



## **The impact of legal advocacy experience within the US Supreme Court on trial decision outcomes**

D'Rosario, Michael. 2017. The impact of legal advocacy experience within the US Supreme Court on trial decision outcomes, *International journal of strategic decision sciences*, vol. 8, no. 1, pp. 65-76.

DOI: [10.4018/IJSDS.2017010105](https://doi.org/10.4018/IJSDS.2017010105)

©2017, IGI Global

Reproduced with permission.

Downloaded from DRO:

<http://hdl.handle.net/10536/DRO/DU:30103726>

# The Impact of Legal Advocacy Experience Within the US Supreme Court on Trial Decision Outcomes

Michael D’Rosario, Department of Finance, Deakin Business School, Deakin University, Geelong, Australia

## ABSTRACT

Can advocacy experience differentials be used in formulating a model to predict trial outcomes in the US Supreme Court? In recent years, a number of studies have considered the role of experience before the Supreme Court in the determination of trial outcomes. The work of Sheehan, Mishler and Songer supports the assertion that trial experience possessed by trial lawyers is associated with disproportionate rates of success. McGuire is a significant study into the impact of the experience of competing trial lawyers on judicial decision making. The study identified the experience differentials of lawyers and sought to determine the impact of these differentials on trial outcomes. The study found that trial experience possessed by trial lawyers was associated with favourable trial outcomes. The current study extends upon McGuire, assessing the robustness of the original study employing a series of more advanced parametric estimation techniques. The study then uses the McGuire logistic model framework to develop a model of prediction, employing a backward propagation, multilayer perceptron network model.

## KEYWORDS

ANN, Lawyer Experience, McGuire Logistic Model Framework, Supreme Court

## INTRODUCTION

While the determination of any outcome to a trial process should be based on merit, there is little doubt that the experience of legal advocates does have a bearing on trial outcomes. But little is known about the degree to which the experience of legal representatives influences the decision making of judges. Lawyers often procure significant experience within the United States Supreme Court system and as such are likely to improve as advocates within the system through time. Evidence suggests that individuals frequently engaging in litigation display different rates of success to those that engage with the system only episodically and infrequently (Sheehan et al., 1992). Kevin T McGuire conducted a significant study into the impact of the experience of competing trial lawyers on judicial decision making (McGuire, 1995). The study identified the experience differentials of lawyers and sought to determine the impact of this differential on trial outcomes. The McGuire study considered trial outcomes of the US Supreme Court over a decade-long period, accounting for the experience of trial lawyers and, critically, the litigant strength differential a measure of the relative economic standing

DOI: 10.4018/IJSDS.2017010105

Copyright © 2017, IGI Global. Copying or distributing in print or electronic forms without written permission of IGI Global is prohibited.

of litigants' on trial outcomes (McGuire, 1995). Adopting a logistic regression model, McGuire found that trial experience and a favourable amicus curie (Please note that while amicus curie are not parties to the litigation, they remain an important source of guidance and influence and as such were included as a parameter in the model) are both positively associated with favourable trial outcomes.

The present study shall extend upon McGuire's worthwhile research by firstly considering the direction and strength of the claimed associations noted by McGuire (1995), employing a highly innovative set of estimation techniques. The study shall employ a series of novel parametric estimation techniques that provide potentially superior estimators while accommodating the challenges of the available data. Pertinently, the study will then consider McGuire's existing predictive model and posit an alternative: a multilayer neural network model. This modelling technique can potentially provide superior predictive outcomes. An artificial neural network is a structure that seeks to replicate a neuron structure. It is simply a series of weighted, aggregative, non-linear values that have the potential to provide more accurate predictive outcomes than traditional parametric estimation techniques, as well as alternative non-parametric techniques. The use of such methods is not uncommon in the social sciences but remains relatively underemployed within interdisciplinary legal research.

Importantly, there remains a dearth of research considering key empirical questions, such as the impact of trial experience of lawyers and the role of amicus curie on trial outcomes; and an even greater dearth of literature positing practically framed deterministic models of trial outcomes utilising non-parametric models. Notable parametric studies within this area of research include those of Matthew Sag, Tonya Jacobi and Barton Beebe. There is little if any non-parametric research within this field. The use of neural networks is not however absent within the broader body of legal research. Warner (1990, 1992, 1993), was amongst the first to posit the benefits of logic based legal expert systems and one of the first legal realists to employ neural networks within legal expert systems. Warner (1990, 1992, 1993), contends that neural networks have the unique capacity to replicate deontic logic. However, the use of neural networks has not been limited to legal realists alone. Zeleznikow and Nolan (2001) have advanced scholarship in the domain of law and artificial intelligence systems focusing on soft computing based intelligent decision support systems. Lothar Philipps (1991), emphasis the benefit of neural networks within the legal domain. Hobson & Slee (1994), emphasis the benefits of neural networks in their application to case analysis and case based reasoning. Rose & Belew (1989, 1991), claim that neural networks can retrieve appropriate data while employing a legal realistic perception of the law.

While some of these statements may overstate the benefits of neural networks within the legal domain, it is undeniable that neural networks possess greater pattern identification and are more consistent with logic and rule base systems of case classification than logistic regression techniques. The current study responds to both the noted dearth of deterministic research and the dearth of empirical work considering these matters.

The present study adopts a relatively uncommon predictive method in framing a judicial decision model, and as such the paper will detail the manner in which such models can be employed in legal research generally, a further contribution of this paper. The structure of the paper is as follows: firstly, the paper will introduce the extant judicial decision research, highlighting both the nature of the findings and the methods employed. The paper will then outline the artificial neural network method and its potential benefits for legal decision research. The following sections will detail the data, methodology, exploratory estimations and the findings of the research, and the potential implications for future research.

## **LEGAL EXPERIENCE, LEGAL ARGUMENT AND JUDICIAL DECISION MAKING**

There is a significant amount of research that supports the conclusion that Supreme Court justices are highly responsive to the information contained in oral arguments (Baker, 1996, Enns 1998, Stern et al 2002).

There have been a number of significant works conducted considering the matter of judicial decision making. Nancy Billica's *Just Leave it to the Courts* is an example of a quantitative study considering the judicial decision making process. Specifically, Billica considered the process by which 'states are pushed into the courts for review and interpretation' (Billica, 1997). The study considered statutes across a 28-year period through the 86<sup>th</sup> and 99<sup>th</sup> congresses from 1959-86. Through the use of a randomly selected subset of statutes from the main dataset, the study tracks statutes that have been challenged by the courts, considering the issues addressed by the courts in interpreting statutes (Billica, 1997). The study then identifies and measures the factors potentially pertinent to whether and to what extent a particular statute may be the subject of review by the courts (Billica, 1997). The study identified a set of correlates relevant to whether a given statute will land in the courts and the quantum of ensuing cases. The study identifies that different courts can offer markedly different statutory interpretations (Billica, 1997).

In 'The Median Justice on the US Supreme Court', Andrew Martin, Kevin Quinn and Lee Epstein consider the matter of median justice identification (Martin, Quinn and Epstein, 2005). Adapting the innovative method framed by Martin and Quinn (2002), the study develops a superior quantitative method for median justice identification (Martin et al, 2005). 'The Median Justice' is an exemplary study of interdisciplinary law and decision sciences research. Nonetheless, there remain relatively few predictive studies within law and economics, particularly in relation to judicial decision making.

In 'The Supreme Court During Crisis', Lee Epstein et al consider the pertinent matter of how judicial decision making is impacted by threats to national security (Epstein et al, 2005). The study employs a large dataset, incorporating every civil rights and liberties case from between 1945 and 2005. The study indicates that during periods of conflict, when national security is threatened the justices are 'substantially more likely' to curtail rights (Epstein et al, 2005).

Though 'paradoxically' the study contends that:

*...in contradiction to virtually every theory of crisis jurisprudence, war appears to affect only cases that are unrelated to the war. For these cases, the effect of war and other international crises is so substantial, persistent, and consistent that it may surprise even those commentators who long have argued that the Court rallies around the flag in times of crisis. On the other hand, we find no evidence that cases most directly related to the war are affected. (Epstein et al, 2005)*

The study offers one uniting postulate:

*Instead of balancing rights and security in high stakes cases directly related to the war, the Justices retreat to ensuring the institutional checks of the democratic branches (Ibid, 2005). Since rights-oriented and process-oriented dimensions seem to operate in different domains and at different times, and often suggest different outcomes, the predictive factors that work for cases unrelated to the war fail for cases related to the war. If this conjecture is correct, federal judges should consider giving less weight to legal principles outside of wartime but established during wartime, and attorneys should see it as their responsibility to distinguish cases along these lines. (Epstein et al, 2005)*

The study is a further example of the benefits of quantitative methods in judicial decision making research. Like much of the research surveyed in this paper, the extant literature employs one of several popular estimation techniques. In the noted studies, logit and probit-based parametric estimation techniques are employed. The present study, as such, is a worthwhile departure from the norm, adopting a perceptron model to robustly identify associations between advocate experience and trial outcomes and develop potentially more accurate predictive models.

Lee Epstein and Jack Knight assert that Justices behave as strategic actors who acknowledge the preferences of other actors, and other actors' potential decisions when making their own decisions

(Epstein and Knight, 1998). The authors describe this as their strategic account of judicial decisions (Epstein and Knight, 1998). Essentially, in achieving their own goals they acknowledge the decision preferences of other rational actors. The postulate comes from rational choice theory and has served as the basis for much research in the social sciences. Most pertinently, the authors assert that 'law — as it is generated by the Supreme Court — is the long-term product of short-term strategic decision making' (Epstein and Knight, 1998).

Paul Collins considers the impact of political lobbying on the decision making of the U.S. Supreme court. The study considered amicus briefings and consequently group influence during the 1946 to 1995 terms, with the results indicating that pressure groups are effective in influencing the Court's outputs. The study findings suggest that judicial decision makers are influenced by persuasive argumentation presented by organized lobby groups and parties.

Calderia and Wright consider the participation of the various parties that contribute to the decision process of the judiciary as Amici curiae within the U.S. Supreme courts, pursuant to petitions for writs of certiorari and jurisdictional statements, in addition to decisions based on merits. The authors employ a large dataset constituted by all cases involving amici since 1982. The authors classify all amici based on the underlying nature of membership, into specific group subsets, such as citizen based, corporations, staff entities (such as law firms), trade/professional associations and other models of membership. The authors contend that the judicial system is remarkably open to amici at all stages of the judicial process.

Collins and Martinek contribute to the judicial decision making literature by considering whether ideology mediates the impact of amicus curiae briefs on the decision making in the U.S. courts of appeals. The authors contend that amicus briefs afford the judges with critical information that influences their voting decisions. Employing conventional econometric methodology, the authors assert that the amicus briefing influences judicial decision making, but the influence is contingent on the judge's ideology. Liberal amicus briefs increase the chances that moderately conservative and conservative judges will cast a liberal vote. Conservative amicus briefs enhance the probability that moderate judges and conservative judges will cast a conservative vote.

The noted studies are broadly representative of the scope of works within the sphere of judicial decision making research, from a methodological perspective. Given the familiarity of the readership in relation to such works and methods of research, it is pertinent to consider the divergent methodologies adopted within this study in greater detail. In the next section, the ANN methodology is introduced, detailing potential applications within empirical legal research.

## STUDY DATA AND METHODOLOGY

The data was procured from McGuire. His seminal study compiled trial experience and outcome data over a 5 year period between 1977 and 1982. This study utilises consolidated court determination data as derived by McGuire, and from the United States Supreme Court Judicial database, as presented in McGuire's seminal study. A total of 916 cases ( $n = 916$ ) were included in the study.

It is postulated herein that the advocacy experience of trial lawyers is positively and significantly associated with favourable judicial outcomes when accounting for varying litigant advantage levels. A further postulate is that favourable amicus curie and litigant advantage factors are positively and significantly associated with favourable judicial outcomes. These postulates are consistent with McGuire (McGuire, 1995). The former study adopted a logistic regression framework to build a simple predictive model. Herein, the predictive model is constructed using an alternative non-parametric estimation technique capable of more accurate predictive outcomes based on neural networks theory. Specifically, a multilayer perceptron network structure based on artificial neural networks theory is adopted, employing backward propagation techniques for model training. This approach lends itself well to the nature and structure of the available data. Given the acknowledged dearth of research considering these matters, the present study seeks to respond to critical omissions in the extant literature

by framing a potentially more robust and accurate predictive model. A precis to the technique noting its benefits over the more popular regression based approaches is provided below.

In interdisciplinary law and economics research, we must frequently contend with situations where we must account for a change in a qualitative factor rather than a quantitative factor. Consistency in the measurement of the existence of litigant advantage factors and judicial experience factors is fundamental to empirical judicial decision research. Similar issues are frequently contended with when conducting empirical research into IPR (Maskus, 2000). Indeed, within interdisciplinary IPR research, measurement and consistency have been the source of much debate (Maskus, 2000). Keith Maskus notes that when measuring IPR impacts, any numerical measure asserting to capture IPRs is generally the subject of criticism (Maskus, 2000).

While interdisciplinary tax analysis research affords exacting and readily measurable variables, the analysis of IPRs is not so straightforward (Maskus, 2000). This is certainly also true of judicial decision research. While the body of research is steadily growing, the relative dearth of research means that there are few consistently adopted measurement norms and practices. The current analysis is potentially afflicted by such issues, because the methodologies employed do seek to account for the magnitude effects through a subjective index or scale variables. This is, however, addressed through the adoption of the methodologies developed by Reginald Sheehan, William Mishler and Donald Songer (Sheehan et al., 1992). The authors framed an approach to determining potential differences in litigant status, and as such potential litigant advantage, using an ordinal scale of 1-10 to classify litigants based on potential or perceived advantage levels (Sheehan et al., 1992). This approach was adopted in the construction of the McGuire database (McGuire, 1995).

Where PS is the dependent variable of petitioner success and LA and AC are the values of litigant advantage and the dummy variable amicus curie, respectively. JE is the critical judicial experience index, accounting for the experience of attorneys in the US Supreme Court system. LI is an ordinal scale variable denoting the lower court's attitude towards matters of liberty. CI is an ordinal scale variable denoting the lower court's attitude towards matters of commerce (see Table 1). While it is self-evident that a number of factors are likely to have influence, aside from those considered herein, the model specifications seek to determine the critical association between litigant advantage, amicus curie and, critically, judicial experience with petitioner success. The manner with which this relationship is explored appears to be the first unique contribution of the current study, extending upon the existing work of McGuire (1995). This study adopts a perceptron model in place of a logistic regression building both a deterministic model, with the intention of understanding the impact of Amici and Lawyer experience on judicial decision making and with the intent of predicting the likelihood of petitioner success, this latter task is where the multilayer perceptron model may afford the most significant benefit. While the employed data is binary built on ordinal scales and as such the expected approach for statistical robustness ceteris paribus is logit regression or multiple discriminant analysis such approaches are relatively less capable of accommodating dynamic and non-

**Table 1. Variable descriptions**

Variable	Variable Description	Variable Type
<i>PS</i>	The success of the petitioner	Binary
<i>LA</i>	The level of advantage of the litigant	Scale
<i>AC</i>	Nature of amicus curie presented to the courts	Scale
<i>JE</i>	The experience level differentials of the trial lawyers	Scale
<i>LI</i>	The attitude of the court to matters of liberty	Scale
<i>CI</i>	The attitude of the court to matters of commerce	Scale

linear relationships and perhaps most pertinently threshold effects relating to independent variables and interaction effects between independent variables.

### **The Multilayer Perceptron Network Model**

The article given both its interdisciplinary nature and focus, introduces a number of technical concepts that are likely to be unfamiliar to non-technical audiences and in particular those with a largely jurisprudential background. In the interest of engaging with the broadest audience, the following precis on neural network methodology is provided. The multilayer perceptron model is simply a form of the artificial neural network model (herewith MP-ANN), a non-parametric predictive model that seeks to replicate the structure of a biological neuron as it occurs in nature. Within a biological neuron, the synaptic signals are received by the soma (cell body) and dendrites and transmitted through the axon. Within an artificial neuron the vector of the variable is passed from one, neuron in the input layer to another neuron with the assignment of synaptic weights (weighted values). Transfer occurs through the assistance of a transfer function, the synthetic axon. More plainly, it is a series of weighted, aggregative, non-linear values. Neural network models emerged out of the body of research on artificial intelligence that should attempt to model the human process of learning through the development of a modelling framework resembling the structure of the human brain (Patterson, 1996). The MP-ANN model has the potential to provide more accurate predictive outcomes than traditional parametric estimation techniques and is therefore well suited to the present research questions.

Rosenblatt (1952) posited the perceptron model, framing the concept of the retina layer with distributing input values (Rosenblatt 1952). The work of Rosenblatt owed much to the initial work of Hebb, and also McCulloch, and Pitts (1943). The work of Marvin Minsky and Seymour Papert asserted a number of the limitations of the perceptron model (Minsky and Papert, 1969), and their work brought about a diminution of interest in perceptron modelling. The key challenge evidenced in the perceptron model as posited by Rosenblatt was the step function approach that made neural network models problematic to train. However, interest in the field was renewed with the advent of techniques to moderate the issues associated with neural network training. A revival of sorts occurred as a result of the work of Rumelhart, Hinton and Williams, whose seminal paper 'Learning Internal Representations by Error Propagation' offered resolution. Their work enabled the effective training for a multilayer neural network. The backpropagation training method and a viable algorithm was first asserted in Rumelhart, Hinton and Williams (Rumelhart et al., 1986) and was the first practical method for training neural networks. Herein, a gradient descent backpropagation training method is used for efficiency. A multilayer perceptron model is selected as an alternative to the radial basis function due to its superiority in estimation. Any potential gains in estimation efficiency associated with the radial basis method are offset by the losses in estimation accuracy.

The MP-ANN is selected as a viable alternative to the McGuire logit (The multi-layer perceptron model shares a level of association with the logistic model, employing a logistic/sigmoid function. A more apt description however would be that the logistic model is a simple single layer perceptron model, that does not employ a second (layer) weighted aggregate input set) model because it offers a more accurate framework for prediction. Consider some of the challenges associated with alternative parametric techniques. Firstly, there are obvious issues associated with predictive analysis where the process is non-linear. Logit and probit models are relatively less capable of modelling non-linear decision processes robustly. MP-ANN overcomes this challenge through the application of weighted, aggregative non-linear values-based estimation. The MP-ANN model appears superior when modelling complex relationships, such as the relationship between judicial experience variables and judicial outcomes. Neural models appear more consistent with judicial decision making (Lothar & Philipps (1981), Hobson & Slee (1994) and Rose & Belew (1989)). Members of the judicial system carefully weigh alternative criteria, balancing competing criteria and implicitly employ these weighted considerations in their decision making. While the process may not be explicit it is

nonetheless essential to judicial decision making. Neural networks are more capable of replicating complex decision criteria and decision weighting than the commonly employed logistic regression.

It is pertinent to acknowledge the notable criticisms of neural network methodology within legal research. Dan Hunter outlines the problems of using statistical sub-symbolic techniques in a symbolic field of research. Hunter, is also particularly critical of the legal realist philosophy that in his view underpins current neural network implementations. Before preceding to the responsorial arguments, it is notable that Hunters criticisms are based largely on what he considers to be flawed realist arguments about the scope and application of networks and not the methodology itself. Hunter suggests that it is likely that we will observe useful applications of the methodology in the future. Notably at the time that Hunter's paper article was authored, computing technologies did not enable the use of complex multi-layer perceptron models, nor was it possible to account for independent variable importance and as such the methods available now are far more useful in accommodating symbolic and sub symbolic languages as well as the open texture within legal research. Herein, I agree with the assertions of Hunter to the extent that legal realists espouse that neural networks can provide normative assessments of legal outcomes and jurisprudence. Neural network methodology does however enable a fuller analysis of judicial outcomes, jurisprudence and the importance of different variables in relation to such determinations.

It is important also to acknowledge that all of Hunter's criticisms are either equally or more valid in their application to traditional regression based methods employed within legal research. Indeed, legal empiricism will always face challenges when authors seek to employ such methods in normative accounts. The great advantage of neural networks derives from their ability to assign importance to variable criteria and to apply a learning framework, something that is decidedly lacking within traditional regression methodologies if we are to view legal reasoning as a parallel process, a view not supported in this paper, out perspective of neural networks is likely to be stronger still. Consider the assertion of Warner (1989), arguably the strongest proponent of this position.

"While our language dictates a sequential description of the (legal reasoning) process, the process is in fact parallel. Many aspects of the problem resolution process are carried out simultaneously. The problem domain is defined by the initial statement of the problem. That initial problem is then resolved into a number of issues... the solution to which will be sought within the problem domain utilizing a sub symbolic paradigm that is not rule based".

Even if the view of Warner does not coincide with our interpretation of legal reasoning as a process, neural networks nonetheless remain vastly superior within legal research incorporating legal reasoning when compared with traditional regression methodology's and specifically for the purpose of clarity in the context of this paper logit models. This is because logit models are incapable of modelling dynamic nonlinear functions and as previously noted accounting for the interaction between variables.

## Model Structure and Parameters

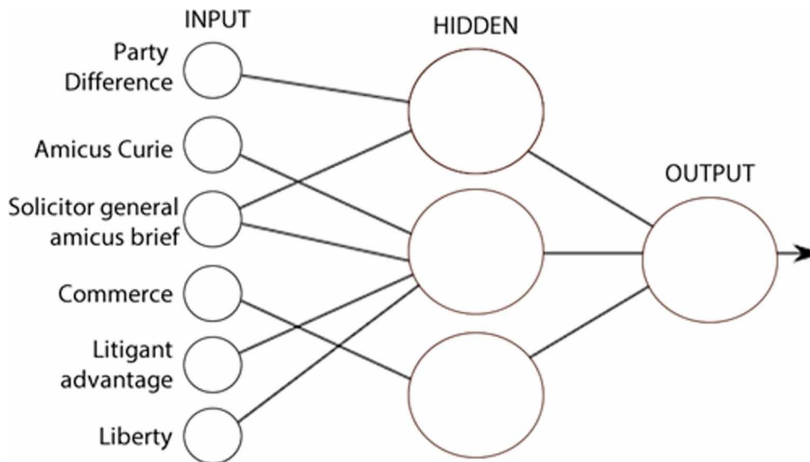
Figure 1 conceptually illustrates the study's posited three-layer perceptron network. This model denotes the layer and node structure of the model. It does not define weights and connections, which are purely illustrative.

The ANN model employed — the MP-ANN model — is a three-layer model with an input layer deriving from the variables provided by McGuire (1995).

Petitioner success is the variable denoting the success of the petitioning party. Litigant Advantage denoted the level of advantage of the litigant. Amicus Curie is a dummy variable pertaining to the nature of the amicus curie position, respectively. The solicitor general amicus curie is the position of the solicitor general in relation to the case. Party difference is the critical judicial experience index, accounting for the experience of attorneys in the US Supreme Court system. These definitions should be provided in the introduction to the article. Liberty is an ordinal scale variable denoting the lower court's attitude towards matters of liberty. Commerce is an ordinal scale variable denoting the lower court's attitude towards matters of commerce.



Figure 1. Conceptual model diagram



## The Layers of a Neural Network

The first layer is the input layer. The input layer processes and standardises a vector of the predictor variables, herein the judicial experience, amici curie and attitudinal variables. These variables then take the range of -1 to 1. These standardised values are then passed through to the hidden layer. This process loosely resembles the process of “saltatory conduction” with a biological neuron.

The standardised vector values of the predictor variables, the judicial experience, amici curie and attitudinal variables, reach the hidden layer. At the hidden layer, their standardised values are multiplied by a weight, with the weighted values aggregated and passed forward to the transfer function. The values from the transfer function are passed to the output layer.

The output layer accepts the values from the hidden layer, multiplying the values by weights, with the weighted values aggregated and passed forward to the transfer function. The values from the transfer function are the model outputs. As the target variable is the outcome for the petitioner, there are 2 neurons in the output layer producing 2 values, one for each of the categories of the target variable.

The hidden layer size was determined based on Jeff Heaton’s work, noting optimality generally exists between the number of input layers and output layers (Heaton, 2011). Moreover, two hidden layers may be employed when modelling data that evidences discontinuities, such as a saw-tooth wave pattern. Given the nature of the underlying dataset, this is not the case. The use of an additional hidden layer reduces estimation efficiency and offers no improvement to the model. Herein, a single hidden layer model is employed to avoid a greater risk of converging to an incorrect local minima. One potential shortcoming of the selected methodology relates to the efficiency of ANN models in accommodating non-numeric variables, commonly referred to as dummy variables in economics. This is, however, only the case where non-numeric variables can take a large number of values, which is not the case in the present study.

Data pruning techniques were not employed to improve model resolution, and the model performed efficiently and robustly absent of any node reduction. As such, the input values remain the same as those employed in McGuire. This is pertinent as it ensures that each predictive modelling framework can be compared on equal terms employing the same parameter set.

## MODEL FINDINGS

The MP-ANN is able to robustly predict the outcome of proceedings 73 per cent of the time. This suggests that the underlying variables are useful in the determination of the outcome of a Supreme

Court deliberation. The model is capable of robustly predicting outcomes more consistently than any of the models in the extant literature. The size of the relative weights employed in the model is detailed below.

The goodness-of-fit measures suggest that the MP-ANN is superior to the logistic model in the McGuire study (McGuire, 1995). While the McGuire model remains conceptually and empirically sound, it does not achieve the same level of predictive accuracy. The findings suggest that this model is more predictively capable, with an improvement in predictive accuracy of 5.7 per cent. This is amongst the most pertinent contributions of the present study. The study conveys the benefits associated with non-parametric methods of prediction in judicial research and the benefit of MP-ANN modelling in the development of predictive models in the judicial sciences.

Employing the methods of Garson (1991) and Goh (1995), the normalised importance analysis suggests that the party difference (the difference in the experience levels of the lawyers representing each party) ranks as the most important determinant of litigant success, followed by the amicus curiae difference position. The data is presented graphically in Table 2. The analysis supports the assertion that judicial experience may be as pertinent, or more pertinent, than the views presented by the ‘friends of the court’. The attitude of the courts in relation to matters of liberty and matters of commerce appear to be largely immaterial to the success of litigants, given the variable importance analysis findings. These findings are broadly consistent with the findings of the McGuire study, with a few notable points of difference (McGuire, 1995).

Firstly, unlike the McGuire study, the current study indicates that lawyer experience differences (denoted by the partydiff variable) are slightly more important to the judicial outcome than amici curie differentials. This finding that is reversed in the McGuire study. This finding is also supported by a substantially higher Pseudo R squared value of 0.323 representing a 114% increase on Pseudo R squared value evidenced within the earlier study, suggesting that the model is able to explain the underlying variation in the data more consistently and to a greater extent than the earlier model (see Figure 2).

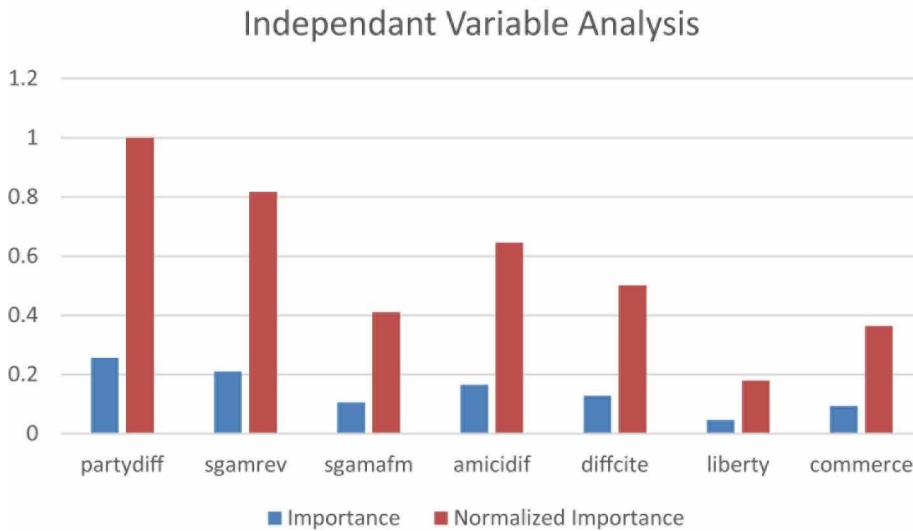
## CONCLUDING REMARKS

The paper has made clear the benefits associated with predictive modelling and specifically non-parametric modelling in legal research. The findings indicate the employed predictive techniques may be more beneficial than logistic models when emphasis is placed on outcome prediction rather than casual association. The study suggests that while logit and probit models are nonetheless useful exploratory tools when engaging in empirical work and can accommodate non-linear problems; they fail to accommodate non-linear problems with complex interactions as efficiently as multilayer perceptron models, at least when the desire is to engage in a predictive exercise. Further research

Table 2. Normalized variable importance

	Importance	Normalized Importance
partydiff	.256	100.0%
sgamrev	.209	81.6%
sgamafm	.105	41.0%
amicidif	.165	64.5%
diffcite	.128	50.1%
liberty	.046	17.9%
commerce	.093	36.3%

Figure 2. Independent variable importance analysis



is nonetheless necessary to determine the potential generalizability of this claim. The notion that perceptron modelling may be beneficial and potentially superior in judicial sciences research represents the first contribution of the study. A further contribution of the study is its response to the dearth of studies employing empirical methods to legal research, and in particular methods beyond logistic regression. While the study was conducted with U.S. data, it appears plausible that the methods employed would be effective when employed with data from other jurisdictions.

It is noteworthy however, that the availability of judicial outcomes data from other jurisdictions makes comparative analysis challenging. This notwithstanding, the present article serves to further advocate for the need of greater judicial data availability. This is particularly true in the Australian context where little judicial data is systematically collected, and disseminated beyond the NSW Department of Justice, data provision, through the Australian government data.gov.au initiative. Moreover, the limited data that is collected within the Australian jurisdiction lacks the granularity evident within the more expansive datasets made available by E.U. authorities and within the United States.

Importantly the current research displays the potential benefits associated with such data collection efforts when employed with modern parametric estimation techniques. The MP-ANN models may be highly effective tools in determining the relative importance of independent variables and the model of estimation to employ when engaging in casual analysis. Moreover, the modelling methodology employed herein evidences the benefit of the presented empirical technique in predicting judicial outcomes. Critically, the role of MP-ANN models in empirical legal research is advanced based on the notion that such models require heuristic knowledge pertaining to data preparation and cataloguing, and sufficient technical knowledge to employ the appropriate ANN for the estimation. The ANN models are more conceptually simple than linear modelling techniques (albeit more mathematically complex); not to mention more appropriate for modelling complex decision models, thus supportive of greater potential adoption by the legal fraternity. It would be apt for future research to consider the potential application of probabilistic neural networks to such empirical legal analysis of judicial experience variables.

## ACKNOWLEDGMENT

The author would like to thank the anonymous reviewers and the editor for their insightful comments and suggestions.

## REFERENCES

- Baker, L. A. (1996). Interdisciplinary Due Diligence: The Case for Common Sense in the Search for the Swing Justice. *Southern California Law Review*, 70, 187–217.
- Billica, N. E. (1997). *Just Leave it to the Courts: How, When, and Why Congress Abdicates the Legislative Power* [PhD Thesis]. Harvard University.
- Enns, K. M. (1998). Note: Can a California Litigant Prevail in an Action for Legal Malpractice Based on an Attorneys Oral Argument before the United States Supreme Court? *Duke Law Journal*, 48(1), 111–146. doi:10.2307/1373009
- Epstein, L., & Knight, J. (1998). *The Choices Justices Make*. CQ Press.
- Epstein, L., Ho, D. E., King, G., & Segal, J. A. (2005). The Supreme Court During Crisis: How War Affects Only Nonwar Cases. *New York University Law Review*, 80, 1-116.
- Garson, G. D. (1991). Interpreting neural network connection weights. *Artificial Intelligence Expert*, 6(4), 46–51.
- Goh, A. T. C. (1995). Back-propagation neural networks for modeling complex systems. *Artificial Intelligence in Engineering*, 9(3), 143–151.
- Heaton, J. (2011). *Introduction to the Math of Neural Networks*. Heaton Research, Inc.
- Hobson, J. B., & Slee, D. (1994). Indexing the Theft Act 1968 for case based reasoning and artificial neural networks. *Proceedings of the 4th National Conference on Law, Computers and AI*, Exeter.
- Karnin, E. D. (1990). A simple procedure for pruning back-propagation trained neural networks. *IEEE Transactions on Neural Networks and Learning Systems*, 1(2), 239–242.
- Martin, A.D., Quinn, K.M., & Epstein, L. (2005). The Median Justice on the U.S. Supreme Court. *North Carolina Law Review*, 83, 1275-1322.
- Martin, A. D., & Quinn, K. M. (2002). Dynamic ideal point estimation via Markov chain Monte Carlo for the US Supreme Court, 1953–1999. *Political Analysis*, 10(2), 134–153. doi:10.1093/pan/10.2.134
- Maskus, K.E. (2000, August 5-6). Intellectual Property Issues for the United States and Japan: Disputes and Common Interests. *Paper presented at Issues and Options for the Multilateral, Regional, and Bilateral Trade Policies of the United States and Japan Conference*, Ann Arbor, MI.
- McAtee, A., & McGuire, K.T. (2007). Lawyers, Justices, and Issue Salience: When and How Do Legal Arguments Affect the U.S. Supreme Court? *Law & Society Review*, 41(2), 259-278.
- McGuire, K.T. (1995). Repeat Players in the Supreme Court: The Role of Experienced Lawyers in Litigation Success. *Journal of Politics*, 57(1), 187-196.
- Minsky, M., & Papert, S. (1969). *Perceptrons: An Introduction to Computational Geometry*. MIT Press.
- Olden, J.D., & Jackson, D.A. (2002). Illuminating the 'black-box': a randomization approach for understanding variable contributions in artificial neural networks. *Ecological Modelling*, 154(1), 135-150.
- Patterson, D. W. (1996). *Artificial neural networks: theory and applications*. Prentice Hall.
- Philipps, L. (1991). Distribution of damages in car accidents through the use of neural networks. *Cardozo Law Review*, 13, 987-1000.
- Rose, D. E., & Belew, R. K. (1989). Legal information retrieval: a hybrid approach. *Proceedings of the International Conference on AI and Law*, Northeastern University, Boston, MA.
- Rose, D. E., & Belew, R. K. (1991). A connectionist and symbolic hybrid for improving legal research. *International Journal of Man-Machine Studies*, 35(1), 1-33.
- Rosenblatt, F. (1958). The perceptron: A probabilistic model for information storage and organization in the brain. *Psychological Review*, 65(6), 386–408. doi:10.1037/h0042519 PMID:13602029

Rumelhart, D.E., Hinton, G.E., & Williams, R.J. (1986). Learning internal representations by error propagation. In D.E. Rumelhart & J.L. McClelland (Eds.), *Parallel distributed processing: explorations in the microstructure of cognition* (Vol. 1, pp. 318-362). MIT Press.

Sankar, A., & Mammone, R.J. (1991). Optimal pruning of neural tree networks for improved generalization. *Proceedings of the Seattle International Joint Conference on Neural Networks* (Vol. 2, pp. 219-224).

Sheehan, R.S., Mishler, W., & Songer, D.R. (1992). Ideology, Status and the Differential Success of Direct Parties Before the Supreme Court. *American Political Science Review*, 86, 464-471.

Songer, D.R., & Sheehan, R.S. (1992). Who Wins on Appeal? Uppercuts and Underdogs in the United States Courts of Appeals. *American Journal of Political Science*, 36(1), 235-258.

Stern, R. H. (Ed.). (2002). *Supreme Court Practice* (8th ed.). Washington, DC: Bureau of National Affairs.

Weigend, A.S., Rumelhart, D.E., & Huberman, B.A. (1990). Back-propagation, weight-elimination and time series prediction. In D. Touretzky, J. Elman, T. Sejnowski et al. (Eds.), *Proceedings of the 1990 Connectionist Models Summer School* (pp. 105 -116). Morgan Kaufman Publishers.

Warner, D.R. (1989). Towards a simple law machine. *Jurimetrics*, 29(4), 451-467.

Warner, D. R. (1990). The role of neural networks in law machine development. *Rutgers Computer & Technology Law Journal*, 16, 129-144.

Warner, D. R. (1992). A neural network based law machine: Initial Steps. *Rutgers Computer & Technology Law Journal*, 18, 51-63.

Warner, D.R. (1993). A neural network-based law machine: The problem of legitimacy. *Law, Computers and Artificial Intelligence*, 2(2), 135-147.

Zeleznikow, J., & Nolan, J. R. (2001). Using soft computing to build real world intelligent decision support systems in uncertain domains. *Decision Support Systems*, 31(2), 263-285. doi:10.1016/S0167-9236(00)00135-4

*Michael D’Rosario is a lecturer and unit chair in post graduate programs at Deakin business school. Prior to working at Deakin business school Michael worked with PriceWaterhouseCoopers, Ford Motor Company, KordaMentha, and the University of Melbourne. Previously, Michael served as a technical advisor to CPA Australia and presently serves as deputy chairman of Stirling Theological College, Melbourne University of Divinity. Michael is an interdisciplinary researcher having published in the domains of law, econometrics, decision sciences, theoretical economics, and information systems.*