CROSS-LINGUAL SENTIMENT CLASSIFICATION USING SEMI-SUPERVISED LEARNING

MOHAMMAD SADEGH HAJMOHAMMADI

A thesis submitted in fulfilment of the
requirements for the award of the degree of
Doctor of Philosophy (Computer Science)

Faculty of Computing
Universiti Teknologi Malaysia

MAY 2015

“To my beloved wife and son”

ACKNOWLEDGEMENT

Thanks to Almighty ALLAH for providing me the knowledge, guidance and patience to achieve this goal.

I wish to express my sincere appreciation to my main thesis supervisor, Dr. Roliana Ibrahim, for encouragement, guidance, and critics. I am also very thankful to my co-supervisor Professor Dr. Ali Selamat for his guidance, advices, and motivation. At the same time, I also appreciate Universiti Teknologi Malaysia for International Doctorate Fellowship (IDF) award.

I am grateful to all my family members, especially my mother, for their prayers and moral support. I am also deeply indebted to my wife, Fatemeh, for her continuous support and the inspiration throughout the journey. She is the best companion. Special thanks to my son, Mohammad Taha who has inspired me to keep on striving to complete the study.

ABSTRACT

Cross-lingual sentiment classification aims to utilize annotated sentiment resources in one language for text sentiment classification in another language. Automatic machine translation services are the most commonly used tools to directly project information from one language into another. However, different term distribution between translated and original documents, translation errors and different intrinsic structure of documents in various languages are the problems that lead to low performance in sentiment classification. Furthermore, due to the existence of different linguistic terms in different languages, translated documents cannot cover all vocabularies which exist in the original documents. The aim of this thesis is to propose an enhanced framework for cross-lingual sentiment classification to overcome all the aforementioned problems in order to improve the classification performance. Combination of active learning and semi-supervised learning in both single view and bi-view frameworks is proposed to incorporate unlabelled data from the target language in order to reduce term distribution divergence. Using bi-view documents can partially alleviate the negative effects of translation errors. Multi-view semi-supervised learning is also used to overcome the problem of low term-coverage through employing multiple source languages. Features that are extracted from multiple source languages can cover more vocabularies from test data and consequently, more sentimental terms can be used in the classification process. Content similarities of labelled and unlabelled documents are used through graph-based semi-supervised learning approach to incorporate the structure of documents in the target language into the learning process. Performance evaluation performed on sentiment data sets in four different languages certifies the effectiveness of the proposed approaches in comparison to the well-known baseline classification methods. The experiments show that incorporation of unlabelled data from the target language can effectively improve the classification performance. Experimental results also show that using multiple source languages in the multi-view learning model outperforms other methods. The proposed framework is flexible enough to be applied on any new language, and therefore, it can be used to develop multilingual sentiment analysis systems.

ABSTRAK

Klasifikasi sentimen silang bahasa bertujuan untuk menggunakan sumber-sumber sentimen beranotasi dalam satu bahasa untuk pengelasan sentimen teks dalam bahasa lain. Perkhidmatan penterjemahan mesin automatik merupakan alat-alat yang paling biasa digunakan untuk pemetaan langsung maklumat daripada satu bahasa kepada bahasa yang lain. Walau bagaimanapun, agihan terma yang berbeza antara dokumen terjemahan dan asal, kesilapan terjemahan dan struktur intrinsik yang berbeza pada dokumen dalam bahasa berbeza adalah masalah yang membawa kepada prestasi yang rendah dalam klasifikasi sentimen. Tambahan pula, disebabkan oleh kewujudan istilah linguistik yang berbeza dalam pelbagai bahasa, dokumen yang diterjemahkan tidak boleh meliputi semua kosa kata yang wujud dalam dokumen asal. Tujuan tesis ini adalah untuk mencadangkan rangka kerja yang dipertingkat bagi klasifikasi sentimen silang bahasa untuk mengatasi semua masalah yang dinyatakan di atas bagi meningkatkan prestasi klasifikasi. Gabungan pembelajaran aktif dan pembelajaran separa-selia dalam kedua-dua rangka kerja pandangan tunggal dan dwi-pandangan telah dicadangkan bagi menggabungkan data tidak dilabel dari bahasa sasaran untuk mengurangkan kesan negatif kesilapan terjemahan. Menggunakan dokumen dwi-pandangan boleh mengurangkan kesan negatif daripada kesilapan terjemahan. Pembelajaran separa-selia pelbagai pandangan juga digunakan untuk mengatasi masalah liputan terma yang rendah melalui penggunaan pelbagai bahasa sumber. Ciri-ciri yang diekstrak dari pelbagai bahasa sumber boleh meliputi lebih banyak perbendaharaan kata dalam data ujian dan membolehkan terma sentimental yang lebih banyak digunakan untuk menyumbang dalam proses pengelasan. Persamaan kandungan dokumen dilabel dan tidak dilabel digunakan melalui pendekatan separa-selia pembelajaran berasaskan graf untuk menggabungkan struktur dokumen dalam bahasa sasaran di dalam proses pembelajaran. Penilaian prestasi yang telah dijalankan pada set data sentimen dalam empat bahasa berbeza membuktikan keberkesanan pendekatan yang dicadangkan berbanding dengan kaedah klasifikasi yang terkenal dan asas. Ujikaji menunjukkan bahawa penggabungan data tidak dilabel dari bahasa sasaran boleh meningkatkan prestasi klasifikasi dengan berkesan. Keputusan ujikaji juga menunjukkan bahawa penggunaan bahasa pelbagai sumber dalam model pembelajaran pelbagai pandangan mengatasi prestasi kaedah-kaedah lain. Rangka kerja yang dicadangkan adalah cukup anjal untuk digunakan dalam apa-apa bahasa yang baru dan oleh itu, boleh digunakan untuk membangunkan sistem analisis sentiment berbilang bahasa.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

ACT	-	Active Co-Training
AL	-	Active Learning
ANN	-	Artificial Neural Network
AST	-	Active Self-Training
BOW	-	Bag Of Word
CRF	-	Conditional Random Fields
CLSC	-	Cross-Lingual Sentiment Classification
DACT	-	Density-based Active Co-Training
DBAST	-	Density-Based Active Self-Training
DF	-	Document Frequency
GBSSL	-	Graph-Based Semi-Supervised Learning
IG	-	Information Gain
IR	-	Information Retrieval
k-NN	-	*k*-Nearest Neighbour
MI	-	Mutual Information

MLMV	-	Multiple source Language Multi-View
MLSV	-	Multiple source Language Single-View
MSLDS	-	Multiple Source Languages Data Set
MT	-	Machine Translation
NB	-	Naïve Bayes
NLP	-	Natural Language Processing
POS	-	Part Of Speech
QBC	-	Query By Committee
SCL	-	Structural Correspondence Learning
SO	-	Sentiment Orientation
SO-CAL	-	Semantic Orientation CALculator
SSL	-	Semi-Supervised Learning
SSLDS	-	Single Source Language Data Set
ST	-	Self-Training
SVM	-	Support Vector Machine
TF-IDF	-	Term Frequency – Inverse Document Frequency
TSVM	-	Transductive Support Vector Machine

LIST OF APPENDICES

CHAPTER 1

INTRODUCTION

1.1 Overview

Over the years, surveys have been the main method for answering the question "*what do people think?*". Careful samplings of the polled population and a standardized questionnaire have become the standard ways of learning about large groups of people. Recently though, the era of widespread internet access and social media has brought a new way of learning about large populations. The advent of Web2.0 and social media contents such as online review web sites and personal blogs have created several opportunities for understanding the opinions of other people about social events, companies, products, news etc. However, because of the proliferation of different web sites, the task of finding and scanning opinion sites on the web and summarizing their information has been a very difficult task. We can find a huge volume of opinionated text at each site and obviously the task of analysing and summarizing this information into a useful format is very difficult. Therefore, an automated opinion mining and summarizing system is needed to overcome this difficulty.

Traditional Natural Language Processing (NLP) applications mostly concentrate on topical text characterization that deals with the communicated facts and objective presentation of the information. In recent years, the natural language community has recognized the value in analysing emotions and opinions expressed in free text. Opinion mining is the task of having computers automatically extract and understand the opinions in a text.

Text sentiment classification refers to the task of determining the sentiment polarity (e.g. positive or negative) of a given text document (Liu and Zhang, 2012). Recently, sentiment classification has received considerable attention in the natural language processing research community due to its many useful applications such as online product review classification (Kang *et al.*, 2012) and opinion summarization (Ku *et al.*, 2006).

Up until now, different methods have been used for sentiment classification. These methods can be categorised into two groups, namely; lexicon-based and machine learning based methods. The lexicon-based methods classify text documents based on the polarity of words and phrases contained in the text. If a text document contains more positive than negative terms, for example, it is classified as positive and vice versa (Turney, 2002; Taboada *et al.*, 2011). A sentiment lexicon is always used to determine the sentiment polarity of each term. In contrast, machine learning methods train a sentiment classifier based on labelled data using some machine learning classification algorithms (Pang *et al.*, 2002; Moraes *et al.*, 2013). The performance of these methods depends intensively on both the quality and quantity of labelled data as the training set for the sentiment classifier. Based on these two groups of methods, sentiment lexicons and annotated sentiment data can be seen as the most important resources for sentiment classification.

Although, this area is under consideration from the last decade for English language (Pang *et al.*, 2002; Turney, 2002), unfortunately, other languages are relatively ignored by the research communities. This has led to a scarcity of labelled corpus and sentiment lexicons in other languages (Wan, 2011; Martín-Valdivia *et al.*, 2013). Further, manual construction of reliable sentiment resources is a very difficult and time-consuming task. Therefore, the challenge is how to utilize labelled sentiment resources in one language (a resource-rich language such as English is always called the source language) for sentiment classification in another language (a resource-scarce language is called the target language). This subsequently leads to an interesting area of research called cross-lingual sentiment classification (CLSC).

The most direct solution to this problem is the use of machine translation systems to directly project the information of data from one language into another (Banea *et al.*, 2008; Wan, 2011; Martín-Valdivia *et al.*, 2013; Balahur and Turchi, 2014). However, because the training set and the test set come from two different languages having differing linguistic terms and writing styles, as well as originating from different cultures with various people interests, these methods cannot attain the performance results of monolingual sentiment classification methods in which the training and test samples are from the same language.

Due to this problem, numerous researchers try to find reliable techniques for cross-lingual sentiment classification. Different term distribution in the original and translated text, translation errors in the resource projection stage and different writing styles and document structures in different languages are some of serious problems which researchers were confronted with.

To overcome these problems, making use of unlabelled data from the target language can be helpful, since this type of data is always easy to obtain and has the same term distribution, same writing style, and same structure as the target language data. Therefore, employing unlabelled data from the target language in the learning process is expected to result in better classification performance in CLSC. This is the main idea behind all proposed approaches in this study. Active learning (AL) (Wang *et al.*, 2012) and semi-supervised learning (SSL) (Ortigosa-Hernández *et al.*, 2012) are two well-known techniques that make use of unlabelled data to improve classification performance. Both techniques are iterative processes. AL aims to reduce manual labelling efforts by finding the most informative examples for human labelling, while SSL tries to automatically label examples from unlabelled data in each cycle. Various types of semi-supervised learning models are proposed to overcome the aforementioned problems in this study.In this research, semi-supervised learning and active learning are utilized in order to incorporate unlabelled data from the target language and several classification models are proposed based on these approachs.

1.2 Background of the problem

User generated reviews are very important in business, e-commerce and education, since they consist of valuable opinions produced from user experiences. For example, in e-commerce sites, a product quality can be assessed by reading customer's reviews about the product. It can help customers to decide whether to select the product or not and it can help companies as well to evaluate their products.

Sentiment classification dates back to the early 2000's. There are two early works trying this task reported by Pang *et al.* (2002) and Turney (2002). Two different approaches were introduced in these two studies. The first paper used machine learning (or supervised) approach (Pang *et al.*, 2002) and the other one exploited a lexicon-based method (Turney, 2002). Supervised approaches rely on a large set of labelled data to train a classifier and then use this classifier to estimate the polarity label of unlabelled test data. Most of the existing studies locate sentiment classification as a supervised classification problem (Pang *et al.*, 2002; Riloff *et al.*, 2006; Prabowo and Thelwall, 2009; Ye *et al.*, 2009; Zhang *et al.*, 2011; Kang *et al.*, 2012). In supervised methods, some researchers considered different feature sets and various feature selection techniques to increase the performance of sentiment classification. The Bag of words (BOW) approach is the most popular techniques for text representation in sentiment classification (Pang *et al.*, 2002; Wang *et al.*, 2014). The main disadvantage of supervised methods is that it is very hard to prepare and annotate a large amount of labelled training data.

In parallel, several works have been performed in this area by using sentiment lexicons to classify documents according their sentiment. All of these works try to calculate the sentiment orientation of words in a document by using a dictionary or by exploiting a search engine to calculate the association of words with a known polarity seed set (Turney, 2002; Harb *et al.*, 2008; Taboada *et al.*, 2011). These types of works are considered as lexicon-based methods and are strongly dependent on sentiment lexicons.

The labelled corpus and sentiment lexicons are the most important resources for sentiment classification task. Since most recent research studies in sentiment classification have been performed in some limited number of languages, there are an insufficient number of annotated corpus and sentiment lexicon in other languages. Recently some researchers focus on cross-lingual sentiment classification, which tries to use sentiment resources in one language for sentiment classification in other languages.

Most approaches focused on resource projection from one language (always English) to another language with few sentiment resources and then used machine learning approach for sentiment classification, based on the projected resources. For example in (Banea *et al.*, 2008; Banea *et al.*, 2010), automatic machine translation engines were used to translate the English resources for subjectivity classification and then machine learning approaches were employed for classification based on translated corpora as training data. In some other works, resource translation was employed to compensate for the lack of training data in supervised sentiment classification in languages other than English (Dasgupta and Ng, 2009; Wan, 2009; Zhao et al., 2010; Wan, 2011). Most existing works in this area have used machine translation systems to translate labelled training data from the source language into the target language and perform sentiment classification in the target language (Banea *et al.*, 2010; Balahur and Turchi, 2014). Some other researchers have employed machine translation in the opposite direction so as to translate unlabelled test data from the target language into the source language and performed the classification in the source language (Prettenhofer and Stein, 2010; Martín-Valdivia *et al.*, 2013). Although machine translation is a reasonable tool for resource projection in the field of sentiment classification, working with translated data implies an increasing number of features, sparseness, and noise in datasets.

Another approach is that of feature translation, which involves translating the features extracted from labelled documents (Shi *et al.*, 2010; Moh and Zhang, 2012). The features, selected by a feature selection technique, are translated into different languages. Subsequently, based on those translated features, a new model is trained for each language. This approach only needs a bilingual dictionary to translate the

selected features. However, it can suffer from the inaccuracies of dictionary translation, in that words may have different meanings in different contexts. Additionally, selecting the features to be translated can be an intricate process.

Some other researchers tries to overcome the problem of CLSC through domain adaptation techniques (Prettenhofer and Stein, 2010; Wei and Pal, 2010). They adapted Structural Correspondence Learning (SCL) (Blitzer *et al.*, 2006) to use unlabelled data and a word translation oracle to induce correspondence among the words from both the source and target languages. However, translation errors and different document structures between two languages have been ignored in these studies.

The previous studies exhibit that relying only on translated resources cannot produce satisfactory result in cross-lingual sentiment classification, because the machine translation engines are still far from satisfactory. Even if the machine translation do well, it might have a systematic bias (Duh *et al.*, 2011). For example, the word “awesome” might be common in English reviews but when a non-English review translate to English, the work “excellent” may be generated instead. From the translation perspective, this is a correct translation but from classifier perspective, there is a domain mismatch due to differences in word distribution. Therefore, researchers try to overcome these limitations in different frameworks.

In recent studies, researchers employed semi-supervised learning to improve the accuracy of cross-lingual sentiment classification. In Wan (2011), two different views were used by exploiting semi-supervised co-training approach to classify Chinese review documents by using English training documents. Because the examples with the highest confidence are selected to add to the training data in each step of co-training and these examples are not necessarily the most informative ones, the improvement in the accuracy of this model is very limited. Additionally, when the initial classifiers in each view are not good enough, there will be an increased probability of adding examples having incorrect labels to the training set. Therefore, the addition of noisy examples not only cannot increase the accuracy of the learning model, but will also gradually decrease the performance of each classifier.

Although recent research works have tried to overcome some problems in resource projection and sentiment classification in CLSC, there are still several research gaps in this research area, which have not been considered in the literature. These gaps can be summarized as considering translation errors and information loss during the resource projection process, the problem of low vocabularies coverage and creation of sparseness in data representation of text documents in the target language and considering different intrinsic structures of text documents in the source and target languages.

Taking into account the existing gaps, this research aims to deal with the problems of cross-lingual sentiment classification under the semi-supervised learning strategy. Unlabelled documents from the target language are employed in the learning process of CLSC using semi-supervised learning approaches in order to narrow down the gaps between the training and test data. These unlabelled documents are always easy to obtain and have the same characteristics with the test documents. Therefore, employing unlabelled documents from the target language is expected to result in better classification performance in CLSC. Various types of semi-supervised learning models are proposed to overcome the aforementioned problems in this study.

1.3 Problem statement

In this study, we intend to overcome the problem of cross-lingual sentiment classification. This problem can be defined as follow:

Suppose we have two different languages: source language and target language and two different document sets: $L^S = \left\{ d_1^S, d_2^S, ..., d_{n_s}^S \right\}$ denotes the labelled text document set in the source language and $U^T = \left\{ d_1^T, d_2^T, ..., d_{n_t}^T \right\}$ denotes the unlabelled text document set in the target language, where n_s and n_t are the number of documents in the source language dataset and the target language dataset respectively. Let $Y^S = \left\{ y_1^S, y_2^S, ..., y_{n_s}^S \right\}$ denotes the label set of text documents in the source

language that $y_i^S = +1$ if the overall sentiment expressed in text document d_i^S is positive, while $y_i^S = -1$ if the overall sentiment expressed in text document d_i^S is negative. Given labelled examples (d_i^S, y_i^S) in the source language and unlabelled examples $(d_i^T, ?)$ in the target language, the problem of CLSC is to train a model, in order to predict unknown labels of d_i^T in the target language by leveraging on labelled examples in the source language. We can use machine translation services to fill the language gap by translating the labelled data from the source language into the target or translating the unlabeled data from the target language into the source. However, using translated data leads the existing classification models to be confronted by new problems as follows:

1.3.1 Different term distribution in the source and the target languages

The first problem is the difference in term distribution between the original and the translated text documents due to the dissimilarity in cultures, writing styles and also linguistic expressions in various languages. It means that a term may be frequently used in one language to express an opinion while the translation of that term is rarely used in another language. This problem leads to create different feature distribution between training and test data. Therefore, a classifier, which trains based on the training text documents from the source language, cannot perform well on the test documents in the target language. Incorporating unlabelled data from the target language into the learning process can reduce feature distribution divergence.

1.3.2 Information loss and translation errors in resource projection

Because machine translation quality is still far from satisfactory, there are several translation errors in resource projection process, which leads to decrease the quality of projected data and loss some critical information. These errors may even change the sentiment polarity of an opinionated text document. Therefore, applying

monolingual sentiment classification techniques directly on the erroneous translation of training or test documents may seriously decrease the sentiment classification performance.

1.3.3 Low coverage of the target language terms by the features extracted from the source language text documents

Because the training data and test data come from two different languages having differing linguistic terms, features extracted from text documents of the source language cannot cover all the vocabularies contained in the text documents of the target language. Consequently, several sentimental words may be ignored when documents in the target language are represented based on the extracted features. This problem also leads to create sparseness in data representation in the target language and consequently decrease the performance of sentiment classification.

1.3.4 Different intrinsic structure of text documents in various languages

Due to the discrepancy in writing style and linguistic terms in various languages, the intrinsic structures of documents in different languages are dissimilar. As a result, the classifier trained based on the training data in one language cannot perform well in another language with different intrinsic structure. In fact, ignoring the intrinsic manifold structure of documents in the target language can degrade the classification performance in CLSC.

1.4 Research question

This study aims to overcome the aforementioned problems by exploiting unlabelled documents from the target language into the classification process. Considering these problems, the main research question of this study is:

"How to improve the performance of cross-lingual sentiment classification through incorporating information of unlabelled data from the target language into the learning process?"

In order to answer the main question, the following research questions that address the problem in detail are defined:

(i) How to effectively exploit unlabelled documents from the target language into the learning process of cross-lingual sentiment classification in order to improve the classification performance?

(ii) How to alleviate the destructive effects of translation errors in cross-lingual resource projection?

(iii) How can the use of labelled data from multiple source languages improve the performance of cross-lingual sentiment classification?

(iv) How to involve the intrinsic structure of document in the target language into the learning process of cross-lingual sentiment classifier?

1.5 Research goal

The aim of the research is to propose an enhanced cross-lingual sentiment classification framework in which the aforementioned problems are considered in order to improve the classification performance. By addressing the existing problems in previous works, the research strives to design and develop learning models into the above-mentioned framework which fill the gaps between the training and test documents in the source and target languages with the ultimate goal of improving the performance of CLSC.

1.6 Research objectives

In order to achieve the research goal, several research objectives have been identified and listed as follows:

(i) To propose a learning model based on the combination of semi-supervised learning and active learning to effectively utilize unlabelled documents into the learning process of cross-lingual sentiment classification.

(ii) To improve the performance of the first proposed model by employing bidirectional translation to create bi-view data in order to alleviate the destructive effects of translation errors.

(iii) To propose a multi-view semi-supervised learning model in which labelled data from multiple source languages are employed to cover more vocabularies from the target language in order to improve the performance of cross-lingual sentiment classification.

(iv) To propose a similarity-based classification model using graph-based semi-supervised learning in which the intrinsic structure of documents in the target language is considered.

1.7 Research scopes

To solve the cross-lingual sentiment classification problem in this research, the following constraints are considered:

(i) This research focuses on classifying book review documents (Prettenhofer and Stein, 2010; Pan *et al.*, 2011) based on the overall

sentiment orientation of each text document due to the availability of this domain in different languages.

(ii) Machine translation is used as projection tool in this study to translate whole text document from one language into another. Google Translate engine has been utilized as machine translation service.

(iii) This research only focuses on increasing the performance of machine learning methods in cross-lingual sentiment classification and lexicon-based approaches will not be considered.

(iv) In this study, two European and two Asian languages are used as the target languages while English is used as the main source language.

(v) In this study, with the exception for Tokenization tool, it is assumed that there are not any NLP tools (i.e. POS tagger or parser) in the target language.

(vi) Content similarity of documents is used as a simple structural similarity measure which introduce the intrinsic structure of documents in the graph-based method and other methods of introducing intrinsic structure (e.g. opinions, methods of expressing sentiment, opinion holder characteristics) are not considered in this study.

1.8 Significance of the research

In the past decade, sentiment analysis has become a hot research field and a booming industry. For instance, IBM SPSS[1] provides quantitative sentiment summaries of survey data to assist businesses in understanding consumer attitudes. LexisNexis[2] compiles consumer confidence and brand perception summaries using news media, while OpSec[3] also mines user-generated data (social media). Wall Street has also started to use sentiment analysis in their trading algorithms with companies like OpFine[4] providing up-to-date sentiment tracking of financial news. Even several major news sources like The Washington Post[5] now provide social media statistics on popular political figures.

As mentioned in (Pang and Lee, 2008), 81% of users in internet have performed online search on a product at least once and 73% to 87% of these users report that product reviews had a significant influence on their purchase. These statistics show that the sentiment classification of reviews is very helpful to customers to select appropriate products, which has motivated researchers to pay more attention to this area. For this classification task there are several method introduced by researchers and most accurate methods are machine learning methods. Unfortunately, in many languages, there are not enough annotated sentiment resources to use in supervised classification and manual construction of labelled corpus is a very hard and time-consuming task.

On the other hand, in many applications, companies want to analyse and compare the opinions of their customers about their services and products in different

---

[1] http://www-01.ibm.com/software/analytics/spss/

[2] http://www.lexisnexis.com/risk/data-analytics.aspx

[3]http://opsecsecurity.com/brand-protection/online-brand-protection/sentimentanalysis

[4] http://www.opfine.com/

[5] http://www.washingtonpost.com/politics/mention-machine/

countries with different languages. Therefore, employing new techniques that use labelled data in a resource–rich language to train sentiment classifier in a resource-scarce language is very useful in actual world. Recently, several methods have been proposed to solve this problem by using machine translation to translate labelled corpus from the source language into the target language or translate unlabelled data from the target language into the source language and applying monolingual sentiment classification on translated data. Since machine translation quality is still far from satisfactory, applying monolingual sentiment classification methods on translated data may apparently decrease the classification accuracy. In addition, even if translation of labelled or unlabelled data is completely correct, the cross-lingual classifier cannot perform as well as monolingual classifier since the data distribution across languages is different due to the difference in culture, writing style, and linguistic expression. In addition, the structure of data in target language should be considered as an important parameter to design classification models. Therefore, this study aims to create the cross-lingual classification models that use the labelled and unlabelled data in the source and the target languages to improve the cross-lingual classification performance, which is an urgent need in today's sentiment analysis applications.

1.9 Thesis outline

This thesis is organized into eight chapters as follows:

Chapter 1, *Introduction,* is started with an introduction to the research topic. After that, the research background and research problems are explained and research questions and objectives of research are introduced. Finally, the importance of research is expressed.

Chapter 2, *Literature review*, provides the background information and reviews the previous studies in this field that leads to find the research gaps and formulate the research problem.

Chapter 3, *Research methodology*, explains the methods and datasets, which are used in this research. The research flow is described systematically in this chapter. Evaluation metrics and evaluation framework also are explained in this chapter.

Chapter 4, *Density-based active self-training model for cross-lingual sentiment classification*, explains the development process of the first proposed model, which combine active learning and self-training to incorporate unlabelled data in the learning process. This model is evaluated and compared with some other baseline methods in this chapter.

Chapter 5, *Density-based active co-training model based on bi-view data*, addresses the design and development steps of second proposed model, which enhances the first proposed model by using bidirectional translation in order to decrease the negative effects of translation errors. Corresponding results and evaluations are also given in this chapter.

Chapter 6, *Multiple source languages in multi-view semi-supervised learning model*, investigates the effects of using multiple source languages on the classification performance of CLSC and introduces the third proposed model, which uses multi-view semi-supervised learning approach.

Chapter 7, *Incorporating intrinsic structure of target language data through graph-based semi-supervised learning model*, describes the implementation process of the last proposed model, which employs the intrinsic structure of documents in the target language into the learning process. This chapter also shows the results obtained from the proposed model and compares the performance of this model with other methods.

Chapter 8, *Conclusion and future works*, concludes the research, provides the list of contributions, states the limitations of proposed models and expresses some recommendations for future study.

1.10 Summary

The principles of the research and the essential parts of this study were introduced in this chapter. An overview of the research topic, the background of the research problem, problem statement along with research questions, research goal, objectives, and scopes of the current research as well as the significant of this research were described as an introduction of this study. The aim of this chapter is to provide an overall description of the main parts of this research.

REFERENCES

Baccianella, S., Esuli, A., and Sebastiani, F. (2010). SentiWordNet 3.0: An Enhanced Lexical Resource for Sentiment Analysis and Opinion Mining. *Proceedings of the Seventh conference on International Language Resources and Evaluation (LREC'10)*, European Language Resources Association (ELRA), 2200-2204.

Balahur, A., Mihalcea, R., and Montoyo, A. (2014). Computational approaches to subjectivity and sentiment analysis: Present and envisaged methods and applications. *Computer Speech & Language* 28(1), 1-6.

Balahur, A., and Turchi, M. (2012). Multilingual Sentiment Analysis using Machine Translation? *Proceedings of the 3rd Workshop on Computational Approaches to Subjectivity and Sentiment Analysis*, 12 July 2012. Jeju, Republic of Korea, 52-60.

Balahur, A., and Turchi, M. (2014). Comparative experiments using supervised learning and machine translation for multilingual sentiment analysis. *Computer Speech & Language* 28(1), 56-75.

Banea, C., Mihalcea, R., and Wiebe, J. (2010). Multilingual subjectivity: are more languages better? *Proceedings of the 23rd International Conference on Computational Linguistics*, 23-27 August. Beijing, China, Association for Computational Linguistics, 28-36.

Banea, C., Mihalcea, R., and Wiebe, J. (2014). Sense-level subjectivity in a multilingual setting. *Computer Speech & Language* 28(1), 7-19.

Banea, C., Mihalcea, R., Wiebe, J., and Hassan, S. (2008). Multilingual subjectivity analysis using machine translation. *Proceedings of the Conference on Empirical Methods in Natural Language Processing (EMNLP '08)*, 25-27 October. Stroudsburg, PA, USA, Association for Computational Linguistics, 127-135.

Blitzer, J., Dredze, M., and Pereira, F. (2007). Biographies, bollywood, boom-boxes and blenders: Domain adaptation for sentiment classification. *Proceedings of the*

45th Annual Meeting of the Association of Computational Linguistics, 23-25 June. Prague, Czech Republic, 440-447.

Blitzer, J., McDonald, R., and Pereira, F. (2006). Domain adaptation with structural correspondence learning. *Proceedings of the 2006 Conference on Empirical Methods in Natural Language Processing*, 22-23 July. Sydney, Australia, Association for Computational Linguistics, 120-128.

Blum, A., and Mitchell, T. (1998). Combining labeled and unlabeled data with co-training. *Proceedings of the eleventh annual conference on Computational learning theory*, 24-26 July. Madison, WI, USA, ACM, 92-100.

Brefeld, U., and Scheffer, T. (2004). Co-EM support vector learning. *Proceedings of the twenty-first international conference on Machine learning*, Banff, Alberta, Canada, ACM, 16-23.

Brooke, J., Tofiloski, M., and Taboada, M. (2009). Cross-Linguistic Sentiment Analysis: From English to Spanish. *Proceeding of International Conference on Recent Advances in NLP*, Borovets, Bulgaria Association for Computational Linguistics, 50-54.

Chen, L. S., Liu, C. H., and Chiu, H. J. (2011). A neural network based approach for sentiment classification in the blogosphere. *Journal of Informetrics* 5(2), 313-322.

Cheng, J., and Wang, K. (2007). Active learning for image retrieval with Co-SVM. *Pattern Recognition* 40(1), 330-334.

Chihli, H., and Hao-Kai, L. (2013). Using Objective Words in SentiWordNet to Improve Word-of-Mouth Sentiment Classification. *Intelligent Systems, IEEE* 28(2), 47-54.

Cohn, D., Atlas, L., and Ladner, R. (1994). Improving Generalization with Active Learning. *Machine Learning* 15(2), 201-221.

Cui, H., Mittal, V., and Datar, M. (2006). Comparative experiments on sentiment classification for online product reviews. *Proceedings of the 21st national conference on Artificial intelligence - Volume 2*, Boston, Massachusetts, USA, AAAI Press, 1265-1270.

Dasgupta, S., and Ng, V. (2009). Mine the easy, classify the hard: a semi-supervised approach to automatic sentiment classification. *Proceedings of the Joint Conference of the 47th Annual Meeting of the ACL and the 4th International*

Joint Conference on Natural Language Processing of the AFNLP: Volume 2, 2-7 Aug. Suntec, Singapore, Association for Computational Linguistics, 701-709.

Demirtas, E., and Pechenizkiy, M. (2013). Cross-lingual Polarity Detection with Machine Translation. *Proceedings of the Second International Workshop on Issues of Sentiment Discovery and Opinion Mining*, August. Chicago, IL, USA, ACM, 9-17.

Denecke, K. (2008). Using sentiwordnet for multilingual sentiment analysis. *IEEE 24th International Conference on Data Engineering Workshop (ICDEW 2008).* 7-12 April. Cancun, Mexico, IEEE, 507-512.

Duh, K., Fujino, A., and Nagata, M. (2011). Is machine translation ripe for cross-lingual sentiment classification? *Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies: short papers - Volume 2*, June. Portland, Oregon, Association for Computational Linguistics, 429-433.

Esuli, A., and Sebastiani, F. (2005). Determining the semantic orientation of terms through gloss classification. *Proceedings of the 14th ACM international conference on Information and knowledge management*, 31 October - 05 November. Bremen, Germany, ACM, 617-624.

Freund, Y., Seung, H. S., Shamir, E., and Tishby, N. (1997). Selective Sampling Using the Query by Committee Algorithm. *Machine Learnining* 28(2-3), 133-168.

Gamon, M. (2004). Sentiment classification on customer feedback data: noisy data, large feature vectors, and the role of linguistic analysis. *Proceedings of the 20th international conference on Computational Linguistics*, Geneva, Switzerland, Association for Computational Linguistics, 841-847.

Gamon, M., Aue, A., Corston-oliver, S., and Ringger, E. K. (2005). Pulse: Mining customer opinions from free text. *Proceedings of the 6th international conference on Advances in Intelligent Data Analysis*, 8-10 September Madrid, Spain, 121-132.

Ghorbel, H., and Jacot, D. (2011). Sentiment Analysis of French Movie Reviews. In Pallotta, V., Soro, A., and Vargiu, E. (Ed.) *Advances in Distributed Agent-Based Retrieval Tools* **(**97-108). Springer Berlin Heidelberg.

Goldberg, A. B., and Zhu, X. (2006). Seeing stars when there aren't many stars: graph-based semi-supervised learning for sentiment categorization. *Proceedings of the*

First Workshop on Graph Based Methods for Natural Language Processing, June. New York, USA, Association for Computational Linguistics, 45-52.

Harb, A., Planti, M., Dray, G., Roche, M., Fran, Trousset, o., and Poncelet, P. (2008). Web opinion mining: how to extract opinions from blogs? *Proceedings of the 5th international conference on Soft computing as transdisciplinary science and technology*, 27-31 Oct. Cergy-Pontoise, France, ACM, 211-217.

Hatzivassiloglou, V., and McKeown, K. R. (1997). Predicting the semantic orientation of adjectives. *Proceedings of the eighth conference on European chapter of the Association for Computational Linguistics*, Madrid, Spain, Association for Computational Linguistics, 174-181.

Hogenboom, A., Heerschop, B., Frasincar, F., Kaymak, U., and de Jong, F. (2014). Multi-lingual support for lexicon-based sentiment analysis guided by semantics. *Decision Support Systems* 63,43-53.

Hu, M., and Liu, B. (2004a). Mining and summarizing customer reviews. *Proceedings of the tenth ACM SIGKDD international conference on Knowledge discovery and data mining*, 22-25 August Seattle, WA, USA, ACM, 168-177.

Hu, M., and Liu, B. (2004e). Mining opinion features in customer reviews. *Proceedings of the 19th national conference on Artifical intelligence*, 25-29 July. San Jose, California, AAAI Press, 755-760.

Huang, S., Niu, Z., and Shi, C. (2014). Automatic construction of domain-specific sentiment lexicon based on constrained label propagation. *Knowledge-Based Systems* 56, 191-200.

Izard, C. E. (1971). *The Face of Emotion*. New York, USA: Appleton Century Crofts.

Jakob, N., and Gurevych, I. (2010). Extracting opinion targets in a single- and cross-domain setting with conditional random fields. *Proceedings of the 2010 Conference on Empirical Methods in Natural Language Processing*, 9-11 Oct. Cambridge, Massachusetts, USA, Association for Computational Linguistics, 1035-1045.

Jiang, Z., Zhang, S., and Zeng, J. (2013). A hybrid generative/discriminative method for semi-supervised classification. *Knowledge-Based Systems* 37, 137-145.

Jin, W., Ho, H. H., and Srihari, R. K. (2009). OpinionMiner: a novel machine learning system for web opinion mining and extraction. *Proceedings of the 15th ACM SIGKDD international conference on Knowledge discovery and data mining*, 28 June - 1 July. Paris, France, ACM, 1195-1204.

Jingbo, Z., Huizhen, W., Tsou, B. K., and Ma, M. (2010). Active Learning With Sampling by Uncertainty and Density for Data Annotations. *IEEE Transactions on Audio, Speech, and Language Processing* 18(6), 1323-1331.

Joachims, T. (1999). Making large-scale support vector machine learning practical. In Scholkopf, B., Burges, C. J. C., and Smola, A. J. (Ed.) *Advances in kernel methods* **(**169-184). Cambridge, MA, USA: MIT Press.

Kang, H., Yoo, S. J., and Han, D. (2012). Senti-lexicon and improved Naïve Bayes algorithms for sentiment analysis of restaurant reviews. *Expert Systems with Applications* 39(5), 6000-6010.

Kessler, J. S., and Nicolov, N. (2009). Targeting sentiment expressions through supervised ranking of linguistic configurations. *Proceedings of the Third International AAAI Conference on Weblogs and Social Media*, San Jose, California, USA, 90-97.

Kim, S. M., and Hovy, E. (2004). Determining the sentiment of opinions. *Proceedings of the 20th international conference on Computational Linguistics (COLING '04)*, Geneva, Switzerland, Association for Computational Linguistics, 1367-1375.

Kim, S. M., and Hovy, E. (2006). Identifying and analyzing judgment opinions. *Proceedings of the main conference on Human Language Technology Conference of the North American Chapter of the Association of Computational Linguistics*, New York, USA, Association for Computational Linguistics, 200-207.

Ku, L. W., Liang, Y. T., and Chen, H. H. (2006). Opinion extraction, summarization and tracking in news and blog corpora. *Proceedings of AAAI Spring Symposium on Computational Approaches to Analyzing Weblogs*, 27-29 March. Palo Alto, California, AAAI, 100-107.

Leng, Y., Xu, X., and Qi, G. (2013). Combining active learning and semi-supervised learning to construct SVM classifier. *Knowledge-Based Systems* 44, 121-131.

Lewis, D. D., and Gale, W. A. (1994). A sequential algorithm for training text classifiers. *Proceedings of the 17th annual international ACM SIGIR conference on Research and development in information retrieval*, Dublin, Ireland, Springer-Verlag, 3-12.

Li, M., and Sethi, I. K. (2006). Confidence-based active learning. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 28(8), 1251-1261.

Li, S., Ju, S., Zhou, G., and Li, X. (2012). Active Learning for Imbalanced Sentiment Classification. *Proceedings of the 2012 Joint Conference on Empirical Methods in Natural Language Processing and Computational Natural Language Learning*, 12-14 July. Jeju Island, Korea, Association for Computational Linguistics, 139-148.

Li, S., Wang, R., Liu, H., and Huang, C.-R. (2013). Active Learning for Cross-Lingual Sentiment Classification. In Zhou, G., Li, J., Zhao, D., and Feng, Y. (Ed.) *Natural Language Processing and Chinese Computing* **(**236-246). Springer Berlin Heidelberg.

Li, Y., and Guo, M. (2012). A new relational Tri-training system with adaptive data editing for inductive logic programming. *Knowledge-Based Systems* 35, 173-185.

Liu, B. (2012). *Sentiment Analysis and Opinion Mining*. California, USA: Morgan & Claypool Publishers.

Liu, B., and Zhang, L. (2012). A Survey of Opinion Mining and Sentiment Analysis. In Aggarwal, C. C., and Zhai, C. (Ed.) *Mining Text Data* **(**415-463). USA: Springer US.

Liu, K., and Zhao, J. (2009). Cross-domain sentiment classification using a two-stage method. *Proceedings of the 18th ACM conference on Information and knowledge management*, 2-6 November. Hong Kong, China, ACM, 1717-1720.

Liu, X., Pan, S., Hao, Z., and Lin, Z. (2014). Graph-based semi-supervised learning by mixed label propagation with a soft constraint. *Information Sciences* 277, 327-337.

Lu, B., Tan, C., Cardie, C., and Tsou, B. K. (2011). Joint bilingual sentiment classification with unlabeled parallel corpora. *Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies - Volume 1*, Portland, Oregon, Association for Computational Linguistics, 320-330.

Martín-Valdivia, M. T., Martínez-Cámara, E., Perea-Ortega, J. M., and Ureña-López, L. A. (2013). Sentiment polarity detection in Spanish reviews combining supervised and unsupervised approaches. *Expert Systems with Applications* 40(10), 3934-3942.

Martin, J. R., and White, P. R. R. (2005). *The Language Of Evaluation: Appraisal In English*. London: Palgrave.

McDonald, R., Hannan, K., Neylon, T., Wells, M., and Reynar, J. (2007). Structured Models for Fine-to-Coarse Sentiment Analysis. *Proceedings of the 45th Annual Meeting of the Association of Computational Linguistics*, 23-30 June. Prague, Czech Republic, 432–439.

Meng, X., Wei, F., Liu, X., Zhou, M., Xu, G., and Wang, H. (2012). Cross-Lingual Mixture Model for Sentiment Classification. *Proceedings of the 50th Annual Meeting of the Association for Computational Linguistics*, 8-14 July. Jeju, Republic of Korea, 572–581.

Mihalcea, R., Banea, C., and Wiebe, J. (2007). Learning multilingual subjective language via cross-lingual projections. *Proceedings of the 45th Annual Meeting of the Association of Computational Linguistics*, 23-30 June. Prague, Czech Republic, 976–983.

Moh, T. S., and Zhang, Z. (2012). Cross-lingual text classification with model translation and document translation. *Proceedings of the 50th Annual Southeast Regional Conference*, Tuscaloosa, Alabama, ACM, 71-76.

Montejo-Ráez, A., Martínez-Cámara, E., Martín-Valdivia, M. T., and Ureña-López, L. A. (2014). Ranked WordNet graph for Sentiment Polarity Classification in Twitter. *Computer Speech & Language* 28(1), 93-107.

Montoyo, A., Martínez-Barco, P., and Balahur, A. (2012). Subjectivity and sentiment analysis: An overview of the current state of the area and envisaged developments. *Decision Support Systems* 53(4), 675-679.

Moraes, R., Valiati, J. F., and Gavião Neto, W. P. (2013). Document-level sentiment classification: An empirical comparison between SVM and ANN. *Expert Systems with Applications* 40(2), 621-633.

Muslea, I., Minton, S., and Knoblock, C. A. (2006). Active learning with multiple views. *Journal of Artificial Intelligence Research* 27(1), 203-233.

Na, J. C., Khoo, C., and Wu, P. H. J. (2005). Use of negation phrases in automatic sentiment classification of product reviews. *Library Collections, Acquisitions, and Technical Services* 29(2), 180-191.

Neviarouskaya, A., Prendinger, H., and Ishizuka, M. (2010). Recognition of affect, judgment, and appreciation in text. *Proceedings of the 23rd International Conference on Computational Linguistics*, Beijing, China, Association for Computational Linguistics, 806-814.

Ortigosa-Hernández, J., Rodríguez, J. D., Alzate, L., Lucania, M., Inza, I., and Lozano, J. A. (2012). Approaching Sentiment Analysis by using semi-supervised learning of multi-dimensional classifiers. *Neurocomputing* 92, 98-115.

Pan, J., Xue, G. R., Yu, Y., and Wang, Y. (2011). Cross-Lingual Sentiment Classification via Bi-view Non-negative Matrix Tri-Factorization. In Huang, J., Cao, L., and Srivastava, J. (Ed.) *Advances in Knowledge Discovery and Data Mining* **(**289-300). Springer Berlin / Heidelberg.

Pang, B., and Lee, L. (2004). A sentimental education: sentiment analysis using subjectivity summarization based on minimum cuts. *Proceedings of the 42nd Annual Meeting on Association for Computational Linguistics (ACL '04)*, Barcelona, Spain, Association for Computational Linguistics, 271-278.

Pang, B., and Lee, L. (2005). Seeing stars: exploiting class relationships for sentiment categorization with respect to rating scales. *Proceedings of the 43rd Annual Meeting on Association for Computational Linguistics*, Ann Arbor, Michigan, Association for Computational Linguistics, 115-124.

Pang, B., and Lee, L. (2008). Opinion Mining and Sentiment Analysis. *Foundations and Trends in Information Retrieval* 2(1-2), 1-135.

Pang, B., Lee, L., and Vaithyanathan, S. (2002). Thumbs up? sentiment classification using machine learning techniques. *Proceedings of the ACL-02 conference on Empirical methods in natural language processing*, Association for Computational Linguistics, 79-86.

Park, S. B., and Zhang, B. T. (2004). Co-trained support vector machines for large scale unstructured document classification using unlabeled data and syntactic information. *Information Processing & Management* 40(3), 421-439.

Ponomareva, N., and Thelwall, M. (2013). Semi-supervised vs. Cross-domain Graphs for Sentiment Analysis. *Proceedings of the International Conference Recent Advances in Natural Language Processing (RANLP)*, September. Hissar, Bulgaria, INCOMA Ltd., 571-578.

Popescu, A. M., and Etzioni, O. (2005). Extracting product features and opinions from reviews. *Proceedings of the conference on Human Language Technology and Empirical Methods in Natural Language Processing*, Vancouver, British Columbia, Canada, Association for Computational Linguistics, 339-346.

Prabowo, R., and Thelwall, M. (2009). Sentiment analysis: A combined approach. *Journal of Informetrics* 3(2), 143-157.

Prettenhofer, P., and Stein, B. (2010). Cross-language text classification using structural correspondence learning. *Proceedings of the 48th Annual Meeting of the Association for Computational Linguistics*, Uppsala, Sweden, Association for Computational Linguistics, 1118-1127.

Prettenhofer, P., and Stein, B. (2011). Cross-Lingual Adaptation Using Structural Correspondence Learning. *ACM Transactions on Intelligent Systems and Technology (TIST)* 3(1), 1-22.

Qiu, G., Liu, B., Bu, J., and Chen, C. (2011). Opinion word expansion and target extraction through double propagation. *Computational linguistics* 37(1), 9-27.

Ren, Y., Kaji, N., Yoshinaga, N., Toyoda, M., and Kitsuregawa, M. (2011). Sentiment Classification in Resource-Scarce Languages by using Label Propagation. *25th Pacific Asia Conference on Language, Information and Computation*, 16-18 December. Singapore, 420--429.

Riloff, E., Patwardhan, S., and Wiebe, J. (2006). Feature subsumption for opinion analysis. *Proceedings of the 2006 Conference on Empirical Methods in Natural Language Processing*, Sydney, Australia, Association for Computational Linguistics, 440-448.

Roy, N., and McCallum, A. (2001). Toward Optimal Active Learning through Sampling Estimation of Error Reduction. *Proceedings of the Eighteenth International Conference on Machine Learning*, Williamstown, MA, USA, Morgan Kaufmann Publishers Inc., 441-448.

Serrano-Guerrero, J., Olivas, J. A., Romero, F. P., and Herrera-Viedma, E. (2015). Sentiment analysis: A review and comparative analysis of web services. *Information Sciences* 311, 18-38.

Shi, L., Mihalcea, R., and Tian, M. (2010). Cross language text classification by model translation and semi-supervised learning. *Proceedings of the 2010 Conference on Empirical Methods in Natural Language Processing*, Cambridge, Massachusetts, Association for Computational Linguistics, 1057-1067.

Sinno Jialin, P., and Qiang, Y. (2010). A Survey on Transfer Learning. *IEEE Transactions on Knowledge and Data Engineering* 22(10), 1345-1359.

Stoyanov, V., and Cardie, C. (2008). Topic identification for fine-grained opinion analysis. *Proceedings of the 22nd International Conference on Computational Linguistics - Volume 1*, Manchester, United Kingdom, Association for Computational Linguistics, 817-824.

Sun, S., and Zhang, Q. (2011). Multiple-View Multiple-Learner Semi-Supervised Learning. *Neural Processing Letters* 34(3), 229-240.

Taboada, M., Brooke, J., Tofiloski, M., Voll, K., and Stede, M. (2011). Lexicon-based methods for sentiment analysis. *Computational linguistics* 37(2), 267-307.

Tan, S., and Wang, Y. (2011). Weighted SCL model for adaptation of sentiment classification. *Expert Systems with Applications* 38(8), 10524-10531.

Tang, M., Luo, X., and Roukos, S. (2002). Active learning for statistical natural language parsing. *Proceedings of the 40th Annual Meeting on Association for Computational Linguistics*, Philadelphia, Pennsylvania, Association for Computational Linguistics, 120-127.

Titov, I., and McDonald, R. (2008). Modeling online reviews with multi-grain topic models. *Proceedings of the 17th international conference on World Wide Web*, Beijing, China, ACM, 111-120.

Turney, P. D. (2002). Thumbs up or thumbs down?: semantic orientation applied to unsupervised classification of reviews. *Proceedings of the 40th Annual Meeting on Association for Computational Linguistics*, Philadelphia, Pennsylvania, Association for Computational Linguistics, 417-424.

Turney, P. D., and Littman, M. L. (2003). Measuring praise and criticism: Inference of semantic orientation from association. *ACM Transactions on Information Systems (TOIS)* 21(4), 315-346.

Wan, X. (2008). Using bilingual knowledge and ensemble techniques for unsupervised Chinese sentiment analysis. *Proceedings of the Conference on Empirical Methods in Natural Language Processing*, Honolulu, Hawaii, Association for Computational Linguistics, 553-561.

Wan, X. (2009). Co-training for cross-lingual sentiment classification. *Proceedings of the Joint Conference of the 47th Annual Meeting of the ACL and the 4th International Joint Conference on Natural Language Processing of the AFNLP*, Suntec, Singapore, Association for Computational Linguistics, 235-243.

Wan, X. (2011). Bilingual co-training for sentiment classification of Chinese product reviews. *Computational Linguistics* 37(3), 587-616.

Wang, G., Sun, J., Ma, J., Xu, K., and Gu, J. (2014). Sentiment classification: The contribution of ensemble learning. *Decision Support Systems* 57, 77-93.

Wang, R., Kwong, S., and Chen, D. (2012). Inconsistency-based active learning for support vector machines. *Pattern Recognition* 45(10), 3751-3767.

Wei, B., and Pal, C. (2010). Cross lingual adaptation: an experiment on sentiment classifications. *Proceedings of the ACL 2010 Conference Short Papers*, Uppsala, Sweden, Association for Computational Linguistics, 258-262.

Whitelaw, C., Garg, N., and Argamon, S. (2005). Using appraisal groups for sentiment analysis. *Proceedings of the 14th ACM international conference on Information and knowledge management*, Bremen, Germany, ACM, 625-631.

Wiebe, J., and Mihalcea, R. (2006). Word sense and subjectivity. *Proceedings of the 21st International Conference on Computational Linguistics and the 44th annual meeting of the Association for Computational Linguistics*, 17-21 July. Sydney, Australia, Association for Computational Linguistics, 1065-1072.

Wiebe, J., and Riloff, E. (2005). Creating subjective and objective sentence classifiers from unannotated texts. In Gelbukh, A. (Ed.) *Computational Linguistics and Intelligent Text Processing* **(**486-497). Berlin: Springer-Verlag.

Wiebe, J. M., Bruce, R. F., and O'Hara, T. P. (1999). Development and use of a gold-standard data set for subjectivity classifications. *Proceedings of the 37th annual meeting of the Association for Computational Linguistics on Computational Linguistics*, Stroudsburg, PA, USA, Association for Computational Linguistics, 246-253.

Wilson, T., Wiebe, J., and Hwa, R. (2004). Just how mad are you? Finding strong and weak opinion clauses. *Proceedings of the 19th national conference on Artifical intelligence (AAAI'04)*, 25-29 July. San Jose, California, AAAI Press, 761-769.

Wu, Q., Tan, S., Zhai, H., Zhang, G., Duan, M., and Cheng, X. (2009). SentiRank: Cross-Domain Graph Ranking for Sentiment Classification. *Proceedings of the 2009 IEEE/WIC/ACM International Joint Conference on Web Intelligence and Intelligent Agent Technology - Volume 01*, 15 – 18 September. Milano, Italy, IEEE Computer Society, 309-314.

Wu, Q., and Tan, S. B. (2011). A two-stage framework for cross-domain sentiment classification. *Expert Systems with Applications* 38(11), 14269-14275.

Xia, R., Zong, C., and Li, S. (2011). Ensemble of feature sets and classification algorithms for sentiment classification. *Information Sciences* 181(6), 1138-1152.

Yan, G., He, W., Shen, J., and Tang, C. (2014). A bilingual approach for conducting Chinese and English social media sentiment analysis. *Computer Networks* 75, Part B, 491-503.

Yang, J. Y., Kim, H. J., and Lee, S. G. (2010). Feature-based product review summarization utilizing user score. *Journal of information science and engineering* 26(6), 1973-1990.

Ye, Q., Zhang, Z., and Law, R. (2009). Sentiment classification of online reviews to travel destinations by supervised machine learning approaches. *Expert Systems with Applications* 36(3, Part 2), 6527-6535.

Yi, J., Nasukawa, T., Bunescu, R., and Niblack, W. (2003). Sentiment Analyzer: Extracting Sentiments about a Given Topic using Natural Language Processing Techniques. *Proceedings of the Third IEEE International Conference on Data Mining*, 19-22 November. Florida, USA, IEEE Computer Society, 427 - 434.

Yu, H., and Hatzivassiloglou, V. (2003). Towards answering opinion questions: separating facts from opinions and identifying the polarity of opinion sentences. *Proceedings of the 2003 conference on Empirical methods in natural language processing (EMNLP 2003)*, 11-12 July. Sapporo, Japan, Association for Computational Linguistics, 129-136.

Yu, N., and Kubler, S. (2011). Filling the gap: semi-supervised learning for opinion detection across domains. *Proceedings of the Fifteenth Conference on Computational Natural Language Learning*, Portland, Oregon, Association for Computational Linguistics, 200-209.

Zhang, Y., Wen, J., Wang, X., and Jiang, Z. (2014). Semi-Supervised Learning Combining Co-Training with Active Learning. *Expert Systems with Applications* 41(5), 2372-2378.

Zhang, Z., Ye, Q., Zhang, Z., and Li, Y. (2011). Sentiment classification of Internet restaurant reviews written in Cantonese. *Expert Systems with Applications* 38(6), 7674-7682.

Zhao, Y. Y., Qin, B., and Liu, T. (2010). Integrating Intra- and Inter-document Evidences for Improving Sentence Sentiment Classification. *Acta Automatica Sinica* 36(10), 1417-1425.

Zhi-Hua, Z., and Ming, L. (2005). Tri-training: exploiting unlabeled data using three classifiers. *IEEE Transactions on Knowledge and Data Engineering* 17(11), 1529-1541.

Zhou, G., Zhou, Y., Guo, X., Tu, X., and He, T. (2015). Cross-domain sentiment classification via topical correspondence transfer. *Neurocomputing* 159, 298-305.

Zhou, S., Chen, Q., and Wang, X. (2013). Active deep learning method for semi-supervised sentiment classification. *Neurocomputing* 120, 536-546.

Zhou, S., Chen, Q., and Wang, X. (2014). Fuzzy deep belief networks for semi-supervised sentiment classification. *Neurocomputing* 131, 312-322.

Zhou, Z. H., and Li, M. (2010). Semi-supervised learning by disagreement. *Knowledge and Information Systems* 24(3), 415-439.

Zhu, J., and Ma, M. (2012). Uncertainty-based active learning with instability estimation for text classification. *ACM Transactions on Speech and Language Processing (TSLP)* 8(4), 1-21.

Zhu, J., Wang, H., Yao, T., and Tsou, B. K. (2008). Active learning with sampling by uncertainty and density for word sense disambiguation and text classification. *Proceedings of the 22nd International Conference on Computational Linguistics - Volume 1*, Manchester, United Kingdom, Association for Computational Linguistics, 1137-1144.

Zhu, X. (2006). Semi-supervised learning literature survey, Technical Report 1530, Computer Sciences, University of Wisconsin-Madison.

Zhu, X., and Ghahramani, Z. (2002). Learning from labeled and unlabeled data with label propagation, Technical Report CMU-CALD-02-107, Carnegie Mellon University.