

Data Science Techniques for Law and Justice: Current State of Research and Open Problems

Alexandre Quemy

IBM Poland Software Lab, Cracow, Poland
Faculty of Computing, Poznań University of Technology, Poznań, Poland
aquemy@pl.ibm.com

Abstract. By comparing the state of research in Legal Analysis to the needs of legal agents, we extract four fundamental problems and discuss how they are covered by the current best approaches. In particular, we review the recent statistical models, relying on Machine Learning coupled to Natural Language Processing techniques, and the Abstract Argumentation applied to the legal domain before giving some new perspectives of research.

Keywords: Legal Analysis, Abstract Argumentation, Case-based Reasoning

1 Introduction

The legal environment is a *messy concept* [41] that intrinsically poses a certain amount of difficulties to analyze: grey areas of interpretation, many exceptions, non-stationarity, deductive and inductive reasoning, non classical logic, etc.

In this paper, we review the large spectrum of studies on the legal domain to extract a list of major problems to cover: statistical studies, helped by Natural Language Processing techniques, Case-based Reasoning and Abstract Argumentation. A major classification factor appears to be the nature and scope of the information. Some methods will rely on legal knowledge and reasoning techniques while others will exploit the available data: past decisions records, non-legal case features, etc. However, to be broadly adopted by practitioners, a model needs to be explicit and to provide an explanation, meaning it must integrate legal knowledge at some point. This is a challenge, in particular for some Machine Learning techniques such as Random Forests that are well-known to propose non-analytical and hard-to-interpret results.

The organization of the paper is as follow: in Section 2, we state the problems to discuss their legitimacy. The next Sections are dedicated to the state of research respectively in statistical models, Case-based Reasoning (CBR) and Abstract Argumentation (AA). Last, we will discuss the current limitations and argue that the fundamental problems did not receive a homogeneous interest from the research community.

2 Law and Justice fundamental problems

To define the Law and Justice problems and orient the research we need to understand the practitioners. A major debate among the legal community is the interpretation problem, namely *legalism vs realism*, i.e. are the judges objectively applying a method contained in the text (legalism) or do they create their own interpretation (realism). For Frydman [17], the interpretation is an intellectual operation to delimit the extent of the rule in the given context of its application. For Troper [49], the interpretation is iterative because the law is expressed in generic terms and in the interpretation lie some discretionary elements. Despite this genericity, new situations will emerge, creating grey zones of interpretation to be answered by an interpretation [47]. The non-omniscience of judges coupled to the abstraction of the law often offer a room for manoeuvre on the applicability of some legal arguments and the details of sentencing. This can lead to several valid legal justifications up to a certain degree. Finding the best justifications among the possible justifications w.r.t. some criteria is a concrete problem.

For some jurists, the law must be strictly considered as an incitation mechanism to tend to the economic efficiency. For a given legal environment, one might want to control it to reach this goal (e.g. by changing the law). Broadly speaking, studying how it complies with society goals and ideals is a necessity. Due to its decentralized and non-stationary nature some drifts might appear or the current implementation of institutions or policies might not match the theoretical expectations. Some actors cannot directly change the law but they have to take decisions according to the current environment (a company needs to value the risk of its actions w.r.t. the uncertainty and take some decisions). Those observations lead to distinguish four fundamental problems:

- Predicting the outcome of a case given the legal environment. (Prediction)
- Building a legal justification, given some facts, a set of law texts with the jurisprudence and an outcome. (Justification)
- Taking the *best* decisions w.r.t. the legal environment dynamics and some criteria. (Decision)
- Modifying the legal environment dynamics to match some criteria. (Control)

The Prediction problem is challenging, even for the best legal experts: 67.4% and 58% accuracy, respectively for the judges and whole case decision, is observed [42] for the Supreme Court of the United States (SCOTUS). Using crowds, the Fantasy Scotus¹ project reached respectively 85,20% and 84,85% correct predictions. The Justification problem is in a way, an extension of the Prediction one: the reasons of a decision have to be given, including the sentence details. The Justification does not consist in explaining the prediction but finding an explanation based solely on legal factors.

¹ <https://fantasyscotus.lexpredict.com/>

One can illustrate the four problems with the European Union directive to set fines for abuses of dominant position²:

“The level of a fine must be sufficiently high both to punish the firms involved and to deter others from practices that infringe the competition rules. [...] The basic amount is calculated as a percentage of the value of the sales connected with the infringement [...]. The percentage of the value of sales is determined according to the gravity of the infringement (nature, combined market share of all the parties concerned, geographic scope, etc.) and may be as much as 30 %.”

The Prediction problem consists in determining if the verdict will sanction or not the infringement, possibly using non-legal factors. The Justification problem resides in estimating the fines compatible with the text. It involves estimating the dynamic quantities mentioned such as the combined market share. A possible Control problem would be to determine if in practice the calculation reaches the goals, namely punishing and preventing the other agents to break the law. Conversely, for a company, a Decision problem would be to determine if given the current state of the legal environment, an aggressive takeover would not be perceived as breaking the law and thus leading to undesirable sanctions.

3 Predictive Models

An ideal court, is a court where the judges are perfectly rational, free of all bias and preferences, and omniscient. In this configuration, all judges must reach the same decisions, and two exactly similar cases would result in the same decision. The decisions are not correlated: it is impossible to predict a judgement from an ideal court using the information from the past cases. Of course, in real life there is no such ideal court and statistical methods try to detect hidden patterns between the sequence of decisions to predict future decisions³.

In [34, 42] a classification tree per Justice is used to predict the votes after being trained on cases described by 6 visible⁴ features such as the lower court circuit or the type of petitioner. The model performs significantly better than the experts, while the difference of prediction at the judge level is not significant. An important result is also that the experts were better at drawing good conclusions on the vote of the most extremely ideologically oriented judges⁵.

The Stochastic Block Model [19] predicted 77% of individual Justice votes given the votes of the other Justices of the SCOTUS and the history of decisions. Despite the model is not fully predictive and cannot be used to build legal explanation, as it does not use any legal arguments or case-based information,

² <http://eur-lex.europa.eu/legal-content/EN/TXT/?uri=URISERV%3A126118>

³ One may notice that the less a court is predictable, the closer it is from an ideal court.

⁴ That is to say also available to legal experts.

⁵ Where the measure is calculated as given by Martin and Quinn [38, 39].

but hidden associations between social actors, it clearly exhibits empirical proofs to support the attitudinalism paradigm⁶.

A general, robust and fully-predictive model has been proposed in [27] using Random Forest [18] and successfully identified 69.7% of the Court decisions over 60 years. The model is built up on more than 300 features, divided into 3 categories: Court and Justice level information, Case information, Historical Justice and Court information. The case features account for 23% of the predictive power while the Court background only for less than 5%. In other words, most of the predictive power holds in the behavioral trend including ideological shifts. More than the exact percentage for this particular model and Court, this comforts the realism paradigm once again. The classifier returns the weights aka the importance of every feature in the prediction. Those weights evolves in time giving additional comprehensive hints on the Court dynamic despite a complex interpretation due to the correlation between the features [48].

Using NLP techniques, the authors of [2] achieve 79% accuracy to predict the decisions of the European Court of Human Rights (ECtHR). They make the hypothesis that the textual content of the European Convention of the Human Rights and the case elements holds hints that will influence the decision of the Judge. They extracted from the case the top 2000 N-grams, calculated their similarity matrix based on cosine measure and partition this matrix using Spectral Clustering to obtain a set of interpretable topics. The binary prediction was made using an SVM with linear kernel. Contrary to the previous studies, they found out the formal facts are the most important predictive factor which tend to favor realism. However, as they used legal documents to extract their features, the non-legal hints are most likely to be less present.

As enlightened by [19,27], one prominent factor in the predictive power of statistical methods is the *ideology* of the judge. To capture this reality, many estimators of the *ideal point*, a latent variable to position the ideology in a continuous space, have been developed. A taxonomy of estimators can be established according to the type of information they rely on: the party affiliation [46], the expert judgements [43,44] or the votes [38,39], possibly including decision-related material [23,30,45]. All of them focus on SCOTUS. The Segal-Cover Score [43,44] is constructed out of editor's assessments published before the nomination of a given judge and whose information are extracted and interpreted manually. The Martin-Quinn Score [33,38,39] uses an Item Response Model [22] with the ideal point modelled as a random walk, thus evolving in time. NLP techniques are used in [30,45] to model the influence of opinion texts (amicus briefs, opinions) on the ideal point of judges while in [23], the ideal point is expressed in the topic space of the legal texts, enabling a more complete *preference profile* per Justice.

⁶ On top the predictions, the authors shown the existence of different predictability between judges, implying a difference of attitude toward the law, as well as a decrease in the SCOTUS predictability during some periods or depending on the political party at the presidency.

4 Case-based Reasoning

To solve a given problem, a case-based system performs the following cycle [1]:

1. Search for the most related past cases, either by filtering the irrelevant cases or selecting the closest ones depending on a metric and a KNN algorithm.
2. Adapt the best case solution to the new case.
3. Evaluate and revise the proposed solution, including at least why the solution is not satisfying.
4. Integrate the solution to the database.

Among the legal CBRs, CATO [3] is one of the most famous. Used to teach argumentation to students, it consists in 8 basics *reasoning moves* over a database of cases described by manually extracted legal factors. The factors are connected together in a pre-established hierarchy. The links are annotated with a plus (+) or minus (−) to indicate if a factor attacks or defend another one. In a sense, CATO can be seen as an ancestor of Abstract Argumentation.

If CBRs are more efficient and reliable than classical rule-based systems when it comes to law oriented problems [28], many drawbacks subsist: similarity and relevance of precedent cases are dynamic, non-stationary as social and governmental laws evolve [11]. A novel approach to reason by analogy is to learn some rules from a set of similar cases and then to use those rules to reason and predict a new case [25]. The rules represent the prevailing norm in a legal environment at a given moment and thus can evolve while the law corpus remains the same.

5 Abstract Argumentation

Abstract Argumentation (AA) [13] is a hot topic in non-monotonic reasoning, that is to say where there is a need to reason with pieces of information that can turn to be in contradiction with classical logic. The versatility of AA is thus it can be used as a descriptive tool, for instance by including agent's beliefs in order to study and understand a given situation, as well as a normative tool where, given a knowledge base, we want to infer the best action or decision to take [4]. In [10] the author defines the usage of AA with the following three steps:

1. Defining the arguments and the relation(s) between them.
2. Valuating the arguments using their relations, a strength, etc.
3. Selecting some arguments using some criteria (a *semantic*).

From this very succinct summary, one may catch a glimpse on the interest of AA applied to the legal domain, either as a modeling or decision aid tool: the judgment of a case in both Common Law or Civil Law countries follow this exact procedure of collecting the evidence and legal facts, evaluating their suitability in the context of the case, take a decision based on those facts.

In a similar fashion of CATO, an Argumentation Framework (AF) where the arguments are taken manually from a set of cases and the attack deduced from

the opinions has been proposed [8, 9]. Given a particular case, the arguments that do not apply are removed to obtain a new AF, a subset of the initial one. The key arguments are calculated to determine if the outcome is admissible, i.e. in favor of the plaintiff.

Oren and al. [35, 36] instantiate arguments as inference rules and use Subjectic Logic [24] to value the strength of the arguments. After giving an algorithm to propagate the strength of the arguments, they define a dialogue game protocol for several agents to argue about the state of the environment. Using Assumption-Based Argumentation [14], Dung and Thang [16] model a pool of agents arguing about some arguments in front of judges with a final jury taking a decision. If the jury can also introduce new arguments, they are limited to considering the probabilities of causal arguments, while the judges are the only one to determine the admissibility of an argument. Other Assumption-Based Argumentation applications to the legal domain can be found in [15, 29]. Meanwhile, several quantitative methods to calculate and propagate the strength of arguments have been developed [6, 40] with e.g. Social AA [31] that intends to model debates and decision-making in social networks using a voting system. Future work should consider using them in legal analysis due to good results obtained in systems with similar characteristics.

Recently, CBR based on AA formalism have been proposed. In the previous approaches the arguments are seen as elements of a case while in AA-CBR the arguments are the cases, and the attacks relations between cases. The cases are defined as a set of features and an outcome. In [50], the outcome of a new case is given depending on the *grounded extension* and a justification⁷ to support the decision is built using a Dispute Tree [14] while in [5] the outcome is deduced from rules learnt from past cases. In [37], the outcome is given after a deliberation between several agents and a fully adaptive and dynamic approach such that the agents learn from each other and are able to resolve conflict through specific a game protocol.

6 Summary, Limitations, and Future Work

As analyzed in Sections 2 to 5, the approaches can be broken down into three groups, namely: the statistical models, the case-based reasoning and the abstract argumentation.

Among the **statistical models**, we distinguish two categories. First, some methods use Machine Learning [19, 27, 34, 42] or NLP [2] to predict the outcome of a case given the past cases or votes. Papers in the second category focus on modeling and estimating the ideal point of judges and predict votes solely using this estimation [23, 30, 33, 38, 39, 43–45]. Those methods tackle the prediction problem but cannot handle the justification problem. All models adopt a binary outcome and further work must integrate sentencing elements for more complex and detailed outcomes.

⁷ To be precise, some previous cases and some of their specific features. Thus, this is not a legal justification for Roman Law.

	Information	LK	General	Robust	Fully Pred.	Extra Data	Just.
Pred. Models							
[34, 42]	Data-driven	no	no	no	yes	past cases	no
[19]	Data-driven	no	yes	yes	no	past votes	no
[27]	Data-driven	no	yes	yes	yes	past cases	no
[2]	Data-driven	no	yes	no	yes	past cases	no
Ideal Point							
[43, 44]	Data-driven	no	yes	no	yes	non-legal	no
[33, 38, 39]	Data-driven	no	yes	yes	yes	past votes	no
[30, 45]	Data-driven	no	yes	no	yes	amicus	no
[23]	Data-driven	no	yes	no	yes	opinions	no
CBR							
[3]	Rule-based	yes	no	no	yes	past cases & legal factors	yes
AA							
[8, 9]	Both	yes	no	no	yes	past cases & legal factors	yes
[35, 36]	Both	yes	yes	yes	yes	norm	yes
[14]	Rule-Based	no	yes	yes	yes	-	yes
[6, 40]	Both	no	yes	no	yes	-	-
[31]	Rule-based	no	yes	no	yes	-	-
AA-CBR							
[50]	Both	no	no	no	yes	past cases	partly
[5]	Both	no	yes	yes	yes	past cases	partly
[37]	Both	no	yes	yes	yes	past cases	partly

Table 1: Comparison of Law and Justice approaches.

The features to compare include: (1) whether the method relies on pre-defined rules or exploit the data (Information), (2) if it uses Legal Knowledge (LK), (3) the capacity to be applied to any case over years (General), (4) to adapt to the environment shifts (Robust), (5) to make a prediction solely based on past information (Fully Pred.), (6) the extra-data used on top of the information about the current case (Extra Data) and (7) the capacity to justify a decision (Just.).

Case-base reasoning systems evolved in their formalism from ad-hoc structures to the Abstract Argumentation structure. The most important factor to distinguish between CBR systems is the way they build the justification. In [3, 8, 9] the justification is based on pre-determined legal factor hierarchy and past cases description, while in [5, 37, 50] the factor hierarchy is deduced from past cases relations. Most CBRs neglect the importance of non-legal factors, and thus, they implicitly work in an ideal court setting, ruining their capacity to handle the Prediction problem.

Finally, CBRs do not account for the temporal dimension of the legal environment. Future work must focus on integrating *dynamic features* into (AA)-CBR systems. The case features are static and if a feature is shared between cases, e.g. the judge, they may be considered close because they are judged by the same person. However, the preferences of the judges change in time and are influenced by local documents such as case amicus. Thus this cannot directly be used as a feature. Conversely, using the value of the ideal point estimation as a feature is not a good idea: the ideal point strongly depends on many underlying features (the judge, the area of law, other case feature).

In **Abstract Argumentation**, apart from AA-CBR, two kinds of approaches emerge. A *positive* one intends to model real-life decision processes or environment

[14, 35, 36]. A *normative* one tries to elaborate methods to select among the best alternatives and discuss arguments [6, 31, 40]. If AA is a promising tool to handle the Prediction and Justification problem, the level it operates, mainly at case level, is not directly suitable for the Decision and Control. Despite a rich literature on dynamic in AA [7, 12, 32] including recent attempts to solve decision process in non-stationary environments [20, 21], as far as we know, there is no attempt to tackle the Control and Decision problem.

A main pitfall of all the approaches is the lack of automation: the Segal-Cover score relies on manually extracted information, the features hierarchy in CATO are manually constructed, etc. Future work should focus on the automation using the recent progresses in NLP and the availability of data. For instance, it is possible to update the Segal-Cover score on a daily basis by extracting the information published over the internet⁸.

To conclude, many elements of law look like finance 25 up to 50 years ago [26]. We do believe tranfering risk assessment techniques from finance to law is the way to handle all the four problems but requires to corretly quantitatively model the underlying dynamics of the legal environment: a new challenge for data science.

References

1. Aamodt, A., Plaza, E.: Case-based reasoning: Foundational issues, methodological variations, and system approaches. *AI Communications* 7(1), pp. 39–59 (1994)
2. Aletras, N., Tsarapatsanis, D., PreoŃiu-Pietro, D., Lampos, V.: Predicting judicial decisions of the European Court of Human Rights: a Natural Language Processing perspective. *PeerJ Computer Science* 2, pp. 93–112 (2016)
3. Aleven, V., Ashley, K.D.: Evaluating a learning environment for case-based argumentation skills. In: *Proc. of Int. Conf. on Artificial Intelligence and Law (ICAIL)*. pp. 170–179 (1997)
4. Amgoud, L.: A unified setting for inference and decision: An argumentation-based approach. In: *Proc. of Int. Conf. on Uncertainty in Artificial Intelligence (UAI)*. pp. 26–33 (2005)
5. Athakravi, D., Satoh, K., Law, M., Broda, K., Russo, A.: Automated inference of rules with exception from past legal cases using ASP. In: *Proc. of Int. Conf. on Logic Programming and Non Monotonic Reasoning (LPNMR)*. pp. 83–96 (2015)
6. Baroni, P., Romano, M., Toni, F., Aurisicchio, M., Bertanza, G.: Automatic evaluation of design alternatives with quantitative argumentation. *Argument & Computation* 6(1), pp. 24–49 (2015)
7. Barringer, H., Gabbay, D., Woods, J.: Temporal dynamics of support and attack networks: From argumentation to zoology. In: *Proc. of Int. Conf. on Mechanizing Mathematical Reasoning*. pp. 59–98 (2005)
8. Bench-Capon, T.J.: Representation of case law as an argumentation framework. In: *Proc. of Int. Conf. on Legal Knowledge and Information Systems (JURIX)*. pp. 103–112 (2002)
9. Bench-Capon, T.J.: Try to see it my way: Modelling persuasion in legal discourse. *Artificial Intelligence and Law* 11(4), pp. 271–287 (2003)

⁸ e.g.: IBM Watson Services offer query services over hundred of thousands of articles indexed every day.

10. Cayrol, C., Lagasquie-Schiex, M.C.: On the acceptability of arguments in bipolar argumentation frameworks. In: Proc. of European Conference on Symbolic and Quantitative Approaches to Reasoning and Uncertainty (ECSQARU). pp. 378–389 (2005)
11. Delgado, P.: Survey of casebased reasoning as applied to the legal domain (2007), unpublished
12. Delobelle, J., Haret, A., Konieczny, S., Maily, J., Rossit, J., Woltran, S.: Merging of abstract argumentation frameworks. In: Proc. of Int. Conf. on Principles of Knowledge Representation and Reasoning (KR). pp. 33–42 (2016)
13. Dung, P.M.: On the acceptability of arguments and its fundamental role in non-monotonic reasoning, logic programming and n-person games. *Artificial Intelligence* 77, pp. 321–357 (1995)
14. Dung, P.M., Kowalski, R.A., Toni, F.: Dialectic proof procedures for assumption-based, admissible argumentation. *Artificial Intelligence* 170(2), pp. 114–159 (2006)
15. Dung, P.M., Thang, P.M.: Towards an argument-based model of legal doctrines in common law of contracts. In: Proc. of Int. Conference on Computational Logic in Multi-Agent Systems (CLIMA). pp. 111–126 (2008)
16. Dung, P.M., Thang, P.M.: Towards (probabilistic) argumentation for jury-based dispute resolution. In: Proc. of Int. Conf. on Computational Models of Argument (COMMA). pp. 171–182 (2010)
17. Frydman, B.: *Le sens des lois: histoire de l’interprétation et de la raison juridique. Penser le droit*, Bruylant (2005)
18. Geurts, P., Ernst, D., Wehenkel, L.: Extremely randomized trees. *Machine Learning* 63(1), pp. 3–42 (2006)
19. Guimerà, R., Sales-Pardo, M.: Justice blocks and predictability of u.s. supreme court votes. *PLOS ONE* 6(11), pp. 1–8 (2011)
20. Hadoux, E., Beynier, A., Maudet, N., Weng, P., Hunter, A.: Optimization of Probabilistic Argumentation With Markov Decision Models. In: Proc. of Int. Joint Conf. on Artificial Intelligence (IJCAI). pp. 2004–2010 (2015)
21. Hadoux, E., Beynier, A., Weng, P.: Sequential decision-making under non-stationary environments via sequential change-point detection. In: Workshop on Learning over Multiple Contexts (LMCE) at ECML-PKDD (2014)
22. Hambleton, R.: *Fundamentals of Item Response Theory. Measurement Methods for the Social Science*, SAGE Publications (1991)
23. Islam, M.R., Hossain, K.S.M.T., Krishnan, S., Ramakrishnan, N.: Inferring multi-dimensional ideal points for us supreme court justices. In: Proc. of Int. Conf. on Artificial Intelligence (AAAI). pp. 4–12 (2016)
24. Jøsang, A.: *Subjective Logic: A Formalism for Reasoning Under Uncertainty. Artificial Intelligence: Foundations, Theory, and Algorithms*, Springer (2016)
25. Kannai, R., Schild, U.J., Zeleznikow, J.: There is more to legal reasoning with analogies than case based reasoning, but what? In: *Language, Culture, Computation. Computing of the Humanities, Law, and Narratives*. pp. 440–451 (2014)
26. Katz, D.M., Bommarito, M.J.: Fin(legal)tech – law’s future from finance’s past (talk) (2017), retr. Mar 21, 2017, from <https://speakerdeck.com/danielkatz/fin-legal-tech-laws-future-from-finance-past-professors-daniel-martin-katz-plus-michael-j-bommarito>
27. Katz, D.M., Bommarito, M.J., Blackman, J.: Predicting the behavior of the supreme court of the united states: A general approach. *SSRN Electronic Journal* (2014)
28. Kowalski, A.: Case-based reasoning and the deep structure approach to knowledge representation. In: Proc. of Int. Conf. on Artificial Intelligence and Law (ICAIL). pp. 21–30 (1991)

29. Kowalski, R.A., Toni, F.: Abstract argumentation. *Artificial Intelligence and Law* 4(3), pp. 275–296 (1996)
30. Lauderdale, B.E., Clark, T.S.: Scaling politically meaningful dimensions using texts and votes. *American Journal of Political Science* 58(3), pp. 754–771 (2014)
31. Leite, J., Martins, J.: Social abstract argumentation. In: *Proc. of Int. Joint Conf. on Artificial Intelligence (IJCAI)*. pp. 2287–2292 (2011)
32. Maily, J.: Dynamic of argumentation frameworks. In: *Proc. of Int. Joint Conf. on Artificial Intelligence (IJCAI)*. pp. 3233–3234 (2013)
33. Martin, A.D., Quinn, K.M., Epstein, L.: The median justice on the united states supreme court. *North Carolina Law Review* 83, pp. 1275–1322 (2004)
34. Martin, A.D., Quinn, K.M., Kim, P.T., Ruger, T.W.: Competing approaches to predicting supreme court decision making. *Perspectives on Politics* pp. 761–767 (2004)
35. Modgil, S., Faci, N., Meneguzzi, F., Oren, N., Miles, S., Luck, M.: A framework for monitoring agent-based normative systems. In: *Proc. of Int. Conf. on Autonomous Agents and Multiagent Systems (AAMAS)*. pp. 153–160 (2009)
36. Nirn, O.: An Argumentation Framework Supporting Evidential Reasoning with Applications to Contract Monitoring. Ph.D. thesis, University of Aberdeen (2007)
37. Ontañón, S., Plaza, E.: An argumentation-based framework for deliberation in multi-agent systems. In: *Proc. of Int. Conf. on Argumentation in Multi-agent Systems (ArgMAS)*. pp. 178–196 (2008)
38. Quinn, K.M., Martin, A.D.: Dynamic Ideal Point Estimation via Markov Chain Monte Carlo for the U.S. Supreme Court, 1953-1999. *Political Analysis* 10(2), pp. 134–153 (2002)
39. Quinn, K.M., Park, J.H., Martin, A.D.: Improving judicial ideal point estimates with a more realistic model of opinion content (2006), unpublished
40. Rago, A., Toni, F., Aurisicchio, M., Baroni, P.: Discontinuity-free decision support with quantitative argumentation debates. In: *Proc. of Int. Conf. on Principles of Knowledge Representation and Reasoning (KR)*. pp. 63–73 (2016)
41. Rissland, E.L.: Ai and similarity. *IEEE Intelligent Systems* 21(3), pp. 39–49 (2006)
42. Ruger, T.W., Kim, P.T., Martin, A.D., Quinn, K.M.: The supreme court forecasting project: Legal and political science approaches to predicting supreme court decisionmaking. *Columbia Law Review* 104(4), pp. 1150–1210 (2004)
43. Segal, J.A., Cover, A.D.: Ideological values and the votes of u.s. supreme court justices. *American Political Science Review* 83(2), pp. 557–565 (1989)
44. Segal, J.A., Epstein, L., Cameron, C.M., Spaeth, H.J.: Ideological values and the votes of u.s. supreme court justices revisited. *The Journal of Politics* 57(3), pp. 812–823 (1995)
45. Sim, Y., Routledge, B.R., Smith, N.A.: The utility of text: The case of amicus briefs and the supreme court. *Computing Research Repository* (2014)
46. Spitzer, M.L., Cohen, L.: Solving the chevron puzzle. *Journal of Law & Contemporary Problems* 57, pp. 65–110 (1994)
47. Sunstein, C.R.: How law constructs preferences. *Georgetown Law Journal* 86, pp. 2637–2652 (1998)
48. Tolosi, L., Lengauer, T.: Classification with correlated features: unreliability of feature ranking and solutions. *Bioinformatics* 27(14), pp. 1986–1994 (2011)
49. Troper, M.: *La théorie du droit, le droit, l'état*. Léviathan, Presses universitaires de France (2001)
50. Čyras, K., Satoh, K., Toni, F.: Abstract argumentation for case-based reasoning. In: *Proc of Int. Conf. on Principles of Knowledge Representation and Reasoning (KR)*. pp. 549–552 (2016)