A Review on Predicting Text Similarity in Social Media Using Ant Colony Optimization Method

مراجعة حول التنبؤ بتشابه النص في وسائل التواصل الاجتماعي باستخدام طريقة تحسين مستعمرة النمل

عبدالله احمد إسماعيل المطري
Abdullah Ahmed Ismael Almatri [1]
عبد العزيز احمد ثوابه
Abdulaziz Ahmed Thawaba[2]
عبدالله سعيد غريب
Abdullah Saeed Ghareb [3]

https://doi.org/10.54582/TSJ.2.2.92

(1) Faculty of IT&CS University of Saba Region, Marib, Yemen
Email: raid.yeme123@gmail.com

(2)Faculty of IT&CS University of Saba Region, Marib, Yemen
Email: azizth@usr.ac

(3) Faculty of IT&CS University of Saba Region, Marib, Yemen
Email: aghurieb@usr.ac

Abstract

This paper aims to highlight the latest and most widely used artificial intelligence methods in text similarity prediction. The study focuses on articles published from 2010 to 2022 in the field of text similarity prediction and algorithms applied in this field. In this paper, systematic scientific comparisons have been made between existing approaches for predicting text similarity to answer the question raised in this study about the most used and accurate approach. Through previous studies and the comparisons made in this paper, the Ant Colony Optimization Algorithm (ACO) approach was found to be the most frequently used in text similarity prediction and solving scheduling problems.

Keywords: Semantic similarity, social media, prediction, Ant Colony Optimization Algorithm.

الملخص:

تهدف هذه الدراسة الي تسليط الضوء على احدث أساليب التعلم العميق وأكثرها استخداما في التنبؤ بتشابه النص، وتركز على المقالات المنشورة من عام 2010م الى عام 2022م في مجال التنبؤ بتشابه النص والخوارزميات المطبقة في هذا المجال، حيث تم اجراء مقارنات علمية منهجية بين الأساليب الحالية للتنبؤ بتشابه النص للإجابة على السؤال المطروح في هذه الدراسة حول النهج الأكثر استخداما ودقة ، ومن خلال الدراسات السابقة والمقارنات التي تم إجراؤها تبين ان منهج خوارزمية تحسين مستعمرة النمل هو الأكثر استخداما في التنبؤ بتشابه النص وحل مشاكل الجدولة.

الكلمات المفتاحية : التشابه الدلالي، التواصل الاجتماعي، التنبؤ، خوارزمية تحسين مستعمرة النمل

I. Introduction

Text similarity is one of the active research and application topics in Natural Language Processing [1]. The Ant Colony Optimization Algorithm (ACO) has become one of the most widely used met heuristics for solving combinatorial problems [2]. To improve its performance, the ACO algorithm has been integrated with other methods and inter-programming techniques. Hybridization of ACO algorithms has shown much improvement in results for various problems. Recent trends have also shown the implementation of ACO algorithms with parallel versions to solve many problems like rule induction classification of feature selection. The availability of multicore CPU architectures and GPUs has made it possible to develop improved parallel versions of ACO algorithms [3].

The objectives of this study are to compare the most used and modern methods of Ant Colony Optimization and also scrutinize the most successful methods of semantic prediction. This study analyzes several previous studies to highlight the methods used to enhance Ant Colony Optimization and investigate the reasons that made it the most widely used technique. It also focuses on the mechanisms that made the Ant Colony Optimization used in many areas and how it supports decisions related to the labor market, companies, and commercial advertisements.

II. Literature Review

Pattern recognition tasks like classification have uses in a variety of industries. It necessitates the creation of a model that roughly represents the link between the input's attributes and the output's categories. A collection of classes to which the entity may belong is represented by the outputs, while the inputs define various qualities of an entity that may be an object, a process, or an event [4]. Marco Dorigo invented ant colony optimization (ACO) techniques in the early 1990s. Ants create a pheromone chemical trail on the ground as they

go. They follow scent cues from pheromones, and frequently select the highlighted pathways [5].

1. *Social Media*

Social media platforms enable regular individuals to produce and publish their content. Social media channels serve as examples. In contrast to traditional media like newspapers, books, and television, social media allow nearly anybody to generate content and access information at a low cost. These seven function blocks—identity, conversations, sharing, presence, connections, reputation, and groups—are present in some or all forms of social media. Social media significantly influences the development of notable events. For instance, following the Tohoku Earthquake in Japan, individuals used social media to connect with friends, share crisis information, and locate essential services [3].

2. *Text Similarity*

Text similarity is a text mining approach when computing the similarity between different texts. In Natural Language Processing (NLP), the answer to “how two words/phrases/documents are similar to each other? Text similarity calculates how two words/phrases/documents are close to each other. That closeness may be lexical or meaning restrictions. Semantic similarity is about the meaning closeness, and lexical similarity is about the closeness of the word set.

Lexical similarity indicates that these two phrases are similar and nearly identical because they have the same word set. Despite the word set’s closeness, they are fundamentally distinct in terms of semantic similarity because each has a unique meaning. The text must first be transformed into a vector of features before the algorithm can choose an appropriate feature representation, such as TF, or IDF. Finally, the representation vectors of texts exhibit commonality. There are numerous methods for determining text similarity and whether they take semantic relations into account. The most well-known methods:

- Jaccard Similarity
- Cosine Similarity
- K-Means
- Latent Semantic Indexing (LSI)
- Latent Dirichlet Allocation (LDA)
- The algorithm for optimizing ant colonies.

Most earlier methods work well when paired with word embedding algorithms (such as Word2Vec). We will begin with a Google search example.

“Terrorism, fraud, theft, and beggaring are all against . . .”

If a user searched for the first phrase on Google, the second phrase would appear in the top 5 results. These two statements are not close to one another regarding lexical similarity. Despite having different word sets, they are quite comparable in terms of semantics because their meanings are so similar: The outcome (second) phrase) will evolve. In the end, events have an impact on search results. The outcome will have a different word set but a very similar meaning, and that much is certain. Their semantic similarity gauges the degree to which two texts’ meanings are similar. Typically, this metric yields a score between 0 and 1. Between 0 and 1, the degree of similarity between the two terms is shown. (Semantic similarity) [6].

3. *Semantic Similarity*

One of the most significant uses of natural language processing has been text similarity. When two texts surface closeness, share terms, and infer the same meaning, they are considered to be similar. The degree of semantic equivalency between two language elements, such as concepts, sentences, or texts, is measured by semantic similarity. Knowledge-Based Similarity is a category that is used to compare concepts of semantic similarity. A node in an ontology graph represents each notion in this category. Because the graph represents the concepts from the corpus, this is also known as the topological method. A smaller number of edges connecting two concepts (nodes) indicates that they are more semantically and conceptually similar. Create a topology for this graph, and you'll see that "coin" is more closely related to "money" than "credit card."

3.1 Statistical-Based Similarity

This type calculates the semantic similarity based on learning features' vectors from the corpus. Vectors representation can depend on many techniques, like count or TF-IDF in Latent Semantic Analysis (LSA), weights of Wikipedia concepts in Explicit Semantic Analysis (ESA), synonyms in Pointwise Mutual Information (PMI), and co-occurring words of a set of predefined words in Hyperspace Analogue to Language (HAL). Because word embedding represents the semantic relationship between words of this type, most prior approaches can be integrated with them for improved results.

3.2 String-Based Similarity

Measuring semantic similarity doesn't depend on this type separately but combines it with other types for measuring the distance between non-zero vectors of features. The most important algorithms of this type are Manhattan Distance, Euclidean Distance, Cosine Similarity, Jaccard Index, and Sorensen-Dice Index.

3.3 Language Model-Based Similarity

The scientific community introduced Language Model-Based Similarity in 2016 as a novel semantic similarity measurement between two English phrases, assuming they are syntactically correct. This type has five main steps:

- Removing stop words
- Tagging the two phrases using any Part of Speech (POS) algorithm
- From the tagging step output, this type forms a structure tree for each phrase (parsing tree)
- Building an undirected weighted graph using the parsing tree
- Finally, the similarity is calculated as the minimum distance path between nodes (words).

Example:

All the algorithms we mentioned in this article are already implemented and optimized in different programming languages, mainly Python and Java Se match is one of the most recent tools in Python for measuring semantic similarity. It depends on the knowledge-based similarity type. The following code snippet shows how simply you can compare the semantic similarity of two English vocabulary words with an output of 0.5. From the match, semantic similarity imports Word Net Similarity

WNS = WordNet Similarity ()

WNS. word similarity (‘dog’, ‘cat’, ‘li’)

3.4 Methods Numerical Similarity

Semantic relatedness, the inverse of semantic distance in computational linguistics, presupposes that two items are semantically related if they have any semantic relation. A unique semantic similarity metric illustrates the similarities between two concepts based on their hierarchical relationships. Semantic relatedness, a more general concept that does not always depend on hierarchical links, is a specific case of semantic similarity. Both semantic relatedness and similarity are covered in this work. In the sections that follow, we refer to them as semantic similarity interchangeably and divide them into corpus-based approaches and knowledge-based methods. Corpus-based approaches mostly rely on the context of words found in the corpus. Thus, they primarily assess the generic semantic similarity of words. Based on the hierarchical relationships recorded in WordNet; knowledge-based algorithms determine the semantic similarity of words. Because they consider all possible semantic relationships between words, corpus-based approaches have more extensive computational applications. In contrast, knowledge-based approaches would be more beneficial when applications are needed to encode word hierarchies. In the sections that follow, we examine both categories of techniques [7].

3.5 Automatic Extraction of Semantic

The quality of search results can be greatly improved by using text and link information from Web pages. However, the coverage of user studies, which do not scale with the Web’s size, heterogeneity, and growth, restricts the evaluation of automatic semantic measurements. Here, we suggest using topical directories, which are human-generated, to quantify semantic links between a huge number of pairings of Web sites or subjects.

In psychological theory, the issue of assessing semantic similarity in a network has a long history. Semantic similarity has become crucial to knowledge representation, where unique networks or ontologies are employed to characterize items and their relationships. By calculating the distance between the nodes in a network representation, many strategies attempt to assess semantic similarity. These frameworks are predicated on the idea that two things will appear closer in a network representation the closer their semantic relationships are. We introduced a brand-new semantic similarity metric for Web pages that extends the well-established information-theoretic tree-based metric to

the general scenario where pages are categorized in arbitrary graph ontologies nodes with hierarchical and non-hierarchical components. This method is easily adaptable to extracting semantic data from topical ontologies and online directories like Yahoo, the ODP, and their offshoots. The resemblance is frequently used to illustrate a gratifying relationship [8].

In Natural Language Processing, one of the difficult and unsolved research challenges is estimating the semantic similarity between text data (NLP). It is challenging to establish rule-based approaches for semantic similarity measurements due to the flexibility of natural language. Several semantic similarity techniques have been used over time to solve this problem. This survey article charts the development of such methods, classifying them according to their underlying principles as knowledge-based, corpus-based, neural network-based, and hybrid methods. It starts with traditional NLP techniques like kernel-based methods and progresses to the most recent research on transformer-based models. Examining the benefits and drawbacks of each technique.

Notably, our technique is used to locate and assign sentiment to 44% of hidden words that are not included in sentiment lexicons. Because of co-occurrence words, 51.46% of tweets change their emotion (from positive to negative, from positive to neutral, or vice versa); our findings demonstrate that the hybrid method performs the best out of the three methods for each level of Maslow's hierarchy of needs. To assign sentiment intensity and direction to words, a novel contextual semantic sentiment representation of words called an opinion circle is the goal. Using Maslow's theory keywords, the study discussed the use of opinion circles and sentiment lexicons at the tweet and entity levels [8].

4. *Prediction*

Prediction is the practice of guessing something by considering the previous activity to forecast a circumstance that necessitates exact computation to provide the solution [9]. The Artificial Neural Network with the Backpropagation Method was utilized for the computation [6]. Even while human agents, particularly specialists, can generally make superior forecasts using social media, there are still good reasons for us to strive to anticipate automatically. Firstly, compared with human labor, automatic prediction with machines has a much lower cost [10]. Secondly, people frequently overvalue negligible probabilities while undervaluing substantial probabilities. Consequently, humans are not very good at predicting unlikely and likely situations. Thirdly, a person's decisions may be influenced— intentionally or unintentionally—by their desires, interests, and benefits rather than only by objective likelihood. Last but not least, automatic prediction techniques could process additional data quickly and respond [11]. Social media has given us a new way to gather, extract, and use the wisdom of crowds objectively with little cost and high efficiency, even though it is still a relatively unexplored area of research and its conclusions are only moderately accurate. [12]

However, much of the literature uses these terms interchangeably, along with terms like semantic distance. Semantic similarity, semantic distance, and relatedness all mean, “How much does term A have to do with term B?” The answer to this question is usually a number between -1 and 1 or between 0 and 1, where 1 signifies extremely high similarity. By combining closely connected terms and separating those that are distantly related, one can intuitively

see how semantically different similar concepts are. Additionally, typical in mind maps and thought maps is this. The Semantic folding method offers a more direct method of comparing the semantic similarity of two linguistic units. A linguistic item, such as a term or a paragraph, can be represented using this method by creating a pixel for each of its active semantic properties, for example, a 128 x 128 grid. By contrasting image representations of each feature set, it is possible to compare the semantics of two objects visually directly. At a low cost, prediction gathers the thoughts and emotions of various groups of people. By analyzing the qualities and content of social media, we can learn about social structure, assess qualitatively and quantitatively activity patterns, and, in certain cases, forecast future occurrences involving people. This study discusses the domains that can be anticipated from current social media, then overviews available predictors and techniques of prediction, and finally discusses challenges and possible future directions [3].

4.1 Prediction Subjects

This section outlines potential areas for making predictions using social media. In general, a topic that could be easily predicted on social media must satisfy the following conditions. First, there must be a human-related event as the prediction subject. Users share their thoughts and beliefs on social media. The information is analyzed, extracted, and integrated by prediction systems, which subsequently create predictions based on the influence of people on the subject being predicted. But let’s say the subject is a phenomenon unrelated to people, like an eclipse. Even if many people talk about something on social media, their opinions do not matter how that event plays out. As a result, it could not anticipate natural catastrophes whose course is unaffected by human activity using social media data. Second, if many individuals are involved, the distribution of those involved on social media should be identical to or comparable to that in the real world. Because not everyone in the world uses social media, social media users can typically be considered representative samples of the relevant populations. However, because sampling is an unavoidable process, biased samples may result. Although we can’t avoid biased samples, we should ensure that their percentage is within a realistic and acceptable range. Lastly, the events involved should be simple enough to discuss in public. Otherwise, social media will have biased content. For instance, there is a social consensus that tipping sensibly is good and that tipping too little is rude and inappropriate. Nearly no one is prepared to say that they left too little in the way of tips when faced with such societal pressure. One could employ the anonymous mode to find a solution to this problem. However, such an anonymous mode would lack knowledge about important social network structures [7].

5. *Regression Method*

Regression approaches, such as the one used in section 2.4.1, examine the relationship between the dependent variable, predicted outcome, and one or more independent factors, such as social network features. There are both linear and non-linear regression models. However, the relation appears to be best described by the linear model. As a result, we frequently use linear regression models instead of non-linear ones, such as polynomial, exponential, and logarithmic ones. The variables in the linear regression model could be either unprocessed or converted data. [7].

6. *Artificial Neural Network*

An artificial neural network is a computer model that mimics the functioning of the human brain. Numerous synthetic neurons make up an artificial neural network. Additionally, these neurons may be a part of many interrelated groups, such as the input layer, hidden layer, and output layer. The input layer is in charge of obtaining unprocessed data and sending it to the following layer. The ultimate prediction outcome will be provided to us by the output layer. Our main duty is selecting the network topology and creating the hidden layer when utilizing an artificial neural network to make predictions. Self-organizing map (SOM), a type of artificial neural network, could be used for features' dimensionality reduction for additional analysis in addition to using them to forecast directly. [7]

In data mining and machine learning, a decision tree is a visual tool. After moving from the root node to the leaf, one entity will receive the prediction outcome. Two fundamental and prevalent decision tree types are classification and regression. When the prediction output consists of discrete classes, classification tree analysis is used. Additionally, regression trees are utilized when the outcome is projected to be continuous. The decision tree is a white box model, which is more readily explicable than the artificial neural network, which is a black box model. Additionally, decision trees perform well with dummy and empty parameters.

7. *Model-based Prediction*

The hardest approach to making predictions might be the Model-based prediction. Before making a forecast, we must first create a mathematical model of the subject, which necessitates in-depth knowledge. We do not yet have sufficient knowledge of social media to create efficient models for them. Despite certain advancements in modeling, model-based prediction is still a controversial and difficult subject. Social media prediction is a new research area with several difficulties. Here, we draw attention to a few crucial and critical upcoming projects [3].

7.1 Interpreting Predictions Using Social Theory

Currently, the trial-and-error method is used by researchers to select predictors. We have no idea why some predictors are superior to others or how they could forecast the outcome. We just take a group of measures to be trained on test data, identify which ones have the greatest coefficients, and use them to make up the prediction model without knowing the underlying reasoning behind these metrics and the final prediction result. Therefore, lacking a credible hypothesis to justify, we cannot be certain that a model that performs well in one situation will also perform well in other contexts. Because of this, some models perform admirably in one election prediction but disastrously in another. We need to understand the theory and logic of the model to ensure that it works well in every situation. Using additional prediction techniques, most scientists employ straightforward techniques like linear regression analysis. Under some circumstances, these techniques are known to be effective. Because a complex system produces social media, there is a non-linear association between the predictors and the predicted outcomes.

Additionally, combining several approaches might result in a breakthrough. A surface learning agent, such as instantly taught neural networks, can quickly adapt to new social media modes and emerging trends when used in such a combination.

Furthermore, an artificial intelligence agent concentrates on enduring patterns. In short, we should experiment with non-linear methods to determine the most effective ones and combinations for each prediction realm.

7.2 Using Social Media to Model Predictions

We still don’t fully understand social media. Prediction objects, for instance, come in various forms and display various attributes.

7.3 Semantic Analysis System for Social Media

Semantic analysis is widely employed even if it is not required for prediction algorithms. As a result, the performance of the predictions depends on how accurate the semantic analysis is. Lexicon or prior statistics may serve as the foundation for semantic analysis. Social media content has a lexical pattern similar to informal and formal English. Social media has given us a new way to gather, extract, and objectively use the wisdom of crowds at a low cost and with high efficiency, even though this is still a relatively unexplored area of research [3].

7.4 *Ant Colony Optimization (ACO)*

ACO algorithms have several distinguishing characteristics that might be seen as stepping stones. Heuristic data, the pheromone update rule, the transition rule, the probability function, the parameter values, and the termination condition are all considered when constructing the solution. It becomes clear that combining the various methods that may be created for each produces a wide range of ant colony algorithms, each of which is better suited to a certain class of issues. Various proposals can be found in the extensive body of literature on the subject to either enhance previous findings or merely address a brand-new kind of issue. Ant Colony Optimization (ACO) algorithms belong to a class of heuristics based on the behavior of nature ants. These algorithms have been used to solve many combinatorial optimization problems and have been known to outperform other popular heuristics such as Genetic Algorithms. Therefore, we believe that the number of ACO based algorithms will continue to grow for a long time [12]. Heuristics are better suited to tackle big instances of these problems since they often require much less computer power, even though certain minor examples of these problems can be addressed with exact approaches. A metaheuristic called Ant Colony Optimization was first developed to address issues about combinatorial optimization difficulty [3]. Performance of the suggested model in comparison to other recent sentiment analysis works. It is not surprising that there may have been some misclassifications because of the diversity of the English language and its countless grammatical variations, misspellings, slang, and complex structures. This model performs better than all existing baseline algorithms, with an accuracy rate of more than 96..[18] %

Additionally, the dataset under consideration mainly contained English words. A dataset with sentences from a different language would provide an intriguing dataset on which to perform the study. Since reaching 100% accuracy is difficult (due to the diversity of the English language), it may be worthwhile to investigate the theory put forth about sentiment analysis as a criterion to distinguish humans from bots [13].

III. Research Methodology

The researchers conducted relevant studies from 2010 to 2022. They used many websites to collect the data. They used Google Scholar, Connected papers1, and Cross Ref. They found 60 related works and then filtered the collected materials by checking them via indexing companies- Scopus and Web of Science. They did that to verify the quality and accuracy of the information. We predict that ACO-based algorithms will become more prevalent over time by giving a variety of approaches to the ACO building blocks found in the literature. The most famous approaches were reviewed in terms of advantages and disadvantages.

IV. Research Analysis and Discussion

After analyzing the literature, one might conclude that ACO is a viable method for resolving scheduling issues. We were able to suggest potential directions for future studies based on the literature review, in addition to some recommendations for using ACO algorithms. Swarm intelligence techniques such as ant colony optimization have successfully resolved various discrete and continuous optimization issues. ACO has been used in several studies to address various discrete and continuous optimization issues, including vehicle routing, quadratic assignment issues, and graph coloring. The main benefit of employing a meta-heuristic to find near-optimal solutions is that, even though a fully optimal solution might not be found, the time needed to solve the problem is typically manageable. It focuses on the four traditional production contexts. It examines and categorizes published research that employs ACO to address scheduling issues (single machine, parallel machine, fellowship, and job shop). It was feasible to confirm the viability of the ant-based algorithms (ACO, ACS, and MMAS) for scheduling issues. Even though the ACO method is simple, the ease with which scheduling-specific heuristics and other metaheuristics can be incorporated raises the complexity and research potential associated with using this metaheuristic for this class of issues [14].

According to the analysis in this study, only the preliminary findings have been published in this area of study, which is relatively new. The initial investigations focused on scheduling issues with a single machine. Environments like fellowships, parallel machines, and job shops weren't addressed until much later. The additional features developed to solve specific issues also demonstrate the progress of knowledge in this field of study (e.g., dominance criteria involving a single machine and flow shop scheduling problems). Some advice for using ACO algorithms to solve scheduling difficulties can be derived from the quantitative analysis.

This might be cruel that the job-to-job approach is ordinarily a basic beginning point for modern ACO applications to planning issues. The job-to-position plot is the cross breed approach which utilized both job-to-job and job-to-position, illustrating that it is conceivable to combine both procedures. Another result is that since it was utilized in about 20 of the distributions we checked on, the definition of preservability as word of the halfway arrangement developed by a single insect could be a reasonable strategy for planning applications. This study further demonstrates the advantages of including scheduling-specific data in ACO implementations. This may be seen in the following:

- The employment of dominance criteria in some articles.
- The application of problem-specific pheromone initialization techniques intended to speed up the algorithm.

Potential topics for future research also emerged. We observed that a single-criterion objective function was used in the four manufacturing environments. This was not the case for multi-criteria objective functions. Therefore, one can conclude that using ACO for scheduling problems to optimize multi-criteria objective functions is a relatively new research field and can be addressed. In addition, our review indicates that the general evolution of ACO applications relating to scheduling problems followed an expected path: single machine, parallel machine, flow shop, and job shop.

The Ant Colony Optimization (ACO) algorithms is a type of heuristic that is based on how ants in nature behave. Numerous combinatorial optimization problems have been solved using these techniques, which have performed better than other well-known heuristics like genetic algorithms. As a result, we think that the number of ACO-based algorithms will increase significantly in the future. By giving various ways that can be found in the literature about the ACO building blocks, this work aims to give the reader a consultation guide for creating ACO algorithms. [12]

In this study, we present all methods in researchers' papers. These papers presented three methods focusing on the text's words and incorporating semantic information. The similarity calculation method focusing on features related to the text's words will only produce less accurate results. In their feature vector of texts and determine semantic similarity. These techniques—cosine similarity using TF-IDF vectors, cosine similarity using word embedding, and soft cosine similarity using word embedding—are based on corpus-based and knowledge-based techniques. Finding similarities between brief news texts using TF-IDF vectors worked best among these three cosine similarities. We compared three methods for deciding how comparable two brief content news stories are to one another in terms of semantics. The comparative writing created by the strategy are straightforward to get it and may be utilized promptly in other data [19]. The experiment was validated using data sets from AG about the news. The three methods are cosine similarity with TF-IDF vectors, word2vec vectors, and soft cosine similarity with word2vec vectors.

These three techniques all produced encouraging outcomes. When the findings were cross-validated, and the newsgroup of a news story and its corresponding most comparable article were compared, cosine with TF-IDF had the highest accuracy among these three vectors. The method's most comparable documents are straightforward, making it simpler to apply in various information retrieval techniques. Using the Doc2Vec model rather than the Word2Vec model,

which represents a document as a vector rather than averaging the word vectors in a document, can improve the accuracy of the other two techniques. This model can be used to improve the procedures further. The ACO-based algorithm ACOAR addresses the rough set theory problem of attribute reduction. These characteristics of this algorithm include updating the pheromone trails of the edges bridging each pair of unique characteristics in the current best solution. The upper and lower trail limits are the maximum pheromone levels. A quick technique is used to create potential solutions. ACOAR can locate solutions with very low cardinality thanks to its pheromone updating rule and solution creation method. However, the accuracy results we obtained fall short of the good ones found in the literature. Approaches to multi-objective optimization might be a highly fruitful route. As a result of the conventional rough set theory's inability to handle real-valued properties, additional study is needed to reduce memory usage and ease the burden on computing.

Table 1: Artificial intelligence methods and algorithms used to predict semantic similarity in social media [1] [15] [16] [17]

Algorithm	Most used	Advantages - disadvantages
Heuristics	Heuristics integrate semantic information into the similarity calculation process, which will produce less accurate findings; cosine similarity using TF-IDF vectors using word embedding, and soft cosine similarity, using word embedding are based on corpus-based and knowledge-based techniques.	are more effective at solving complex problems because they often use much less computational power
Ant colony optimization (ACO)	ACO technique is used in literature to resolve scheduling issues. ACO has been used in several studies to address various discrete and continuous optimization issues, including vehicle routing, quadratic assignment issues, graph coloring,	This meta-heuristic was first developed to solve problems within the class of ACO. Many combinatorial optimization problems have been solved using these techniques, which have been shown to perform better than other well-known heuristics like genetic algorithms. ACO is used for trust Calculation, and particle swarm optimization is used for searching swarm pheromones in the online social network.

F-1 score		According to this study, the suggested model performs better than all existing baseline algorithms, with an accuracy rate of more than 96%. The research also demonstrates that hyperparameter adjustment improves model performance.
rule-based modeling	More novel algorithms and modeling approaches are applied	Driven algorithms to increase prediction accuracy with the development of machine
	.for building load prediction	.learning and information science The building load prediction under demand response and building-grid interaction is .becoming more complicated and challenging
machine -learning based load prediction	The real building applications of load prediction models	The realization of automation and the reduction of the engineering costs are the key advantages of machine learning-based load ,prediction but it is very challenging to embed load prediction algorithms and models into the existing BAS and IoT .systems in the industry

From Table 1, the researchers found that the ACO technique is the most used in the literature to resolve scheduling issues. ACO has been used in several studies to address various discrete and continuous optimization issues, including vehicle routing, quadratic assignment issues, and graph coloring. This could mean that the job-to-job approach is typically a simple starting point for new ACO applications to scheduling issues. Regarding the job-to-position scheme, a hybrid method that employs both job-to-job and job-to-position suggests that it is possible to mix both tactics when necessary. ACOAR can quickly identify solutions with very tiny cardinality. A method is chosen to construct the solution, heuristic information, pheromone updating rule, transition rule, and probability function, parameter values, and termination condition are all included in the definition of Ant Colony Optimization, a meta-heuristic that was first developed to solve problems within the class of combinatorial optimization; this is regarded as the ideal approach. Another conclusion from this study is that, given that it was employed in over 60 of the publications we reviewed, the definition of visibility as a function of the partial solution constructed by a single ant is a promising tactic for scheduling applications.

Adopting dominance criteria in some articles and using problem-specific pheromone initialization techniques are benefits of introducing scheduling-specific information into ACO implementations. Future study areas may have also surfaced. We found that a single-criterion objective function was applied in the four production contexts examined in this research. For objective functions with several criteria, this was not true. As a result, it can be inferred that employing ACO for scheduling issues to optimize multi-criteria objective functions is a relatively new study area that can be addressed (no work with this characteristic was discovered for single or parallel environments, and only one paper regarding job shops). Our analysis also shows that the overall development of ACO applications for scheduling issues followed an expected route, including single machine, parallel machine, flow shop, and job shop. Future research in these areas is also anticipated. Although some of the problems handled deal with setup-dependent timeframes, we see that ACO does not have a solution for job-shop situations with this restriction. A similar observation may be made about the potential for creating production batches, which is currently not considered in job shop issues.

V. Conclusion

All in all, the ACO technique is the most used in reviewed literature to resolve scheduling issues. Regarding the job-to-position scheme, a hybrid method that employs both job-to-job and job-to-position suggests that it is possible to mix both tactics when necessary. ACOAR can quickly identify solutions with very tiny cardinality. It has visibility as a function of the partial solution constructed by a single ant and is a promising tactic for scheduling applications. According to this study, the adoption of dominance criteria in some articles and the usage of problem-specific pheromone initialization techniques are all benefits of introducing scheduling-specific information into ACO implementations. Through previous studies and comparisons conducted in this research, it was found that the ant colony optimization (ACO) algorithm approach is suitable for social media in dealing with text similarity and solving scheduling problems. The researchers recommend having much more analysis of the hybrid ACO-GA algorithm to go much deeper in analyzing the features of ACO. Also, it is preferred to investigate other areas, such as the cost and architecture of these approaches and the ease of installation and utilization.

References

[1] Neto, R. T., & Godinho Filho, M. (2013). Literature review regarding Ant Colony Optimization applied to scheduling problems: Guidelines for implementation and directions for future research. *Engineering applications of artificial intelligence*, *26*(1), 150-161.

[2] Akhtar, A. (2019). Evolution of Ant Colony Optimization Algorithm--A Brief Literature Review. *arXiv preprint arXiv:1908.08007*.

[3] Yu, S., & Kak, S. (2012). A survey of prediction using social media. *arXiv preprint arXiv:1203.1647*.

[4] Bouktif, S., Hanna, E. M., Zaki, N., & Khousa, E. A. (2014). Ant colony optimization algorithm for interpretable Bayesian classifiers combination: application to medical predictions. *PloS one*, *9*(2), e86456.

[5] Colorni, A., Dorigo, M., & Maniezzo, V. (1992, September). An Investigation of some Properties of an” Ant Algorithm”. In *Ppsn* (Vol. 92, No. 1992).

[6] Shawabkeh, A., Faris, H., Aljarah, I., Abu-Salih, B., Alboaneen, D., & Alhindawi, N. (2021). An evolutionary-based random weight networks with taguchi method for Arabic web pages classification. *Arabian Journal for Science and Engineering*, *46*, 3955-3980.

[7] Araque, O., Zhu, G., & Iglesias, C. A. (2019). A semantic similarity-based perspective of affect lexicons for sentiment analysis. *Knowledge-Based Systems*, *165*, 346-359.

[8] Maguitman, A. G., Menczer, F., Roinestad, H., & Vespignani, A. (2005, May). Algorithmic detection of semantic similarity. In *Proceedings of the 14th international conference on World Wide Web* (pp. 107116-).

[9] Helbing, D., & Balietti, S. (2011). From social data mining to forecasting socio-economic crises. *The European Physical Journal Special Topics*, *195*, 3-68.

[10] Bothos, E., Apostolou, D., & Mentzas, G. (2010). Using social media to predict future events with agent-based markets. *IEEE Intelligent Systems*, *25*(06), 50-58.

[11] Corley, C. D., & Mihalcea, R. (2005, June). Measuring the semantic similarity of texts. In *Proceedings of the ACL workshop on empirical modeling of semantic equivalence and entailment* (pp. 1318-).

[12] Monteiro, M., Fontes, D. B., & Fontes, F. A. (2012). *Ant Colony Optimization: a literature survey*. FEP-UP, School of Economics and Management, University of Porto.

[13] Priyadarshini, I., & Cotton, C. (2021). A novel LSTM–CNN–grid search-based deep neural network for sentiment analysis. *The Journal of Supercomputing*, *77*(12), 13911-13932.

[14] Kakad, S., & Dhage, S. (2021). Cross domain-based ontology construction via Jaccard Semantic Similarity with hybrid optimization model. *Expert Systems with Applications*, *178*, 115046.

[15] Ke, L., Feng, Z., & Ren, Z. (2008). An efficient ant colony optimization approach to attri-

bute reduction in rough set theory. *Pattern Recognition Letters*, *29*(9), 1351-1357.

[16] Alsuliman, F., Bhattacharyya, S., Slhoub, K., Nur, N., & Chambers, C. N. (2022, June). Social Media vs. News Platforms: A Cross-Analysis for Fake News Detection Using Web Scraping and NLP. In *Proceedings of the 15th International Conference on PErvasive Technologies Related to Assistive Environments* (pp. 190196-).

[17] Goyal, R., Updhyay, A. K., & Sharma, S. (2019). Trust Prediction Using Ant Colony Optimization and Particle Swarm Optimization in Social Networks. In *Emerging Trends in Expert Applications and Security: Proceedings of ICETEAS 2018* (pp. 485491-). Springer Singapore.

[18] Priyadarshini, I., & Cotton, C. (2021). A novel LSTM–CNN–grid search-based deep neural network for sentiment analysis. *The Journal of Supercomputing*, *77*(12), 13911-13932.

[19] Sitikhu, P., Pahi, K., Thapa, P., & Shakya, S. (2019, November). A comparison of semantic similarity methods for maximum human interpretability. In *2019 artificial intelligence for transforming business and society (AITB)* (Vol. 1, pp. 14-). IEEE.

https://doi.org/10.54582/TSJ.2.2.92