Jing Yang 1536 words

Assessment of Project: From Case Law to Ratio Decidendi

1. Executive Summary

The concept of *ratio decidendi* or just ratio, is an important principle in common law which guides decision for legal case or the doctrine of *stare decisis* or *precedence*, whereby the ratio applied on one case is expected to be applied in future similar cases in determining their outcomes. In the process of ratio, it is essential to identify the ratio, and this involves significant work characterized by extensive reading of several related cases. However, due to the large nature of cases, reading the cases and identifying just ratio is a time-consuming process and error prone activity. It is for this reason that automating the process can be a major opportunity to enhance and create value in the work of judges and lawyers. Through automation, it can become easy to search and retrieve just ratio from numerous previous and related cases to the one that is of interest. The current project seeks to create automatic classifiers/programs based on Machine Learning (ML) and Natural Language Processing (NLP) towards identifying ratio from case laws. In this quest, the project is based on four key objectives which include:

1. Compare the correlation between principles in cited and citing paragraphs to ratio. To do this they have manually identified ratio using the legally sound Wambaugh's test
2. Improve over Shulayeva et. al.'s ML model performance in identifying principles in case law
3. Adapt Adedjouma et al.'s NLP work for cited paragraph identification
4. Automate ratio identification by combining the above classifiers

The study finds that, already, there are legal research tools that already exist, which seek to make searching for information and just ratio easier. Some of the tools that can be considered as search engines include Westlaw search index, LexisNexis, and British and Irish legal information institute.

From the analysis, while the tools present the most advanced search engines for legal research and can provide enough information to guide professionals in their selection of relevant and important cases, they leave the identification of what law actually is to the researcher, and thus a major gap that needs to be covered with NLP and ML. To address the above challenges, the project recommends a system which is based on a methodology of identify ratio automatically guided by two tasks which include identifying principles and identifying cited paragraphs in previous cases. In the new approach, the study improves by adding new features to Shulayeva's et al's framework resulting in major improvements in the performance of the system. On the other hand, in implementing cited paragraph identifier, the project offers new way of accessing and searching paragraph citations which is based on attribution of citations to cases of interest. Finally, the project implements an automatic ratio identifier, which allows it easier to search and identify just ratio in the new system with a high level of accuracy.

2. Strengths and Weaknesses of the Project

The project presents a major and highly in-depth analysis of the current practice of just ratio, and subsequently, presents an important alternative to the existing models of research in previous cases for better decision-making. From the study's extensiveness and contribution to knowledge on automating legal research regarding *ratio decidendi* is highly critical and a possible revolution into how legal practice is undertaken. To this end, the project has various strengths and weaknesses that emerge.

Firstly, one of the major strengths of the project is that it builds on already existing concepts regarding the application of ML and NLP for legal research in *ratio decidendi.* The models suggested by Adedjouma et al. and Shulayeva et. al. is widely revered in the legal practice as highly relevant and practical in legal research. Hence, the use of the models as a basis for the

current project helps make the new model to be more relevant and relatable to what the legal field considers as important approach to legal research in just ratio. Another important strength is that the project has gone to an in-depth analysis of existing legal research and identified their strengths and gaps, and thus ensuring that it presents a strong case as to the need for a new system. The other strength of the project is that it does not just describe but goes on to show the programming that is behind some of the key concepts towards the development of a new system that is anchored on ML and NLP. However, one of the major limitations is that the system does not recommend a way of identifying cases that are not yet cited. Hence, this is a major shortcoming of making the system useful in identifying recently decided cases which are not yet cited in other case laws. Furthermore, the project has extensively limited itself on the works of Adedjouma et al. and Shulayeva et. al., and thus any errors in the previous models could represent a doom to the new system. Hence, applying and relying on even more possible models would have presented an important addition to making the work even more in-depth and one able to elicit more debate regarding the best way of automating the search for just ratio. Overall, the work is highly insightful and a major contributor to new ways of applying just ratio in the legal field.

3. High Level Summary and Context in which the Project has Been Taken

The current research has been based on the understanding that the existing models of undertaking just ratio are highly limiting due to their time consuming and error-prone processes. Firstly, the study notes that much of the research regarding *ratio decidendi* is manual where legal practitioners must take significant time reading numerous cases to identify specific segments in judgments including paragraphs and principles that can be relied on. The project builds upon already existing models based on ML and NLP as developed by Shulayeva et. al. and Adedjouma et al. et al respectively. In Shulayeva et. al. (2017), the authors recommended a system based on Naïve

Bayesian Multinomial Classifier, which is based on various features including features of classification and a machine learning framework. In their ML based system, the introduction of a feature selection capability allowed the system to narrow number of features drastically, making their system more accurate and reliable. On the other hand, the project takes important insights from Adedjouma et al. (2015), which recommended an automated system based on NLP for easier navigation and handling of cross references and their interpretation in their contexts. While the two systems by the two different studies recommend important ways of finding information for legal purposes, they each have unique approaches, that is, based on NLP and ML. It is from these two systems that the current project has recommended a new system that builds upon the two processes or systems by Adedjouma et al. and Shulayeva et. al. Hence, the new system can by the current project can be considered as an improvement as it brings together the principles of NLP and ML, and thus a better outcome and applicability in legal research automation. Consequently, the new system has been inspired by the desire to develop a better one that can utilize both NLP and ML, and greatly improve the ability of legal practitioners to find legal information that can make their work in applying just ratio in decisions.

4. Ethical Issues Associated with the Project

The use of online systems based on machine learning and natural language process as the one recommended in the current study presents various ethical issues of concern. Firstly, one of the major issues regards search engine bias, which may arise when some legal cases are preferred by the system's algorithms as compared to others. Such a bias may present a major injustice when legal practitioners and judges rely on such information only in guiding their arguments to the detriment of individuals seeking justice. Another major ethical issue concerns the entry of data regarding the various cases. In this case, the use of systems and over reliance on it means that legal

practitioners can only access information that is has been entered into such a system. If some cases are not there, they may be ignored or not utilized at all, and thus leaving out important *just ratios* out of consideration to the detriment of justice. The fact that it relies on concepts such as cross reference means that cases that are not referenced in a manner that a system can identify would be skipped or not found, and thus leaving them out of consideration in the just ratio process. The other ethical issue regards the ownership and protection of such data. If the data is stored by private entities, it presents a major risk for the public whose information has been stored in trust to few individuals. Even when the information is stored by public entities such as governments, it still means that few individuals have privileges of accessing and modifying it, and thus a major risk for the public if malicious individuals access and alter it in a manner to deny justice. Hence, it is essential to find ways of overcoming the above ethical issues to ensure the credibility of the system.

References

Adedjouma, M., Sannier, N., Sabetzadeh, M., & Briand, L. (2015). An automated framework for detection and resolution of cross references in legal texts. Requirements Engineering, 22(2), 215–237. doi:10.1007/s00766-015-0241-3

Shulayeva, O., Siddharthan, A., & Wyner, A. (2017). Recognizing cited facts and principles in legal judgements. *Artificial Intelligence and Law, 25*(1), 107-126.