A Hybrid Neurofuzzy System for Legal Text Comparison and Analysis

Jorge Martinez-Gil

Abstract

The daily tasks of legal professionals are often hindered by the rapid pace at which new legislation is produced. Moreover, this legislation is typically presented in unstructured formats, unsuitable for automated computer processing. This leads to a vast amount of heterogeneous information being created in a disorganized way, resulting in information overload. To address this issue, we have developed a new model for comparing legal texts that integrates the latest advances in language processing through neural architectures with traditional fuzzy logic techniques. We have tested this model using the lawSentence200 benchmark dataset, and the initial results are promising.

Keywords: Neurofuzzy systems, Legal Information Processing

1. Introduction

The rapid daily generation of information poses severe difficulties for the legal industry, mainly when dealing with the range, quantity, and speed at which this information appears. Legal professionals are often overwhelmed by managing vast amounts of data, which can reduce efficiency and increase the chances of mistakes. The sheer volume of documents and the need to process them quickly makes it difficult to maintain accuracy and thoroughness in legal work.

Legal Intelligence (LI) provides potential solutions by automating repetitive and time-consuming tasks to combat this issue. Through advanced techniques in database management, decision-making, information retrieval, and natural language processing (NLP), LI can help sift through large datasets and retrieve valuable legal data more effectively [9]. This significantly reduces the workload for legal professionals and allows for more efficient handling of legal documentation.

The main objective of this research is to develop new approaches to improve Legal Intelligence systems, focusing on the ability to mimic human decision-making within the legal field. One specific aspect of this research is comparing sections of legal texts by analyzing the similarity of meaning between sentences and paragraphs. Legal texts are often structured with formal language and technical terms, making the task more complicated.

We use transformer models because they have been shown to provide reliable results. However, these models have limitations, such as making their outputs more interpretable for human operators. Another challenge is that they require large amounts of data. Still, our working hypothesis is that combining these models with other techniques like fuzzy logic can lead to good results [11, 12].

Legal professionals often need to understand how models arrive at their conclusions. Many explainable artificial intelligence (XAI) solutions are designed with a focus on users who have an advanced understanding of technical subjects. However, symbolic artificial intelligence forms the basis for alternative methods that make systems more interpretable. Following this strategy, we explore neurofuzzy models. Below is a summary of this study's contributions:

- A neurofuzzy model is proposed, combining a fuzzy system with a neural network to address challenges posed by the legal text.
- We empirically evaluate our method using well-known legal datasets, such as lawSentence200, comparing it to state-of-the-art techniques.

The rest of this paper is organized as follows: Section 2 reviews prior work related to neurofuzzy systems for text processing. Section 3 presents the technical details of our system. Section 4 describes the experimental setup and results from our evaluation. The final section summarizes the key findings.

2. Related Work

This section describes a neurofuzzy computational model aimed at addressing the problem of determining semantic similarity in legal texts, with the dual objective of accuracy and interpretability [8]. Legal documents are particularly challenging to process using traditional methods. This type of approach could also be extended to other fields, including biomedicine [10, 13] and e-recruitment [7, 15].

We use models like BERT [5], ELMo [16], and USE [2], combined with Mamdani inference [6]. Our research is inspired by the foundational work of Angelov and Buswell [1], which informs the setup of the fuzzy component. There may also be value in exploring models of the Takagi Sugeno type [18].

3. A Neurofuzzy Approach for Legal Analysis

Our contribution is a concurrent neurofuzzy system that accounts for specific aspects that make legal text processing difficult. The system is built with neural and fuzzy components, designed to function together [17].

The neural component employs transformer models, which effectively translate abstract representations between different forms. These models use an encoder-decoder architecture, where the encoder learns to represent the input data, and the decoder generates the output.

In this study, we define 20 fuzzy rules and allow logical operators as outlined in [3]. Additionally, multi-objective algorithms could help balance the system's accuracy and ease of interpretation [4]. We now turn to our empirical study to assess the approach's effectiveness and provide a comparison to current methods.

4. Experimental Study

This section describes the experimental setup and benchmark dataset used in our evaluation. We provide an in-depth analysis of the results and compare them to other methods.

4.1. Datasets and Evaluation Criteria

We work with the benchmark dataset *lawSentence200*, which includes 200 pairs of paragraphs from legal documents. Legal experts have manually labeled these paragraphs with their degree of semantic similarity, on a scale from 1 (not similar) to 5 (equivalent). Below is an example of one of the paragraph pairs:

> *This undertaking shall be governed by the laws of New South Wales and shall terminate upon cessation of obligations under the Confidentiality Agreement in accordance with clause 6 (Term) of the Confidentiality Agreement.*
>
> *This agreement is governed by the laws of New South Wales, Australia, and each party irrevocably and unconditionally submits to the non-exclusive jurisdiction of the courts of New South Wales and the Commonwealth of Australia.*

The experts rated the similarity between these paragraphs as 3 on a scale of 1 to 5.

4.2. Results

The following section presents the experimental results obtained from applying our neurofuzzy system to the legal text dataset. Figure 1 provides a clear visual comparison between the results generated by our model and the ground truth provided by human experts. The red line in the figure represents the human judgment or expert evaluation of the semantic similarity between legal text pairs, serving as the benchmark for assessing the accuracy of our approach. The blue line, on the other hand, illustrates the best performance achieved by our neurofuzzy system after extensive training and optimization.

The following tables present the results of our neurofuzzy system compared to various state-of-the-art approaches using two different evaluation metrics: Pearson correlation and Spearman correlation. These metrics measure the degree of similarity between the model's output and the ground truth provided by human experts [14].

In Table 1 (left), the Pearson correlation coefficient is used to evaluate the performance of the models. This metric captures the linear relationship between the model's predictions and the human-annotated scores. The neurofuzzy system's performance is highlighted in two variations: the median performance, which demonstrates the model's typical accuracy across trials, and the maximum performance, representing the best result achieved. Compared to other models, such as BERT and ELMo, our neurofuzzy system shows competitive results, particularly in the maximum performance category, where it outperforms the baseline approaches.

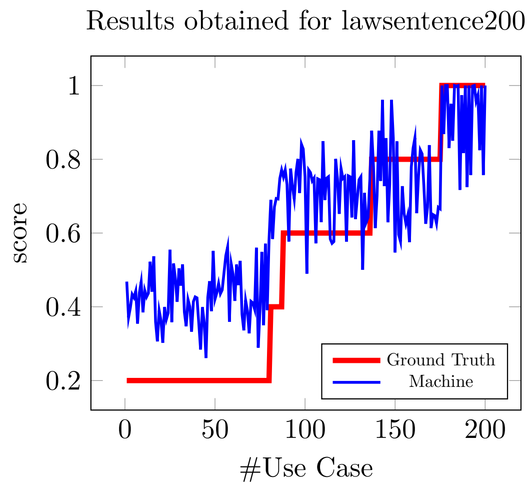

Figure 1: Overall view of the results obtained for the experiment

Table 2 (right) shows the results using the Spearman rank correlation coefficient, which measures the rank-based similarity between the predicted scores and the ground truth. This evaluation focuses more on the ordinal relationship between data points rather than their absolute values. Once again, the neurofuzzy system delivers strong results, with both its median and maximum performance surpassing those of several other approaches. This demonstrates the model's ability to maintain accuracy not only in terms of linear relationships but also when focusing on rank-based similarities.

The results from both tables indicate that the neurofuzzy system is well-suited for handling complex legal texts and provides robust performance in comparison to traditional models.

Approach	**Score** (σ)
USE Angular	0.398
ELMo Euclidean	0.546
ELMo Manhattan	0.553
ELMo Angular	0.618
ELMo Cosine	0.624
USE Euclidean	0.649
USE Manhattan	0.660
USE Cosine	0.705
BERT Pairwise	0.736
BERT Manhattan	0.740
BERT Euclidean	0.743
BERT Cosine	0.780
Neurofuzzy (median)	0.788
BERT Inner Product	0.803
Neurofuzzy (maximum)	0.826

Table 1: Results using Pearson Correlation

Approach	**Score** (ρ)
ELMo Euclidean	0.488
ELMo Manhattan	0.497
ELMo Angular	0.559
ELMo Cosine	0.585
USE Angular	0.612
USE Euclidean	0.693
USE Cosine	0.693
USE Manhattan	0.694
BERT Euclidean	0.741
BERT Manhattan	0.742
BERT Cosine	0.758
BERT Pairwise	0.766
Neurofuzzy (median)	0.779
BERT Inner Product	0.793
Neurofuzzy (maximum)	0.808

Table 2: Results using Spearman Correlation

4.3. Further Analysis

We now present a detailed analysis of the convergence behavior observed during the training of our neurofuzzy system. The goal is to highlight the relationship between accuracy and interpretability, two key metrics in evaluating the performance of machine learning models, especially in fields where understanding how decisions are made is critical, such as the legal domain. The convergence analysis is essential to demonstrate how well the neurofuzzy system adapts over time to optimize its prediction accuracy while maintaining a balance with the interpretability of its output, a feature often missing in purely neural-based models.

Figure 2a illustrates the progression of the training process for achieving optimal results for the Pearson correlation coefficient. Since we employ stochastic methods during training, the model does not always follow a predictable or uniform trajectory toward convergence. To account for this variability, we conducted 20 independent experiments and averaged their outcomes. The figure provides a comprehensive view of how the model's performance improves over time. The red line indicates the minimum values encountered during these trials, representing cases where the model struggled to achieve higher accuracy. The blue line tracks the median performance across the experiments, serving as an indicator of the most typical or expected behavior of the system. Lastly, the black line highlights the maximum performance, showcasing the potential peak accuracy the model can achieve under optimal conditions.

Figure 2b, on the other hand, focuses on the evolutionary process used to optimize the Spearman Rank Correlation. Similar to the approach taken with the Pearson correlation, the results here are based on 20 independent runs, which were necessary to capture the full range of the model's performance under different training conditions. As with Figure 2a, the minimum, median, and maximum values are depicted using red, blue, and black lines, respectively. This evolutionary approach allows us to fine-tune the neurofuzzy system, gradually improving its performance by continuously adjusting the model parameters based on feedback from the evaluation metrics.

The use of stochastic methods ensures that the model explores a wide range of potential solutions rather than converging prematurely on a suboptimal solution. However, it also introduces variability in the results, which is why multiple runs are necessary to accurately capture the behavior of the system. The median line in both figures provides a reliable indicator of the model's overall performance, while the minimum and maximum lines serve to highlight the variability that can occur during training.

The convergence analysis is particularly important because it demonstrates that the neurofuzzy system is not only capable of achieving high accuracy but also does so consistently across different trials. Moreover, by comparing the evolution of Pearson and Spearman correlations, we can observe how the system behaves when tasked with different objectives—one focused on linear relationships and the other on rank-based similarities. This dual optimization reflects the system's versatility in handling various types of data and evaluation metrics, making it well-suited for complex applications like legal text analysis.

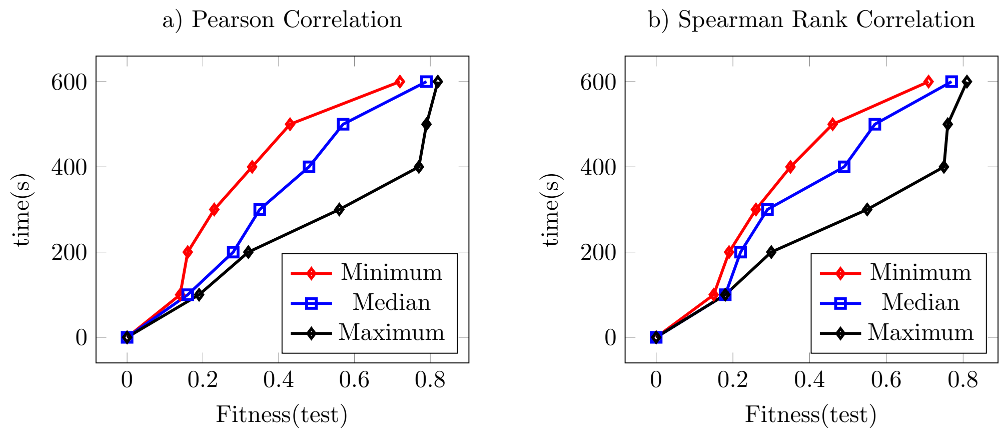

Figure 2: Convergence analysis during the execution of the evolutionary strategy for Pearson and Spearman correlations

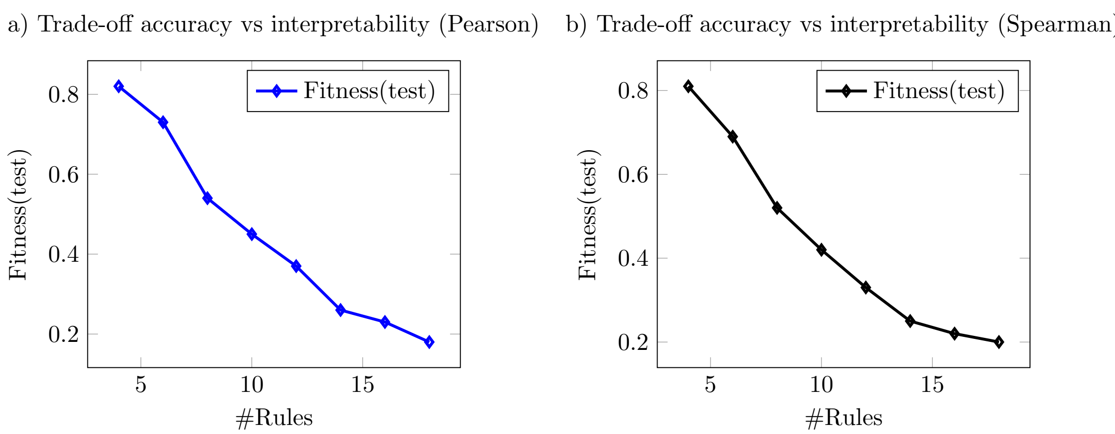

Figure 3: Analysis of accuracy vs interpretability trade-off for Pearson and Spearman correlations

4.4. Discussion

Legal text analysis is a challenging domain, and neurofuzzy systems are well-suited for this task. Due to the unique characteristics of legal documents, direct word-for-word comparison is not always effective. While neurofuzzy systems have been widely used in areas such as control systems or industrial applications, our approach represents a novel application of this computational method to legal text processing. The results show that this approach is viable and can offer significant advantages. The neural part of the model was trained using general-purpose text data, while the fuzzy part was specifically trained on legal terminology patterns. This combined strategy opens new possibilities for applying hybrid models to other language-related problems.

5. Conclusions

Neural networks and fuzzy logic have strengths and weaknesses when applied to complex problems. Neural networks are good at pattern recognition but can lack transparency in decision-making, while fuzzy logic systems offer interpretability but are more complex to automate in generating rules. Our work addressed these limitations by creating a hybrid system that merges both approaches.

Neurofuzzy systems have been extensively studied in engineering and industrial contexts but have limited use in natural language processing. Advances in neural-based methods now allow text to be converted into numerical vectors while preserving positional information. This capability makes neurofuzzy systems an effective tool for analyzing complex legal sentences and paragraphs. The fuzzy component then helps calculate similarity scores according to the task. Our findings suggest that this hybrid approach can offer new ways of processing legal texts, combining the benefits of both neural and fuzzy systems. We recommend further exploration of this methodology across different text-processing tasks.

References

[1] Plamen P. Angelov and Richard A. Buswell. Automatic generation of fuzzy rule-based models from data by genetic algorithms. *Inf. Sci.*, 150(1-2):17–31, 2003.

[2] Daniel Cer, Yinfei Yang, Sheng-yi Kong, Nan Hua, Nicole Limtiaco, Rhomni St. John, Noah Constant, Mario Guajardo-Cespedes, Steve Yuan, Chris Tar, Brian Strope, and Ray Kurzweil. Universal sentence encoder for english. In Eduardo Blanco and Wei Lu, editors, *Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing, EMNLP 2018: System Demonstrations, Brussels, Belgium, October 31 - November 4, 2018*, pages 169–174. Association for Computational Linguistics, 2018.

[3] Pablo Cingolani and Jesús Alcalá-Fdez. jfuzzylogic: a java library to design fuzzy logic controllers according to the standard for fuzzy control programming. *Int. J. Comput. Intell. Syst.*, 6(sup1):61–75, 2013.

[4] Kalyanmoy Deb, Samir Agrawal, Amrit Pratap, and T. Meyarivan. A fast and elitist multiobjective genetic algorithm: NSGA-II. *IEEE Trans. Evol. Comput.*, 6(2):182–197, 2002.

[5] Jacob Devlin, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova. Bert: Pre-training of deep bidirectional transformers for language understanding. *Preprint arXiv:1810.04805*, 2018.

[6] E. H. Mamdani and S. Assilian. An experiment in linguistic synthesis with a fuzzy logic controller. *Int. J. Hum.-Comput. Stud.*, 51(2):135–147, 1999.

[7] Jorge Martinez-Gil. An overview of knowledge management techniques for e-recruitment. *J. Inf. Knowl. Manag.*, 13(2), 2014.

[8] Jorge Martinez-Gil. An overview of textual semantic similarity measures based on web intelligence. *Artif. Intell. Rev.*, 42(4):935–943, 2014.

[9] Jorge Martinez-Gil. Automated knowledge base management: A survey. *Comput. Sci. Rev.*, 18:1–9, 2015.

[10] Jorge Martinez-Gil. Coto: A novel approach for fuzzy aggregation of semantic similarity measures. *Cogn. Syst. Res.*, 40:8–17, 2016.

[11] Jorge Martinez-Gil, Enrique Alba, and José Francisco Aldana-Montes. Optimizing ontology alignments by using genetic algorithms. In Christophe Guéret, Pascal Hitzler, and Stefan Schlobach, editors, *Proceedings of the First International Workshop on Nature Inspired Reasoning for the Semantic Web, Karlsruhe, Germany, October 27, 2008*, volume 419 of *CEUR Workshop Proceedings*. CEUR-WS.org, 2008.

[12] Jorge Martinez-Gil and José Francisco Aldana-Montes. Evaluation of two heuristic approaches to solve the ontology meta-matching problem. volume 26, pages 225–247, 2011.

[13] Jorge Martinez-Gil and José Francisco Aldana-Montes. An overview of current ontology meta-matching solutions. *Knowl. Eng. Rev.*, 27(4):393–412, 2012.

[14] Jorge Martinez-Gil and José Francisco Aldana-Montes. Semantic similarity measurement using historical google search patterns. *Inf. Syst. Frontiers*, 15(3):399–410, 2013.

[15] Jorge Martinez-Gil, Alejandra Lorena Paoletti, and Klaus-Dieter Schewe. A smart approach for matching, learning and querying information from the human resources domain. In Mirjana Ivanovic, Bernhard Thalheim, Barbara Catania, Klaus-Dieter Schewe, Marite Kirikova, Petr Saloun, Ajantha Dahanayake, Tania Cerquitelli, Elena Baralis, and Pietro Michiardi, editors, *New Trends in Databases and Information Systems - ADBIS 2016 Short Papers and Workshops, BigDap, DCSA, DC, Prague, Czech Republic, August 28-31, 2016, Proceedings*, volume 637 of *Communications in Computer and Information Science*, pages 157–167. Springer, 2016.

[16] Matthew E Peters, Mark Neumann, Mohit Iyyer, Matt Gardner, Christopher Clark, Kenton Lee, and Luke Zettlemoyer. Deep contextualized word representations. In *Proceedings of NAACL-HLT*, pages 2227–2237, 2018.

[17] Leszek Rutkowski and Krzysztof Cpalka. Flexible neuro-fuzzy systems. *IEEE Trans. Neural Networks*, 14(3):554–574, 2003.

[18] Tomohiro Takagi and Michio Sugeno. Fuzzy identification of systems and its applications to modeling and control. *IEEE Trans. Syst. Man Cybern.*, 15(1):116–132, 1985.