütze (2009) explains that because of grammatical reasons, each document may use different forms of a word and also there are families of derivationally related words that may have similar meanings. Hence, in different situations, it might be useful to search for one of these words as a root word. For instance, it may help to return some other documents that contain another word related to the root word (Manning, Raghavan and Schütze 2009). Accordingly, by using this technique, all of the roots of the words recorded in the 'word_list' will be recognised and recorded in a new list called 'root_list' for the comparison process in the next step.

3.2.5 Comparison process

There are four other lists that have been defined for the software, whereby each one of these lists are allocated to one specific representational system. This includes a collection of words, which are the associated predicates for the specified system. These lists are named (1) 'visual_list', (2) 'auditory_list', (3) 'kinesthetic_list', and (4) 'auditory_digital_list'. A collection of various documents has been used in order to create a collection of relevant vocabulary. These include 'NLP Home Study Programme (V2.0)' published by Juiced Concepts Limited (2012), 'Representational Systems' published by Brefi Group Limited (2004) and 'The Power of Words' written by Katy McAfee (2014). These sources provide a list of popular words associated to each representational system. As a result, all these mentioned words have been used in this research and added to the relevant lists for visual, auditory, kinesthetic and auditory digital representational systems. Tables 3.1 to 3.4 below are showing some of the most common predicates recorded in the relevant lists.

| Visual predicates list | | | |
|-------------------------------|-------------------|-------------------|--------------------|
| See | Saw | Seen | Look |
| Appear | Observing | Appearance | View |
| Show | Shown | Dawn | Reveal |
| Envision | Illuminate | Twinkle | Clear |
| Foggy | Focus | Hazy | Crystal |
| Flash | Image | Picture | Spark |
| Frame | Snapshot | Vivid | Imagine |
| Clarity | Witness | Illustrate | Vague |
| Outlook | Inspect | Sight | Light |
| Scene | Watch | Perceive | Perspective |
| Observe | Vision | Angle | Sign |
| ... | ... | ... | ... |

Table 3.1: Example of visual representational system predicates

| Auditory predicates list | | | |
|---------------------------------|------------------|--------------------|--------------------|
| Hear | Listen | Sound | Music |
| Harmonize | Tune in | Tune out | Ear |
| Ring | Bell | Silence | Heard |
| Resonate | Deaf | Mellifluous | Dissonance |
| Dissonant | Overtones | Attune | Outspoken |
| Tell | Announce | Remark | Overtones |
| Unhearing | Audible | Voice | Interview |
| Talk | Speak | Rumor | State |
| Whine | Babble | Echo | Orchestrate |
| Whisper | Oral | Hum | Speechless |
| ... | ... | ... | ... |

Table 3.2: Example of auditory representational system predicates

| Kinaesthetic predicates list | | | |
|-------------------------------------|-------------------|------------------|-----------------|
| Feel | Touch | Grasp | Catch |
| Hold | Contact | Throw out | Hard |
| Feeling | Concrete | Scrape | Handle |
| Suffer | Impression | Flow | Lukewarm |
| Slip | Tap | Shift | Throw |
| Turn around | Unfeeling | Callous | Solid |
| Unjudging | Softly | Soft | Rub |
| Unsettles | Smooth | Pushy | Push |
| Panicky | Stumble | Muddled | Relaxed |
| Relax | Loose | Sore | Bearable |
| Cool | Tepid | Charge | Heavy |
| ... | ... | ... | ... |

Table 3.3: Example of kinaesthetic representational system predicates

| Auditory digital predicates list | | | |
|---|-------------------|----------------------|-----------------|
| Sense | Experience | Understand | Catch |
| Learn | Process | Decide | Hard |
| Consider | Change | Perceive | Handle |
| Distinct | Conceive | Know | Lukewarm |
| Conscious | Recall | Communicate | Throw |
| Plan | Advice | Function | Solid |
| Create | Activate | Repeat | Rub |
| Logically | Reasonable | Statistically | Push |
| ... | ... | ... | ... |

Table 3.4: Example of auditory digital representational system predicates

These four lists mentioned in section 3.2.5, were expanded and validated through identifying all appropriate synonyms for each one of the representational system predicates recorded in them. NLTK or Natural Language Toolkit in Python was used for this purpose and the identified synonyms were recorded in four separate lists named (1) 'visual_synonym_list', (2) 'auditory_synonym_list', (3) 'kinesthetic_synonym_list', and (4) 'auditory_digital_synonym_list'. As a result, a comprehensive collection of all possible representational system predicates was acquired. Thus, in this step the software is able to compare the obtained list, explained in section 3.2.4, containing the roots of a word (root_list), with the four lists associated with each representational system, explained in section 3.2.5 and their relevant synonym lists. If any of the words in the root_list being analysed is included in any of the four representational system predicates lists or their synonym lists, the counter for that specific representational system is increased. The list of predicates for each representational system is exclusive and there are no overlaps. Words which are unspecified as a predicate are ignored and not allocated to any representational system. Finally, all counters will be compared together and the application informs the person about his or her preferred representational system. Moreover, some solutions will be suggested to improve the communication effectiveness and learning methodologies of the individual. The relevant solutions have been collected from the aforementioned documents.

3.3 Automating the process of using the Meta model

3.3.1 Software structure

In this phase of software development process, three language patterns from the deletion process: (1) unspecified nouns, (2) comparative deletions (3) ‘Ly’ adverbs; two language patterns from the distortion process: (1) mind reading (2) linguistic presuppositions; and two language patterns from the generalisation process: (1) universal quantifiers and (2) modal operators are considered.

The software starts the conversation by asking the first of ten set questions. The theme of the questions in this study are regarding the user’s work environment. The software continues the conversation with the user based on the user’s answers. The software is able to identify the language patterns used in the user’s response and follow up by asking the relevant Meta model recovery questions to clarify any obscured information. The user will be then informed about the missing information and issues identified in the conversation. Additionally, clarifications or explanations about the presented issues will be provided by software.

3.3.2 Programming language and the relevant library

Python, a powerful programming language for processing linguistic data and NLTK, a useful library for natural language processing in Python, were used to develop

this software. NLTK provides basic classes for representing the data relevant to the natural language as well as convenient interfaces for performing tasks such as text classification, syntactic parsing and part-of-speech tagging (Bird, Klein and Loper, 2009).

3.3.3 Asking the questions

The software starts the conversation by asking the first of the ten set questions. The theme of the questions in this study are regarding the user's work environment and the software continues the conversation with the user, based on the user's answers. The user's answers to any of the set questions will be analysed for any ambiguity. This is carried out through the application of the Meta model in order to clarify the statement for the computer. Following clarification, using the Meta model questions, the user will be presented with the next of the ten set questions. The steps of the Meta model analysis are executed inside the body of a function which is used in a 'for loop' to be repeated for all answers provided by the user, ensuring clarification of every ambiguous statement. The set of 10 questions used in the Meta model is as follows:

- 1- First of all, could you please describe your work environment with a few sentences? It can be anything about your feeling about your job, your relationship and communication with your colleagues and your manager. You could also discuss your tasks and responsibilities and the way you are doing your job.

2- Let's talk about it from a different point of view and about some specific issues.

When you get a task at work, how is it easier to understand it and carry it out?

For example, having a written task and a clear plan. Would you prefer it is someone explains it to you instead? Is getting a sense of purpose for your job important to you?

3- What about meetings and presentations in your company. When someone is presenting something in a meeting, how do you find it easier to follow the presentation? For example slides and visual aspects, Logic and structure of presentation, tone of voice or even body language.

4- What do you think about your discussions with your manager or your colleagues? What are you most often influenced by? For example the other person's point of view, their logic, their tone of voice or maybe their feeling.

5- Imagine you are in a situation where you need to make an important decision. How would you go about making that decision? What are your important decisions based on?

6- What is your most important strength in communicating with your colleagues?

7- How do you assess how well you are doing at work?

8- When you recall a time you were immensely drawn to someone (can be a

manager, colleague, friend or etc) what was the very first thing that attracted you to them? For example the way they looked or something they said to you or the way they touched you or something you felt.

9- Think about when you walk or drive to work. How do you navigate when you are walking or driving? For example look at road signs, following a map, listening to a familiar sound like GPS or even your feeling or sense of where you are. Please explain your navigation method to find your way in a few sentences.

10- When you have some problems at work and your problems get you down, how do you help yourself to understand the problems or to find a solution? For example do you usually write them down to see them clearly? Or you prefer to talk or listen to another person or you usually sort them out in your head until they make sense?

An important aspect in designing these questions were the importance of encouraging depth and breadth in the information sought about their work. To ensure this, the questions encourage thinking about one concept from different aspects and prompt ideas by providing ways to consider answering the questions. Moreover the same concept about their job is discussed in a range contexts through a variety of questions. The main concepts focused on in these questions are the performance of the participant or their relationship with their role at work and also the concept of communication at their workplace. Questions 2, 5, 7, 9 and 10 explore the former through asking about general themes of carrying out one's role in the workplace, these were, executive functions such

as decision making, problem solving and method of navigation. The perspective of the participants about their vision of carrying out their job and what can optimise their performance was also explored in these questions. The second concept of communication at the workplace was investigated through questions 3, 4, 6 and 8. The variability of establishing communication at the workplace was attempted to be covered as widely as possible by asking about variety in presentation of information and also about face to face discussions and exchange of ideas. The notion of one-to-one communication was further explored by asking about the participants' experience of communicating their ideas and their experience of being responsive to communication directed to them in questions 6 and 8 respectively. The first question, on the other hand focuses on neither of these concepts and asks a general open-ended question about the experience of the participant at their work place to allow for any specific aspects or concerns of the participants role to be disclosed and discussed with the follow up questions. The variability of questions are also important in being able to acquire a substantial amount of information for the NLP analysis. This increases the accuracy of the results obtained as increasing the amount of information tested would in turn decrease the significance of error.

3.3.4 Defining the key words

Eight lists have been created for this software which include specific key words or identifiers that are used in different steps of the Meta model process. They are regarding personal pronouns, determiners, necessity identifiers, impossibility identifiers and universal quantifiers, explained previously. Table 3.5 shows the content of these lists.

| List | Content of the list | |
|----------------------------------|--|--|
| Personal Pronouns 1 | he she him her | They them his their |
| Personal Pronouns 2 | It | |
| Determiners 1 | this that | These those |
| Determiners 2 | There | |
| Necessity Identifiers | has to have to had to | need to must should |
| Unnecessity Identifiers | do not have to did not have to don't have to didn't have to does not have to doesn't have to should not shouldn't | do not need to don't need to did not need to didn't need to does not need to doesn't need to must not mustn't |
| Impossibility Identifiers | cannot can't impossible is not possible isn't possible | could not couldn't not possible |
| Universal quantifiers | never ever | always all |

Table 3.5: Specific key words to be used in different steps of the Meta model

3.3.5 Tokenisation process

In response to the user's answers, the software uses the "Tokenisation" technique. In this step, the user's answer will be recorded as a string which will be then divided into different sentences and each sentence will be analysed separately. All sentences will be recorded in a list named 'meta_sentence_list' and each sentence will be divided into different words, this forms a second list named 'meta_word_list'.

3.3.6 Lexical and Syntactic analysis

Following the tokenisation process, Part-Of-Speech tagging or the POS tagging technique is used and the role of each word in each sentence, in other words, all verbs, nouns, adjectives, adverbs and other relevant elements in each sentence will be recognised. The same POS tagger which was built in the previous phase of the software development process was also used in this part. Penn treebank tag set was employed in this part as well.

Following the POS tagging process, the software creates two different lists; the first list consists of the pronouns in each sentence as they may be indicative of missing information. This list is named 'meta_pronouns_list'. The second list is of the adverbs in each sentence. This lists is named 'meta_adverb_list'. Each one of these lists will be created by using a loop and checking the POS tags for each word in each sentence. Thus, if the relevant POS tag existed in the sentence, that specific word will be recorded in the

relevant list and these lists will be used for the comparison process in the next step.

3.3.7 Comparison process

There are four different lists related to specific pronouns and determiner words that have been defined for the software previously and were explained in section 3.3.4. These lists will be used during the comparison process; they will be compared with the created list of pronouns in each sentence in the previous step (`meta_pronoun_list`). Hence, the created lists in the lexical and syntax analyses, explained in section 3.3.6, will be compared to each one of those four lists one by one. Detection of similarity between each of the two lists, leads the specific words to be recorded in a new list named ‘`similarity_list`’, as the final list.

The software also creates four other lists consisting of necessary identifiers, unnecessary identifiers, impossibility identifiers and universal quantifiers. The strategy for creating these lists is the same, as they are being created using a loop. However, the POS tags are not necessary in this case and instead, each word in each sentence will be compared to the words recorded in the relevant list, as defined and explained previously. As a result, the final list consists of six separate lists, which will be used in the following step of checking the conditions and the decision making process.

Five lists including two personal pronouns list and two determiners lists explained in section 3.3.4, and `meta_pronouns_list` explained in section 3.3.6, are related to the deletion process. On the other hand, `meta_adverb_list` explained in section 3.3.6 is

related to distortion process. Finally, four lists including necessity identifiers, unnecessity identifiers, impossibility identifiers and universal quantifiers explained in section 3.3.4, are related to generalisation process. As it was mentioned, these lists will be used during the comparison process in order to create the final lists to be used during the decision making process.

3.3.8 Decision making process

In this step, the software checks the conditions and in the case of any words recorded in any of the final lists, the software asks a specific relevant recovery question from the user. For instance, if the user has written one paragraph, the format of a recovery question would be as follows:

You said: “.....(The sentence that includes missing information will be repeated in here).....”.

The relevant question word (Who/What/Which/Where) exactly? Could you explain further?

Hence, the software encourages the user to expand on the missing information and to clarify the meaning of the statement made. The user's answers to the recovery questions will be recorded, in order to be used in the following steps.

On the other hand, the list consisting of adverbs (*meta_adverb_list*), created during the lexical and syntax analyses will be used in checking the condition process. If this list

was empty and there were no ‘Ly’ adverbs used in the user’s sentences, the software would continue the conversation in the standard format. If the list was not empty, however, a recovery question would be asked from the user, such as:

You said: “.....(The sentence that includes the intended adverb will be repeated in here).....”.

(The intended adverb) than what? / Why (The intended adverb)? / How (The intended adverb)?

Thus, the user’s expansion on their statement via the recovery question would be recorded to be used in future steps. This process may be repeated for all answers to the recovery questions provided by the user. If there were any remaining missing information, the software will continue asking recovery questions to clarify the statement. There is a counter for each one of deletion, distortion and generalisation processes which will be increased after asking each relevant recovery question. This is how the number of deletion, distortion and generalisation processes identified by the software, are to be recognised. This information will be used in chapter 4 in order to discuss the results.

3.3.9 Informing the person

After each recovery question, the user will be informed about any issues in their sentence and the clarification or explanation that they made, after responding to the

recovery question. The format will be as follows:

You said: “.....(The sentence that includes missing information, changed information or generalised information, will be repeated in here).....”.

The issue in your sentence was (The relevant issue. For instance, unspecified noun which is an element of deletion process in the Meta model) Your clarification or explanation for this issue is: “.....(The user’s answer to the recovery question).....”.

3.3.10 Repetition process

The user’s answers to any of the set questions based on their work environment will be analysed for any ambiguity which the Meta model could be used for in order to clarify the statement for the computer. Following clarification, with the use of the Meta model questions. In cases that do not require clarification, the user will be presented with the next of the ten set questions, explained in section 3.3.3.

The steps of the Meta model analysis are executed inside the body of a function which is used in a ‘for loop’ to be repeated for all answers provided by the user, ensuring clarification of every statement. Figure 3.2 shows the development procedure for automating the Meta model process.

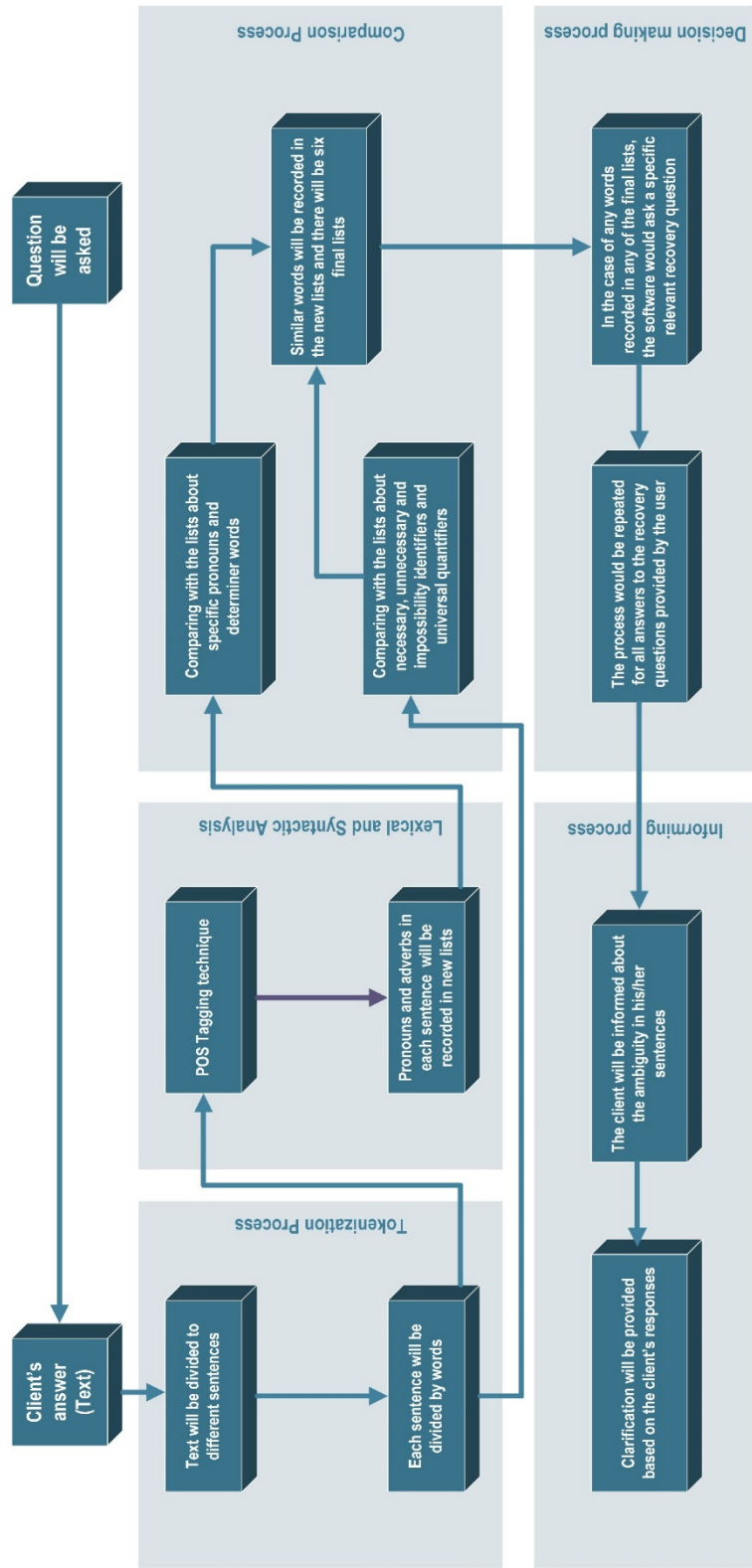


Figure 3.2: Development procedure of the second phase of implementation

3.4 Automating the process of personality type prediction

3.4.1 Development procedure

NLTK and XGBoost were used in this phase of software development process. NLTK is a powerful natural language processing toolkit for developing Python programs to work with human language data. XGBoost is an optimised distributed gradient boosting library in python that implements machine learning algorithms under the Gradient Boosting framework. Pandas, numpy, re, seaborn, matplotlib and sklearn are other python libraries that were used in this part.

The software is able to have a conversation with human beings and communicate through an interactive environment starting with a brief introduction followed by some questions. The theme of this conversation is related to the work environment. The answers will be recorded in a file. Based on an overall analysis, the software will then predict the type of personality according to MBTI personality categories and also suggest the most appropriate positions for that employee in the organisation. In the second step, the employee will respond to a set of multiple-choice questions which have been computerised. Based on the answers, the software will recognise the relevant Meta programs used by the employee. Based on the relevant characteristics of those Meta programs, the software identifies the personality type according to MBTI personality categories. In the final step, the result of machine learning prediction will be compared with the result of multiple choice questionnaire to analyse the accuracy of the result.

3.4.2 Dataset for training the model

Myers-Briggs personality type dataset, a publicly available kaggle dataset containing 8600 rows of data, was used in this research. In this dataset, each row consists of two columns. The first column is for the MBTI personality type of a given person, and the second column includes fifty posts obtained from the individual's social media. Each post has been separated by three pipe characters (Kaggle, 2018). This data has been collected from the users of an online forum where users in the first step take a questionnaire that recognises their MBTI type and in the second step chat with other users (Hernandez and Knight, 2017).

3.4.3 Proportionality in dataset

In this step, seaborn which is a Python data visualisation library and matplotlib which is a Python 2D plotting library, were used for data preview and to check the distribution of the MBTI personality types in the dataset. Figure 3.3 and table 3.6 show the number of occurrences for each MBTI personality type in the dataset.

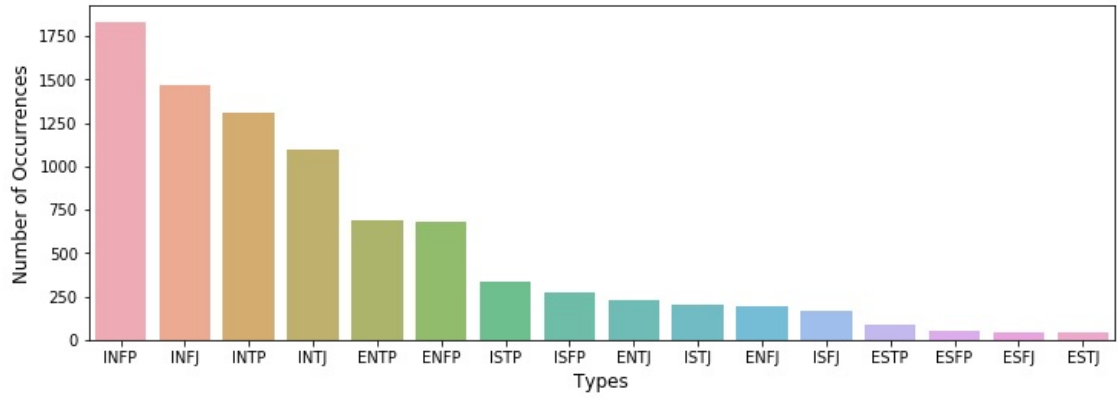


Figure 3.3: Number of occurrence for each MBTI personality type in the dataset

| Personality Type | Number of people |
|------------------|------------------|
| INFP | 1832 |
| INFJ | 1470 |
| INTP | 1304 |
| INTJ | 1091 |
| ENTP | 685 |
| ENFP | 675 |
| ISTP | 337 |
| ISFP | 271 |
| ENTJ | 231 |
| ISTJ | 205 |
| ENFJ | 190 |
| ISFJ | 166 |
| ESTP | 89 |

| | |
|-------------|----|
| ESFP | 48 |
| ESFJ | 42 |
| ESTJ | 39 |

Table 3.6: The breakdown of the total number of people of each type in the dataset

Similarly, Figure 3.4 and table 3.7 show the percentage of occurrences for each MBTI personality type in the dataset.

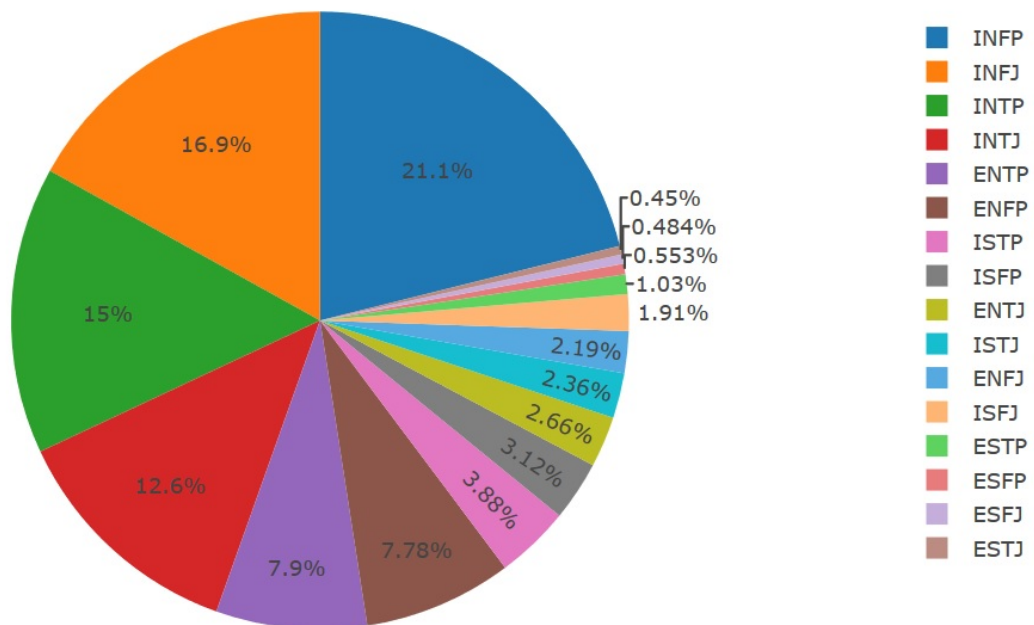


Figure 3.4: Percentage of occurrence for each MBTI personality type in the dataset

| Type | Frequency in Population |
|-------------|-------------------------|
| INFP | 21.1% |
| INFJ | 16.9% |
| INTP | 15% |
| INTJ | 12.6% |
| ENTP | 7.9% |
| ENFP | 7.78% |
| ISTP | 3.88% |
| ISFP | 3.12% |
| ENTJ | 2.66% |
| ISTJ | 2.36% |
| ENFJ | 2.19% |
| ISFJ | 1.91% |
| ESTP | 1.03% |
| ESFP | 0.55% |
| ESFJ | 0.48% |
| ESTJ | 0.45% |

Table 3.7: Percentage of occurrence for each MBTI personality type in the dataset

Figure 3.3 And 3.4, and Table 3.6 and 3.7 Show a non-uniform representation of MBTI types in the dataset which is not commensurate with the actual proportions of MBTI types

in the general population shown in table 3.8. As a result, it was clear that some cleaning in the dataset would be necessary in order to improve the accuracy of the proportional representation of each MBTI type.

| Type | Frequency in Population |
|-------------|-------------------------|
| ISFJ | 13.8% |
| ESFJ | 12.3% |
| ISTJ | 11.6% |
| ISFP | 8.8% |
| ESTJ | 8.7% |
| ESFP | 8.5% |
| ENFP | 8.1% |
| ISTP | 5.4% |
| INFP | 4.4% |
| ESTP | 4.3% |
| INTP | 3.3% |
| ENTP | 3.2% |
| ENFJ | 2.5% |
| INTJ | 2.1% |
| ENTJ | 1.8% |
| INFJ | 1.5% |

Table 3.8: Personality type distribution in the general population (Myers, 1998)

3.4.4 Categorising the type indicators in four dimensions

Four different categories were created for the type indicators in order to understand the distribution of types indicators in the dataset. The first category was for Introversion (I) / Extroversion (E), the second category was for Intuition (N) / Sensing (S), the third was for Thinking (T) / Feeling (F) and the fourth category was for Judging (J) / Perceiving (P). As a result, for each category, one letter will return and at the end there will be four letters which represents one of the 16 personality types in MBTI. For instance, if the first category is returning I, the second category is returning N, the third category is returning T and the fourth category is returning J, the relevant personality type would be INTJ. Table 3.9 And Figure 3.5 show the distribution across type indicators.

| Type indicator | | Distribution |
|----------------|------------------|--------------|
| I/E | Introversion (I) | 1999 |
| | Extroversion (E) | 6676 |
| N/S | Intuition (N) | 1197 |
| | Sensing (S) | 7478 |
| T/F | Thinking (T) | 4694 |
| | Feeling (F) | 3981 |
| J/P | Judging (J) | 5241 |
| | Perceiving (P) | 3434 |

Table 3.9: Distribution across types indicators

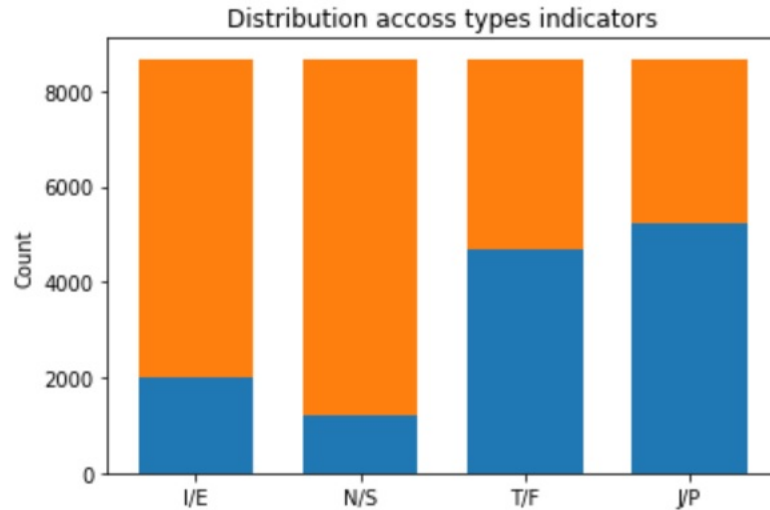


Figure 3.5: Distribution across types indicators

According to Table 3.10 and Figure 3.5, for the first category of Introversion (I) / Extroversion (E), the distribution of Extroversion (E) is much greater than Introversion (I). Similarly, for the second category which is Intuition (N) / Sensing (S), the distribution of Sensing (S) is much higher than Intuition (N). Figure 3.5 and table 3.10 also show that for the third category which is Thinking (T) / Feeling (F), the distribution of Thinking (T) is slightly more than Feeling (F). Finally, for the fourth category which is Judging (J) / Perceiving (P), the distribution of Judging (J) is greater than Perceiving (P).

Table 3.10 is a correlation matrix that shows correlation efficient between personality type identifiers. Each random variable (X_i) in a correlation matrix is correlated with each of the other values in the table (X_j), this can help to understand which pairs have the highest correlation.

| | IE | NS | TF | JP |
|----|-----------|-----------|-----------|-----------|
| IE | 1.000000 | -0.045899 | -0.069573 | 0.161939 |
| NS | -0.045899 | 1.000000 | -0.080954 | 0.014922 |
| TF | -0.069573 | -0.080954 | 1.000000 | -0.004673 |
| JP | 0.161939 | 0.014922 | -0.004673 | 1.000000 |

Table 3.10: Correlation efficient between personality type indicators

Figure 3.6 which is created using Matplotlib library in Python, also shows the Pearson correlation coefficient that can measure the strength between variables and relationships. In order to understand how strong tthe relationship is between two variables, the coefficient value must be found, which can range between -1.00 and 1.00.

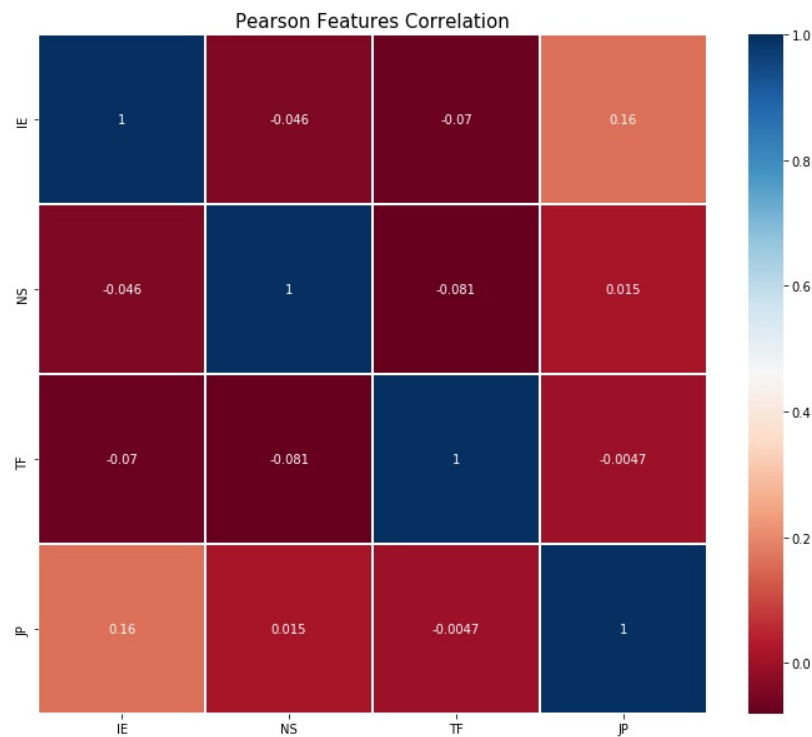


Figure 3.6: Pearson correlation coefficient between personality type indicators

3.4.5 Pre-processing the dataset

As discussed earlier, data in this dataset was collected from an Internet forum and after analysing the content of the dataset, it was clear that some word removal is necessary. This is significant because the model should be able to generalise the English language. As a result, NLTK was used to remove all urls and stop words from the dataset. As the data was collected from an Internet forum created for discussion about personality type, the MBTI types were removed from the dataset as well, because they appear too many times in the posts and might affect the accuracy of the model. Finally, in order to make the dataset more meaningful, the text was lemmatised, this means inflected forms of the words were transformed into their root words.

3.4.6 Vectorise with count and Term Frequency–Inverse

Document Frequency (TF-IDF)

Sklearn library was used to recognise the words appearing in 10% to 70% of the posts. In the first step, posts were put into a matrix of token counts. In the next step, the model learns the vocabulary dictionary and returns a term-document matrix. The count matrix then transforms into a normalised tf-idf representation which will be used for the gradient boosting model. Finally, 791 words appear in 10% to 70% of the posts.

3.4.7 Classification task

In machine learning there are two different types of classification. In the first type, based on a set of observations, the aim is to establish the existence of classes or clusters in the data. In the second type, a certain number of classes may exist and the aim is to establish a rule or a set of rules to classify a new observation into one of the existing classes (Punnoose and Ajit, 2016). The first type is known as Unsupervised Learning and the second type as Supervised Learning (Michie and et al, 1994).

The classification task was break down with 16 classes into four smaller binary classification tasks. The reason is that each MBTI type is made of four binary classes. Each one of these binary classes represents an aspect of personality according to the MBTI personality model. As a result, four different binary classifiers were trained, whereby each one specialises in one of the aspects of personality.

Thus, in this step, a model for each type indicator was built individually. Term Frequency-Inverse Document Frequency (TF-IDF) was performed and MBTI type indicators were binarised. Variable X was used for posts in tf-idf representation and variable Y was used for binarised MBTI type indicator.

3.4.8 Developing Gradient boosting model for the dataset

Numpy, XGBoost and sklearn were used in this step to create the gradient boosting model. MBTI type indicators were trained individually and the data was then split into

training and testing datasets. The SciKit library in Python provides a tool, called the Model selection library. This tool was used in order to make sure the dataset splits in a random manner. There is a class in the library named ‘train_test_split’. Using this class the dataset was split into the training and testing datasets in various proportions. ‘test_size’ parameter was set to 0.3. It is not necessary to specify ‘train_size’ parameter if ‘test_size’ parameter has been specified. Then ‘random_state’ parameter, which will act as the seed for the random number generator during the split, was set to 7. The model was fit onto the training data and the predictions were made for the testing data. Predictions were evaluated and Table 3.11 shows the results.

| Binary class | MBTI personality type | Accuracy |
|--------------|-------------------------------------|----------|
| IE | Introversion (I) – Extroversion (E) | 78.17% |
| NS | Intuition (I) – Sensing (S) | 86.06% |
| FT | Feeling (F) – Thinking (T) | 71.78% |
| JP | Judging (J) – Perceiving (P) | 65.70% |

Table 3.11: Accuracy of MBTI personality type prediction by software

After this step, the performance of the XGBoost model on the testing dataset during training was evaluated and early stopping was monitored. The first step of configuration strategy was to run the default configuration and review plots of the learning curves on the training and validation datasets. In the second step, if the system was overlearning, the learning rate was decreased and/or the number of trees were increased. Finally, in the

third step, if the system was under learning, the learning rate was increased and/or the number of trees were decreased in order to speed up the learning. Learning rate in XGBoost should be set to 0.1 or lower, and the addition of more trees will be required for smaller values. Moreover, there is not much benefit observed with deeper trees, the depth of trees should be configured in the range of 2 to 8. Furthermore, row sampling should be configured in the range of 30% to 80% of the training dataset. Thus, `tree_depth` in the created XGBoost was configured and parameters for XGBoost were setup as follow:

```
n_estimators = 200
max_depth = 2
nthread = 8
learning_rate = 0.2
```

MBTI type indicators were trained individually and then the data was split into training and testing datasets. The model was fit onto the training data and the predictions were made for the testing data. Predictions were evaluated and Table 3.12 shows the results.

| Binary class | MBTI personality type | Accuracy |
|--------------|-------------------------------------|----------|
| IE | Introversion (I) – Extroversion (E) | 79.01% |
| NS | Intuition (I) – Sensing (S) | 85.96% |
| FT | Feeling (F) – Thinking (T) | 74.19% |
| JP | Judging (J) – Perceiving (P) | 65.42% |

Table 3.12: Prediction accuracy for each MBTI binary class

Table 3.13 shows the result of comparing Table 3.11 and Table 3.12. According to Table 3.13 after configuration, the performance of the model and the accuracy has been slightly improved in Introversion (I) – Extroversion (E) category and considerably improved in Feeling (F) – Thinking (T) category. The accuracy in Intuition (I) – Sensing (S) and Judging (J) – Perceiving (P) categories, however, has been slightly decreased.

| Binary class | MBTI personality type | Accuracy after configuration | Accuracy before configuration | Difference |
|--------------|-------------------------------------|------------------------------|-------------------------------|------------|
| IE | Introversion (I) – Extroversion (E) | 79.01% | 78.17% | 0.84 |
| NS | Intuition (I) – Sensing (S) | 85.96% | 86.06% | - 0.1 |
| FT | Feeling (F) – Thinking (T) | 74.19% | 71.78% | 2.41 |
| JP | Judging (J) – Perceiving (P) | 65.42% | 65.70% | - 0.28 |

Table 3.13: Comparison of accuracy prediction before and after configuration

The scikit-learn library provides the ability of searching combinations of parameters and this capability was used in order to discover the optimal way to configure the model for achieving top performance. This is called Hyperparameter tuning in XGBoost model. As a result, the parameters including (1) the number and size of trees,

(2) the learning rate and number of trees, and (3) the row and column subsampling rates, are the parameters to consider when tuning.

The accuracy of prediction after configuration was also compared to one of the latest and most successful existing methods which used the same dataset. This method was introduced by Hernandez and Knight in 2017 and it has been discussed in section 3.4.1 of this thesis. They used various types of recurrent neural network (RNN) such as simple RNN, GRU, LSTM, and Bidirectional LSTM to build their classifier. For evaluation, they used two different methods, which were (1) post classification methodology and (2) user classification methodology. Table 3.14 compares the results of post classification methodology and user classification methodology.

| Binary class | MBTI personality type | Accuracy of post classification methodology | Accuracy of user classification methodology | Difference |
|--------------|--|---|---|------------|
| IE | Introversion (I) – Extroversion (E) | 54.0% | 67.6% | 13.6% |
| NS | Intuition (I) – Sensing (S) | 52.9% | 62.0% | 9.1% |
| FT | Feeling (F) – Thinking (T) | 57.8% | 77.8% | 20% |
| JP | Judging (J) – Perceiving (P) | 52.9% | 63.7% | 10.8% |

Table 3.14: Comparison of accuracy of post classification and user classification methods for recurrent neural network

According to Table 3.14, the accuracy of recurrent neural network model using user classification methodology was better than the recurrent neural network model using post classification methodology. Thus, the accuracy of Extreme Gradient Boosting was compared to their Recurrent Neural Network classifier using user classification methodology. Table 3.15 shows the results of this comparison.

| Binary class | MBTI personality type | Accuracy of Extreme Gradient Boosting | Accuracy of Recurrent Neural Network | Difference |
|--------------|-------------------------------------|---------------------------------------|--------------------------------------|------------|
| IE | Introversion (I) – Extroversion (E) | 79.01% | 67.6% | 11.41% |
| NS | Intuition (I) – Sensing (S) | 85.96% | 62.0% | 23.96% |
| FT | Feeling (F) – Thinking (T) | 74.19% | 77.8% | -3.61% |
| JP | Judging (J) – Perceiving (P) | 65.42% | 63.7% | 1.72% |

Table 3.15: Comparison of accuracy Extreme Gradient Boosting model and recurrent neural network model

Table 3.15 shows that the Extreme Gradient Boosting classifier in three dimensions of MBTI personality types has a greater degree of accuracy than recurrent neural network. Regarding the Intuition (I) – Sensing (S) and Introversion (I) – Extroversion (E) categories, the accuracy of the Extreme Gradient Boosting is significantly greater than recurrent neural network and for Judging (J) – Perceiving (P) category, the accuracy is slightly better. This is while the accuracy of recurrent neural network for Feeling (F) – Thinking (T) is considerably better than the Extreme Gradient

Boosting classifier. Thus, the overall performance of the Extreme Gradient Boosting classifier is better than the recurrent neural network for this specific dataset.

3.4.9 Employees' MBTI personality type prediction

As mentioned earlier, each employee's conversation with the software will be recorded in a separate file. The software will read each file separately and apply the created gradient boosting model to predict each employee's personality type. As a result, in the first step, the file will be pre-processed. Then the type indicators from each one of the four dimensions will be predicted and the result, which is one of the 16 personality types in MBTI model, will be presented. All employees' personality types will be recorded in a separate file for further analysis.

3.4.10 multiple choice questionnaire

An official MBTI questionnaire including 70 questions, which is able to identify the type of personality, was used in this part of the research. A sample of this questionnaire was used in Dartmouth Hitchcock Medical Centre and it was computerised in this research. This questionnaire is available in Appendix 2. As a result, the software starts another conversation with an employee by asking the first of 70 set questions. In general, 10 questions out of these 70 questions are trying to identify the dominant Meta program

in the first category which is Introversion (I) / Extroversion (E). Then for each one of the other three categories including Intuition (N) / Sensing (S), Thinking (T) / Feeling (F) and Judging (J) / Perceiving (P), there are 20 questions. After responding to all these questions, the software will identify the four basic Meta programs for this employee. Based on these Meta programs, the software will identify the employee's personality type which is one of the 16 MBTI personality types. In this step, four letters will be recorded in a list as the identifier for personality type of the employee. For instance, if the four basic Meta programs for this employee are Extroversion, Sensing, Thinking and Perceiving, the identifier list for the personality type of the employee will include 'E', 'S', 'T' and 'P'. This list will be used later in the comparison process. The result of this computerised questionnaire is discussed in the results and discussion section in chapter 4.

3.4.11 Designing questions

There have been 8 counters defined for four different categories of Meta programmes, which are Extroversion-Introversion (E-I), Sensation-Intuition (S-N), Thinking-Feeling (T-F), and Judgment-Perception (J-P). The relevant counter is increased based on the employee's answer to each question. Counters are compared to each other to understand the four basic Meta programmes and the identified Meta programme is recorded in a list. Figure 3.7 shows the scoring strategy for this computerised questionnaire.

| | Col 1 | | Col 2 | | Col 3 | | Col 4 | | Col 5 | | Col 6 | | Col 7 | | |
|----|-----------|---|-------|---|-------|-----------|-------|---|-------|---|-----------|---|-------|---|---|
| | A | B | A | B | A | B | A | B | A | B | A | B | A | B | |
| 1 | | | 2 | | 3 | | 4 | | 5 | | 6 | | 7 | | |
| 8 | | | 9 | | 10 | | 11 | | 12 | | 13 | | 14 | | |
| 15 | | | 16 | | 17 | | 18 | | 19 | | 20 | | 21 | | |
| 22 | | | 23 | | 24 | | 25 | | 26 | | 27 | | 28 | | |
| 29 | | | 30 | | 31 | | 32 | | 33 | | 34 | | 35 | | |
| 36 | | | 37 | | 38 | | 39 | | 40 | | 41 | | 42 | | |
| 43 | | | 44 | | 45 | | 46 | | 47 | | 48 | | 49 | | |
| 50 | | | 51 | | 52 | | 53 | | 54 | | 55 | | 56 | | |
| 57 | | | 58 | | 59 | | 60 | | 61 | | 62 | | 63 | | |
| 64 | | | 65 | | 66 | | 67 | | 68 | | 69 | | 70 | | |
| | Copy to → | | | | | Copy to → | | | | | Copy to → | | | | |
| | E | I | | | | S | N | | | | T | F | | J | P |

Figure 3.7: Scoring strategy for the computerized MBTI questionnaire

According to Figure 3.7 each question is related to one category of Meta programmes and attempts to identify the dominant Meta programme in that category. For instance, Figure 3.7 shows that question 8 is based on the Extroversion-Introversion (E-I) category. For each column, the number of A and B tick marks are counted, the total will be counted at the bottom. For columns 2, 4, and 6 the subtotals need to be combined. The totals for column 2 are copied to the spaces below the totals for column 3 and the same process occurs for columns 4 and 6. Totals will be added downwards to calculate the final sums. Finally, the letter corresponding to the highest number in the grand total is considered as the preference. Thus, the four letters combined constitute the personality type.

3.4.12 Comparison process for predicted personality types

There are 16 different lists defined for the software which are related to each one of the MBTI personality types. Each one of these lists includes four letters which are used as an identifier for each personality type. As a result, the identifier list for the personality type of the employee which was explained in section 3.4.10, will be compared with these 16 different lists which have been defined for the software, in order to see which matches the employee's list. Following this, the relevant information about that list is printed. This information is regarding the description of the personality type and the most appropriate careers for this type of personality.

3.4.13 File processing procedure for predicted personality types

The result of the questionnaire which is the personality type for each employee will be recorded as a separate file. The content of this file is compared with another file containing the result of gradient boosting model predictions for employees' personality types, in order to analyse the accuracy of the result. The system also analyses these files and compares the number of personalities which are recorded in these files. Based on this information, the most popular personality type in the organisation is identified and the manager is informed. This facilitates the process of understanding the organisational

culture, as the most popular personality type in the organisation is one of the most important factors in this process.

3.5 Participants and Data gathering procedure

The first phase of implementation, the automation of identifying the preferred representational system, has been tested on a group of 55 students at London Metropolitan University. The students are at different levels of study including 14 PhD researchers, 17 Masters students and 24 Bachelors students. Some of these students had additional work experience in industry. 31 participants were female and 24 were male, their ages ranging from 18 to 34. 23 participants were between 18 to 21 which is 41.8% of all participants. 14 participants were between 22 to 25, 12 participants were between 26 to 29 and finally 6 participants were between 30 to 34 years old. The majority of participants were in the first group with an age range of 18 to 21.

Before interaction with the software, the participants were fully informed about the software and its application. Participants were also notified that all the information provided will be private and remain confidential and not shared with any third parties. Additionally, that their answers would not be used for any other purposes other than analysing the accuracy and performance of the software. They were also made aware that there are no consequences such as financial loss, mental or physical harm as a result of their participation and their answers will not have any effect on their study or personal life. The duration of the conversation with the application ranged between 10 to 15

minutes, this depended on the length of their answers and their typing speed. They were also asked to respond to a multiple choice questionnaire for representational system preference, designed by Steve Antcliff in a book named 'Life coaching-made simple' published in 2009. This questionnaire contains 12 multiple-choice questions providing 4 possible answers and it is available in Appendix 1. Each answer indicates preference of a specific representational system but the order of answers is different for each question. For instance, in question 1, the first option is associated with the kinesthetic, the second to the auditory, the third to the auditory digital and the fourth answer to the visual representational system. In question 2, however, the order of association of the 4 answers to representational systems is the visual, the auditory, the kinesthetic and the auditory visual representational system respectively. Participants are asked to rate these options by choosing a number from 1 (As the least descriptive of them) to 4 (As the closest answer to describe them). This questionnaire has an answer sheet to calculate the final score for each representational system and the preferred representational system can be identified based on the highest score between the categories. Figure 2 shows the answer sheet for this questionnaire.

| Question Number | Visual | Auditory | Kinaesthetic | Auditory Digital |
|-----------------|--------|----------|--------------|------------------|
| 1 | d | b | a | c |
| 2 | a | b | c | d |
| 3 | a | d | b | c |
| 4 | b | d | a | c |
| 5 | c | d | b | a |
| 6 | c | a | d | b |
| 7 | c | a | b | d |
| 8 | d | b | c | a |
| 9 | b | c | d | a |
| 10 | a | c | d | b |
| 11 | d | a | c | b |
| 12 | b | c | a | d |
| total | V = | A = | K = | A _d = |

Figure 3.8: Answer sheet for the preferred representational system questionnaire

This questionnaire was chosen after considering four other popular representational system questionnaires offered in various documents including ‘NLP Home Study Programme (V2.0)’ published by Juiced Concepts Limited (2012), ‘Representational Systems’ published by Brefi Group Limited (2004), ‘Introduction to Neuro-Linguistic Programming’ published by Transformed Destiny (2015) and ‘The Power of Words’ written by Katy McAfee (2014). It was realised that these questionnaires are a simplified version of the chosen questionnaire with less number of questions with some not considering the auditory digital representational system. The results of this questionnaire were compared with the results of the intelligent software for similarities and examination of the software’s performance and accuracy.

The second phase of implementation, based on Meta model automation, has been tested on 50 participants with varying age, profession and lifestyle. Participants were fully informed about the function of the software and they were notified that the information provided will not be shared with any third parties and will remain private and confidential. They were also made aware that there are no mental or physical harm through participation in this study, and that they were not at risk of financial loss or a negative impact on their professional or personal life. The estimated time for this test ranged between 20 to 30 minutes depending on their typing speed and the length of their answers.

The third phase of implementation, consisted of automating the personality type prediction, was tested in UKSTUDY Company which is located in Brighton. This company consists of 14 employees located in the office and 10 other temporary project based employees. All employees contributed to the data gathering process. In the first stage, they were asked to communicate with the software and in the second stage, they were asked to respond to a computerised MBTI personality type indicator questionnaire. Before starting the data gathering process, all employees were informed that the provided information will remain completely private and confidential and not shared with any third parties. They were notified that their answers will not be used for any other purposes other than the analysis of the software performance. They were also made aware that there are no consequences as a result of their participation, such as financial loss or any mental or physical harm, nor any effect on their personal life. The duration of the conversation with the software was between 12 to 16 minutes, depending on their typing speed and the length of their answers. Moreover, the time taken for responding to the multiple choice questionnaire was between 8 to 11 minutes.

3.6 Data analysis procedure

After the data gathering process, the accuracy and the performance of the software were evaluated. Answers for each individual was recorded in a separate file and all representational system predicates were extracted from the text manually. These extracted words were then compared with the extracted words by the software that have been obtained automatically. Finally, the results for all 55 samples were calculated and these results were presented in figures shown in the results section. Moreover, the preferred representational system identified by the software were compared to the results of the manual questionnaires for similarity and to analyse accuracy.

3.7 Demonstration of the software

In this section, a worked example showing conversational extracts will be demonstrated using screenshots from the software. As it was explained in 3.2.2, the software starts the conversation with a brief introduction followed by response-based questions. Figure 3.9 shows the introduction and the beginning of the conversation.

```
Hi! My name is INPO!  
I'm an Intelligent NLP practitioner for Organisations!  
It's my pleasure to have a quick chat with you.I'm gonna ask a few questions about you and your organisation.  
  
Please try to say the first thing that comes to your mind and then explain your answers clearly and in full details.  
Remember that the more information you give me, the more accurate result you will get.  
  
These questions will help me to analyze and understand the current model of your personality.  
I will help you to understand how you process and store information in your mind.  
I will help you to understand your preferred representational systems and learning style.  
  
By understanding this, you can improve your communication with your colleagues or customers.  
It can also help you to learn different things faster and have a more effective conversation.  
It will help to transfer or teach your knowledge to other colleagues in a more effective way as well.  
  
Well, Lets start!  
First of all, could you please describe your work environment with a few sentences?  
It can be anything about your feeling about your job, your relationship and cocommunication with your colleagues  
and your manager, your tasks and responsibilities, the way that you are doing your job or etc.
```

Figure 3.9: Software beginning the conversation

Figure 3.9 shows that the software introduces itself in the first step. It has been named ‘INPO’ short for intelligent NLP practitioner for organisations. Then, the software gives an instruction to the user about the structure of the conversation and answering the questions, so the user will get the most accurate result. The software is asking the first question in the next step. This question is as follows:

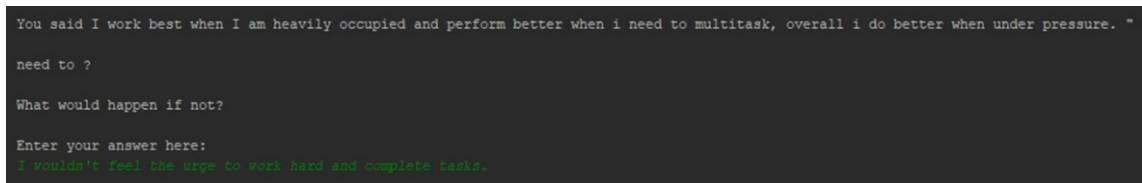
“First of all, could you please describe your work environment with a few sentences? It can be anything about your feeling about your job, your relationship and communication with your colleagues and your manager, your tasks and responsibilities, the way that you are doing your job or etc.”

In this example, the user responded to this question as follows:

“At times, it can get very busy and overwhelming and other times it is quite

relaxed. I work best when I am heavily occupied and perform better when I need to multitask, overall I do better when under pressure.”

The software recognised that there is a generalisation in the user’s response because of detecting a necessity identifier. As a result, the related recovery question was asked as shown in figure 3.10.



```
You said I work best when I am heavily occupied and perform better when i need to multitask, overall i do better when under pressure. "
need to ?
What would happen if not?
Enter your answer here:
I wouldn't feel the urge to work hard and complete tasks.
```

Figure 3.10: Software detecting generalisation

Figure 3.10 shows that after recognising distortion in the user’s response, the software repeated that sentence for the user and asked the following question:

“You said I work best when I am heavily occupied and perform better when I need to multitask, overall I do better when under pressure.

Need to?

What would happen if not?”

The user responds to this question for more clarification about the generalised information. Figure 3.10 shows the user’s response was as follows:

“I wouldn’t feel the urge to work hard and complete tasks.”

After the clarification by the user, the software continued the conversation. In another part of the conversation, the user stated that:

“Because there is diversity and flexibility but not a lot. There are different types of tasks that are required to be completed but there are only a handful.”

Figure 3.10 also shows that the software recognised another generalisation in the first sentence of the user’s response because of detecting an impossibility identifier. Thus, the relevant recovery question was asked which was as follows:

*“You said because there I diversity and flexibility but not a lot.
Why not? Can you explain more?”*

The user responded to this question as follows:

“Because there is only a number of things that need to be done, so it is not highly varied but a little.”

Following the previous recovery question, the software asked another recovery question related to the distortion in the user’s response:

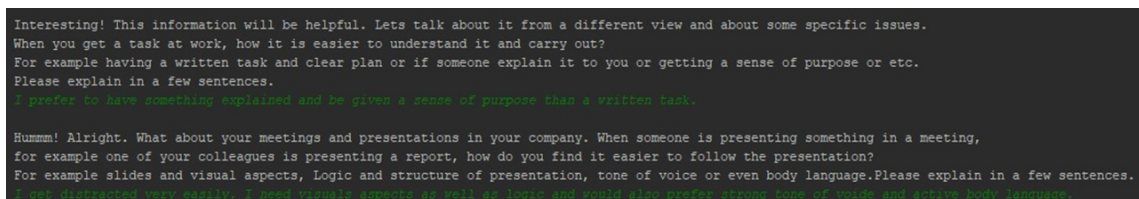
“You said there different types of task that are required to be completed but they are only a handful.

Only? How do you know?”

The user responded to this question as follows:

“Because there is not a hug number of tasks but a small number.”

In this step, the software realised that there is no need for more clarification. As a result, the conversation was continued through asking the next question from the 10 set questions. This is shown in figure 3.11.



```
Interesting! This information will be helpful. Lets talk about it from a different view and about some specific issues.
When you get a task at work, how it is easier to understand it and carry out?
For example having a written task and clear plan or if someone explain it to you or getting a sense of purpose or etc.
Please explain in a few sentences.
I prefer to have something explained and be given a sense of purpose than a written task.

Hummm! Alright. What about your meetings and presentations in your company. When someone is presenting something in a meeting,
for example one of your colleagues is presenting a report, how do you find it easier to follow the presentation?
For example slides and visual aspects, Logic and structure of presentation, tone of voice or even body language.Please explain in a few sentences.
I get distracted very easily, I need visual aspects as well as logic and would also prefer strong tone of words and active body language.
```

Figure 3.11: Software detecting no deletion, distortion or generalisation

Figure 3.11 shows that the software asked the next question as follow:

“Interesting! This information will be helpful. Lets talk about it from a different view and about some specific issues.

When you get a task at work, how it is easier to understand it and carry out?

For example, having a written task and clear plan or someone explain to you or getting a sense of purpose or etc.”

The user responded as follows:

“I prefer to have something explained and be given a sense of purpose than a written task.”

In this step, the software did not recognise any deletion, distortion or generalisation. That is why there is no recovery question asked. The software asked the next question from the 10 set questions to continue the conversation and collect more information. Figure 3.11 shows that the next question asked by the software was as follows:

*“Hummm! Alright. What about your meeting and presentations in your company. When someone is presenting something in a meeting, for example one of your colleagues is presenting a report, how do you find it easier to follow the presentation?
For example, slides and visual aspects, logic and structure of presentation, tone of voice or even body language? Please explain in a few sentences.”*

Following this question the user responded as follows:

“I get distracted very easily. I need visual aspects as well as logic and would also prefer strong tone of voice and active body language.”

The software continued the conversation through asking the next question from the 10 set

questions as follows:

“Well, what do you think about your discussion with your manager or your colleagues? What are you most often influenced by?

For example the other person’s point of view, their logic, their tone of voice or maybe their feeling. Please explain in a few sentences.”

The user responded as follows:

“Their logic and feeling.”

In this step, a deletion was detected by the software and the relevant recovery question was asked as follow:

“You said their logic and feeling.

Their?

Who exactly?”

Figure 3.12 shows that the user clarified the statement by saying that the previous response was about his/her manager. After this response, the conversation was continued through asking the next question from the 10 set questions, as the software did not recognise any more clarifications to be needed.

```

Well, what do you think about your discussions with your manager or your colleagues? What are you most often influenced by?
For example the other person's point of view, their logic, their tone of voice or maybe their feeling. Please explain with a few sentences.
their logic and feeling
You said " their logic and feeling "

their ?

Who exactly?
The manager.
The manager

Alright! Imagine you are in a situation that you need to make an important decision.
How do you usually make important decisions? What are your important decisions based on?
Please explain in a few sentences.
I write a lists of options, they are based on my priorities.

```

Figure 3.12: Software detected delition

During the conversation, the software was recording the representational systems predicates at the same time with application of the Meta model in conversation. Figure 3.13 shows that after finishing the conversation, the software informed the user about his/her preferred representational system as well as a description and recommendations to improve his/her communication and learning in the organisation.

```

You are an Auditory Digital person.
As an Auditory Digital person you will learn best by:

Working things out in your mind.
Having as many facts available as possible.
Having figures available.
Having as much information as possible to consider and make conclusions about.
Like handouts.
Enjoy additional reading material.
Links to additional information and resources to research for yourself.
Enjoy project work and research.
Memorize by steps, procedures, sequences.
Talk to yourself and carry on conversations with other people in your mind.

auditory 4
digital 3
kinesthetic 1
visual 1
Realists who are quick t make practical decisions.
This person can be a supervisor.
Some examples of appropriate jobs for this type of personality:
1- Auditor2- Accountat3- Chief financial officer4- Web developer engineer5- Government employee

```

Figure 3.13: Final report provided by software

According to figure 3.13, this user was an auditory digital person. The software also reports the number of identified predicates related to each one of the representational systems. Figure 3.12 also shows in another part of the report, that the user was informed about his/her personality type and was given a description about this personality type and the most appropriate jobs related to this type of personality. Figure 3.13 shows that the software suggested the following about the personality type of this user:

“Realists who are quick to make practical decisions.

The person can be a supervisor.

Some examples of appropriate jobs for this type of personality:

1- Auditor 2- Accountant 3- Chief financial officer 4- Web developer engineer 5- Government employee”

3.8 Summary

The methodology used in this research have been presented in this chapter. The implementation process in this research was divided into three different phases and each phase has been explained separately. The first phase was regarding the automation process of identifying the preferred representational system and it was presented in 5 different sections in this chapter. The second phase of implementation was based on the Meta model automation and it was presented in 10 different sections. Finally, the third phase was about automating the personality type prediction. Prior to the explanation of this phase of implementation, other related methodologies for personality type prediction were

discussed and different classification techniques were described. Followed by this, the methodology and the machine learning technique used in this research were presented and described in 13 different sections. Then, the data gathering procedure and data analysis strategy in this research were explained and a demonstration of the software was presented at the end of this chapter.

4 Results and discussion

4.1 Introduction

Different parts of the developed system have been tested separately and the results will be presented in this chapter. First, the results of representational system identification will be described and then these results will be compared with the results of a questionnaire which was used to identify the preferred representational system manually. Next, the results of automated Meta model system will be presented. A NLP practitioner did the process of applying the Meta model and in the next step, the system achievements were compared with the human achievements. The third part of the system which is personality type prediction, was tested in a company and the results will be presented in the next section. A computerized MBTI personality type indicator questionnaire was used in the company as well and the results of questionnaire will compare with the result of automated personality type predictor system. The representational system identification and the Meta model system were also used in that company and the results will be described in this chapter.

4.2 Results of automated representation system identification

As it was explained in section 1.2, one of the main questions in this research was how the process of identifying the preferred representational system of a person and the most popular preferred representational system in an organisation can be completed

automatically using a software, using Neuro Linguistic Programming techniques. This research question was defined because of two objectives mentioned in section 1.1, that this research was trying to achieve. These objectives include:

- Automating the recognition of the preferred representational system of each employee and the most popular preferred representational system in the organisation for application in improving communication and teamwork in organisations.
- Eliminating the contributing human factors and errors such as lack of skill and experience, personal judgment and opinion, inaccuracy or mistakes of NLP practitioners, from the process of applying Neuro Linguistic Programming.

Thus, this experiment was designed to compare the performance of the software and human in this aspect to make sure the objectives were achieved. As previously mentioned in data gathering process, participants were asked to respond to a questionnaire before use of the software. The result of questionnaire is shown in figure 4.1.

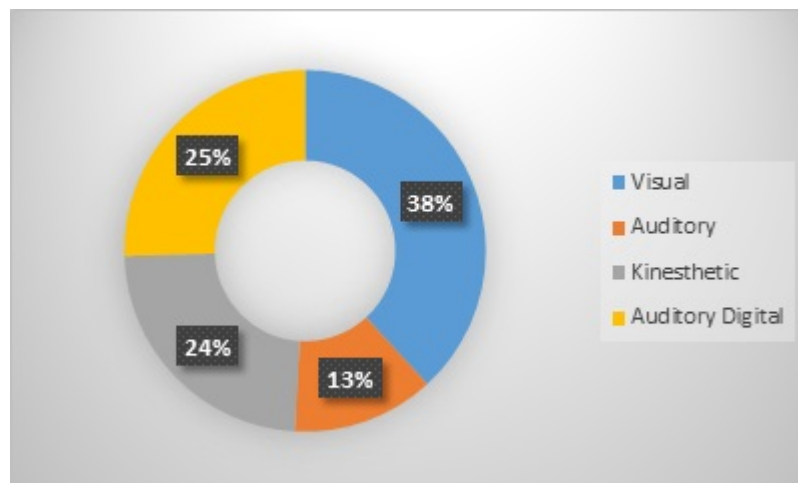


Figure 4.1: Questionnaire results for the preferred representational system of participants

Analysing the questionnaire results, it was noted that the preferred representational system for 21 participants was visual, for 7 participants auditory, for 13 participants kinaesthetic and for 14 participants auditory digital. According to figure 4.1, 38% of participants were visual, 13% were auditory, 24% were kinaesthetic and 25% were auditory digital. The achieved result by the software is shown in figure 4.2.

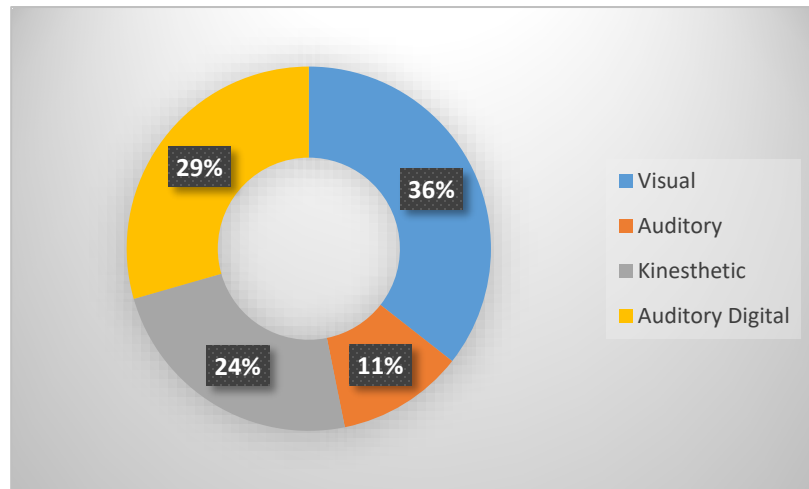


Figure 4.2: Software results for the preferred representational system of participants

After analysing the software results, 20 participants have been identified with visual preference, 6 participants with auditory preference, 13 participants with kinaesthetic preference and 16 participants with auditor digital preference. According figure 4.2, 36% were visual, 11% auditory, 24% kinaesthetic and 29% auditory digital. A comparison of the results of the manual questionnaire and the results obtained by the software was carried out. This is shown in figure 4.3.

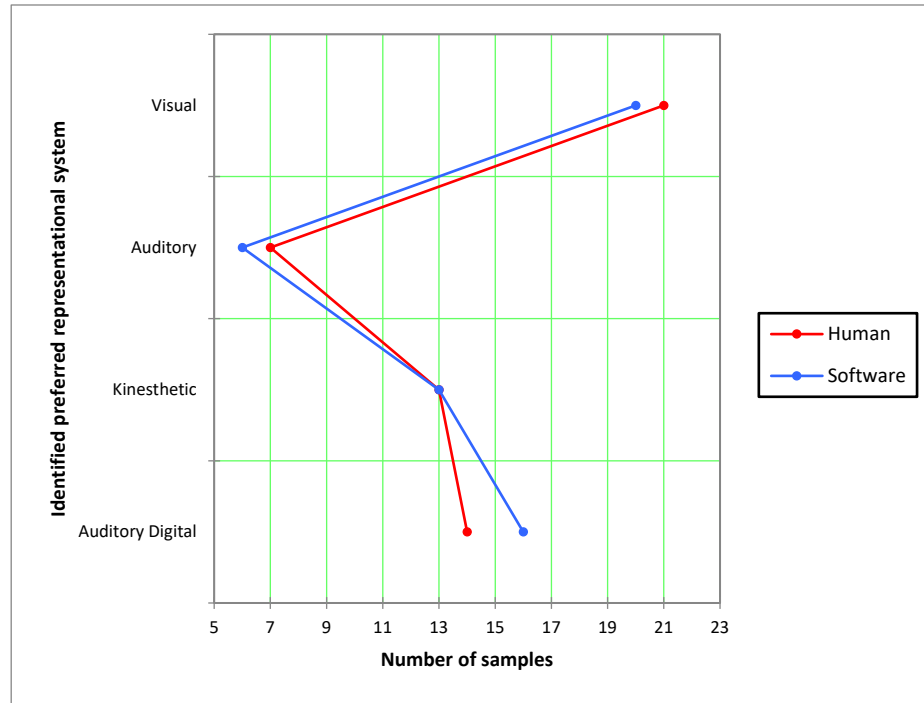


Figure 4.3: Comparing the number of preferred representational systems identified by a questionnaire and software

As shown in figure 4.3, the numbers of participants identified as preferring a specific representational system via the two methods is very close. Only one subject being in different categories of representational systems comparing the manual questionnaire with the software. While the questionnaires' results identify 21 participants with preference of the visual representational system while the software's results identify 20 participants as visual people. Figure 4.3 also shows that the number of kinaesthetic participants is the same for the questionnaire and the software. Furthermore, there is a difference of 2 people between the questionnaire's results and the software's results for auditory digital participants.

According to table 4.1, the manual questionnaire has recognised 38.18% of participants with visual preference, 12.72% with auditory preference, 23.63% with kinaesthetic preference and 25.45% with auditory digital preference. While the software has recognised 36.36% with visual preference, 10.90% with auditory preference, 23.63% with kinaesthetic and 29.09% with auditory digital preference. This reveals -1.82% difference between the manual questionnaire and the software results for the visual preference and -1.82% difference for auditory preference. For kinaesthetic preference, the percentage for the software is exactly the same as manual questionnaire. Finally, the difference between the software and the manual questionnaire for auditory digital preference is 3.64%. These percentages could provide further proof that the software performs well.

| | Visual | Auditory | Kinaesthetic | Auditory Digital |
|---------------|--------|----------|--------------|------------------|
| Questionnaire | 38.18% | 12.72% | 23.63% | 25.45% |
| Software | 36.36% | 10.90% | 23.63% | 29.09% |
| Difference | -1.82% | -1.82% | 0% | 3.64% |

Table 4.1: Difference between questionnaire and automated representational system identifier results

In summary, these findings demonstrated an improved performance in identification of the auditory digital representational system and matched the questionnaire in identifying the kinaesthetic representational system preference. The

visual and auditory representational systems had the difference of one participant being identified differently. These results are significant in demonstrating the optimal performance of the software.

There are a variety of theories regarding the proportions of people preferring each one of the preferred representational systems. For instance, Matthews (2016) predicts that 50% of people are visual, 30% are kinaesthetic and 20% are both auditory and auditory digital. While Rayner Institute (2015) estimates that 30% of people are auditory digital, 40% are kinaesthetic, 10% are auditory and 20% are visual. It can be assumed, therefore, that the community considered may have an impact on the results obtained. According to Kory (2001) auditory digital is used most often by people with a more academic or professional impression. In this research, all participants were from an academic community and the achieved result is confirmative of Kory's statement, where the most popular representational system was the auditory digital in comparison to other representational systems.

In analysis of the obtained results, the number of predicates or words that have been referring to each one of the representational systems in the participants' answers were also evaluated. There were 974 words identified as predicates in the questionnaire responses and the result for each one of representational systems is shown in figure 4.4.

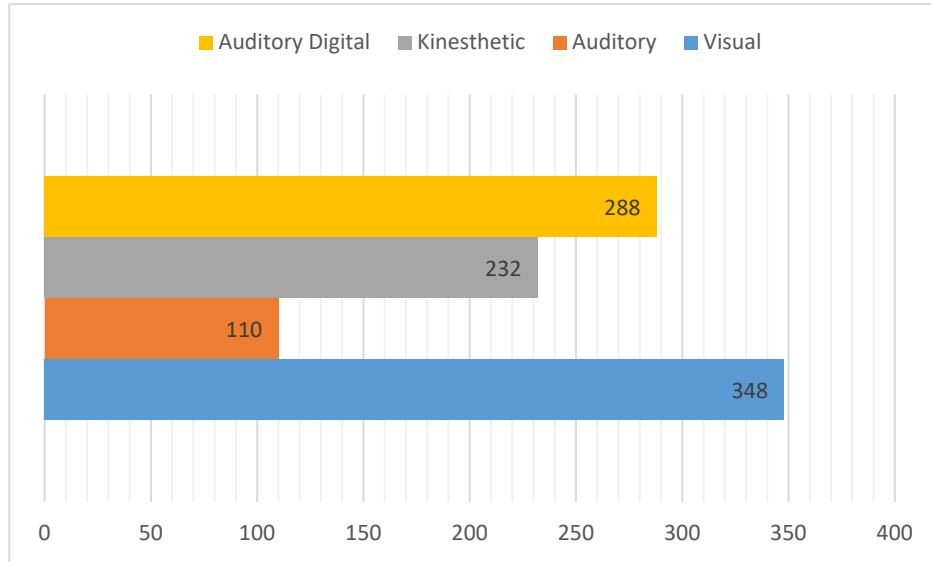


Figure 4.4: Number of predicates related to each representational system identified by human

Results show that out of 974 predicates, 351 predicates were associated to the visual representational system, 107 predicates were to the auditory, 234 predicates to the kinaesthetic and 282 predicates to the auditory digital representational system. Table 4.2 shows that 36.03% of words have been associated to visual preference, 10.98% to auditory, 24.02% to kinaesthetic and 28.95% to auditory digital.

| | Visual | Auditory | Kinaesthetic | Auditory Digital |
|-----------------------|--------|----------|--------------|------------------|
| Number of word | 351 | 107 | 234 | 282 |
| Percentage | 36.03% | 10.98% | 24.02% | 28.95% |

Table 4.2: Recognised predicates by human

The number of predicates identified by the software were 978 words. The results are shown in figure 4.5.

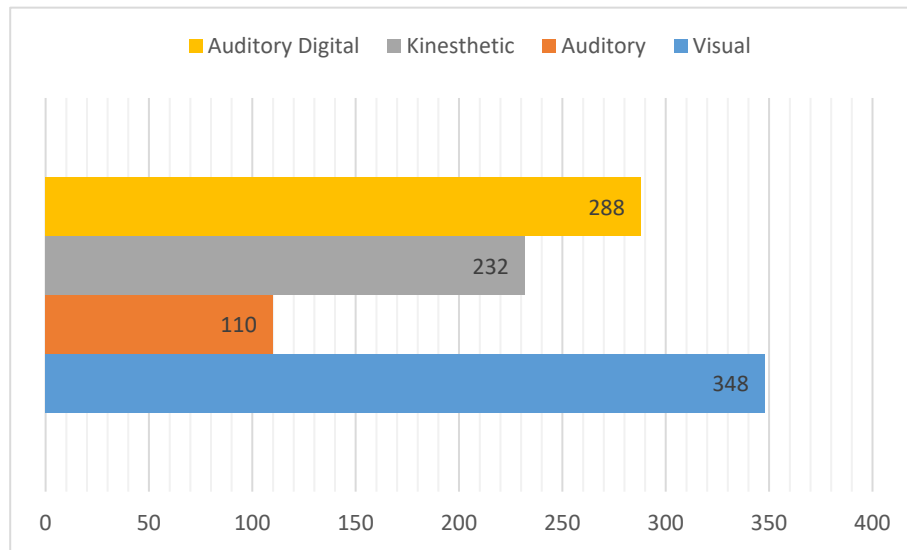


Figure 4.5: Number of predicates related to each representational system identified by software

As presented in figure 6, out of 978 predicates, 348 were associated to the visual representational system, 110 predicates to auditory, 232 to kinaesthetic and 288 predicates to the auditory digital representational system. Table 4.3 shows that 35.58% of words have been associated to visual preference, 11.24% to auditory, 23.72% to kinaesthetic and 29.44% to auditory digital.

| | Visual | Auditory | Kinaesthetic | Auditory Digital |
|-----------------------|--------|----------|--------------|------------------|
| Number of word | 348 | 110 | 232 | 288 |
| Percentage | 35.58% | 11.24% | 23.72% | 29.44% |

Table 4.3: Recognised predicates by software

Finally, the number of predicates identified manually were compared to the number of predicates identified by software and the results shown in figure 4.6.

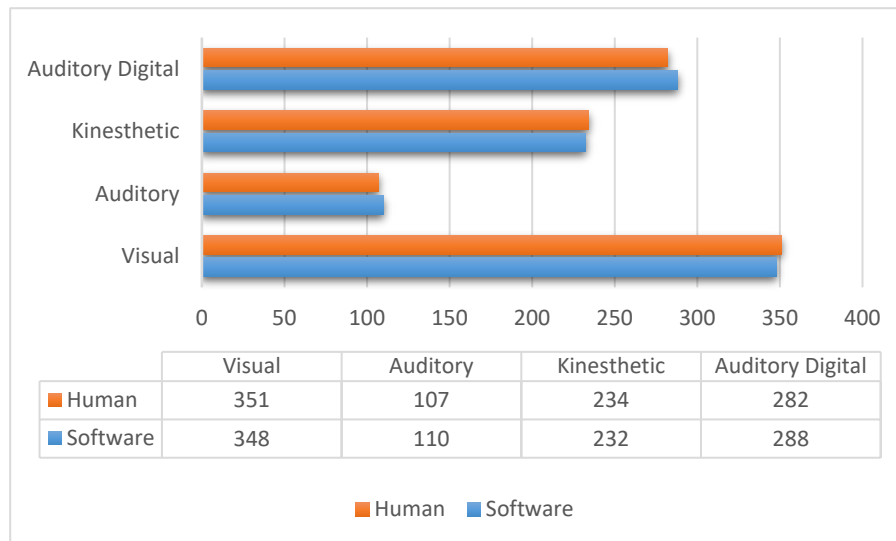


Figure 4.6: Comparing the number of predicates identified by human and software

The most notable difference presented in figure 4.6 is the increased number of predicates recognised by the software in comparison to the manual questionnaire conveying an increased accuracy and competence by the intelligent software. The definition of competence being the capability of the software to satisfy the objectives of

its purpose. The figure also shows that in recognition of preference of the visual representational system, the manual route has been relatively better than the software in terms of identifying the relevant predicates where 351 predicates were identified compared to the 348 predicates identified by the software. For the auditory representational system, 107 predicates have been identified manually and 110 predicates were identified by the software showing better performance by the latter. For the kinaesthetic representational system, the data is similar to the visual representational system and the manual system has been slightly better than the software whereby 234 predicates have been identified manually while 232 predicates have been identified by the software. Finally, for the auditory digital representational system, the software has been more successful, identifying 6 more predicates. The number of identified predicates by the questionnaire was 282 while the software recognised 288. Table 4.4 shows the percentages and the difference between the manual performance and the software.

| | Visual | Auditory | Kinaesthetic | Auditory Digital |
|------------|--------|----------|--------------|------------------|
| Human | 36.03% | 10.98% | 23.63% | 28.95% |
| Software | 35.58% | 11.24% | 23.63% | 29.44% |
| Difference | -0.45% | 0.26% | 0% | 0.49% |

Table 4.4: Comparing the human and software performance

Moreover, the number of visual predicates for each person, recognised by software and human were calculated separately and a ‘time series visualisation’ diagram has been

used to compare the results. This is shown in Figure 4.7.

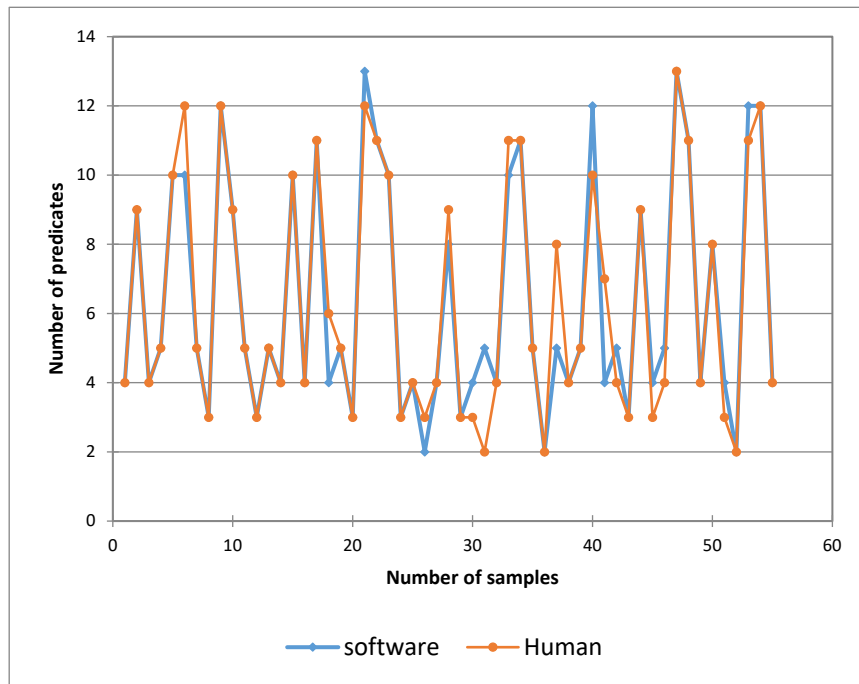


Figure 4.7: Comparing the number of visual predicates, identified by the human and software

Figure 4.7 shows 39 out of 55 people to have been identified with the same number of visual predicates. For 10 cases, the difference in the number of recognised visual predicates was only 1 word. This difference for 3 of the cases was 2 words, and for the remaining three, by 3 words.

As mentioned before, figure 4.3 shows that out of 55 people, the software identifies 20 people with the visual preferred representational system while results acquired by humans recognised 21 cases. Considering this number and the data presented in figure 4.7, it could be assumed that both human and software have identified the visual

preference for the same cases, even if there is a small difference between the number of recognised visual predicates for each person.

Figure 4.8 shows the result of comparing the number of recognised auditory predicates for each person, identified by software and human.

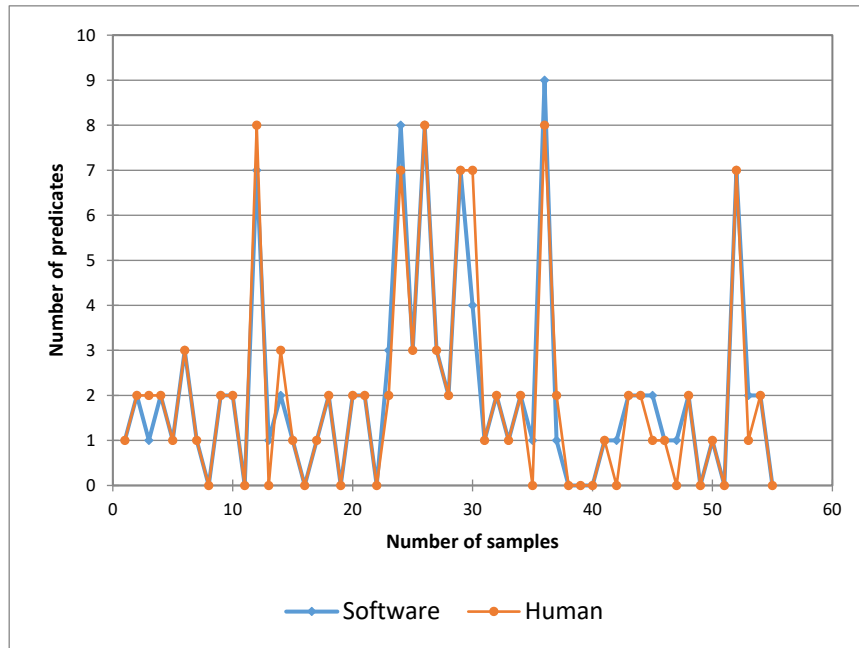


Figure 4.8: Comparing the number of auditory predicates, identified by the human and software

Figure 4.8 shows that 41 out of 55 people have been identified with the same number of the recognised auditory predicates. 13 people have been recognised with the difference of only one word and only one person was identified with the difference of 3 words.

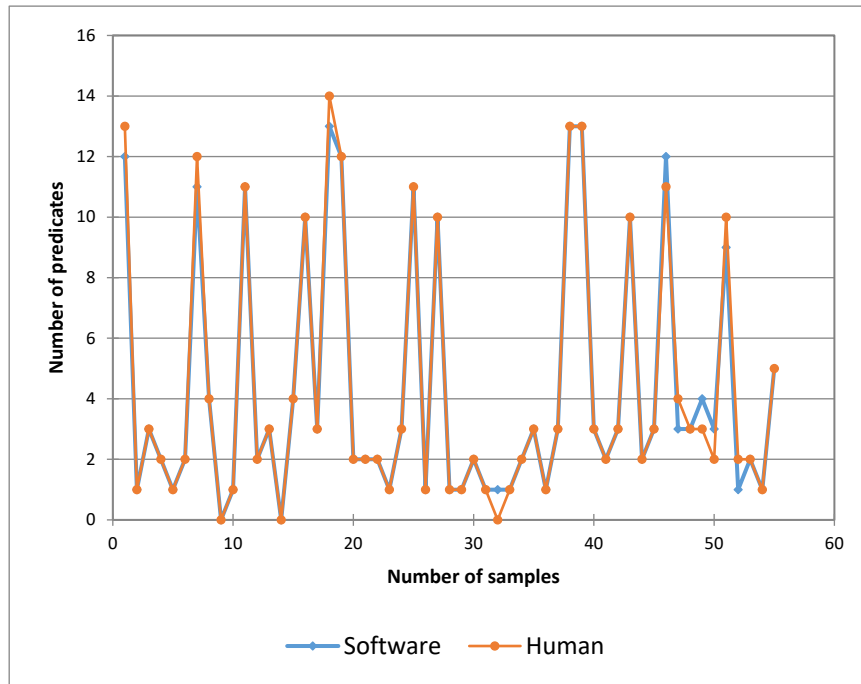


Figure 4.9: Comparing the number of kinaesthetic predicates, identified by the human and software

According to figure 4.9, the performance of software and human in regards to identifying the kinaesthetic preference has been very similar and the number of people who have been identified with the same number of recognised kinaesthetic predicates is 45 out of 55. Figure 4.9 shows that the other 10 people have been identified with the difference of only one word. Finally, figure 4.10 shows 35 out of 55 people to have been identified with the same number of recognised kinaesthetic predicates. There are 17 people who have been identified with the difference of one word between the software results and the human results. One person has been identified with the difference of two words, one person with the difference of 3 words and one person with the difference of 4 words.

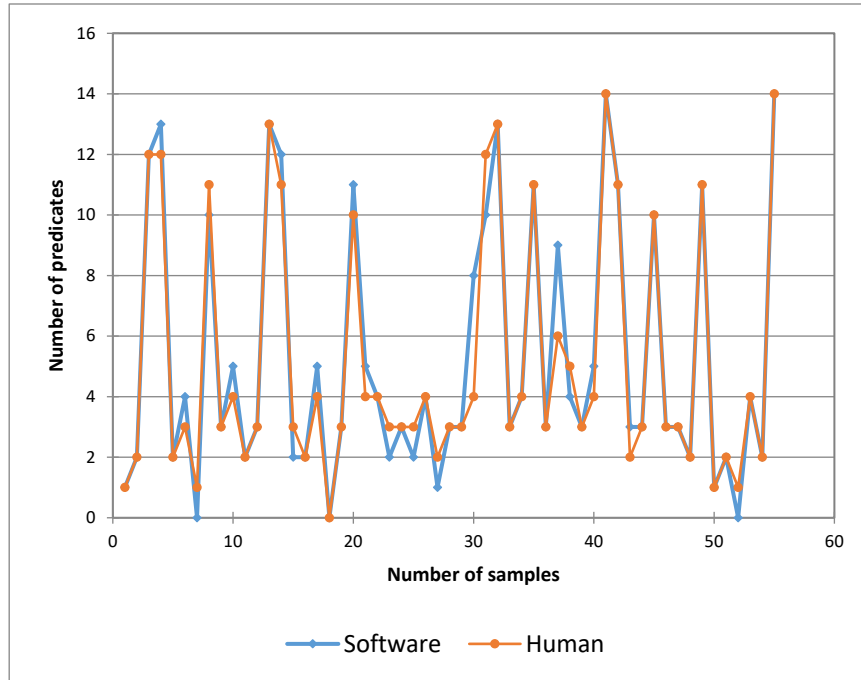


Figure 4.10: Comparing the number of auditory digital predicates, identified by the human and software

After analysing all predicated recognised by both software and human, three main factors have been recognised to contribute to the difference between the number of predicates associated to each representational system, identified by the software and human. The first reason for this is the limitation of words in the software dictionary, which impacted the efficiency of the software. The second factor contributing to the difference of performance is the human errors involved leading to some predicates being identified by the software but not by the human. Finally, synonyms of words used in some cases had been recorded in association to other representational systems. As a result, the recognised predicate has been considered as an indicator for one of the other representational systems

and not for the one which has been identified by human. The identification of the preferred representational system is based on detection of many predicates from multiple answers to questions to ensure accuracy. This ensures that the recognition of a predicate with error does not have an overall impact on correct identification of the preferred representational system.

Overall, results suggest that the software is able to replicate the human performance and recognise the preferred representational system correctly based on the language used by an individual. The results also present more accuracy in performance in some aspects of this process in comparison to the manual alternative. However, there was an important matter to be considered during the data gathering process by the software. The efficiency of the automatic software increases by a higher word count of answers where the response includes more detail, and consequently more words. Participants were informed about this and asked to respond to each question clearly, explaining their answers vividly. Accordingly, the questions were designed in a way to encourage participants to give full answers, using the maximum number of words. Nonetheless, some of our participants were responding to some of questions in short sentences with the minimal words. This was expected to cause an issue with potentiating the software to be accurate in its analysis.

Interestingly, despite some participants responding to some of the questions with very few words, the overall result was accurate nonetheless, and very similar to the result of the manual questionnaire. The reason for this could be the number of questions presented for the purpose of the same examination in order to ensure an accurate prediction. Moreover, a variety of questions considering different aspects of personality were available to the participants. Hence, the participant was invited to thinking about one

concept from different views and also attempt to explain it in distinctive ways. As a result, the overall result achieved was correct, even if some of their answers were restrictive for the software's analysis.

The results of this experiment demonstrate achieving the objective of eliminating human error by replacing the dependence on the judgement of the human practitioner by an intelligent software. This experiment also demonstrates that this was achieved without the compromise of the accuracy of results since they are very close in outcome with minimal shortcoming of the software and even demonstrated enhanced performance in one area. Another aspect to be considered is the possible justification for the difference observed between the software and human especially in the case of the difference being so small. It could be argued that the close resemblance of performance could validate the capability of the software in replicating the performance of a qualified NLP practitioner. This, therefore, leaves the question of which being the true analysis considering the close resemblance of the results and the only significant difference being the existence of the possibility of human error in the case of the human NLP practitioner.

4.3 Results of automated Meta model system

According to section 1.1, two of the research objectives included:

- Providing a tool for employees and managers to use in order to identify any personal, communicational and organisational problems in the organisation.

- Eliminating the contributing human factors and errors such as lack of skill and experience, personal judgment and opinion, inaccuracy or mistakes of NLP practitioners, from the process of applying Neuro Linguistic Programming.

According to these objectives, the following research question was defined in section 1.2:

- How to apply the NLP Meta model automatically during a conversation between a human and computer using Natural Language Processing techniques

As a result, this experiment was designed to compare the performance of the software and human practitioner to make sure the mentioned objectives were achieved. After the data gathering process, the conversations between the software and participants were analysed by a NLP practitioner (human). The results were compared to the software for examining the accuracy of the software's results and evaluating its performance. The software identified 904 deletions, 328 distortions and 452 generalisations. The number of deletions identified by the NLP practitioner, on the other hand, were 703, in addition to 542 distortions and 351 generalisations. In other words, 54% of the recovery questions by the software were associated to deletion, 19% to distortion and 27% to generalization, as demonstrated in Fig. 4.11.

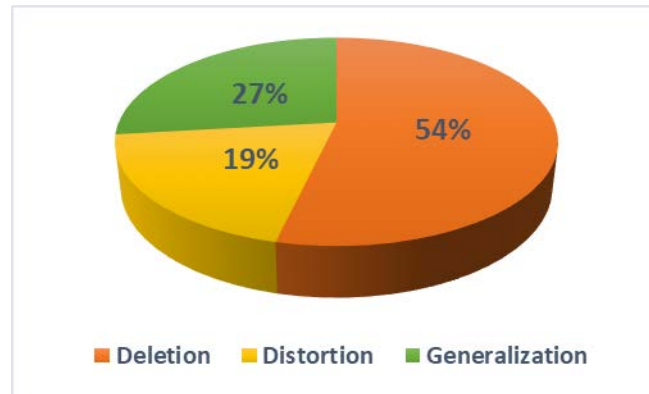


Figure 4.11: Number of recovery questions about deletion, distortion and generalization, asked by the software

For the identified processes by the NLP practitioner, 23% were related to deletion, 29% were related to distortion and 48% were related to generalization. This is shown in Fig. 4.12.

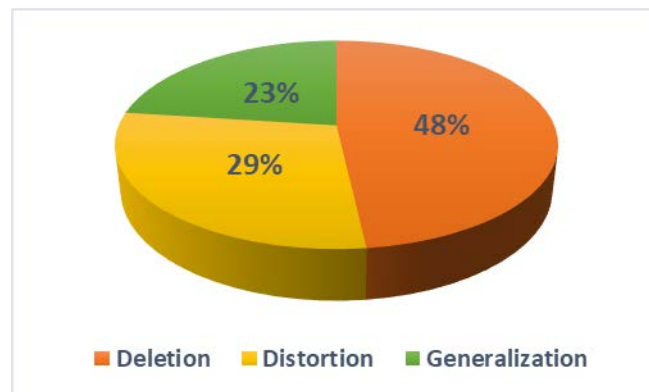


Figure 4.12: Number of recovery questions about deletion, distortion and generalization, asked by human

The number of identified deletions, distortions and generalizations by the software were compared to the NLP practitioner, as shown in Fig. 4.13.

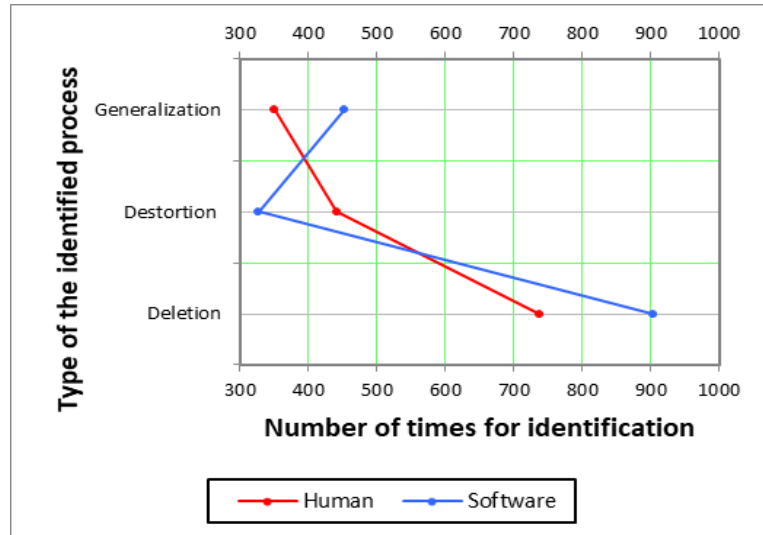


Figure 4.13: Comparing the number of Deletion, Distortion and Generalization, identified by the human and software

According to Fig. 4.13, the software had a better performance than the NLP practitioner, in identifying the deletion processes. Table 4.5 shows that the software’s performance in this regard was 6% more competent than that of the NLP practitioner. Fig. 4.13 also shows, however, that the software was not as successful as the NLP practitioner in identifying the distortion processes by 10%, as seen in Table 4.5. Finally, the software was also more effective with regards to identifying the generalization processes. Table 4.5, demonstrates this difference to be by 4%.

| | The identified process | | |
|-------------------|------------------------|-------------------|-----------------------|
| | <i>Deletion</i> | <i>Distortion</i> | <i>Generalization</i> |
| Software | 54% | 19% | 27% |
| Human | 48% | 29% | 23% |
| Difference | 6% | -10% | 4% |

Table 4.5: Comparing the performance of the software and human

The number of recovery questions related to each category of the deletion, distortion and generalisation processes were also recorded. Figure. 4.14 shows that 398 questions were regarding unspecified nouns, 202 questions were regarding to comparative deletions and 304 questions were regarding to ‘Ly’ adverbs in the user-software conversation. On the other hand, 278 questions were regarding to unspecified nouns, 167 questions were regarding comparative deletions and 293 questions were ‘Ly’ adverbs, in the case of the NLP practitioner.

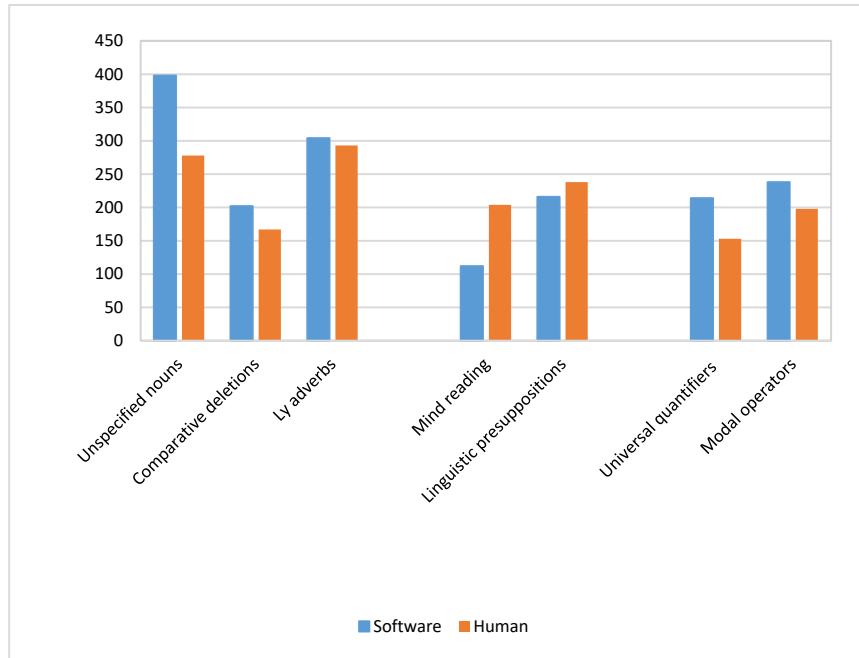


Figure 4.14: Comparing the number of recovery questions related to each category of deletion, distortion and generalisation processes, asked by the software and human.

According to Table 4.6, there is a 6% difference between the performance of the software and the NLP practitioner, in favour of the software. Table 4.6 also shows that there is no difference between the result of the software and the practitioner with regards to the identification of comparative deletions, the performance of the practitioner however was 6% better than the software regarding the recognition of ‘Ly’ adverbs.

| | Deletion | | |
|-------------------|--------------------------|------------------------------|-------------------|
| | <i>Unspecified nouns</i> | <i>Comparative deletions</i> | <i>Ly adverbs</i> |
| Software | 44% | 22% | 34% |
| Human | 38% | 22% | 40% |
| Difference | 6% | 0% | -6% |

Table 4.6: Comparing the number of recovery questions related to each deletion category asked by the software and human

The number of recovery questions relating to the distortion process were analysed. Figure. 4.14 and Table 4.7 demonstrate that 112 (34%) recovery questions asked by the software were about mind reading in the distortion process, whilst 216 (66%) questions were about the linguistic presuppositions. This is while the practitioner asked 204 (46%) questions in relation to mind reading and 238 (54%) questions related to linguistic presuppositions. According to Table 4.7, the software performed better than the practitioner regarding identification of the linguistic presuppositions but the practitioner performed better than the software in relation to the identification of mind reading. This is because of the work coverage in the lists related to the distortion process. Thus, the human practitioner was considering additional elements for the identification of mind reading in the distortion process and as a result, had a better performance in this aspect in comparison to the software. There is also another reason which is related to the experience of the practitioner. A master practitioner contributed to this research who has extensive experience and is qualified for training of regular practitioners of NLP. Thus,

the possibility for having a performance affected by any human error, personal opinion, personal judgment or lack of experience is very low and the performance is better than the software.

| | Distortion | |
|-------------------|---------------------|-----------------------------------|
| | <i>Mind reading</i> | <i>Linguistic presuppositions</i> |
| Software | 34% | 66% |
| Human | 46% | 54% |
| Difference | -12% | 12% |

Table 4.7: Comparing the number of recovery questions related to each distortion category asked by the software and human

Finally, the generalisation recovery questions were analysed, which demonstrated 214 questions related to universal quantifiers and 238 questions related to modal operators. This is while the practitioner asked 153 questions about universal quantifiers and 198 questions about modal operators. Table 4.8 shows that the performance of the practitioner was 12% better than the software in recognising modal operators, whereas the performance of the software was 12% better than the practitioner in recognising universal quantifiers.

| | Generalization | |
|-------------------|------------------------|------------------------------|
| | <i>Modal operators</i> | <i>Universal quantifiers</i> |
| Software | 53% | 47% |
| Human | 56% | 44% |
| Difference | -12% | 12% |

Table 4.8: Comparing the number of recovery questions related to each generalization category asked by the software and human

The average time for the software to process and analyse the participants' statements and respond accordingly did not surpass 1 second. This reflects the increased efficiency of the software in comparison to the manual alternative, where the practitioner would require more time to read and comprehend the participants' statements in order to respond appropriately.

Availability of some participants was a limitation that may have influenced this study. As described, the software was tested on 50 participants. Although the outlined outcome is comprehensive, 100 or more participants may have further improved the results (Faber and Fonseca, 2014; Martinez-Mesa et al, 2014).

Overall, the objectives addressed by this experiment were achieved through demonstrating improved performance of the software in comparison to the human practitioner with the occurrence of diminished performance being justified. This reflects the establishment of reliability in substituting the human practitioner with an intelligent software, where the inherent human errors of the practitioner can be eliminated without

conceding on the application of the Metal model with accuracy. It should also be considered that although the potential of a certain improvement exists in substituting the human practitioner with a software, the shortcomings of identifying distortions by the software has to be addressed for this substitution to be an advantageous adjustment.

4.4 Results of automated personality type prediction

In section 1.1, two objectives were defined as follow:

- Automating the recognition of the personality type of each employee and the most popular personality type in the organisation for application in improvement of the task allocation process in organisations
- Increasing the accuracy, reliability and efficiency of the current methods for personality type prediction.

Based on these objectives the following research questions were defined in section 1.2:

- How the current behavioural patterns of an employee or a group in an organisation can be understood using a software instead of human (a person who works as a consultant, NLP practitioner or psychologist)
- How to predict the personality type of a person and the most popular personality type in an organisation using an intelligent software with improved performance in comparison to the previous automation attempts.

As a result, this experiment was designed to compare the performance of the software

with the previous automation attempts for personality type prediction. In data gathering process, participants were asked to respond to the questionnaire and then have a conversation with the software. Figure 4.15 Shows the results of questionnaire.

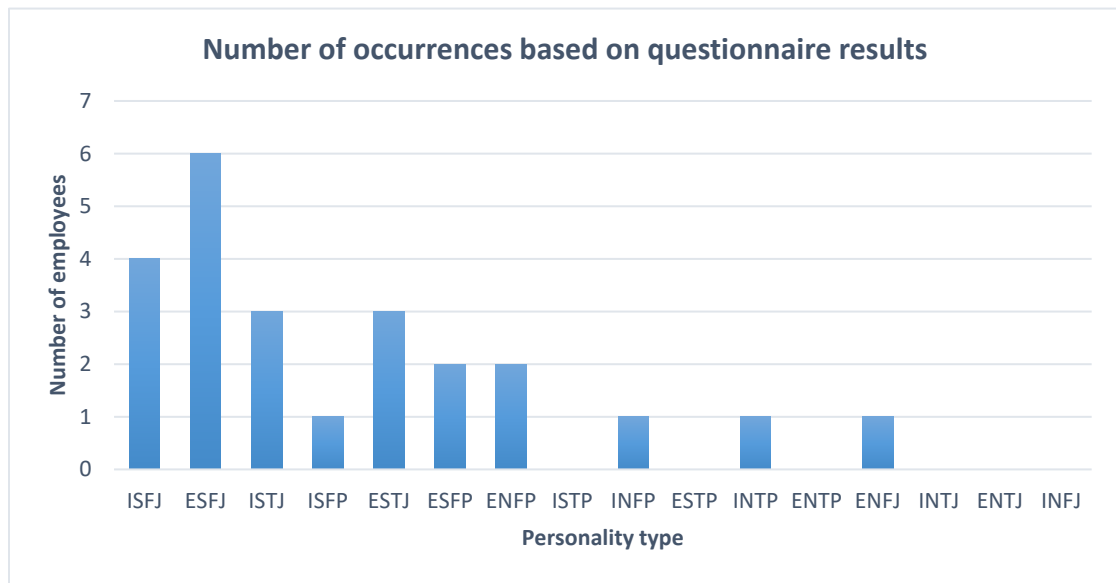


Figure 4.15: Number of personality type in the organisation based on the questionnaire results

Figure 4.15 shows that 6 employees were identified with ESFJ personality type and ESFJ was the biggest group in the company. ISFJ was the second popular personality type in the company and 4 employees were identified with this personality type. The number of employees who were identified with ISTJ and ESTJ were equal and 3 employees were categorized in each on these groups. According to Figure 4.15, ISTJ and ESTJ jointly can be considered as the third most popular personality type in the organisation. Moreover, the number of employees with ESFP and ENFP personality type were equal and 2

employees were categorized in each group. Furthermore, only 1 employee was identified with each one of ISFP, INFP, INTP and ENFJ personality types. Figure 4.15 also shows that nobody was identified with ISTP, ESTP, ENTP, INTJ, ENTJ and INFJ personality types.

On the other hand, Figure 4.16 shows the results of the software predictions for employee personality types.

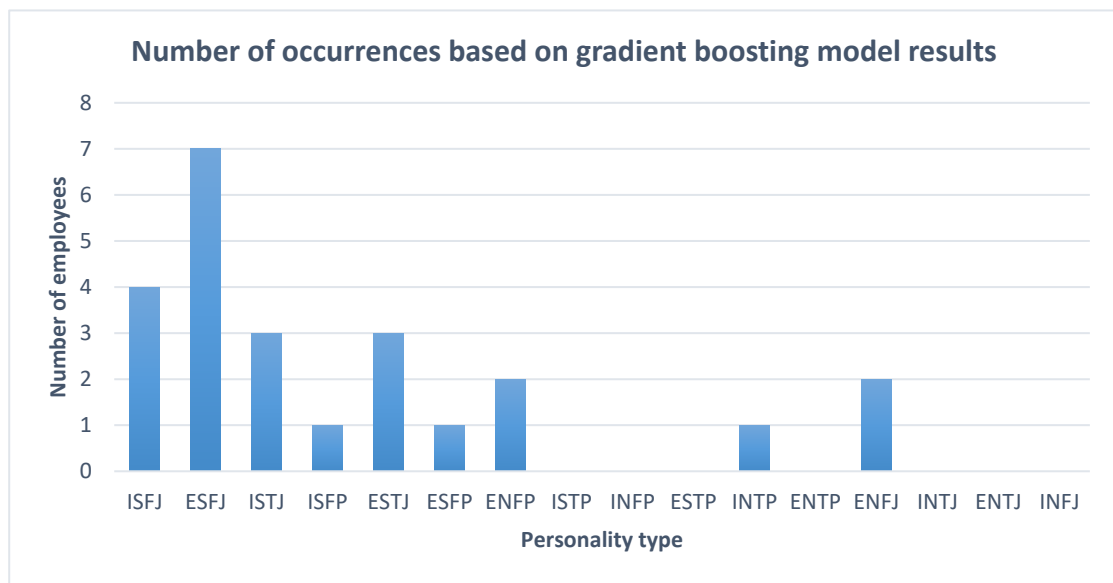


Figure 4.16: Number of personality type in the organisation based on the software results

Figure 4.16 shows that the software identified 7 employees with ESFJ personality type. ISFJ personality type was identified for 4 employees, 3 employees were identified with each one of ISTJ and ESTJ personality types. There were also 2 employees that were identified with ENFP and 2 employees with ENFJ. Furthermore, one employee was identified with each one of ISFP, ESFP and INTP personality types. In this company, no

employee was identified with ISTP, INFP, ESTP, ENTP, INTJ, ENTJ and INFJ personality types. Figure 4.17 shows the results of comparing the number of identified personality types via questionnaire and software in this company.

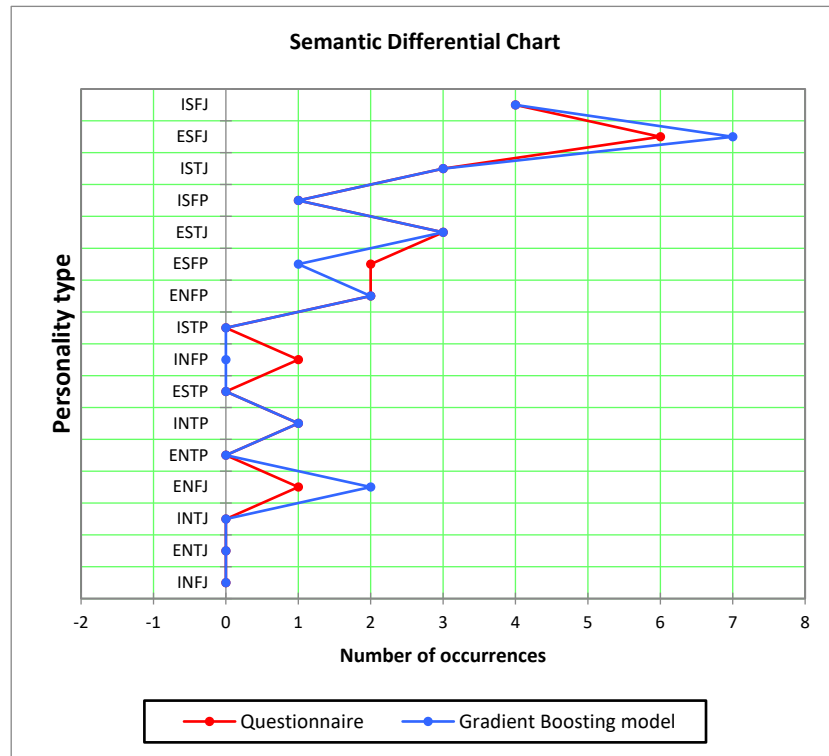


Figure 4.17: Comparing the number of identified personality types via questionnaire and software

Figure 4.17 Shows that the software identified 7 employees with a ESFJ personality type, while 6 employees were identified with this personality type through the questionnaire. Thus, the software identified one more employee in this category. Figure 4.17 also shows that the software identified one more employee than the questionnaire with ESFP personality type and one less employee than the questionnaire with ENFJ

personality type. Moreover, the questionnaire identified no employee with INFP personality type while the software identified one employee with this personality type. For other personality types, the results of the software were exactly identical to the questionnaire and they both identified the same number of employees for each personality type. Correspondingly, both the software and the questionnaire identified no employee with ISTP, ESTP, ENTP, INTJ, ENTJ and INFJ personality types. Figure 4.18 and 4.19 show the frequency of each personality type in the company identified by the software and the questionnaire.

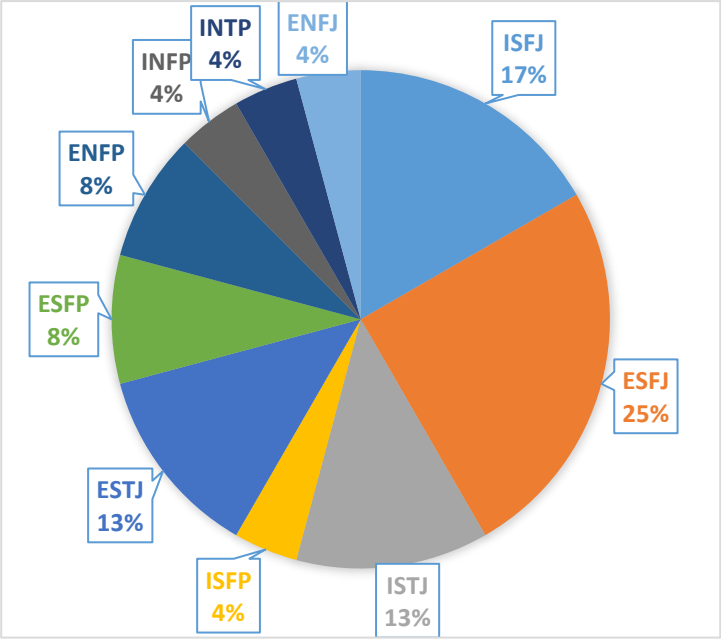


Figure 4.18: Frequency of each personality type in the company identified by the software

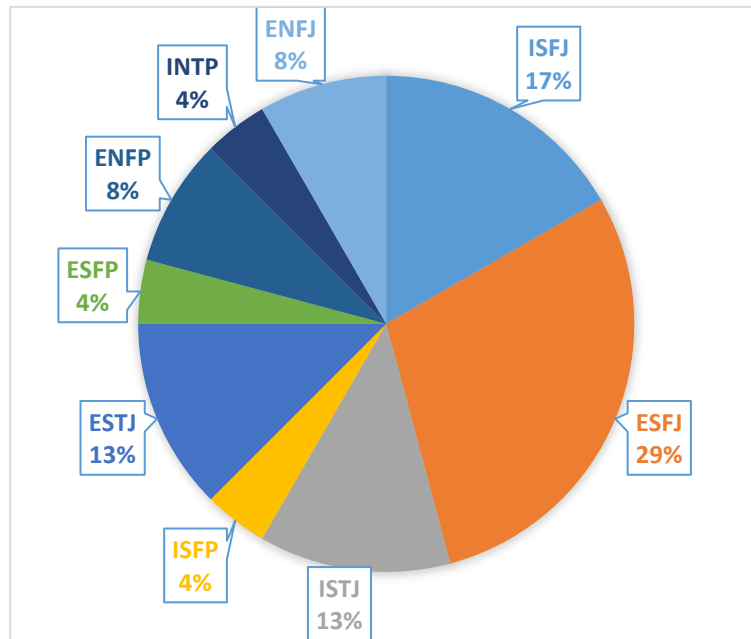


Figure 4.19: Frequency of each personality type in the company identified by the software

According to Table 4.9, showing the results of comparing Figure 4.18 and Figure 4.19, the results of the software were exactly the same as the questionnaire for 12 out of 16 personality type categories. There is +3% difference between the software results and the questionnaire results for ESFJ personality type and the software identified more employees with this personality type than the questionnaire. Furthermore, there is -4% difference between the software and questionnaire results for ESFP and INFP personality type and for this category the software identified less employees than the questionnaire. There is also +4% difference between the software and questionnaire results for ENFJ personality type and the software identified one more employee than the questionnaire in this personality type category. As a result, the accuracy of the software was slightly less

than the computerised questionnaire in prediction of 4 out of 16 personality type categories. In section 3.4.8, the accuracy of prediction using XGBoost was compared to one of the latest and most successful existing methods which used the same dataset. This method was introduced by Hernandez and Knight in 2017. They used various types of recurrent neural network (RNN) such as simple RNN, GRU, LSTM, and Bidirectional LSTM to build their classifier. The comparison in section 3.4.8 shows that the performance of XGBoost has been better and as a result, other methods cannot achieve a better result in compare to the computerised questionnaire.

| Personality type | Frequency in organization | | |
|------------------|---------------------------|---------------|------------|
| | Gradient Boosting model | Questionnaire | Difference |
| ISFJ | 17% | 17% | 0% |
| ESFJ | 29% | 25% | +3% |
| ISTJ | 13% | 13% | 0% |
| ISFP | 4% | 4% | 0% |
| ESTJ | 13% | 13% | 0% |
| ESFP | 4% | 8% | -4% |
| ENFP | 8% | 8% | 0% |
| ISTP | 0% | 0% | 0% |
| INFP | 0% | 4% | -4% |
| ESTP | 0% | 0% | 0% |
| INTP | 4% | 4% | 0% |
| ENTP | 0% | 0% | 0% |

| | | | |
|-------------|----|----|----|
| ENFJ | 8% | 4% | +4 |
| INTJ | 0% | 0% | 0% |
| ENTJ | 0% | 0% | 0% |
| INFJ | 0% | 0% | 0% |

Table 4.9: Personality type frequency in organisation based on Gradient Boosting model results

A filled radar chart was used to show the values relative to a centre point and Figure 4.20 and 4.21 show the results for the system and questionnaire.

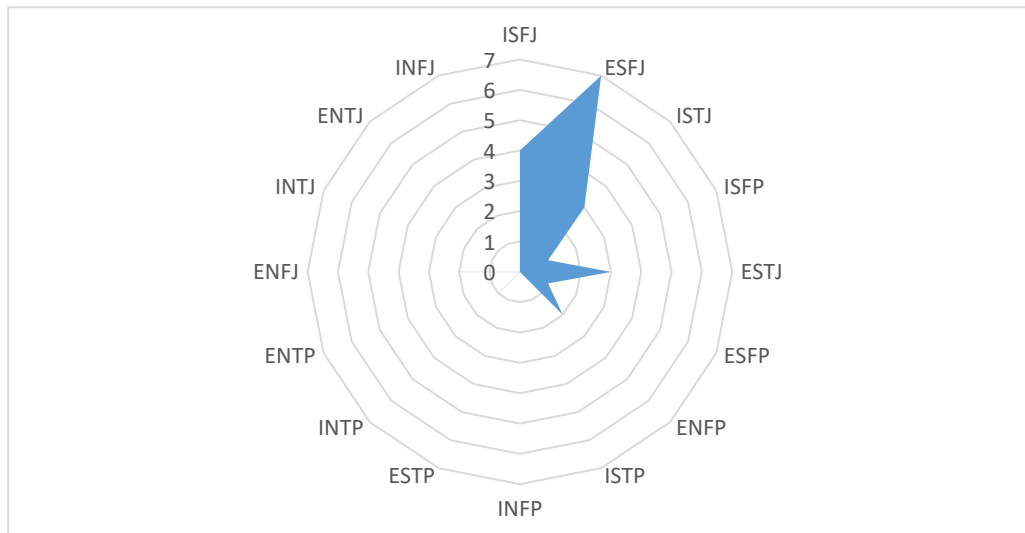


Figure 4.20: Most popular personality type in the company based on the system results

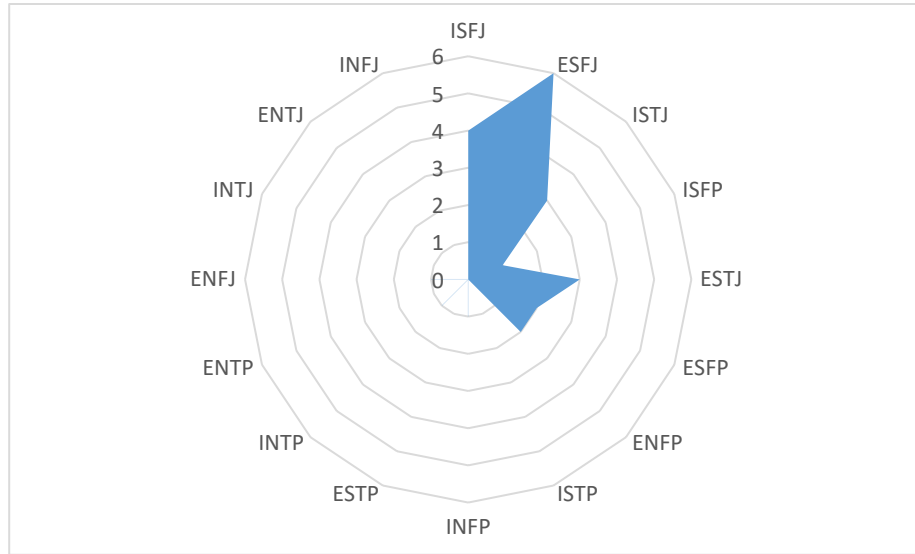


Figure 4.21: Most popular personality type in the company based on the questionnaire results

Figure 4.20 And 4.21 both show that the software identified the most popular personality type in this company was ESFJ and the second most popular personality type was ISFJ. According to Table 2.2 and 2.3 in chapter 2, regarding the most appropriate jobs for each personality type, it can be determined that most of the employees were assigned to an appropriate position and tasks allocations were at a satisfactory level. This was established based on the nature of the positions, the task allocations and the most popular personality types in this company. However, the performance of the company can be improved by considering the task allocation for those employees who are dealing with costumers but their personality type is not categorized in in one of these two popular personality type categories in the company.

Overall, the objective of improving the efficiency, reliability and accuracy of the manual application of personality type prediction was successfully tested for with a

significant majority of results being comparable between the software and MBTI questionnaire. The identification of the most popular personality type amongst participants matched for the traditional and the automated method. Moreover, the difference in the recognised types of personality being 4 categories in the results has to be reconsidered as a true difference of 2. This is due to the recognition of one individual to have a different personality type when tested by the software and the MBTI questionnaire is reflected in the results as a decrease in one category and an increase in another when comparing methods. This suggests the true difference in performance to be better than apparent in the data presented in Table 4.9. This reflects high reliability in being able to implement this alternative in organisations with the capacity to make the appropriate recommendations about task allocation in an organisation.

4.5 Results of using Representational system identifier in the company

The software also recognised the preferred representational system for each employee. Figure 4.22 shows the number of employees in each category.

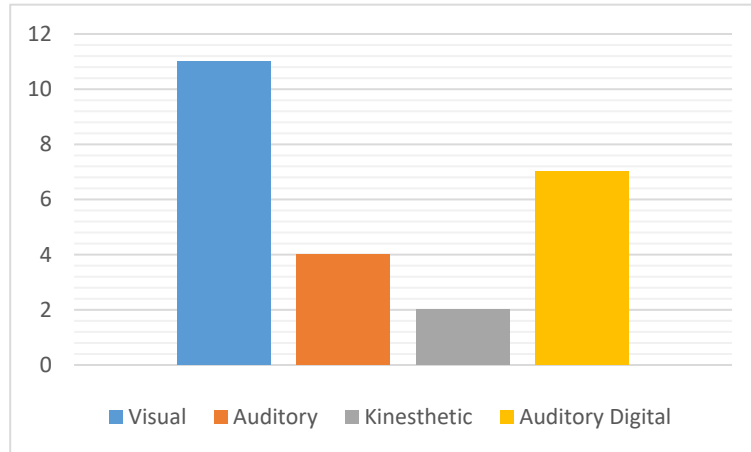


Figure 4.22: Number of employees in each category of preferred representational system

According to Figure 4.22 from the 24 employees in the company, 11 employees were identified with visual representational system as their preferred system, 4 employees were identified with auditory, 2 employees with kinaesthetic and 7 employees with auditory digital representational system. Figure 4.23 Shows the percentage of each preferred representational system in the company.

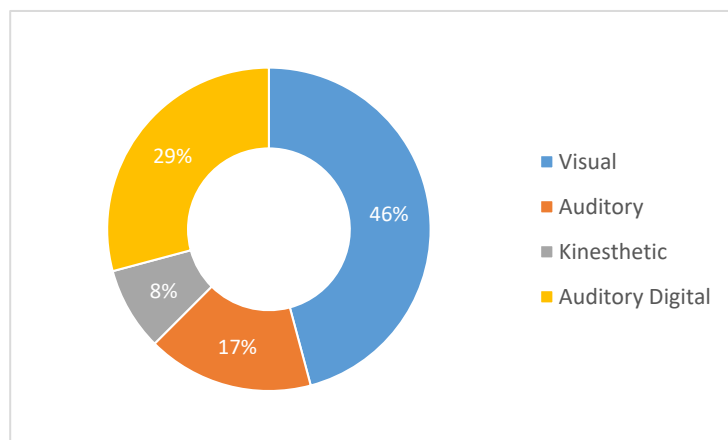


Figure 4.23: percentage of each preferred representational system in the company

According to Figure 4.23 a visual representational system was the most popular representational system in the company and 46% of the employees were identified with this representational system as their preferred representational system. The second popular representational system was auditory digital, this was the preferred representational system for 29% of the employees in the company. Figure 4.23 also shows that the preferred representational system for 17% of the employees was auditory and for 8% of the employees was kinaesthetic which had the lowest popularity in the company. The software successfully provided the relevant solutions to each employee, in order to improve their communications with their colleagues. The software also provided the relevant solutions for visual representational system as the most popular representational system, in order to improve the communication through the company and improve the staff learning and understanding level in the meetings.

4.6 Results of using Meta model system in the company

The software was using the Meta model during the conversation with the employees in the company in order to clarify the statements obscured by deletion, distortion and generalisation and to help them efficiently identify and address their issues in the work environment. After the data gathering process, the conversations between the software and participants were analysed. The software identified 393 deletions, 142 distortions and 196 generalisations. In other words, 54% of the recovery questions by the software were

in relation to deletion, 19% were in relation to distortion and 27% were in relation to generalisation, as demonstrated in Fig 4.24.

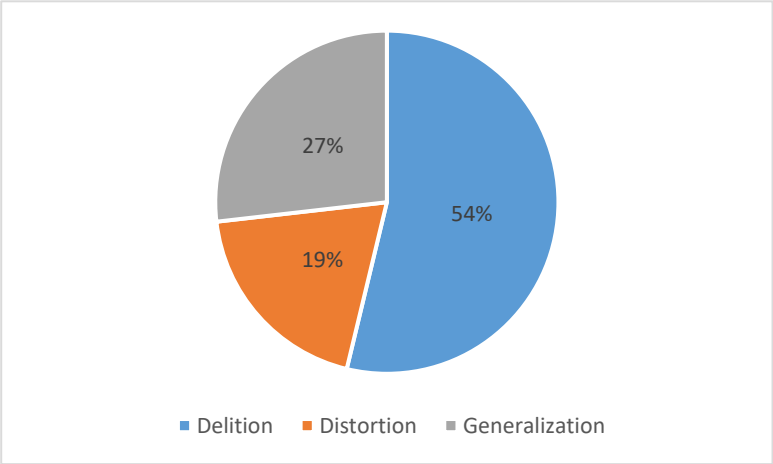


Figure 4.24: percentage of each category of recovery questions asked b the software

The number of recovery questions regarding each category of deletion, distortion and generalisation processes were also recorded and Figure 4.25 shows the results.

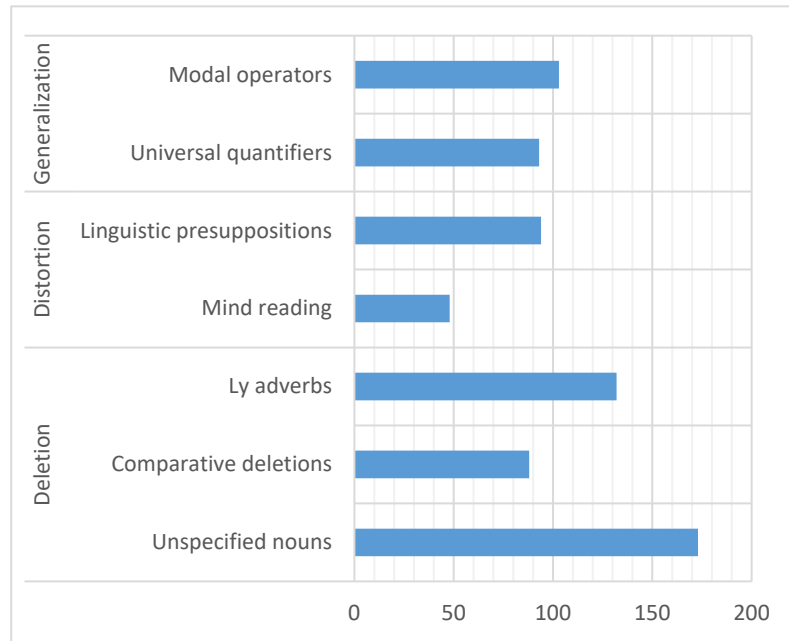


Figure 4.25: number of recovery questions related to each category of the deletion, distortion and generalization

According to Figure 4.25, 173 questions were regarding unspecified nouns, 88 questions were regarding comparative deletions and 132 questions were regarding “Ly” adverbs in the user-software conversation. Figure 4.25 also shows that 48 recovery questions asked by the software related to mind reading in the distortion process while 94 questions were related to the linguistic presuppositions. Finally, in the generalisation process, 93 recovery questions were related to universal quantifiers and 103 questions relating to modal operators.

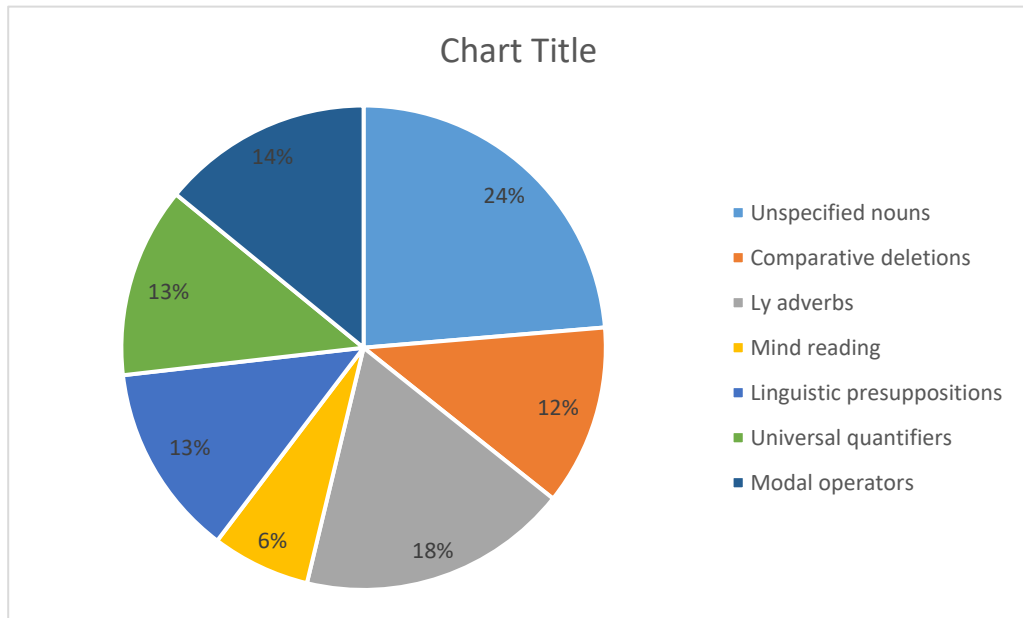


Figure 4.26: Percentage of recovery questions related to each process of the deletion, distortion and generalization

Figure 4.26 displays the percentage of each recovery question used by the software during the conversation with the employees in the company. According to Figure 4.26 the recovery questions regarding unspecified nouns were asked more than other recovery questions and the percentage for this category was 24%. The recovery questions related to ‘Ly’ adverbs had the second highest percentage that is 18%. Both of these categories are related to the deletion process. The third category with the highest percentage is modal operators, which is related to generalisation process and the percentage for this category is 14%. The recovery question related to universal quantifiers which is associated with the generalisation process and linguistic presupposition which is associated with the distortion process both have the same percentage of 13%. Figure 4.26 also shows that 12% of the recovery questions were related to comparative deletions which are associated

with the deletion process and 6% of the recovery questions were related to mind reading which is associated with the distortion process.

4.7 Summary

Following this, the results of the automated Meta model system were presented. A NLP practitioner applied the Meta model, the system achievements were then compared with the human achievements. Overall results show that the software performs more efficiently with a high level of accuracy and reliability in comparison to the practitioner. Based on the results, the proposed software is more successful with regards to the deletion and generalisation processes in comparison to an experienced NLP practitioner. However, the software is slightly less successful in clarifying the distortion processes compared to the practitioner.

Moreover, the third part of the system which is personality type prediction, was tested in a company and the results are presented in Section 4.4 of this chapter. A computerised MBTI personality type indicator questionnaire was used in the company and the results of this questionnaire were compared with the results of the automated personality type predictor system. Overall, results show that the performance of the software was identical to the questionnaire in predicting 12 out of 16 personality type categories. The accuracy of the software was slightly less than the computerised questionnaire in predicting 4 personality type categories. Furthermore, the most popular personality types in the company were recognised accurately by the software. Based on this information, it was

determined that most of the employees were assigned to an appropriate position and the task allocations were at a satisfactory level. In addition, the representational system identification and the Meta model system were also used in the company and the results were described in this chapter. The software successfully recognised the most popular representational system in the company and the relevant solutions were also provided. This can assist in further improving the communication, staff learning and understanding level in meetings. The results show that the automated Meta model was also effectively used and the relevant recovery questions were asked to recover the hidden information in conversations.

5 Conclusion

5.1 Conclusion of the research

Neuro Linguistic Programming (NLP) is one of the most utilised approaches for personality development and it consists of a variety of techniques and escalating levels of processes to aid personal development in clients and oneself. On the other hand, the working environment in an organisation is affected by executive behaviour of employees. As NLP is focusing on personal excellence, it is positively correlated with organisational success via improving executive behaviour. As a result, this research aimed to automate the most important aspects of Neuro Linguistic Programming in order to be used in organisations for the purpose of organisational performance improvement.

This research in the first step, has automated the process of identifying the preferred representational system which is one of the most important aspects of Neuro Linguistic Programming and vital during the personality development process. In the second step, the process of using the Meta model in a conversation was automated and performs in a human-computer interaction. In the third step, Meta programs detection and personality type prediction based on MBTI personality type indicator was automated.

In the first phase of implementation, Natural Language Processing, a subfield of artificial intelligent was used as a tool for automation process. As a result, the first part of a new intelligent software which is able to act like an experienced psychologist or NLP practitioner has been developed based on Python and Natural Language Processing Tool-Kit (NLTK). The Software has been tested on a group of 55 students at London Metropolitan University who have been studying different subjects at PhD research,

undergraduate and postgraduate levels. They were asked to respond to a manual questionnaire, designed in order to understand their preferred representational system. Finally, the obtained results by the software were compared with the attained results by the questionnaire, demonstrating a superior performance and a high level of accuracy and reliability of the software against the manual questionnaire. Furthermore, in recognizing the language and identifying the preferred representational system, the performance of the software was shown to be slightly more accurate than and the results from the questionnaires. Based on the results, it can be concluded that the proposed software is more robust in identifying auditory digital and auditory representational systems than an experienced NLP practitioner. In contrast, it could also be concluded that the developed software is slightly less effective for visual and kinaesthetic representational system analysis compared to a human NLP practitioner. Therefore, the novel methodology presented in this research could successfully improve the accuracy and reliability of the identification process for the preferred representational system with the advantage of significantly decreasing the inaccuracies associated with the manual processes such as, the lack of experience, personal judgment, different level of skills and other human errors.

In the second phase of implementation, Natural language processing was used again, as a tool for the automation process of this part of the system. As a result, the second part of the new intelligent software has been developed which is able perform as a competent NLP practitioner or psychologist. The software has been tested on 50 participants with a good variety backgrounds. The conversations and answers from participants were recorded in separate files and given to an experienced NLP practitioner to be analysed. Finally, obtained results by the software were compared to the obtained results by the

practitioner. A more efficient performance of the software, with a high level of accuracy and reliability, was observed in comparison to the practitioner. Based on the results, it can be concluded that the proposed software is more successful with regards to the deletion and generalization processes in comparison to an experienced NLP practitioner. The software, however, is slightly less successful for clarifying the distortion processes compared to the practitioner. The methodology presented in this research paper could successfully improve the accuracy and reliability of using the Meta model in a conversation through automation of the process. Human errors such as lack of experience, personal judgment, effect of the practitioners' level of skill and other human errors were effectively eliminated from the process and the relevant inaccuracies significantly decreased. This is because instead of a human practitioner, an intelligent software is doing the analysis, which eliminates impact of opinion, judgement, lack of experience or human error. An intelligent analysis is also consistent in nature, which a human practitioner may not be. This is also contributive to decrease of human error.

In the third phase of implementation, Natural language processing Tool-Kit (NLTK) and XGBoost which is an optimized distributed gradient boosting library in Python for implementing machine learning algorithms under the Gradient Boosting framework, were used for automation process. Moreover, Pandas, numpy, re, seaborn, matplotlib and sklearn were other python libraries that were used in this part of the implementation. This part of the software has been tested in a company with 24 employees. They employees were asked to respond to a computerized MBTI personality type indicator questionnaire and then have a conversation with the system. The results of the computerized questionnaire were recorded on a separate file. On the other

hand, the information about each employee during the conversation with the software was recorded on separate files and these files were used for the automated personality type predictor to identify the personality type of each employee. The results of the automated personality type indicator were recorded on a different file. The content of this file was compared with another file containing the result of computerized questionnaire in order to analyse and check the accuracy of the results. The most popular personality type in the organization was also identified by the system. Based on the results, the performance of the software was exactly the same as questionnaire for predicting 12 personality type categories out of 16 personality type categories. The accuracy of the software was slightly less than the computerized questionnaire for predicting 4 personality type categories. Moreover, the software accurately recognized the most popular personality types in the company and based on this information, it was realized that the most of employees were assigned to an appropriate position and the tasks allocation was in a satisfactory level. However, the performance of the company could be improved by considering the task allocation for those employees who were dealing with costumers but their personality type was not categorized in one of these two popular personality type categories in this company.

The software was also used to identify the preferred representational system of each employee and the most popular representational system in the company. According to the results, Visual representational system was the most popular representational system in the company and 46% of employees were identified with this representational system as their preferred representational system. The software successfully provided the relevant solutions to each employee to improve their communications with their colleagues. The

relevant solutions for the most popular representational system in the company were also provided, in order to improve the communication through the company and improve the staff learning and understanding level in the meetings. Furthermore, the automated Meta model was effectively used during the conversation between the employees and the software. According to the results, the software identified 393 deletions, 142 distortions and 196 generalizations and the relevant recovery questions were asked to recover the hidden information in conversations and to help the employees to have a better understanding of their issues in the work environment.

5.2 Recommendations and future work

Another method can be recommended for the first phase of implementation, which was automating the process of identifying the preferred representational system. During this research, it was realized that a specific dataset including the sample of texts for each person and the relevant preferred representational system for that person can be created. In order to create this dataset, Twitter can be used to collect the sample of texts and then participants can be asked to complete a questionnaire to identify their preferred representational system. As a result, the dataset can be including two columns. The first column would be related to the preferred representational systems and the second column would be related to the sample of texts. In the next step, this dataset can be used to train an Extreme Gradient Boosting model or other models created by other machine learning techniques. Finally, the similar method can be used to have a conversation between the

user and software to collect the sample of texts from employees and the trained model can be used to identify the preferred representational system of the employee.

Moreover, speech processing techniques can be used in order to improve the developed system in this research and make it able to work based on the voice, instead of text. As a result, participants can talk to the system instead of communicating via typing and this will save the time of identification and improve the performance of the system.

References

- Aldayel, M. (2013) K-Nearest Neighbor classification for glass identification problem, International conference on computer systems and industrial informatics. IEEE, Sharjah, United Arab Emirates.
- Andreas, S. and Faulkner, C. (1996) NLP: The new technology of achievement, 1st ed., Nicholas Brealey Publishing.
- ANLP (2016) Improving your communication using Manteya Email. Available at: <https://anlp.org/business-resources-for-nlp-professionals/manteya-email>
- Avery, L. (2015) Meta model distinctions and generalisations [Online]. Available at: <http://nlp4uonline.com/blog/2015/05/12/meta-model-distinctions-generalisations/>
- Banchs, R. E., and Li, H. (2012) Iris: a chat-oriented dialogue system based on the vector space model. In *Proceedings of Annual Meeting of the Association for Computational Linguistics (ACL) System Demonstrations*, 37–42.
- Bandler, R. and Grinder, J. (1975) The structure of magic 1: A book about language and therapy, 1st ed., Clifornia: Sience and behaviour books.
- Bishop, C. (2006) Pattern recognition and machine learning, Springer, New York, USA.
- Brefi Group Limited (2004) Representational system. http://www.brefigroup.co.uk/acrobat/nlp_representational_systems.pdf

- Beech, J. (2013) The cognitive functions of each personality type. Available at: <https://siteassets.pagecloud.com/greeleymosaic/downloads/Myers-Briggs-ID-7fbeb3b-f94d-468a-ce4f-7c488703c102.pdfm>
- Bensted C (2014) Representational systems. Available at: <http://badis.co.uk/resources/Repsys.pdf>
- Briscoe, T. (2013) Introduction to linguistics for Natural Language processing, Computer Laboratory, University of Cambridge. Available at: <https://www.cl.cam.ac.uk/teaching/1314/L100/introoling.pdf>
- Bird, S., Klein, E. and Loper, E. (2009) Natural Language Processing with Python. 1st ed., United States: O'Reilly Media, Inc.
- Biau, G. (2012) Analysis of Random Forests model, Journal of machine learning research. Vol. 13. Pages 1063-1095.
- Briggs-Myers, I. (1998) MBTI manual: A guide to the development and use of the Myers – Bridge type indicator, consulting psychologists press.
- Benyamin, D. (2012) A gentle introduction to Random Forests, Ensembles and performance metrics in a commercial system. CitizenNet Blog, the latest innovations and in social media advertising. Available at: <http://blog.citizenet.com/blog/2012/11/10/random-forests-ensembles-and-performance-metrics>
- Bohus, D., and Rudnicky, A. I. 2009. The ravenclaw dialog management framework: Architecture and systems. *Computer Speech & Language* 23(3):332–361.

- Bird S, Klein E, Loper E (2009) Natural Language Processing with Python: Analyzing Text with the Natural Language Toolkit. USA: O'Reilly Media, Inc.
- Cortes, C. and Vapnik, V. (1995) Support-vector networks. Machine learning, 20(3), 273-297.
- Cui, B., and Qi, C. (2017). Survey Analysis of Machine Learning Methods for Natural Language Processing for MBTI Personality Type Prediction. Available at: <http://cs229.stanford.edu/proj2017/final-reports/5242471.pdf>
- Charvet SR (1997) words that change minds: mastering the language of influence, Kendall/Hunt Publishing Co, U.S. 2nd Revised edn
- Chowdhury, G. (2003) Natural Language Processing. Annual review of Information Science and Technology. 37(1): 51-89.
- Casale, P. (2012) NLP secrets: Upgrade your mind with Neuro Linguistic Programming, Creative Media NZ Ltd.
- Champa, H. N. and Anandakumar, K. R. (2010) Artificial neural network for human behaviour prediction through handwriting analysis. International Journal of Computer Applications.
- Chopra, A., Prashar, A. and Sain, A. (2013) Natural Language Processing, International journal of technology enhanceent and emerging engineering research, vol. 1, issue 4, pp. 131-134.
- Carroll, M. (2016) An overview of the Meta model and explanation of the five distortion categories [Online]. Available at: http://www.nlpacademy.co.uk/articles/view/An_overview_of_the_Meta_Model_and_explanation_of_the_5_distortion_categorie/

- Davis. K. (2015) The meta model problem solving strategies [Online]. Available at: <http://nlp-mentor.com/meta-model/>
- Denil, M., Matheson, D. and Freitas, N. (2014) Narrowing the Gap: Random Forests in theory and in practice, Proceedings of the 31st International Conference on Machine Learning, Beijing, China.
- Domínguez-Almendros, S. (2011) “LOGISTIC REGRESSION MODELS”. Allergol Immunopathol (Madr). doi:10.1016/j.aller.2011.05.002.
- Dale R (2010) Classical Approaches to Natural Language Processing, Handbook of Natural Language Processing. Boca Raton: CRC Press.
- Dong Nguyen, A. Seza Dögrüöz, Carolyn P. Rosé, and Franciska de Jong. (2016) Computational sociolinguistics: A survey. Computational Linguistics 42(3):537–593.
- Elston, T. (2017) The Meta model in NLP [Online]. Available at: <https://www.nlpworld.co.uk/meta-model-nlp/>
- Ellerton, R. (2007) NLP’s Auditory Digital Representational Systems. Available: <http://www.renewal.ca/nlp48.html>
- Elsalamony, H. A. (2014) bank direct marketing analysis of data mining techniques, International journal of computer applications, Vol. 85, No. 7, pages 12-22.
- Enayet, O. (2010) Natural Language Processing – The big picture [Online]. Available at: <https://omarsbrain.wordpress.com/tag/natural-language-processing-linguistics-phonology-morphology-discourse-pragmatic-summarization/>

- Ellerton R (2015) Modalities and Representational Systems. Available at: <http://www.renewal.ca/nlp10.htm>
- Freeth, P. (2016) The NLP practitioner manual, 3rd edition, CGW publishing.
- Friedman, J. H. (2001) “Greedy function approximation: a gradient boosting machine”, *Annals of statistics*, 1189-1232.
- Fukunaga, K. (1990) *Introduction to Statistical Pattern Classification*. Academic Press, San Diego, California, USA.
- Freund, Y. and Schapire, R. E. (1997) “A decision-theoretic generalization of on-line learning and an application to boosting”, *Journal of computer and system sciences*, 55(1), 119-139.
- Faber, J., and Fonseca, L. (2014) How sample size influences research outcomes, *Dental Press Journal of Orthodontics*, 19(4): 27–29. Available at: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4296634/>
- Gregory, A. (2011) What the Myers-Briggs Personality Test Can and Can’t Tell You. Available at: <https://www.sitepoint.com/what-the-myers-briggs-personality-test-can-and-cant-tell-you/>
- Gudwin, R. (2000) “Evaluating intelligence: A computational semiotics perspective”, *IEEE International conference on systems, man and cybernetics*, Nashville, Tennessee, USA.
- Gjurkovic, M. and Snajder, J. (2018) Reddit: A gold mine for personality prediction, *Proceedings of the second workshop on computational modelling of people’s opinions, personality and emotions in social media*, New Orleans,

Louisiana. Pages 87-97. Available at:
<https://peopleswksh.github.io/pdf/PEOPLES12.pdf>

- Golbeck, J., Robles, C., Edmondson, M. and Turner, K. (2011) Predicting personality from Twitter, IEEE Third International Conference on Privacy, Security, Risk and Trust and IEEE Third International Conference on Social Computing. Boston, USA. Available at:
<https://ieeexplore.ieee.org/document/6113107/>
- Hervás R, Bautista S, Rodríguez M, Salas T, Vargas A, Gervás P (2012) Integration of lexical and syntactic simplification capabilities in a text editor. *Procedia Computer Science*. 27:94-103.
- Horn, D. (2008) Syntactic structures in language and biology. *Cognitive processing*. 9(14): 153-158.
- Hernandez, R. and Knight, S. (2017), Predicting Myers-Bridge Type Indicator with text classification. 31st conference on Neural Information processing Systems (NIPS). Available at: <https://web.stanford.edu/class/cs224n/reports/6839354.pdf>
- Hoag, J. D. (2017) The NLP Meta model [Online]. Available at:
<http://www.nlpls.com/articles/NLPmetaModel.php>
- Hyeoun-Ae, P. (2013) An introduction to logistic regression: from basic concepts to interpretation with particular attention to nursing domain, *Journal of Korean Academy of Nursing*, Vol. 43, No. 2, pages 154-164.
- Isabel Briggs Myers, Mary H. McCaulley, and Allen L.Hammer. (1990) *Introduction to Type: A description of the theory and applications of the Myers-Briggs type indicator*. Consulting Psychologists Press.

- Janicki, S. (2010) 'Neuro Linguistic Programming', an initial introduction to the wide field of NLP. GRIN Verlag.
- Jakkula, V. (2011) Tutorial on support vector machine (SVM), School of EECS, Washington State University, USA.
- John, E. and Barbuto, Jr. (1997) A critique of the Myers-Briggs Type Indicator and its operationalization of Carl Jung's psychological types. *Psychological Reports* 80(2):611–625.
- Jurafsky, D. and Martin, J. (2014) *Speech and Language processing*. Prentice Hall, Pearson Education International.
- Joey, L. and Yazdanifard, R. (2015) "Can Neuro Linguistic Programming (NLP) be used as contemporary and effective skill for an exceptional manager in an organisation?" *International journal of management, accounting and economics*, vol. 2, issue 5, pp. 457-466.
- Juiced Concepts Limited (2012) *NLP Home Study Programme (V2.0)*. Available at: <https://s3-eu-west-1.amazonaws.com/nlpraconlinemp3/RepSystems/RepresentationalSystems.pdf>
- Komisin, M. and Guinn, C. (2012) Identifying personality types using document classification methods. *Proceedings of the 25th International Florida Artificial Intelligence Research Society Conference, FLAIRS-25*, page 232-237.
- Kleinbaum, D. G. and Klein, M. (2010) *Logistic Regression, Statistics for Biology and Health*, DOI 10.1007/978-1-4419-1742-31, Springer Science Business Media, LLC.

- Lazarus, J. (2010) Successful NLP: for the results you want. Surrey: Crimson Publishing.
- Lewis, R. Goldberg. (1990) An alternative “description of personality”: The big-five factor structure. Journal of personality and social psychology 59(6):1216.
- Linder-Pelz, S. (2010) ‘NLP Coaching’, An evidence-based approach for coaches, leaders and individuals. London: Koganpage.
- Liddy, E. D. (2001) Natural Language Processing, In Encyclopedia of library and information science, 2nd ed., New York: Marcel Dcker, Inc.
- Li, C., Wan, J. and Wang, B. (2017) 16th International Symposium on Distributed Computing and Applications to Business, Engineering and Science (DCABES), Anyang, China. Available at: <https://ieeexplore.ieee.org/document/8253041/?part=1>
- Manteya (2016) Manteya Email, perfect online persuasion. Available at: <http://www.manteya.com>
- McAfee, K. (2014) ‘The power of words’, an introduction to NLP representational systems, becoming a more effective communicator. <http://www.americasmarketingmotivator.com/wp-content/uploads/2014/09/Workbook-The-Power-of-Words-NLP-Representational-Systems.pdf>
- Michie, D., Spiegelhalter, D., and Taylor, C. (1994) Machine Learning, Neural and Statistical Classification. Ellis Horwood Limited. Available at: <https://www1.maths.leeds.ac.uk/~charles/statlog/whole.pdf>

- Monkeypuzzle training and consultancy (2016) ‘The power of your senses’. Using NLP Representational Systems to improve how you communicate, relate and learn. Available at: <http://www.monkeypuzzletraining.co.uk/free-documents/The-Power-of-Your-Senses-NLP.pdf>
- Manning, C. D., Raghavan, P. and Schütze, H. (2009) Introduction to Information Retrieval, Cambridge: Cambridge University Press.
- Martinez-Mesa, J., Gonzalez-Chica, D. A., Bastos, J. L., Bonamigo, R. R., and Duquia, R. P. (2014) Sample size: how many participants do I need in my research? Official publication of Brazilian society of Dermatology. 89(4): 609–615. Available at: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4148275/>
- Mitchell, T. (1997) Machine learning. 2nd ed. USA: McGraw Hill.
- NLP centre (2006) NLP history and development. Available at: <http://www.nlp-center.net/articles/nlp-history-and-development.html>
- Ntekin, A. and Knoll, A. (2013) Gradient Boosting machines, a tutorial, frontiers in neurorobotics. Vol. 3, Article 21.
- Nagadevara, V., Srinivasan, V. and Valk, R. (2008) “Establishing a link between employee turnover and withdrawal behaviours: Application of data mining techniques”, Research and Practice in Human Resource Management, 16(2), 81-97.
- NLP Dynamics Ltd (2013) Representational Systems. Available at: <http://www.distancelearning.academy/wp-content/uploads/2015/02/Representational-Systems.pdf>

- Nugues, P. M. (2006) An introduction to language processing with Perl and Prolog, Springer-Verlag Berlin Heidelberg.
- NLTK official website (2016) NLTK 3.0 documentation. Available at: <http://www.nltk.org/#natural-language-toolkit>
- Oberholzer C (2003) The role of Neuro Linguistic Programming in improving organizational leadership through interpersonal communication development. Master's dissertation. Faculty of Economic and Management Sciences, University of Pretoria. Available at: http://repository.up.ac.za/bitstream/handle/2263/41285/Oberholzer_Role_2014.pdf?sequence=1
- O'Connor J, Seymour J (1993) Introducing Neuro Linguistic Programming: Psychological skills for understanding and influencing people, Revised Edition. London, California: The Aquarian Press.
- O'Connor, J. and McDermott, I. (1997) The art of systems thinking: Essential skills for creativity and problem solving," 1st ed., Thorsons.
- O'Connor J (2013) 'NLP workbook', a practical guide to achieving the results you want. San Francisco: Conari Press.
- O'Brien, D. (2009) Reverse Meta model: Simple deletions [Online]. Available at: <http://ericksonian.com/reverse-meta-model-simple-deletions>
- Palmiero M, Di Matteo R, Belardinelli M (2014) The representation of conceptual knowledge: visual, auditory, and olfactory imagery compared with semantic processing. Cognitive processing 15(2): 143-157.

- Personalitymax™ (2018) Career matches by personality type. Available at: <https://personalitymax.com/personality-types/careers/>
- Punnoose, R. and Ajit, P. (2016) Prediction of employee turnover in organisations using machine learning algorithms, A case for Extreme Gradient Boosting. International Journal of advanced research in Artificial Intelligence. Vol. 5, No. 9, Pages 22-26. Available at: https://thesai.org/Downloads/IJARAI/Volume5No9/Paper_4-Prediction_of_Employee_Turnover_in_Organizations.pdf
- Pegasus NLP (2011) A version of the history of NLP. Available at: <http://www.pe2000.com/nlp-history.htm>
- Rayner institute (2015) The representational systems. Available at: http://www.raynerinstitute.com/uploads/9/8/6/1/9861170/nlp_rep_system.pdf
- Singer M T, Lalich J (1996) 'Crazy therapies'. What are they? Do they work? San Francisco: Jossey-Bass.
- Stoker, J. (2014) Ly adverbs [Online]. Available at: <http://juvenate.org/ly-adverbs/coaching/nlp>
- Shouman, M., Turner, T. and Stocker, R. (2012) "Applying k-Nearest Neighbour in Diagnosing Heart Disease Patients," *International Journal of Information and Education Technology*, vol. 2.
- Saharia, N., Das, D., Sharma, U. and Kalita, J. (2009) Part Of Speech tagger for Assamese Texts, Proceeding of the ACL-IJCNLP 2009 conference short papers, Singapore, pp. 33-36.

- Success Strategies Company (2016) What is the LAB Profile®? Available at: <https://www.successtrategies.com/lab-profile/what-is-lab-profile/>
- Su, P. H., Gasic, M., Mrki, N., Barahona, L. M. R., Ultes, S., Vandyke, D., Wen, T. H., and Young, S. (2016) On-line active reward learning for policy optimisation in spoken dialogue systems. In *Meeting of the Association for Computational Linguistics (ACL)*.
- Tosey, P. and Mathison, J. (2003) Neuro-linguistic programming and learning theory: a response. *Curriculum Journal* 14: 371-388.
- Tandra, T., Suhartono, D., Wongso, R. and Prasetio, Y. (2017) Personality prediction system from Facebook users, 2nd International conference on computer science and computational intelligence, Bali, Indonesia. Available at: https://ac.els-cdn.com/S1877050917320537/1-s2.0-S1877050917320537-main.pdf?_tid=f2e33fcc-1d90-46cb-8ea5-c111bcb0c8ae&acdnat=1534438960_944100586097575861d3b3489cdb6033
- Tosey, P. and Mathison, J. (2006) *Introducing Neuro-Linguistic Programming* Centre for Management Learning & Development. School of Management, University of Surrey.
- Tan J (2003) *Neuro Linguistic Programming, The art and science of excellence*. REV Training and Coaching Company. Available at: http://www.revinspires.me/uploads/1/1/7/5/11750863/what_is_nlp.pdf
- Tosey P, Mathison J, Michelli D (2005) Mapping transformative learning: the potential of Neuro-linguistic Programming. *Journal for Transformative Education*. 3(2): 140-167.

- Tosey P (2010) *Neurolinguistic Programming for leaders and managers*. Gower Handbook of Leadership and Management Development. Surrey: Gower.
- Transform Destiny (2015) *Introduction to Neuro-linguistic programming*. Available at: <http://www.freenlphomestudy.com/membersonly/iNLP/iNLPManual.pdf>
- Vivencio, D. P., Hruschka, E., Nicoletti, M., Dos Santos, E. and Galvao, S. (2007) “Feature-weighted k-Nearest Neighbor Classifier,” *Foundations of Computational Intelligence, 2007. FOCI 2007. IEEE Symposium*, pp. 481–486.
- Witkowski, T. (2010) *Thirty-Five Years of Research on Neuro-Linguistic Programming. NLP Research Data Base. State of The Art or Pseudoscientific Decoration?* *Polish Psychological Bulletin* 41(2): 58-66.
- Webster, J. and Kit, C. (1992) *Tokenization as the initial phase in NLP*. The 15th International Conference on Computational Linguistics, France.
- Wan, D., Zhang, C., Wu, M. and An, Z. (2014) *Personality prediction based on all characters of user social media information*, Chinese national conference on social media processing, page 220-230.
- Youtube (2016) *Using Manteya in outlook*. Available at: <https://www.youtube.com/watch?v=qfBWGQnOBda>
- Yan, Z., Duan, N., Chen, P., Zhou, M., Zhou, J., and Li, Z. (2017) *Building Task-Oriented Dialogue Systems for Online Shopping*, *Proceedings of the Thirty-First AAAI Conference on Artificial Intelligence*. Available at: <https://www.google.com/url?sa=t&rct=j&q=&esrc=s&source=web&cd=8&cad=rja&uact=8&ved=2ahUKEwi12tHRIKPgAhWeXRUIHVUGCNsQFjAHegQIB>

BAC&url=https%3A%2F%2Fwww.aaai.org%2Focs%2Findex.php%2FAAAI%
2FAAAI17%2Fpaper%2Fdownload%2F14261%2F13975&usg=AOvVaw36QY
SBMIhaRq1esIfUC72j

- Yan, Z.; Duan, N.; Bao, J.; Chen, P.; Zhou, M.; Li, Z.; and Zhou, J. 2016. Docchat: An information retrieval approach for chatbot engines using unstructured documents. In *Meeting of the Association for Computational Linguistics*.

Appendix 1- The preferred representational system questionnaire

1. When vacationing at the beach, the first thing that makes me glad to be there is:

- a ___ The feel of the cool sand, the warm sun or the fresh breeze on my face.
- b ___ The roar of the waves, the whistling wind or the sound of birds in the distance.
- c ___ This is the type of vacation that makes sense or the cost is reasonable.
- d ___ The scenery, the bright sun, and the blue water.

2. When overwhelmed, I find it helps if:

- a ___ I can see the big picture.
- b ___ I can hear what's going on.
- c ___ I can get in touch with what is happening.
- d ___ I make sense of things in my head.

3. When given an assignment at work, it is easier to carry out if:

- a ___ I can picture what is required.
- b ___ I have a feeling for what is required.
- c ___ I have an understanding of what is required.
- d ___ I have tuned into what's required.

4. I find it easier to follow a presentation if:

- a ___ I feel in touch with the presenter and the material is within my grasp.
- b ___ There is a visual display so that I can visualize the concepts.
- c ___ The presentation is based on facts and figures and is logically presented.
- d ___ The presenter speaks clearly with varying tonality or uses sound to emphasize message.

5. When buying a car, I make my decision on:

- a ___ The purchase price, gas mileage, safety features, etc.
- b ___ How comfortable the seats are or the feeling I get when I test drive it.
- c ___ The colour, styling or how I would look in it.
- d ___ The sound of the engine or stereo system or how quiet it rides.

6. I communicate my thoughts through:

- a ___ My tone of my voice.
- b ___ My words.
- c ___ My appearance.
- d ___ My feelings.

7. When I am anxious, the first thing that happens is:

- a ___ Things begin to sound different.
- b ___ Things begin to feel different.
- c ___ Things begin to look different.
- d ___ Things begin to not make sense.

8. During a discussion, I am most often influenced by:

a ___ The other person's logic.

b ___ The other person's tone of voice.

c ___ The energy I feel from the other person.

d ___ Seeing the other person's body language or being able to picture the other person's viewpoint.

9. I assess how well I am doing at work based on:

a ___ My understanding of what needs to be done.

b ___ How I see myself making progress.

c ___ How things sound.

d ___ How satisfied I feel.

10. One of my strengths is my ability to:

a ___ See what needs to be done.

b ___ Make sense of new facts and data.

c ___ Hear what sounds right.

d ___ Get in touch with my feelings.

11. I enjoy:

a ___ Choosing a piece of music to listen to.

b ___ Making a logical, compelling point.

c ___ Choosing clothes that are comfortable.

d ___ Choosing clothes that look good.

12. If you agree with someone, you are more likely to say:

a ___ That feels right.

b ___ That looks right.

c ___ That sounds right.

d ___ That makes sense.

Appendix 2- MBTI Personality Type questionnaire

- 1- At a party do you:
 - a. Interact with many, including strangers
 - b. Interact with a few, known to you

- 2- Are you more:
 - a. Realistic than speculative
 - b. Speculative than realistic

- 3- Is it worse to:
 - a. Have your “head in the clouds”
 - b. Be “in a rut”

- 4- Are you more impressed by:
 - a. Principles
 - b. Emotions

- 5- Are you more drawn to the towards the:
 - a. Convincing
 - b. Touching

- 6- Do you prefer to work:
 - a. To deadlines
 - b. Just “whenever”

7- Do you tend to choose:

- a. Rather carefully
- b. Somewhat impulsively

8- At parties do you:

- a. Stay late, with increasing energy
- b. Leave early with decreased energy

9- Are you more attracted to:

- a. sensible people
- b. Imaginative people

10- Are you more interested in:

- a. What is actual
- b. What is possible

11- in judging others are you more swayed by:

- a. Laws than circumstances
- b. Circumstances than laws

12- in approaching others is your inclination to be somewhat:

- a. Objective
- b. Personal

13- Are you more:

- a. Punctual
- b. Leisurely

14- Does it bother you more having things:

- a. Incomplete
- b. Completed

15- In your social groups do you:

- a. Keep abreast of other's happenings
- b. Get behind on the news

16- In doing ordinary things are you more likely to:

- a. Do it the usual way
- b. Do it your own way

17- Writers should:

- a. "say what they mean and mean what they say"
- b. Express thing more by use of analogy

18- Which appeals to you more:

- a. Consistency of thought
- b. Harmonious human relationships

19- Are more comfortable in making:

- a. Logical judgments
- b. Value judgments

20- Do you want things:

- a. Settled and decided
- b. Unsettled and undecided

21- Would you say you are more:

- a. Serious and determined
- b. Easy going

22- In phoning do you:

- a. Rerly question that it will all be said
- b. Rehearse what you'll say

23- Facts:

- a. "speak for themselves"
- b. Illustrate principals

24- Are visionaries:

- a. Somewhat annoying
- b. Rather fascinating

25- Are you more often:

- a. A cool-headed person
- b. A warm hearted person

26- Is it worse to be:

- a. Unjust
- b. Merciless

27- Should one usually let events occur:

- a. By careful selection and choice
- b. Randomly and by chance

28- Do you feel better about:

- a. Having purchased
- b. Having the option to buy

29- In company do you:

- a. Initiate conversation
- b. Wait to be approached

30- Common sense is:

- a. Rarely questionable
- b. Frequently questionable

31- Children often do not:

- a. Make themselves
- b. Exercise their fantasy enough

32- In making decisions do you feel more comfortable with:

- a. Standards
- b. Feelings

33- Are you more:

- a. Firm and gentle
- b. Gentle and firm

34- Which one is more admirable:

- a. The ability to organize and be methodical
- b. The ability to adopt and make do

35- Do you put more value on:

- a. Infinite
- b. Open-minded

36- Does new and non-routine interaction with others:

- a. Stimulate and energize you
- b. Tax your reserves

37- Are you more frequently:

- a. A practical sort of person
- b. A fanciful sort of person

38- Are you more likely to:

- a. See how others are useful
- b. See how others see

39- Which is more satisfying

- a. To discuss an issue thoroughly
- b. To arrive at agreement on an issue

40- Which rules you more:

- a. Your head
- b. Your heart

41- Are you more comfortable with work that is:

- a. Contracted
- b. Done on a casual basis

42- Do you tend to look for:

- a. The orderly
- b. Whatever turns up

43- Do you prefer:

- a. Many friends with brief contact
- b. A few friends with lengthier contact

44- Do you go more by:

- a. Facts
- b. Principals

45- Are you more interested in:

- a. Production and distribution
- b. Design and research

46- Which is more of a compliment:

- a. "There is a very logical person."
- b. "There is a very sentimental person."

47- Do you value in yourself more that you are:

- a. Unwavering
- b. Devoted

48- Do you more often prefer the:

- a. Final and unalterable statement
- b. Tentative and preliminary statement

49- Are you more comfortable:

- a. After a decision
- b. Before a decision

50- Do you:

- a. Speak easily and at length with strangers
- b. Find little to say to strangers

51- Are you more likely to trust your:

- a. Experience
- b. Hunch

52- Do you feel:

- a. More practical than ingenious
- b. More ingenious than practical

53- Which person is more to be complimented:

- a. Clear reason
- b. Strong feeling

54- Are you inclined more to be:

- a. Fair-minded
- b. Sympathetic

55- Is it preferable mostly to:

- a. Make sure things are arranged
- b. Just let things happen

56- In relationships should most things be:

- a. Re-negotiable
- b. Random and circumstantial

57- When the phone rings do you:

- a. Hasten to get to it first
- b. Hope someone else will answer

58- Do you prize more in yourself:

- a. A strong sense of reality
- b. A vivid imagination

59- Are you drawn more to:

- a. Fundamentals
- b. Overtones

60- Which seems the greater error:

- a. To be too passionate
- b. To be too objective

61- Do you see yourself as basically:

- a. Hard-headed
- b. Soft-hearted

62- Which situation appeals to you more:

- a. The structured and scheduled
- b. The unstructured and unscheduled

63- Are you a person that is more:

- a. Routinized than whimsical
- b. Whimsical than routinized

64- Are you more inclined to be:

- a. Easy to approach
- b. Somewhat reserved

65- In writings do you prefer:

- a. The more literal
- b. The more figurative

66- Is it harder for you to:

- a. Identify with others
- b. Utilize others

67- Which do you wish more for yourself:

- a. Clarity of reason
- b. Strength of compassion

68- Which is the greater fault:

- a. Being indiscriminate
- b. Being critical

69- Do you prefer the:

- a. Planned event
- b. Unplanned event

70- Do you tend to be more:

- a. Deliberate than spontaneous
- b. Spontaneous than deliberate

