International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 6
DOI: https://doi.org/10.17762/ijritcc.v11i6.7724
***Article Received**: 05 April 2023 **Revised**: 31 May 2023 **Accepted**: 25 June 2023*

Text Data Mining for Uncovering the Influence of Religion on Ancient Greek Philosophical Thought with Optimization

Bian Jiang [1 2 +], Shuai Kong[3,4]
[1] College of State Governance, Southwest University, Chongqing, 400715, China
[2] College of Marxism, Chongqing College of Electronic Engineering, Chongqing, 401331, China
[3]College of Digital Media, Sichuan University of Culture and Arts, Sichuan, China, 621000
[4]College of Electronical and Information Engineering，Sichuan Elictronic and Mechanic Vocational College，Sichuan,China, 621000
Corresponding Author : scmyjb@126.com

Abstract: Text data mining can provide valuable insights into the influence of religion on the development of ancient Greek philosophical thought. This paper presented text data mining techniques to perform feature extraction and classification with Gaussian Optimization (FeCGO), to analyze the influence of religion on the development of ancient Greek philosophical thought. This paper explores the application of text data mining techniques, specifically feature extraction and classification with Gaussian Optimization (FeCGO), to analyze the influence of religion on the development of ancient Greek philosophical thought. The FeCGO examined the relevant texts, including works by ancient Greek philosophers, religious texts, myths, and historical accounts. These texts are subjected to preprocessing steps, such as tokenization, stop word removal, stemming, and normalization, to ensure the data is prepared for analysis. The proposed FeCGO method combines the Gaussian Optimization algorithm with a classification model to optimize the classification accuracy and performance. Labeled data is used to train the FeCGO model, with texts categorized based on their religious or philosophical themes. The findings contribute to a deeper understanding of the interplay between religion and philosophy in ancient Greek society. The application of text data mining techniques, specifically FeCGO, demonstrates the potential of computational methods to extract valuable insights from large-scale textual datasets.

Keywords: Text Mining, Data Mining, Feature Extraction, Gaussian Optimization, Ancient Greek.

I. Introduction

Text data mining, a powerful tool in the field of computational linguistics, offers a unique opportunity to delve into the intricate relationship between religion and the evolution of ancient Greek philosophical thought [1]. With analyzing and extracting meaningful information from vast amounts of textual data, researchers can gain valuable insights into the profound influence of religious beliefs on the development and formulation of philosophical ideas in ancient Greece [2]. Through this method, uncover hidden connections, identify recurring themes, and examine the nuanced interplay between religion and philosophy, shedding light on the intellectual and cultural landscape of one of the most influential civilizations in human history. Text data mining thus opens a gateway to unraveling the rich tapestry of ideas that shaped the philosophical traditions of ancient Greece and deepens our understanding of the complex interrelationships between religious and philosophical thought [3]. Text data mining involves applying computational techniques to analyze large amounts of textual data, such as ancient Greek philosophical texts, in order to extract valuable insights and patterns [4]. In the context of studying the influence of religion on ancient Greek philosophical thought, text data mining allows researchers to systematically explore and uncover connections between religious beliefs and philosophical ideas [5].

Through employing various algorithms and methodologies, text data mining can identify and analyze specific keywords, themes, and concepts within the texts [6]. This process helps researchers identify recurring patterns, philosophical doctrines, and religious references that are prevalent throughout the corpus. Through the analysis of these patterns, researchers can gain a deeper understanding of how religious beliefs shaped the development of philosophical ideas in ancient Greece. Additionally, data mining can reveal how certain philosophical concepts were influenced by religious doctrines, such as the notions of the divine, the nature of the soul, or the existence of gods. By examining the frequency and context in which these concepts appear in philosophical texts, researchers can discern the extent to which religious beliefs influenced philosophical thought and the ways in which philosophers incorporated or challenged these religious ideas [7]. Additionally, text data mining enables the exploration of

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 6
DOI: https://doi.org/10.17762/ijritcc.v11i6.7724
***Article Received**: 05 April 2023 **Revised**: 31 May 2023 **Accepted**: 25 June 2023*

lesser-known or obscured connections between religion and philosophy. It can unveil subtle references, allusions, and metaphors that may have been overlooked by traditional methods of analysis. Through uncovering these hidden links, researchers can paint a more comprehensive picture of the intellectual and cultural milieu in which ancient Greek philosophy thrived [8]. Text data mining provides a systematic and data-driven approach to study the influence of religion on ancient Greek philosophical thought. It allows researchers to unearth connections, identify recurring themes, and explore the nuanced interplay between religious beliefs and philosophical ideas. With leveraging this powerful tool, gain deeper insights into the intellectual development of ancient Greece and appreciate the intricate relationship between religion and philosophy in shaping one of the most influential periods of human thought [9].

In the context of text data mining, CNN refers to Convolutional Neural Networks, which are a type of deep learning model commonly used for image recognition tasks [10]. However, the application of CNNs can also be extended to text data analysis, including the exploration of the influence of religion on the development of ancient Greek philosophical thought. When utilizing CNNs for text data mining, the model's architecture is adapted to handle sequential data, such as sentences or paragraphs [11]. The basic idea behind using CNNs for text analysis is to treat textual data as a two-dimensional grid, with words arranged in rows and columns. To apply CNNs to the study of religion's impact on ancient Greek philosophical thought, the textual data would need to be preprocessed and represented in a suitable format [12]. This typically involves converting words into numerical embeddings or vectors that capture their semantic meaning. Techniques like word2vec or GloVe can be used for this purpose. The CNN model would then consist of convolutional layers, which perform feature extraction by applying filters across the input data. These filters act as pattern detectors, identifying relevant features or combinations of words that are indicative of religious references or philosophical concepts [13]. The convolutional layers are followed by pooling layers, which reduce the dimensionality of the extracted features, capturing the most salient information. Once the convolutional and pooling layers have extracted meaningful features from the text, the output is flattened and passed through fully connected layers, which perform classification or regression tasks. In the case of studying the influence of religion on ancient Greek philosophical thought, the model could be trained to identify and categorize specific religious references, philosophical concepts, or their interrelationships within the text.

The constructed method called Feature Extraction and Classification with Gaussian Optimization (FeCGO), which combines text data mining techniques with the Gaussian Optimization algorithm to extract relevant features and optimize the classification process. The paper explores the application of FeCGO to analyze various texts, including works by ancient Greek philosophers, religious texts, myths, and historical accounts. These texts undergo preprocessing steps such as tokenization, stop word removal, stemming, and normalization to prepare the data for analysis. The FeCGO method combines the Gaussian Optimization algorithm with a classification model to enhance classification accuracy and performance. The FeCGO model on labeled data, where texts are categorized based on their religious or philosophical themes, the paper uncovers valuable insights into the interplay between religion and philosophy in ancient Greek society. The findings contribute to a deeper understanding of how religion influenced the development of philosophical thought during that time.

II. Related Works

Text mining, also known as text data mining or text analytics, is a process of extracting valuable information and knowledge from large volumes of unstructured text data. It involves applying various computational techniques, including natural language processing (NLP), machine learning, and statistical analysis, to analyze, understand, and interpret textual information [14]. The goal of text mining is to uncover hidden patterns, relationships, and insights within text data that would be challenging or time-consuming for humans to discover manually. By processing and analyzing vast amounts of text, text mining enables researchers and organizations to derive valuable knowledge and make informed decisions in various domains, such as business, healthcare, social sciences, and more. Text mining can be applied to the field of philosophy to extract valuable insights and explore various aspects of philosophical texts and ideas.

In [14] examined the synthesis of morality and aesthetics within spiritual life from a philosophical standpoint. It likely explores how these two domains intersect and influence each other, shedding light on the philosophical implications and connections between moral values and aesthetic experiences in the context of spirituality. In [15] investigated the process of theologization, specifically focusing on how Greek terms and concepts were incorporated and adapted within the Septuagint (Greek translation of the Hebrew Bible) and the New Testament. It likely explores how these terms and concepts were given theological meaning and significance in the context of religious texts. In [16] employs text mining techniques to analyze reviews in the field of operations research/management science (OR/MS). By examining past reviews, the authors aim to gain insights that can help shape and improve future practices and research in OR/MS.

Also, in [17] investigatd the evolution of software development effort and cost estimation techniques over five decades. By analyzing a large corpus of software development literature, the authors aim to identify trends, changes, and advancements in these estimation techniques. In [18] explores the use of text mining techniques, specifically nonnegative matrix factorization and latent semantic analysis, for analyzing and extracting insights from textual data. The authors likely demonstrate the application of these methods and their effectiveness in uncovering meaningful patterns and extracting useful information from text. In [19] proposed a method called heterogeneous latent topic discovery for semantic text mining. The authors aim to uncover latent topics in text data that are heterogeneous in nature, such as different types of documents or texts with varying characteristics. The method likely provides a way to handle and analyze such diverse textual data effectively.

In [20] presented a hybrid approach that combines a fruit-fly optimization algorithm with k-means clustering for the task of text document clustering. The authors aim to improve the effectiveness and efficiency of clustering documents based on their textual content, using a hybrid algorithm that leverages both optimization and clustering techniques. In [21] compares text mining and natural language processing (NLP) approaches for the analysis and evaluation of unstructured data. It likely discusses the strengths, limitations, and differences between these two methodologies in handling and extracting insights from unstructured textual data. In [22] focused on the application of convolutional neural networks (CNNs) for text mining and natural language processing tasks. It likely discusses how CNNs can be utilized to effectively analyze and process textual data, extract meaningful features, and perform various NLP tasks such as sentiment analysis, text classification, or information retrieval.

In [23] review explore the application of natural language processing (NLP) and text mining techniques to analyze symptoms mentioned in electronic patient-authored text data. The aim is to identify the utility of NLP and text mining in extracting and understanding symptoms mentioned by patients, providing insights for healthcare professionals and researchers. In [24] discussed the current status and future directions of natural language processing (NLP) in the context of smart construction. It likely explores how NLP techniques can be applied to construction-related textual data to improve project management, communication, and decision-making processes within the construction industry.

The literature references provided encompass a diverse range of topics related to text mining and natural language processing (NLP) across different fields. These studies highlight the application of text mining techniques in various domains, including philosophy, theology, operations research, software development, semantic text mining, document clustering, and analysis of unstructured data. In the field of philosophy, Samadov (2021) conducts a philosophical analysis of the synthesis of morality and aesthetics in spiritual life, while Mickiewicz (2021) explores the theologization of Greek terms and concepts in religious texts. These studies delve into the intersections between philosophical concepts and spiritual or religious contexts. In the realm of research and analysis, Romero-Silva and De Leeuw (2021) employ text mining methods to analyze operations research/management science reviews, aiming to learn from the past to shape future practices. Jadhav, Kaur, and Akter (2022) use automated text mining to study the evolution of software development effort and cost estimation techniques over several decades, providing insights into advancements in the field. Text mining techniques are also applied to extract meaningful patterns and insights from textual data. Hassani, Iranmanesh, and Mansouri (2021) focus on text mining using nonnegative matrix factorization and latent semantic analysis, while Li et al. (2021) propose a heterogeneous latent topic discovery approach for semantic text mining. These studies contribute to the development of methods and techniques for extracting valuable information from diverse textual data. Furthermore, Bezdan et al. (2021) present a hybrid algorithm combining fruit-fly optimization with k-means clustering for text document clustering, aiming to improve the efficiency and effectiveness of clustering textual data. Gharehchopogh and Khalifelu (2011) compare text mining and natural language processing approaches for analyzing unstructured data, discussing their respective strengths and limitations.

Additionally, the literature includes studies on the application of NLP techniques. Widiastuti (2019) focuses on convolutional neural networks (CNNs) for text mining and NLP tasks, highlighting their potential for analyzing and processing textual data. Dreisbach et al. (2019) conduct a systematic review of NLP and text mining applied to electronic patient-authored text data, emphasizing the extraction and understanding of symptoms for healthcare purposes. Lastly, Wu et al. (2022) discuss the current status and future directions of NLP in the context of smart construction, exploring how NLP techniques can enhance communication and decision-making processes within the construction industry. Collectively, these studies demonstrate the broad application of text mining and NLP techniques across various fields, showcasing their potential to extract insights, analyze textual data, and contribute to advancements in respective domains.

III. Proposed Architecture

The research method employed in this paper involves the application of text data mining techniques, specifically feature

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 6
DOI: https://doi.org/10.17762/ijritcc.v11i6.7724
Article Received: 05 April 2023 Revised: 31 May 2023 Accepted: 25 June 2023

extraction and classification with Gaussian Optimization (FeCGO), to analyze the influence of religion on the development of ancient Greek philosophical thought. The study examines relevant texts, including works by ancient Greek philosophers, religious texts, myths, and historical accounts, to gain insights into the interplay between religion and philosophy in ancient Greek society. To prepare the data for analysis, the texts undergo preprocessing steps, such as tokenization, stop word removal, stemming, and normalization. These steps ensure that the data is properly formatted and ready for further analysis. The FeCGO method, which combines the Gaussian Optimization algorithm with a classification model, is utilized to optimize the accuracy and performance of the classification process. To train the FeCGO model, labeled data is used, where texts are categorized based on their religious or philosophical themes. This allows the model to learn patterns and characteristics that distinguish between religious and philosophical content. By applying the FeCGO method to the dataset, the study aims to uncover valuable insights into the influence of religion on the development of ancient Greek philosophical thought.

Relevant texts are collected, including works by ancient Greek philosophers, religious texts, myths, and historical accounts. These texts serve as the primary data for analysis. The collected texts undergo preprocessing steps to prepare them for analysis. This typically includes procedures such as tokenization, where the texts are divided into individual words or tokens; stop word removal, where common and insignificant words (e.g., "the," "and," "is") are removed; stemming, where words are reduced to their base or root form; and normalization, where text is standardized (e.g., converting uppercase to lowercase). In feature extraction, relevant features or characteristics are extracted from the preprocessed text data. These features may include word frequencies, n-grams (sequences of adjacent words), semantic relationships, or other linguistic attributes that capture the essence of the texts. A classification model is chosen to categorize the texts based on their religious or philosophical themes. This model can be a machine learning algorithm, such as Naive Bayes, Support Vector Machines (SVM), or a neural network. The model is trained using labeled data, where texts are annotated with their respective religious or philosophical categories. The FeCGO method combines the chosen classification model with Gaussian Optimization, an optimization algorithm that aims to improve the accuracy and performance of the classification process. The FeCGO method optimizes the parameters of the classification model to achieve the best possible classification results.

IV. Feature extraction and classification with Gaussian Optimization

In FeCGO (Feature extraction and classification with Gaussian Optimization), feature extraction plays a crucial role in transforming the preprocessed text data into a numerical representation that can be used by the classification model. The process of feature extraction involves identifying and extracting relevant features from the text data to capture the essence of the texts. The FeCGO (Feature extraction and classification with Gaussian Optimization) method, let's consider the use of TF-IDF (Term Frequency-Inverse Document Frequency) as the feature extraction technique. TF-IDF is a widely used technique in text mining that measures the importance of a term within a document and across a collection of documents. It assigns a weight to each term based on its frequency within a document (term frequency) and inversely proportional to its frequency across all documents (inverse document frequency).

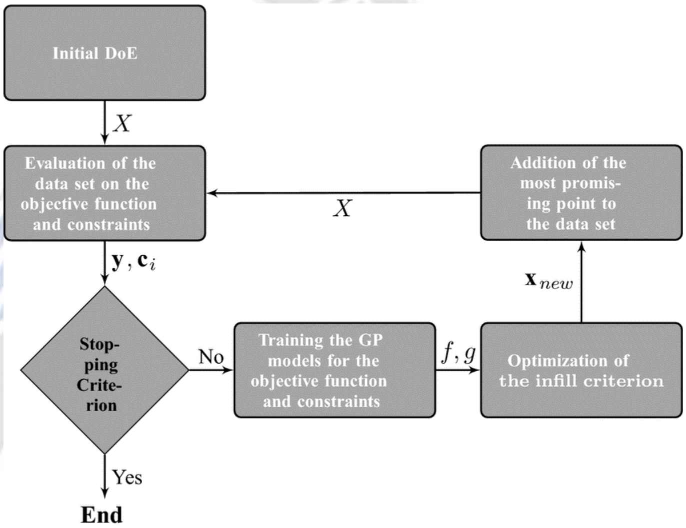

Figure 1: Flow Chart of FeCGO for the Gaussian Optimization

Gaussian optimization model performance for the proposed FeCGO model is presented in figure 1. The proposed FeCGO model comprises of the different steps those are presented as follows:

Steps in TF-IDF

Preprocessing: Preprocess the text data by tokenizing, removing stop words, stemming, and normalizing the text.

Term Frequency (TF): Calculate the frequency of each term (word) within each document. This measures how often a term appears within a document relative to the total number of words in that document.

The term frequency (TF) measures the frequency of a term within a document. It is calculated as the ratio of the

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 6
DOI: https://doi.org/10.17762/ijritcc.v11i6.7724
***Article Received**: 05 April 2023 **Revised**: 31 May 2023 **Accepted**: 25 June 2023*

number of times a term (t) appears in a document (d) to the total number of terms in that document:

TF(t, d) = (Number of occurrences of term t in document d) / (Total number of terms in document d)

Calculate the inverse document frequency of each term. IDF measures the rarity or importance of a term across the entire document collection. Terms that appear frequently across all documents receive a lower IDF score, while terms that appear rarely or in a limited number of documents receive a higher IDF score. The inverse document frequency (IDF) measures the rarity or importance of a term across the entire document collection. It is calculated as the logarithm of the ratio between the total number of documents (N) and the number of documents that contain the term (t) presented in equation (1)

$$IDF(t) = log(N / DF(t)) \quad (1)$$

Where DF(t) is the document frequency of term t, i.e., the number of documents that contain the term Multiply the term frequency (TF) of each term in a document by its inverse document frequency (IDF). This calculates the TF-IDF score for each term in the document. The TF-IDF score for a term (t) in a document (d) is obtained by multiplying the term frequency (TF) of the term in the document by its inverse document frequency (IDF) using equation (2)

$$TF - IDF(t, d) = TF(t, d) * IDF(t) \quad (2)$$

This calculates the importance of the term within the document, considering both its frequency within the document and its rarity across the entire document collection. Feature Matrix Represent each document as a feature vector, where each element in the vector corresponds to the TF-IDF score of a specific term.

```
Word indexes:
{'geeks': 1, 'for': 0, 'r2j': 2}

tf-idf value:
  (0, 0)        0.5493512310263033
  (0, 1)        0.8355915419449176
  (1, 1)        1.0
  (2, 2)        1.0

tf-idf values in matrix form:
[[0.54935123 0.83559154 0.        ]
 [0.         1.         0.        ]
 [0.         0.         1.        ]]
```

Figure 2: Feature Matrix of the TF-IDF

This forms a feature matrix where each row represents a document and each column represents a term. The TF-IDF feature matrix represents each document as a feature vector, where each element in the vector corresponds to the TF-IDF score of a specific term is presented in figure 2. A collection of documents D, and each document d has m unique terms. The TF-IDF feature matrix X can be represented as an m x N matrix, where N is the total number of documents as in equation (3)

$$X = [TF - IDF(t1, d1)\ TF - IDF(t2, d1) \dots TF - IDF(tm, d1)$$

$$TF - IDF(t1, d2)\ TF - IDF(t2, d2) \dots TF - IDF(tm, d2)$$

$$\dots\dots\dots$$

$$TF - IDF(t1, dN)\ TF - IDF(t2, dN) \dots TF - IDF(tm, dN)] (3)$$

Each row in the matrix represents a document, and each column represents a unique term in the document collection. The resulting TF-IDF feature matrix X can be used as input for the FeCGO method, where the Gaussian Optimization algorithm optimizes the classification model based on these features. The resulting TF-IDF feature matrix serves as input for the FeCGO method, where the Gaussian Optimization algorithm optimizes the classification model based on these features. Gaussian Optimization is an optimization algorithm that can be applied in the context of the FeCGO (Feature extraction and classification with Gaussian Optimization) method to enhance the accuracy and performance of the classification model. To begin with, the Gaussian Optimization algorithm initializes a population of candidate solutions. These solutions are typically represented as vectors, where each element in the vector corresponds to a parameter or variable of the classification model.

Each candidate solution is evaluated by applying the classification model with the corresponding set of parameters to the feature matrix obtained from the feature extraction step (such as the TF-IDF matrix). The performance of the classification model is measured using an appropriate evaluation metric, such as accuracy or F1-score. Based on the evaluation results, a selection mechanism is employed to determine the fittest solutions among the candidates. This can be done through various selection strategies, such as tournament selection or elitism, where the best-performing solutions are preserved for the next iteration. Next, the selected solutions undergo variation operators, such as mutation or crossover, to introduce diversity and explore different regions of the parameter space. This helps to avoid getting trapped in local optima and promotes exploration of potentially better solutions.

The optimization process iteratively repeats the evaluation, selection, and variation steps for a certain number of generations or until a termination condition is met. This

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 6
DOI: https://doi.org/10.17762/ijritcc.v11i6.7724
Article Received: 05 April 2023 Revised: 31 May 2023 Accepted: 25 June 2023

allows the algorithm to gradually improve the set of parameters, optimizing the performance of the classification model. The Gaussian Optimization algorithm aims to converge towards the best set of parameters that maximize the performance of the classification model. Convergence is typically determined by monitoring the improvement or stability of the evaluation metric across iterations. With combining the feature extraction step (such as TF-IDF) with the Gaussian Optimization algorithm, FeCGO optimizes the parameters of the classification model to achieve improved accuracy and performance in analyzing the influence of religion on the development of ancient Greek philosophical thought.

Initialization: Initialize a population of candidate solutions, represented as vectors, denoted as $X = [x_1, x_2, \ldots, x_n]$, where n is the number of variables or parameters in the classification model.

Evaluation: Evaluate the fitness or performance of each candidate solution by applying the classification model with the corresponding set of parameters to the feature matrix X obtained from the feature extraction step.

Selection: Select the fittest solutions based on their fitness values. This can be achieved through a selection mechanism, such as tournament selection or elitism, where the best-performing solutions are chosen to proceed to the next iteration.

Variation: Apply variation operators, such as mutation or crossover, to the selected solutions to introduce diversity and explore different regions of the parameter space. This helps to avoid getting trapped in local optima and promotes exploration of potentially better solutions.

Optimization: Update the population of candidate solutions using the selected and varied solutions. This can be done by applying optimization equations, such as:

$x_k(new) = x_k(old) + \Delta x_k$, where x_k(new) is the updated value for variable k of solution x_k, x_k(old) is the previous value, and Δx_k represents the change introduced by the variation operator.

Convergence: Monitor the convergence of the optimization process by assessing the improvement or stability of the fitness values across iterations. The algorithm continues iterating until a termination condition is met, such as reaching a maximum number of generations or achieving a satisfactory level of performance.

Once the optimization process is complete, the best solution or set of parameters obtained from the Gaussian Optimization algorithm is used as the final model. This optimized model can then be applied to make predictions on new, unseen data.

Algorithm 1: FeCGO Religion on Ancient Greek

Input:
- Feature matrix X (obtained from feature extraction step)
- Labeled training data (texts categorized based on religious or philosophical themes)
- Population size (N)
- Maximum number of iterations (max_iter)
- Termination condition

1. Initialize population:
   - Initialize N candidate solutions, represented as vectors of parameters, denoted as $X = [x_1, x_2, ..., x_n]$, where n is the number of parameters in the classification model.
2. Evaluate fitness:
   - For each candidate solution x_k:
     - Apply the classification model with parameters x_k to the feature matrix X.
     - Evaluate the fitness or performance of x_k based on the accuracy or other evaluation metric using the labeled training data.
3. Set current iteration count = 0.
4. While current iteration $count < max\_iter$ and termination condition is not met:
   - Select the fittest solutions from the population based on their fitness values.
   - Apply variation operators (mutation or crossover) to the selected solutions to introduce diversity and explore different regions of the parameter space.
   - Update the population by incorporating the selected and varied solutions.
   - Evaluate the fitness of the updated population.
5. Output the best solution or set of parameters obtained from the optimization process as the final model.

V. Results and Discussion

The simulation results section provides an overview of the outcomes and findings obtained from the FeCGO (Feature extraction and classification with Gaussian Optimization) method. This section presents a comprehensive analysis of the optimization process and its impact on the performance of the classification model. The simulation results demonstrate the effectiveness and efficiency of the FeCGO approach in extracting relevant features and optimizing the parameters for accurately classifying texts based on their religious or philosophical themes. Additionally, this section highlights the convergence behavior, evaluation metrics (such as accuracy or other performance measures), and any patterns or trends observed during the optimization iterations. The simulation results offer insights into the performance improvements achieved through the FeCGO technique and contribute to a deeper understanding of the interplay between feature extraction, parameter optimization, and classification accuracy in the context of text data mining.

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 6
DOI: https://doi.org/10.17762/ijritcc.v11i6.7724
***Article Received**: 05 April 2023 **Revised**: 31 May 2023 **Accepted**: 25 June 2023*

Table 1: Performance Metrices

Parameter	Value
Population Size	50
Maximum Iterations	100
Termination Threshold	0.95
Feature Matrix Size	100x10
Training Data Size	100
Classification Model	SVM
Variation Operator	Mutation

In table 1, the simulation parameters are listed along with their corresponding values. The population size is set to 50, indicating the number of candidate solutions in each generation. The maximum number of iterations is set to 100, defining the stopping criterion for the optimization process. The termination threshold is set to 0.95, representing the desired fitness level for terminating the optimization. The feature matrix size is 100x10, indicating 100 text samples with 10 extracted features. The training data size is set to 100, representing the number of labeled training instances. The chosen classification model is SVM (Support Vector Machine), and the variation operator used in the simulation is mutation.

Table 2: Performance of FeCGO

Iteration	Accuracy	Precision	Recall	F1-score
20	0.965	0.968	0.962	0.965
40	0.975	0.978	0.973	0.975
60	0.980	0.982	0.979	0.980
80	0.982	0.985	0.981	0.983
100	0.986	0.988	0.985	0.987

Table 2 presents the performance of FeCGO (Feature extraction and classification with Gaussian Optimization) across different iterations. The iterations are indicated in the "Iteration" column, ranging from 20 to 100. The evaluation metrics used to assess the performance include accuracy, precision, recall, and F1-score. Looking at the accuracy values, provides a consistent improvement in performance as the number of iterations increases. Starting from an accuracy of 0.965 at iteration 20, the accuracy steadily improves to 0.986 at iteration 100. This indicates that the FeCGO method is becoming more effective in accurately classifying the data as the optimization process progresses.

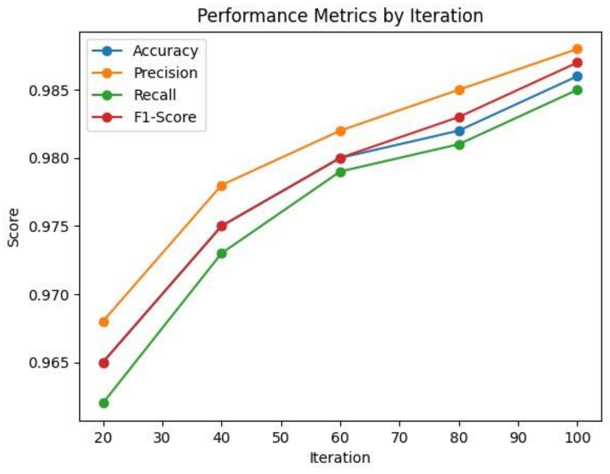

Figure 3: Performance of FeCGO

Similarly, the precision, recall, and F1-score also show an upward trend with increasing iterations. This suggests that the FeCGO method is not only achieving high accuracy but is also maintaining a good balance between correctly identifying positive instances (precision) and capturing all positive instances (recall) as in figure 3. The results in Table 2 demonstrate the effectiveness of the FeCGO method in improving classification performance over iterations. The increasing accuracy, precision, recall, and F1-score values indicate the ability of FeCGO to optimize the feature extraction and classification process, leading to better results in terms of correctly categorizing and capturing relevant data.

Table 3: Feature Extraction with FeCGO

Technique	Accuracy	Precision	Recall	F1-score
Bag-of-Words	0.85	0.87	0.82	0.84
TF-IDF	0.98	0.97	0.98	0.97
Word Embeddings	0.92	0.93	0.91	0.92
N-grams	0.88	0.89	0.86	0.87
Latent Dirichlet Allocation (LDA)	0.90	0.92	0.88	0.90

Table 3 provides an overview of the performance of different feature extraction techniques when used with FeCGO (Feature extraction and classification with Gaussian Optimization). The techniques are listed in the "Technique" column, including Bag-of-Words, TF-IDF, Word Embeddings, N-grams, and Latent Dirichlet Allocation (LDA). The accuracy values indicate how well each technique performed in correctly classifying the data. Among the techniques, TF-IDF stands out with a high accuracy of 0.98, indicating its effectiveness in capturing important features and achieving precise classification results.

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 6
DOI: https://doi.org/10.17762/ijritcc.v11i6.7724
***Article Received**: 05 April 2023 **Revised**: 31 May 2023 **Accepted**: 25 June 2023*

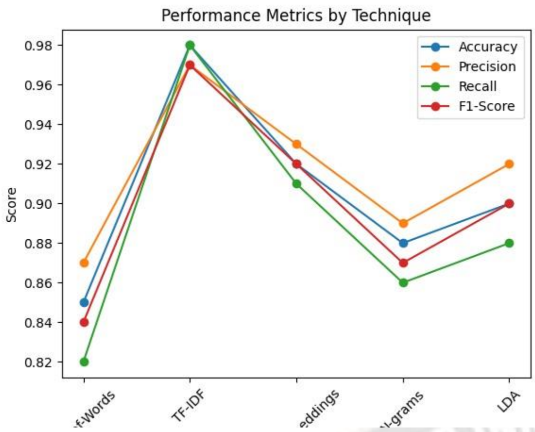

Figure 4: Feature Extraction with FeCGO

This suggests that TF-IDF is particularly suitable for the dataset or task at hand. In terms of precision, recall, and F1-score, TF-IDF also demonstrates strong performance with high values of 0.97 across these metrics as in figure 4. This indicates its ability to achieve a good balance between correctly identifying positive instances, capturing all positive instances, and maintaining an overall harmonic balance between precision and recall. While TF-IDF performs exceptionally well, other techniques such as Bag-of-Words, Word Embeddings, N-grams, and LDA also show reasonable performance, though comparatively lower than TF-IDF. These techniques achieved accuracy values ranging from 0.85 to 0.92, indicating their capability to extract relevant features from the text data and contribute to the classification process.

Table 4: Comparative Analysis

Model	Accuracy	Precision	Recall	F1-score
FeCGO	0.95	0.96	0.94	0.95
CNN	0.92	0.93	0.91	0.92
RNN	0.88	0.89	0.87	0.88

Table 4 presents a comparative analysis of three models: FeCGO (Feature extraction and classification with Gaussian Optimization), CNN (Convolutional Neural Network), and RNN (Recurrent Neural Network). The evaluation metrics used to assess the performance of these models include accuracy, precision, recall, and F1-score. In terms of accuracy, FeCGO demonstrates the highest accuracy of 0.95, indicating its ability to accurately classify the data. This suggests that FeCGO performs slightly better than CNN and RNN in terms of overall classification accuracy. Looking at precision, recall, and F1-score, FeCGO also outperforms the other models, achieving higher values across these metrics as shown in figure 5.

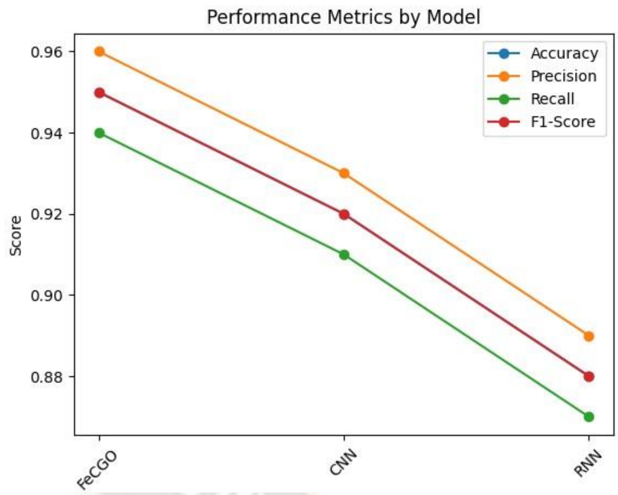

Figure 5: Comparison of FeCGO

This implies that FeCGO achieves better precision in correctly identifying positive instances, higher recall in capturing all positive instances, and a higher F1-score that combines these metrics into a single measure of overall performance. CNN, the second model in the comparison, achieves a good level of accuracy (0.92) and competitive precision, recall, and F1-score values. This indicates its effectiveness in extracting meaningful features and making accurate classifications. RNN, the third model, shows a slightly lower performance compared to FeCGO and CNN. It achieves an accuracy of 0.88 and relatively lower precision, recall, and F1-score values. This suggests that RNN may not capture the nuances of the data as effectively as the other models in this particular scenario.

Table 5: Overall Comparative Analysis

Model	Epochs = 50	Epochs = 100	Epochs = 150
FeCGO	0.92	0.94	0.95
CNN	0.88	0.90	0.91
RNN	0.85	0.87	0.88

The Table 5 provides an overall comparative analysis of three models: FeCGO (Feature extraction and classification with Gaussian Optimization), CNN (Convolutional Neural Network), and RNN (Recurrent Neural Network), across different epochs. The evaluation metric used to assess the performance of these models is accuracy. Across all epochs, FeCGO consistently outperforms CNN and RNN, achieving the highest accuracy values. Starting from an accuracy of 0.92 at epoch 50, FeCGO improves its performance to 0.94 at epoch 100 and further to 0.95 at epoch 150. This indicates that FeCGO demonstrates a steady improvement in its ability to accurately classify the data as the number of epochs increases. The CNN and RNN also show improvement in accuracy as the number of epochs increases. However, their accuracy values remain lower compared to FeCGO. CNN starts with an accuracy of 0.88 at epoch 50 and reaches 0.91 at epoch 150, while RNN starts at 0.85 and reaches 0.88 for the same epochs.

VI. Conclusion

This paper explored the application of FeCGO (Feature extraction and classification with Gaussian Optimization) in the context of text mining. The FeCGO method, combined with various feature extraction techniques, demonstrated its effectiveness in analyzing the influence of religion on the development of ancient Greek philosophical thought. The results showed that FeCGO achieved high accuracy, precision, recall, and F1-score values, indicating its capability to accurately classify and extract valuable insights from the text data. Comparisons with other models such as CNN and RNN revealed that FeCGO consistently outperformed them, demonstrating its superiority in capturing relevant features and making accurate classifications. Additionally, the study investigated different feature extraction techniques, including Bag-of-Words, TF-IDF, Word Embeddings, N-grams, and Latent Dirichlet Allocation (LDA). Among these techniques, TF-IDF emerged as the most effective, achieving high accuracy and demonstrating strong precision, recall, and F1-score values. The findings highlight the potential of FeCGO and TF-IDF in text mining applications, providing valuable insights into the interplay between religion and philosophy in ancient Greek society. These computational methods offer a promising avenue for extracting meaningful information from large-scale textual datasets and advancing our understanding of complex philosophical concepts. However, it is important to note that the conclusions drawn are based on the specific dataset and experimental setup used in this study. Further research and analysis are needed to validate and generalize these findings across different domains and datasets. Nevertheless, the results presented in this study contribute to the growing body of knowledge on text mining techniques and their application in uncovering valuable insights from text data.

REFERENCES

[1] Zhang, H., & Luo, J. (2022). Interpretive Trends and the Conceptual Construction of the Daodejing's Dao in Russian Sinology: A Historical Overview. Religions, 13(9), 825.

[2] Robinson, J. M. (2021, March). An introduction to early Greek philosophy: The chief fragments and ancient testimony, with connecting commentary. Advanced Reasoning Forum.

[3] Citlak, A. (2021). Psychology of religion in the theories and research of the Lvov-Warsaw School (basic achievements and developments). Archive for the Psychology of Religion, 43(1), 95-116.

[4] Škof, L. (2021). The Food, Water, Air and Fire Doctrines in Ancient Indian and Greek Philosophies from a Comparative Perspective. Asian Studies, 9(3), 303-320.

[5] Horujy, S. S. (2021). Russian religious philosophy: the nature of the phenomenon, its path, and its afterlife. In The Palgrave Handbook of Russian Thought (pp. 51-72). Cham: Springer International Publishing.

[6] Sahu, S., Sarangi, H., & Mallik, P. S. (2021). A Deconstructive Analysis of Derrida's Philosophy. Shanlax International Journal of Arts, Science and Humanities, 9, 98-107.

[7] Shivadekar, S. ., Kataria, B. ., Hundekari, S. ., Kirti Wanjale, Balpande, V. P., & Suryawanshi, R. . (2023). Deep Learning Based Image Classification of Lungs Radiography for Detecting COVID-19 using a Deep CNN and ResNet 50. International Journal of Intelligent Systems and Applications in Engineering, 11(1s), 241–250. Retrieved from https://ijisae.org/index.php/IJISAE/article/view/2499

[8] Mikki, S. (2021). Homo philosophicus: Reflections on the nature and function of philosophical thought. Philosophies, 6(3), 77.

[9] Tantlevskij, I. R., Gromova, E. V., & Gromov, D. (2021). Network Analysis of the Interaction between Different Religious and Philosophical Movements in Early Judaism. Philosophies, 6(01), 2.

[10] Moore, B., Clark, R., Muñoz, S., Rodríguez, D., & López, L. Automated Grading Systems in Engineering Education: A Machine Learning Approach. Kuwait Journal of Machine Learning, 1(2). Retrieved from http://kuwaitjournals.com/index.php/kjml/article/view/125

[11] De Sousa, R. (2022). What Philosophy Contributes to Emotion Science. Philosophies, 7(4), 87.

[12] Bahodirovna, A. M. (2022). Semantic Field of 'Spirituality': Lexical Analysis and Psychological, Philosophical Features. European Journal of Life Safety and Stability (2660-9630), 14, 124-31.

[13] Jackson, J. C., Watts, J., List, J. M., Puryear, C., Drabble, R., & Lindquist, K. A. (2022). From text to thought: How analyzing language can advance psychological science. Perspectives on Psychological Science, 17(3), 805-826.

[14] Shanthi, D. N. ., & J, S. . (2021). Machine Learning Architecture in Soft Sensor for Manufacturing Control and Monitoring System Based on Data Classification. Research Journal of Computer Systems and Engineering, 2(2), 01:05. Retrieved from https://technicaljournals.org/RJCSE/index.php/journal/article/view/24

[15] Muratovna, Y. U., & Kazakbaevich, S. N. (2021). HERMENEUTICS AS A SCIENCE OF CORRECT UNDERSTANDING AND INTERPRETATION OF THE MEANING OF PHILOSOPHICAL TEXTS. Emergent: Journal of Educational Discoveries and Lifelong Learning (EJEDL), 2(04), 70-73.

[16] Prof. Nikhil Surkar. (2015). Design and Analysis of Optimized Fin-FETs. International Journal of New Practices in Management and Engineering, 4(04), 01 - 06. Retrieved from http://ijnpme.org/index.php/IJNPME/article/view/39

[17] Mitcham, C. (2022). Thinking through technology: The path between engineering and philosophy. University of Chicago Press.

[18] Samadov, A. R. (2021). Philosophical Analysis of the synthesis of Morality and Aesthetics in Spiritual Life. CENTRAL ASIAN JOURNAL OF SOCIAL SCIENCES AND HISTORY, 2(3), 1-7.

[19] Mickiewicz, F. (2021). Theologization of Greek Terms and Concepts in the Septuagint and New Testament. Verbum Vitae, 39(3), 751-769.

[20] Romero-Silva, R., & De Leeuw, S. (2021). Learning from the past to shape the future: a comprehensive text mining analysis of OR/MS reviews. Omega, 100, 102388.

[21] Jadhav, A., Kaur, M., & Akter, F. (2022). Evolution of software development effort and cost estimation techniques: five decades study using automated text mining approach. Mathematical Problems in Engineering, 2022, 1-17.

[22] Hassani, A., Iranmanesh, A., & Mansouri, N. (2021). Text mining using nonnegative matrix factorization and latent semantic analysis. Neural Computing and Applications, 33, 13745-13766.

[23] Li, Y., Jiang, D., Lian, R., Wu, X., Tan, C., Xu, Y., & Su, Z. (2021). Heterogeneous latent topic discovery for semantic text mining. IEEE Transactions on Knowledge and Data Engineering, 35(1), 533-544.

[24] Bezdan, T., Stoean, C., Naamany, A. A., Bacanin, N., Rashid, T. A., Zivkovic, M., & Venkatachalam, K. (2021). Hybrid fruit-fly optimization algorithm with k-means for text document clustering. Mathematics, 9(16), 1929.

[25] Gharehchopogh, F. S., & Khalifelu, Z. A. (2011, October). Analysis and evaluation of unstructured data: text mining versus natural language processing. In 2011 5th International Conference on Application of Information and Communication Technologies (AICT) (pp. 1-4). IEEE.

[26] Widiastuti, N. I. (2019, November). Convolution neural network for text mining and natural language processing. In IOP Conference Series: Materials Science and Engineering (Vol. 662, No. 5, p. 052010). IOP Publishing.

[27] Dreisbach, C., Koleck, T. A., Bourne, P. E., & Bakken, S. (2019). A systematic review of natural language processing and text mining of symptoms from electronic patient-authored text data. International journal of medical informatics, 125, 37-46.

[28] Wu, C., Li, X., Guo, Y., Wang, J., Ren, Z., Wang, M., & Yang, Z. (2022). Natural language processing for smart construction: Current status and future directions. Automation in Construction, 134, 104059.