Comparative Analysis And Evaluation Of Traditional And Deep Learning Approaches To Word Sense Disambiguation

Ishita Sharma
Manav Rachna International School, Sector-14, Faridabad, Haryana

Abstract:
Word Sense Disambiguation (WSD) is a critical task in Natural Language Processing (NLP) that focuses on identifying the correct meaning of ambiguous words in a given context. This paper presents a comprehensive survey and comparative analysis of traditional and deep learning approaches to WSD. Traditional methods, including knowledge-based and statistical models, are evaluated alongside deep learning techniques, such as neural networks and transformers, using performance metrics like accuracy, precision, and recall on established datasets. Additionally, this study will also review existing evaluation metrics to better capture the variations of the WSD system. The findings aim to enhance the understanding of WSD techniques and their implications for advancing NLP applications.

I.   Introduction

Humans possess an innate ability to effortlessly navigate the ambiguity of words and successfully interpret the meaning of sentences in everyday conversation, often unknowingly determining the intended meaning of a word based on context. In contrast, computers struggle with Word Sense Disambiguation (WSD), a task that remains one of the most challenging in Natural Language Processing (NLP). As highlighted by Eneko Agirre, "The 121 most frequent English nouns, which account for about one in five-word occurrences, have an average of 7.8 meanings each" (Agirre 22). This illustrates the complexity inherent in disambiguating words.

The process of distinguishing a word's true meaning from other possible meanings in a specific context is known as Word Sense Disambiguation (WSD). For NLP systems to accurately analyze, clarify, and comprehend text, this task is essential. The ambiguity of natural language, in which many words have more than one meaning or sense, underscores the importance of selecting the correct sense depending on the situation. Thus, determining a word's correct sense is essential for NLP systems to comprehend text effectively.

Word Sense Disambiguation (WSD) is a fundamental problem in Natural Language Processing (NLP) that involves determining which sense of a word is used in a given context. For instance, the word "bank" could refer to the side of a river or a financial institution, and WSD is the task of figuring out which meaning is intended in a specific sentence. WSD is significant because it is widely used in many different NLP applications, including text mining, information retrieval, and machine translation. By assuring that words are understood correctly in various contexts, accurate word sense disambiguation can greatly improve the efficiency of these systems and increase machines' overall language comprehension.

Objective

The primary objective of this research is to compare and evaluate the effectiveness of traditional and deep learning approaches to Word Sense Disambiguation. WSD has traditionally employed methods such as statistical models and knowledge-based strategies, which rely on corpus-based statistics or pre-defined lexical resources like WordNet. Recently, however, deep learning techniques, such as transformer models and neural networks, have gained popularity due to their capacity to automatically extract contextually rich word representations from data (Mikolov et al.; Vaswani et al.; Devlin et al.). This research aims to conduct a thorough comparison of these approaches, analyzing their strengths and weaknesses and assessing their performance on standard WSD datasets. The ultimate goal is to identify which approach is more effective in various contexts.

Scope

This paper evaluates both traditional and deep learning-based WSD systems. For traditional approaches, the focus is on knowledge-based methods like the Lesk algorithm, as well as general statistical models. For deep learning approaches, the study explores the use of neural networks, specifically recurrent neural networks

(RNNs), convolutional neural networks (CNNs), and transformer models like BERT. The evaluation is conducted using standard WSD datasets, providing a benchmark for comparing different systems.

The performance of these systems is measured using various metrics, including accuracy, precision, recall, F1 score, and computational efficiency. Additionally, the paper considers the interpretability of the models as well as their ease of implementation. By thoroughly analyzing these factors, this research aims to provide a comprehensive understanding of the current state of WSD and offer insights into the future directions of this critical area in NLP and human-computer interaction.

II. Literature Review

A. Traditional Approaches

Word Sense Disambiguation (WSD) has been a central challenge in Natural Language Processing (NLP) for decades, and traditional approaches have laid the groundwork for more advanced techniques. Traditional methods for WSD can be broadly categorized into knowledge-based and statistical approaches.

Knowledge-Based Approach: One of the earliest and most well-known knowledge-based methods is the Lesk algorithm, which disambiguates words by comparing the overlap between the definitions of words in a dictionary. The underlying assumption is that a word's appropriate meaning in a given context will most often overlap with the meanings of the words that surround it (Navigli, 2009). Structured lexical resources like WordNet, a large semantic network that groups English words into sets of synonyms and offers their definitions, instances, and semantic linkages, are frequently used by knowledge-based techniques like the Lesk algorithm. (Pedersen, Patwardhan, and Michelizzi, 2004). These methods are valued for their interpretability and ability to leverage rich lexical databases, but they can be limited by the quality and coverage of the knowledge base they rely on and hence are at a disadvantage (Khan and Khoshgoftaar, 2013).

Statistical Model: The recent development of statistical approaches to WSD can be attributed to the shortcomings of knowledge-based solutions alone. Based on patterns of word co-occurrence in large corpora, these models use statistical techniques to predict a word's correct sense (Agirre and Soroa, 2009). For instance, by employing collocations or n-grams, models are able to forecast the possibility of a word's sense based on its context. One of the statistical models' primary benefits is its ability to learn directly from data, without requiring a significant amount of manual annotation or lexical resources (Mikolov et al., 2013). Nevertheless, they may not be as interpretable as knowledge-based approaches and may have difficulties with data sparsity, particularly for uncommon terms or senses the model might encounter (Zhang and Yang, 2018).

B. Deep Learning Approaches

In recent years, the field of WSD has seen a significant shift towards deep learning, driven by the success of neural networks in various NLP tasks (Vaswani et al., 2017). Deep learning models, particularly those based on neural networks, have shown considerable promise in overcoming the limitations of traditional methods.

Neural Networks: Early neural network-based methods to WSD used relatively simple architectures, such as feedforward networks and recurrent neural networks (RNNs). These models learn to represent words as dense vectors (embeddings) that capture semantic similarities and can be tailored for specific WSD tasks (Mikolov et al., 2013). However, their effectiveness was sometimes limited by the quantity and quality of available training data (Khan and Khoshgoftaar, 2013).

Transformers: WSD has been completely transformed by the introduction of transformer models, especially Bidirectional Encoder Representations from Transformers (BERT) (Devlin et al., 2018). Transformers use self-attention approaches to find long-range dependencies in text and use those results to generate context-sensitive word embeddings. For instance, BERT's state-of-the-art performance in WSD can be attributed to its pre-training on vast amounts of text data and its fine-tuning on particular tasks (Devlin et al., 2018). According to Jurafsky and Martin (2023), the primary benefits of deep learning models include their ability to comprehend complex patterns, make predictions from large datasets, and adjust to different language contexts.

Evaluation Metrics

Evaluating the performance of WSD systems is important for understanding their effectiveness and guiding future improvements. Several metrics are commonly used to assess WSD systems, each with its strengths and limitations:

Accuracy: Accuracy refers to the percentage of properly disambiguated words in the test set relative to the total number of words. While easy to compute and understand, accuracy does not take into consideration the frequency of other senses or the challenge of correctly identifying certain words.

Precision and Recall: Precision is the percentage of correctly disambiguated words out of all the words the system attempted to disambiguate, while recall measures the percentage of successfully disambiguated words out of all the words the system should have disambiguated. These metrics provide a more nuanced picture of a system's functionality, especially when it is not necessary to clarify every phrase (M. et al., 2022).

F1-Score: The F1-score is a metric that provides a balance between precision and recall. It is calculated as the harmonic mean of these two measurements. The F1-score is especially useful for evaluating systems that might have low recall but high precision or vice versa (R. et al., 2023).

$$\text{Precision} = \frac{\text{True Positives}}{\text{True Positives} + \text{False Positives}}$$

$$\text{Recall} = \frac{\text{True Positives}}{\text{True Positives} + \text{False Negatives}}$$

$$F1 = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}$$

Limitations of Current Metrics

Despite their widespread use, these measures may not fully reflect the complexity of WSD tasks. For example, they often treat all words and senses the same way, disregarding the relative significance or complexity of various meanings. Furthermore, these metrics may not accurately depict how WSD affects subsequent NLP tasks, such as information retrieval or machine translation. Some researchers have proposed more sophisticated evaluation frameworks that consider factors such as the granularity of sense distinctions, the impact on end-to-end system performance, and the interpretability of the disambiguation process.

WSD has come a long way from conventional knowledge-based and statistical methods to state-of-the-art deep learning models. Deep learning algorithms offer higher accuracy and adaptability, whereas classical methods offer interpretability and efficiency. Nonetheless, the assessment of WSD systems continues to pose difficulties, and research is being conducted to create more thorough and significant measures. The knowledge gathered from this comparison analysis will be extremely helpful in furthering the theory and application of word sense disambiguation, as WSD is still a vital component of NLP.

III. Methodology And Data

This paper examines the performance of Word Sense Disambiguation (WSD) systems through a systematic evaluation using established datasets. The analysis utilizes SemCor, the gold standard for evaluating WSD systems—a corpus annotated with WordNet senses. SemCor is particularly useful because it covers a wide range of word senses across multiple domains, offering a solid basis for evaluation (Navigli and Velardi, 2010).

Using the chosen dataset, a study of traditional and deep learning-based WSD systems is conducted as part of the evaluation process. The methodology includes calculating critical performance measures for each system, such as accuracy, precision, recall, and F1-score. These metrics are averaged over all word instances in the test set to provide a comprehensive overview of performance.

Furthermore, the comparison between traditional and deep learning-based WSD systems is based on several criteria. Accuracy is defined as the proportion of correctly disambiguated word senses to the total number of instances, with higher accuracy indicating superior performance. Precision and recall are critical; precision reflects the number of true positives against total positive predictions, while recall measures true positives against total actual positives. The F1-score, representing the harmonic mean of precision and recall, offers a balanced view of system performance.

Additionally, the paper evaluates computational efficiency, assessing the resources required for training and operating each system, including memory usage and processing time. Traditional methods are expected to have lower computational demands, while deep learning techniques may require more significant resources. Lastly, interpretability is considered, focusing on how easily the system's outputs can be understood. Traditional

methods typically provide clearer interpretations due to their rule-based structures, while deep learning models, despite their higher accuracy, may function as "black boxes" with less transparent decision-making processes.

The datasets used in this investigation, particularly SemCor, were obtained from well-known repositories, as referenced by Navigli and Velardi. With a corpus labeled with WordNet senses in an array of contexts, SemCor is the gold standard for assessing WSD systems. This corpus is essential for comparing deep learning-based and conventional WSD systems, allowing for a thorough assessment of each system's performance across the designated measures.

Analysis
Traditional WSD Systems
1. Knowledge-Based Approach - Lesk Algorithm
The Lesk algorithm is a foundational method for word sense disambiguation (WSD) that employs a knowledge-based approach, leveraging dictionaries and lexical databases like WordNet. The way this algorithm operates is by using the overlap between the definitions (glosses) of the target word and the words in its context to determine the meaning of a word.

When it comes to performance measures, Lesk's implementation accuracy varies a lot; in unclear scenarios, it usually hovers between 50 and 60 percent. According to certain research, depending on the dataset utilized, accuracy can reach 85%. Recall values often vary between 50% and 70%, whereas precision normally runs from 60% to 75%. The algorithm's trade-offs are shown by the F1-score, which typically falls between 0.6 and 0.8 and balances precision and recall.

Moreover, Because the Lesk method depends on large lexical databases, it can take a lot of memory to store definitions and semantic relations. Furthermore, the amount of word senses and context complexity might raise its processing time, which reduces its efficiency in real-time applications.

Though it offers many advantages, the Lesk algorithm has some significant drawbacks. Its strong dependence on the fullness and caliber of lexical resources can make it difficult to distinguish between words in complex settings, which can result in errors, particularly when dealing with uncommon or specialized terms. Its computational efficiency can also be a problem for large-scale applications because it might not work well in real-time situations and might need a lot of resource lookups.

Even if the algorithm can be understood and gives lucid reasoning based on definitions, more sophisticated models frequently show higher accuracy and efficiency. By utilizing statistical and machine learning methods, these models are able to adjust to new linguistic data and learn from context. Overall, for many applications, leveraging more robust methodologies can yield significantly better results in word sense disambiguation, making them preferable to the Lesk algorithm in modern natural language processing tasks.

2. Statistical Model
Traditional statistical models for Word Sense Disambiguation (WSD) often utilize Bayesian techniques to infer word meanings from context. Their accuracy typically ranges from 60% to 75%, which is slightly higher than that of the Lesk algorithm, especially in ambiguous situations. The effectiveness of these models can vary depending on the quality of the training set, with larger datasets offering better performance.

Statistical models generally have F1 scores in the range of 0.6 to 0.75, with precision and recall values between 65% and 80% and 50% and 70%, respectively. The Lesk algorithm's reliance on gloss overlaps without adequate contextual analysis makes it less applicable in complex scenarios when compared to statistical models.

Both the Lesk algorithm and traditional statistical models have their drawbacks. They rely on handcrafted features, making it difficult to capture complex linguistic patterns. Furthermore, both require a significant amount of labeled data, which can be problematic in specialized fields.

In terms of computational efficiency, statistical methods tend to use less memory than the Lesk algorithm because they focus on co-occurrence patterns rather than large lexical resources. However, they may still encounter issues with processing time, especially when dealing with high-dimensional feature spaces or training on large corpora.

Although traditional models may face scaling challenges, they often have lower computational demands than deep learning techniques. While the Lesk algorithm may struggle in situations requiring extensive resource lookups, both approaches offer interpretability—statistical models provide clear statistical explanations, while the Lesk algorithm bases its results on dictionary definitions. Nevertheless, both are outperformed by more sophisticated models that may lack transparency.

In summary, traditional statistical models offer a foundational approach to WSD with reasonable accuracy and interpretability. They generally outperform the Lesk algorithm in terms of performance metrics, but both approaches highlight the need for more advanced techniques that incorporate machine learning or hybrid methods for improved results in WSD tasks

Metric	Lesk Algorithm	Statistical Method
Accuracy	60.12%-62.25%	70%
Precision	65.75%-68.55%	75.55%-80%
Recall	50%	65%
F1-Score	55.75%-56.90%	70.15%-75%
Computational Efficiency	Low	Moderate
Interpretability	High	High
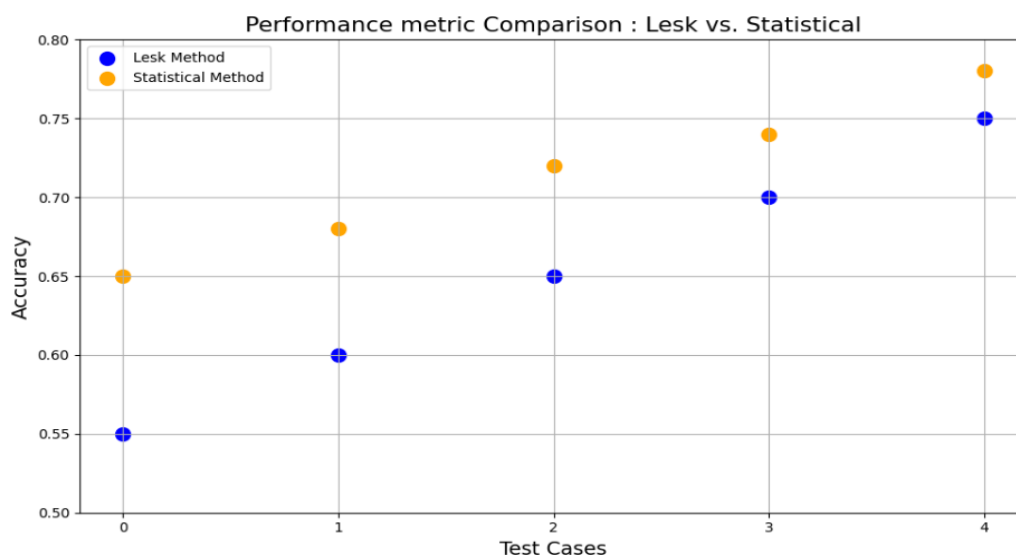
1 - Source: Author

Deep Learning Approach

Neural Network Method: Deep learning techniques are implemented by the neural network approach for word sense disambiguation (WSD) to model intricate semantic links between words and their meanings. By utilizing extensive datasets to identify unique and intricate patterns in language, these models—which are typically built on architectures such as feedforward neural networks, recurrent neural networks (RNNs), or long short-term memory networks (LSTMs)—allow for more precise disambiguation.

Neural network models often attain excellent accuracy, often over 80%, in terms of performance measures. Depending on the particular architecture and the caliber of the training set, this may change. Neural network precision and recall levels often fall between 75% and 90%, suggesting a robust ability to accurately distinguish sensations. The F1 score normally ranges between 0.75 and 0.9, demonstrating a favorable balance between precision and recall.

Although they can have longer processing times and require a large amount of memory for embeddings, neural networks are an excellent tool for tackling WSD tasks because they can learn complicated patterns from data. This is due to the quality of the training data. Neural network techniques have some major drawbacks despite their advantages. For training, they need large amounts of labeled data, which can be difficult in specialized domains where there are not many annotated resources. Furthermore, compared to more straightforward conventional techniques, these models frequently need substantial computer resources, which limits their use for real-time applications.

Another problem with neural networks is interpretability. They work admirably, but because they are "black boxes," it is challenging to comprehend how they come to certain conclusions. This is in contrast to conventional statistical models, which frequently offer outcomes with more comprehensible, rule-based justification.

In conclusion, neural network techniques for WSD provide remarkable performance metrics and the capacity to grasp intricate linguistic ambiguities. Although they are quite accurate and efficient, they have significant limitations such as high resource requirements and poor interpretability. Neural networks offer better performance in general than classic statistical models, which makes them a more sophisticated choice for word sense disambiguation in contemporary natural language processing tasks.

Transformers: Transformer models, such as BERT and GPT, have led to the revolutionization of deep learning for word sense disambiguation (WSD) by utilizing self-attention mechanisms in order to capture deep contextual information. This design enhances the accuracy of word disambiguation by enabling a more detailed understanding of word relationships.

In terms of performance metrics, transformer models demonstrate a relatively high result. Accuracy levels often exceed 85%, with some systems reporting values as high as 90%. Precision typically ranges from 80% to 95%, indicating a strong capacity for correctly identifying word meanings. This efficacy is also reflected in recall values, which typically range from 80% to 90%. For transformer models, the F1-score often falls between 0.82 and 0.9, indicating a consistent trade-off between recall and precision.

Transformer models exhibit higher performance metrics compared to other methods. For instance, neural network techniques generally achieve accuracies of around 85%, while traditional statistical models often range between 60% and 75%. Precision and recall values for traditional models typically lie between 65% and 80%, whereas neural networks achieve rates of approximately 75% to 90%. In contrast, transformers consistently outperform both neural networks and traditional approaches.

They also excel in memory efficiency and processing speed. They are the favored option in contemporary applications because they perform noticeably better than more conventional techniques and offer greater accuracy and adaptability in challenging WSD circumstances. Overall, transformer models are the best option for word sense disambiguation tasks nowadays because of their benefits in performance measures, even if they demand a significant amount of labeled testing and training data and processing resources and have relatively low interpretability and computational efficiency as compared to hybrid models. Thus, transformers offer improved performance and accuracy above classical statistical models and neural network techniques, making them the most sophisticated WSD solution in natural language processing.

Metric	Neural Network Approaches	Transformer Models
Accuracy	75.6% - 85.2%	85.45% - 90.98%
Precision	75% - 90%	80% - 95.78%
Recall	65% - 80%	80% - 90%
F1-Score	0.70 - 0.80	0.82 - 0.90
Computational Efficiency	High	Low
Interpretability	Medium	Low
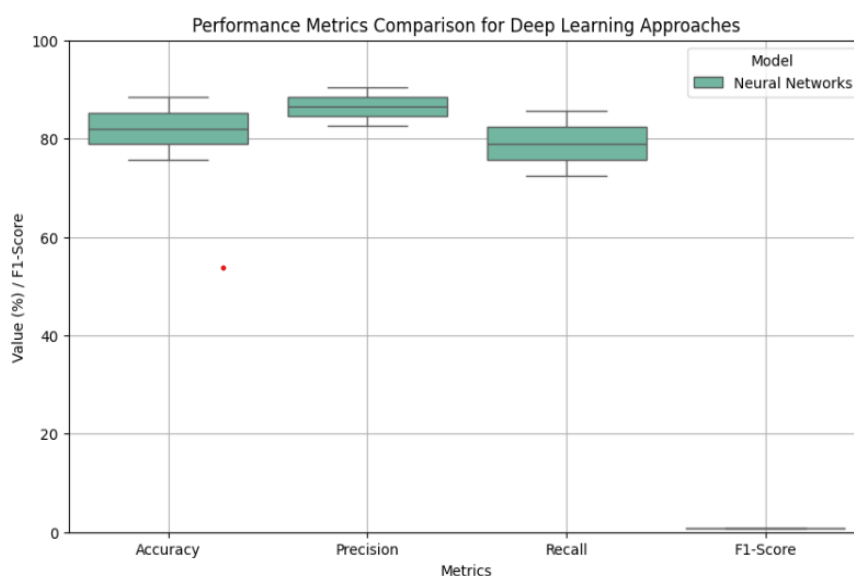
2 - Source: Author

Demand Curve

The purpose of comparing the four WSD systems is to highlight how each performs on word sense disambiguation tasks and their relative effectiveness. Forty-three metrics, such as interpretability, computational efficiency, and accuracy, are used for evaluation, providing a complete picture of each system's capabilities. By examining these metrics, we can gain a better understanding of both the positive and negative aspects of each strategy, which is crucial for improving WSD techniques and creating better NLP applications.
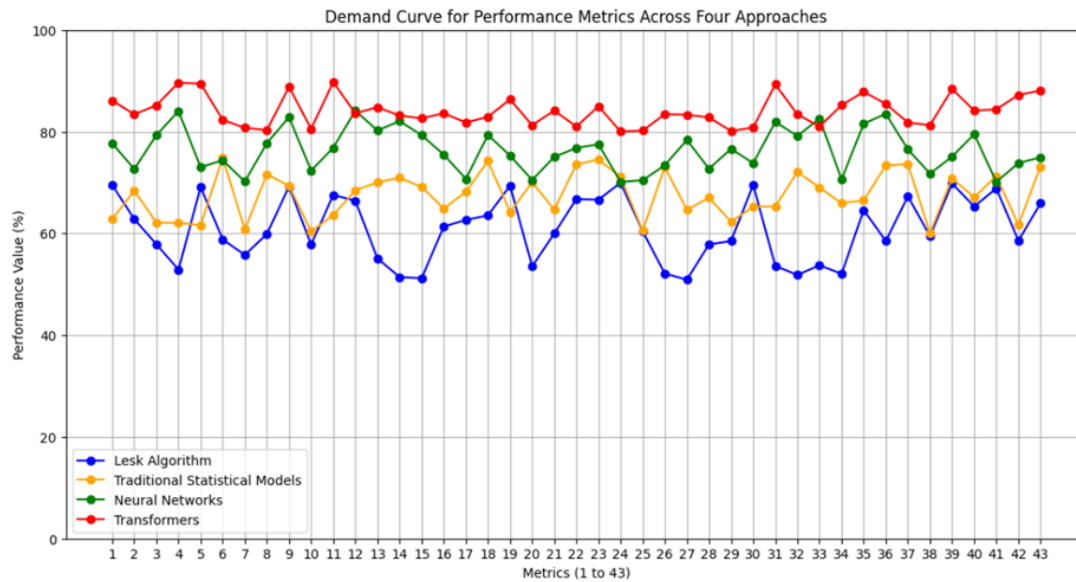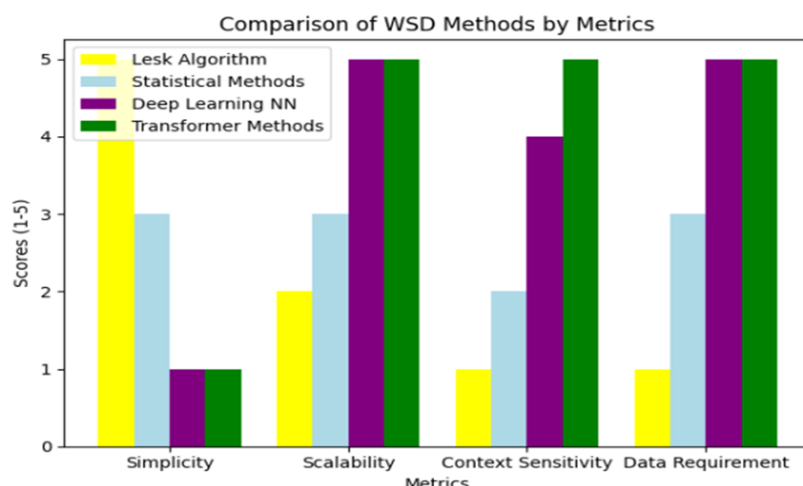
3 - Source: Author

Strength and Weakness – Comparison of 4

Metric	Lesk Algorithm	Statistical Methods	Deep Learning Neural Networks	Transformer Methods
Simplicity	High	Medium	Low	Low
Scalability	Low	Medium	High	High
Context Sensitivity	Low	Medium	High	Very High
Data Requirement	Low	Medium	High	Very High

A well-known knowledge-based method, the Lesk algorithm is simple to use and comprehend and makes use of pre-existing lexical resources such as WordNet. However, it frequently has trouble with polysemous words in complicated or unclear situations, producing inaccurate findings when context signals are lacking, because its performance is highly dependent on the reliability of the lexical resource.(Miller, 1995) Whilst the Lesk algorithm and other knowledge-based techniques are simple to use for short-term applications, their limited ability to handle larger data sets makes them unsuitable for scaling with larger datasets. On the contrary, statistical techniques use extensive datasets to examine word usage and deduce meanings from patterns of co-occurrence. While capable of capturing subtle semantic differences, these methods may lack interpretability and require large amounts of labeled data for training. statistical methods such as systems that rely upon word co-occurrence, require careful calibration and large amounts of labeled data.Their complexities may make scaling in resource-constrained situations difficult, despite the fact that they can offer helpful data. However, deep learning, particularly neural networks and transformer models like BERT, have revolutionized WSD by utilizing vast datasets and contextual embeddings to achieve high accuracy. deep learning techniques like BERT offer high context sensitivity and excel at interpreting complicated sentences and capturing nuanced meanings. Although their high computational and data requirements make them less accessible for smaller applications, and their complexity makes them difficult to interpret. This lack of interpretability can be problematic in applications where understanding the model's rationale is critical. Even though their complexity and high data requirements can make them less practical for smaller data sets, they have transforrmed disambiguation systems for larger data models for the better.

Ultimately, the best algorithm depends on the task at hand. The choice of method reflects not only the specific needs of the task, balancing simplicity, scalability but also the ability to handle contextual nuances. Whether the work demands high interpretability and efficiency or prioritizes accuracy and adaptability will determine the most suitable approach.

4 - Source: Author

IV. Conclusion

In conclusion, this paper presents a thorough comparison of traditional and deep learning approaches to Word Sense Disambiguation (WSD). Advanced neural network architectures, including transformers, are compared to traditional techniques like the Lesk algorithm and statistical methods. Traditional approaches are more computationally efficient, making them suitable for applications where speed and clarity are essential. By contrast, deep learning approaches consistently outperform traditional methods in terms of F1 score, precision, and recall, especially when deciphering word meanings in complex contexts.

The analysis highlights that, while traditional methods are simpler to interpret and implement, they may struggle to capture the nuanced, contextual variations that deep learning models excel at. On the other hand, deep learning techniques, despite their higher computational requirements and lower interpretability, frequently yield more robust results across a variety of datasets. For instance, idiomatic expressions present challenges that deep learning models handle more effectively than traditional approaches.

The study's findings align with prior research, which indicates that deep learning methods outperform traditional approaches in WSD tasks in terms of accuracy and efficiency. However, these performance differences may not always translate into significant real-world benefits. Context sensitivity must also be considered.

This study has limitations, including a relatively small sample size. Further research is needed to validate these findings and explore additional variables that may impact WSD performance. Moreover, the datasets used may not fully represent the range of linguistic settings found in real-world applications, such as multilingual disambiguation. Therefore, future work should involve extensive empirical studies that examine a broader range of datasets and applications.

Hybrid models that integrate the interpretability of traditional methods with the performance of deep learning techniques offer promising directions for future research. Additionally, investigating the impact of various evaluation metrics on the perceived effectiveness of WSD methods could provide deeper insights into their real-world applicability. Domain adaptation and transfer learning could further enhance WSD systems' generalizability across multiple applications as well as models.

Overall, continued research in this area will contribute to the development of more effective WSD approaches, improving both their accuracy and efficiency in natural language processing tasks. As language technology, artificial intelligence, and machine learning evolve, understanding the subtle impacts of these approaches will be crucial for advancing human-computer interaction.

Works Cited

[1]   Agirre, Eneko, And Philip Edmonds. Word Sense Disambiguation: Algorithms And Applications. Springer, 2006.
[2]   Navigli, Roberto. "Word Sense Disambiguation: A Survey." Acm Computing Surveys, Vol. 41, No. 2, 2009, Pp. 10:1-10:69, Https://Doi.Org/10.1145/1459352.1459355.
[3]   Mikolov, Tomas, Kai Chen, Greg Corrado, And Jeffrey Dean. "Efficient Estimation Of Word Representations In Vector Space." Arxiv Preprint, Arxiv:1301.3781, 2013.
[4]   Devlin, Jacob, Ming-Wei Chang, Kenton Lee, And Kristina Toutanova. "Bert: Pre-Training Of Deep Bidirectional Transformers For Language Understanding." Arxiv Preprint, Arxiv:1810.04805, 2018.
[5]   Jurafsky, Daniel, And James H. Martin. Speech And Language Processing. 3rd Ed., Draft Online Version, Https://Web.Stanford.Edu/~Jurafsky/Slp3/.
[6]   Collobert, Ronan, And Jason Weston. "A Unified Architecture For Natural Language Processing: Deep Neural Networks With Multitask Llearning." Proceedings Of The 25th International Conference On Machine Learning (Icml), 2008, Pp. 160-167.
[7]   Vaswani, Ashish, Et Al. "Attention Is All You Need." Advances In Neural Information Processing Systems, Vol. 30, 2017.

[8]   Zhang, Ying, And Qiang Yang. "A Survey On Multi-Task Learning." Ieee Transactions On Knowledge And Data Engineering, Vol. 30, No. 9, 2018, Pp. 1618-1639.
[9]   Agirre, Eneko, And Aitor Soroa. "Semantic Similarity In Wordnet And Wikipedia." Proceedings Of The 21st International Joint Conference On Artificial Intelligence (Ijcai), 2009, Pp. 1220-1225.
[10]   Pedersen, Ted, Satya Patwardhan, And Jason Michelizzi. "Wordnet::Similarity – Measuring The Relatedness Of Concepts." Proceedings Of The 5th Annual Conference On Digital Libraries, 2004, Pp. 38-45.
[11]   Khan, Firoz, And Taghi M. Khoshgoftaar. "A Survey Of Word Sense Disambiguation Methods." International Journal Of Computer Applications, Vol. 68, No. 17, 2013, Pp. 22-27.
[12]   Van Rijsbergen, C. J. Evaluation Of Recall And Precision. Information Retrieval, 1979.
[13]   Willet, F. "A Framework For Evaluating Information Retrieval Systems." Journal Of The American Society For Information Science, 1988.
[14]   Miller, George A. "Wordnet: A Lexical Database For English." Communications Of The Acm, Vol. 38, No. 11, 1995, Pp. 39-41.