A Novel Neurofuzzy Model for the Comparison of Legal Texts[1]

Jorge Martinez-Gil
August 17, 2022

Abstract

The daily work of legal professionals is often hampered by characteristics such as the high speed with which new legislation is generated. In addition, the generation of such legislation is almost always done using unstructured formats that are not prepared for automatic processing by computers. As a result, a large amount of heterogeneous information is generated in a highly chaotic manner, leading to an information overload. We have designed a new model for comparing legal texts that combine the latest advances in language processing through neural architectures with classical fuzzy logic techniques to overcome this problem partially. In this regard, we have evaluated such a model with the lawSentence200 benchmark dataset, and the first results we have obtained seem promising.

Keywords: Neurofuzzy systems, Legal Intelligence, Semantic Textual Similarity

1. Introduction

A significant amount of information constantly being produced daily poses substantial challenges for the legal industry, particularly regarding that information's variety, volume, and velocity. This massive data stream results in many issues for legal professionals, who are overburdened by the continuous flow of information that impedes their tasks and makes them more prone to making mistakes. In order to address this issue, one possible set of solutions can be found within the domain of Legal Intelligence (LI). The aim is to automate mundane and time-consuming processes by utilizing methods at the confluence of database management, decision-making processes, information retrieval, and natural language understanding.

In the context of this effort, our primary focus will be on developing innovative methodologies to introduce new opportunities in LI. Our objective is to develop a system that can simulate human behavior and provide help for decision-making processes within the context of the law. This research aims to evaluate the semantic similarity between different text sections by comparing sentences and paragraphs of a textual character. The legal field is characterized by its use of a formal language dense with modifiers, and legal technicalities present a hurdle in this situation.

[1]This report is deliverable number 3 of the NEFUSI project `https://www.ngi.eu/funded_solution/nefusi/`

We are using transformers models because it has been demonstrated that these models produce the best results. However, these models also have several unique flaws, such as the inability to make it easier for human operators to comprehend their insights. Other flaws include that these models have an excessive requirement for a considerable amount of data. On the other hand, our working hypothesis is that it is possible to get good results if these methods are appropriately combined with other methods better equipped to facilitate explainability, such as fuzzy logic [12].

Legal professionals frequently require the interpretability of models in order to understand the information that has been obtained. Most of the solutions that implement some form of explainable artificial intelligence primarily focus on technical solutions geared toward users who have an extensive understanding of mathematics. Nevertheless, symbolic artificial intelligence is the foundation for yet another family of alternative methodologies. As a result, the research field known as XAI is the one that works to make systems clearer and easier to comprehend. Using this strategy as a guide, we investigate neurofuzzy models. As a result, the following is a concise summary of the contributions that this study has made:

- We propose using a neurofuzzy model, the outcome of a fuzzy system working in conjunction with a neural network. The purpose of this model is to mitigate the shortcomings of both of these systems when dealing with pieces of legal text.
- We empirically evaluate and compare our approach concerning the state-of-the-art in comparing texts of a legal nature using widely used datasets such as lawSentence200.

The remaining parts of this work are organized in the following manner: The works relevant to applying neurofuzzy systems to processing textual information are presented in Section 2. Section 3 presents the technical details for using neurofuzzy systems in legal applications. Section 4 shows the experimental setup and results after submitting our proposal to an exhaustive empirical evaluation. In the final part of this paper, we summarize the most important conclusions.

2. Related works

The following sections present a novel neurofuzzy computational model to address the problem of semantic similarity in texts of a legal nature with the double goal of being accurate and interpretable simultaneously. The main reason is that the legal context has specific characteristics that make it challenging to operate classic techniques such as [7]. An approach of this kind can be applied in other fields such as biomedicine [18] or e-recruitment [9, 13].

In the context of this work, we have focused on using popular models such as BERT[6], ELMo [14] or USE [2] using Mamdani inference [8]. Furthermore, this research is based on the seminal work of Angelov and Buswell [1] since this is how the different parameters of the fuzzy component will be set up. One might also investigate the level of success that could be accomplished using Takagi Sugeno-type models [17].

3. A neurofuzzy approach for legal analysis

Our contribution is a concurrent neurofuzzy system that considers some features that make it challenging to process text of a legal nature automatically. Our system is composed of a neural and fuzzy part designed independently but must be coupled to work together [16].

The neural component uses transformers, models suitable for transitioning abstract representations into another [15]. An encoder-decoder architecture serves as the basis for the transformer models. This architecture enables the encoder to learn how to represent the input data, which then allows this representation to be transferred onto the decoder. The decoder is responsible for obtaining the representation and then providing the output data for the user [3].

In this study, we will consider twenty fuzzy rules. In addition, logical operators will be permitted as explained in [4]. It is also feasible to use multi-objective algorithms [5] if it becomes necessary to model a trade-off between the model's accuracy and its interpretability [10]. Following this, we will discuss the empirical study we have conducted to validate our methodology and provide a quantitative and qualitative analysis of the current state of the art.

4. Experimental Study

This section details our strategy's experimental setup and the benchmark dataset we use. Following that, we thoroughly examine the various methodologies explored and the empirical outcomes. Finally, we provide a discussion of the results that we have achieved.

4.1. Datasets and Evaluation criteria

We work here with the well-known benchmark dataset *lawSentence200*[2] which is composed of 200 pairs of paragraphs extracted from documents of a legal nature, on which a group of legal experts has manually labeled their degree of semantic similarity using a scale between 1 (not similar at all) and 5 (totally equivalent). Below, we can see, as an example, one of the pairs to be compared:

> *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 have declared a similarity between them of 3 (on a scale of 1 to 5).

[2]https://github.com/Huffon/sentence-similarity

4.2. Results

We present the experimental results that we have obtained. Figure 1 represents the best overall results obtained using our neurofuzzy strategy. While the thick red line shows how the results should have been according to human judgment, the final blue line shows the best results obtained using our new approach.

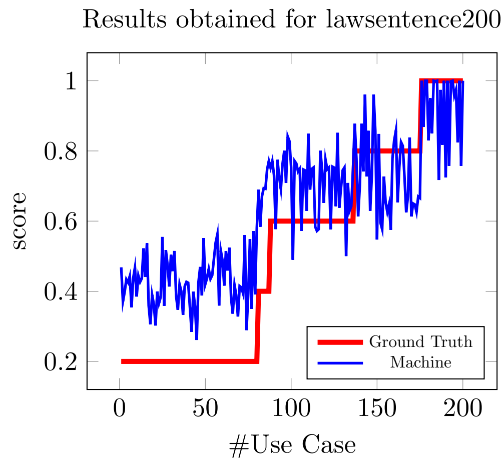

Figure 1: Overall view of the results obtained for the experiment

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 over the lawSentence200 dataset using Pearson Correlation

4.3. Further analysis

We present an analysis of the convergence when training our neurofuzzy systems and the trade-off between accuracy and interpretability [11]. Figure 2a depicts the training that was carried out

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 over the lawSentence200 dataset using Spearman Correlation

in order to determine the Pearson correlation coefficient. Because we are using stochastic methods, the results shown are an average of 20 independent experiments from which we depict the minimum (red), median (blue), and maximum (black). Figure 2b depicts the evolutionary process it took to correctly set up our technique to solve the Spearman Rank Correlation. As in the preceding scenario, the plotted results are the outcome of 20 independent experiments in which we show again the minimum, median, and maximum values obtained.

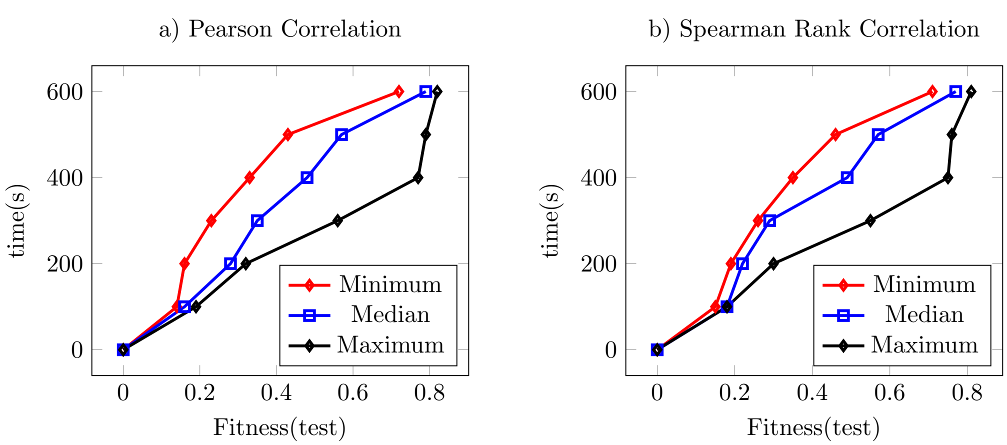

Figure 2: Study of how both solutions converge throughout the execution of the evolutionary strategy for its correct configuration

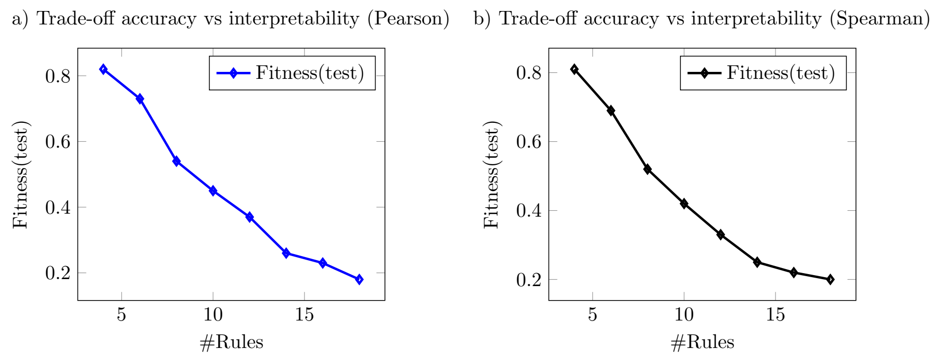

Figure 3: Modeling of the solution to achieve a trade-off between accuracy and interpretability

4.4. Discussion

The processing of legal texts is very suitable for a neurofuzzy system. Because of the unique qualities that the legal text has, it is not possible the comparison word for word. Because neurofuzzy systems are typically utilized in control applications or industrial settings, our method represents a first of its kind in introducing this computational paradigm to this context. In addition, it appears that the outcomes achieved are appropriate. Furthermore, the neural component of the model has only been trained with general-purpose text corpora, while the fuzzy component is the only one trained to recognize patterns originating from legal terminology. In light of these findings, we think it would be worthwhile to examine the use of hybrid systems in various language-related applications as it opens up a new field of possibilities.

5. Conclusions and Future Work

Both neural networks and fuzzy logic have specific qualities that make them suitable for solving particular issues but not others. Neural networks, on the one hand, are valuable tools for identifying patterns. On the other hand, they do not make it easier to comply with the decisions. At the same time, interpretability is possible within fuzzy logic systems. However, automatically deriving the rules for such decisions is rather complicated. These constraints have been the reason behind developing a novel hybrid system. The idea is that two approaches are integrated to overcome the limits of both approaches when considered on their own at the individual level.

Neurofuzzy systems have been the subject of significant research in engineering and other industrial applications. However, their application in the field of natural language understanding has seen little research on it. But the truth is that text can be converted into numerical vectors using new techniques based on neural-based solutions. This transformation can save the positional information about words associated with the text. Because of this, they are an excellent choice for digesting complex sentences and paragraphs (such as legal sentences). Then, a fuzzy logic component can help configure a similarity score according to the user's needs.

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] Cen Chen, Xinjing Huang, Feng Ji, Chengyu Wang, Minghui Qiu, Jun Huang, and Yin Zhang. CAT-BERT: A context-aware transferable BERT model for multi-turn machine reading comprehension. In Christian S. Jensen, Ee-Peng Lim, De-Nian Yang, Wang-Chien Lee, Vincent S. Tseng, Vana Kalogeraki, Jen-Wei Huang, and Chih-Ya Shen, editors, *Database Systems for Advanced Applications - 26th International Conference, DASFAA 2021, Taipei, Taiwan, April 11-14, 2021, Proceedings, Part II*, volume 12682 of *Lecture Notes in Computer Science*, pages 152–167. Springer, 2021.

[4] 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.

[5] 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.

[6] 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.

[7] Lushan Han, Abhay L. Kashyap, Tim Finin, James Mayfield, and Jonathan Weese. Umbc_ebiquity-core: Semantic textual similarity systems. In Mona T. Diab, Timothy Baldwin, and Marco Baroni, editors, *Proceedings of the Second Joint Conference on Lexical and Computational Semantics, *SEM 2013, June 13-14, 2013, Atlanta, Georgia, USA*, pages 44–52. Association for Computational Linguistics, 2013.

[8] 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.

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

[10] Jorge Martinez-Gil and Jose M. Chaves-Gonzalez. Automatic design of semantic similarity controllers based on fuzzy logics. *Expert Syst. Appl.*, 131:45–59, 2019.

[11] Jorge Martinez-Gil and Jose M. Chaves-Gonzalez. Semantic similarity controllers: On the trade-off between accuracy and interpretability. *Knowl. Based Syst.*, 234:107609, 2021.

[12] Jorge Martinez-Gil, Riad Mokadem, Josef Küng, and Abdelkader Hameurlain. A novel neuro-fuzzy approach for semantic similarity measurement. In Matteo Golfarelli, Robert Wrembel, Gabriele Kotsis, A Min Tjoa, and Ismail Khalil, editors, *Big Data Analytics and Knowledge Discovery - 23rd International Conference, DaWaK 2021, Virtual Event, September 27-30, 2021, Proceedings*, volume 12925 of *Lecture Notes in Computer Science*, pages 192–203. Springer, 2021.

[13] 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.

[14] 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.

[15] Barbara Rychalska, Katarzyna Pakulska, Krystyna Chodorowska, Wojciech Walczak, and Piotr Andruszkiewicz. Samsung poland NLP team at semeval-2016 task 1: Necessity for diversity; combining recursive autoencoders, wordnet and ensemble methods to measure semantic similarity. In *Proceedings of the 10th International Workshop on Semantic Evaluation, SemEval-NAACL-HLT 2016, San Diego, CA, USA, June 16-17, 2016*, pages 602–608, 2016.

[16] Himanshu Singh and Yunis Ahmad Lone. *Deep Neuro-Fuzzy Systems with Python.* Springer, 2020.

[17] 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.

[18] Xingsi Xue, Haiyan Yang, Jie Zhang, Jing Zhang, and Dongxu Chen. An automatic biomedical ontology meta-matching technique. *J. Netw. Intell.*, 4(3):109–113, 2019.