

Compliance with International Soft Law: Is the Adoption of Soft Law Predictable?

Michael D’Rosario, Deakin Business School, Deakin University, Sydney, Australia

John Zeleznikow, College of Business, Victoria University, Melbourne, Australia

ABSTRACT

The present article considers the importance of legal system origin in compliance with ‘international soft law,’ or normative provisions contained in non-binding texts. The study considers key economic and governance metrics on national acceptance an implementation of the first Basle accord. Employing a data set of 70 countries, the present study considers the role of market forces and bilateral and multi-lateral pressures on implementation of soft law. There is little known about the role of legal system structure-related variables as factors moderating the implementation of multi-lateral agreements and international soft law, such as the 1988 accord. The present study extends upon research within the extant literature by employing a novel estimation method, a neural network modelling technique, with multi-layer perceptron artificial neural network (MPANN). Consistent with earlier studies, the article identifies a significant and positive effect associated with democratic systems and the implementation of the Basle accord. However, extending upon traditional estimation techniques, the study identifies the significance of savings rates and government effectiveness in determining implementation. Notably, the method is able to achieve a superior goodness of fit and predictive accuracy in determining implementation.

KEYWORDS

Artificial Intelligence, Artificial Neural Network, Basle Accord, Multi-Layer Perceptron, Soft Law

INTRODUCTION

The decision to implement international soft law maybe motivated by a number of different factors as noted by Ho (2002) the implementation of economic soft law, specifically in the case of the Basle Accord of 1998, appears to be driven by the strength of the underlying democratic system of governance in place within a particular jurisdiction affording credence to the democratic legalists theories of international law. While the extant research does appear to support this proposition, there appears to be little, if any, credence afforded to the interaction between democracy and government effectiveness and their impact on the implementation of international soft law. Artificial intelligence methods have been largely absent from discussions of international soft law implementation. This appears problematic and potentially inapt given the plausible interaction effects that exist between government effectiveness democracy and other macro-economic variables in determining the likelihood of soft law implementation.

The concept of soft law, pertains to quasi-legal instruments which have no formal legally binding capability, or where the enforcement capability is assumed to be weaker relatively speaking than the force that underlines traditional law, sometimes in this context referred to as hard law. While

DOI: 10.4018/IJSDS.2018070101

the term soft law is largely synonymous with discussions with international law, it may also be referred to when commenting on domestic legal and doctrinal matters. Herein, our focus remains on international law. An apt example of what may be described as soft law would include resolutions and declarations of the United Nations general assembly components of bilateral and multi-lateral treaties, such as codes of practice and other non-treaty obligations. It is nonetheless problematic identifying a consistent definition of soft law given the fervent debate that exists between those denying that such a legal paradigm exists and those that deem it an additional sphere of international law. Shelton (2000), offers the following guiding definition that soft law is in essence the normative provisions contained in non-binding texts. Arguably, the works of Baxter (1980), and Weil (1983), are seminal within the sphere of research. Baxter (1980), contends that soft law is representative of the different intensity of agreement that exists within the expression of international law. Weil (1983), responds to this assertion and argues for caution when seeking to derive concepts of relative normativity in international law. Shelton (2000), and Abbott & Snidal (2000), offer a unifying set of theses that serve as the conceptual framework for the present study. Specifically, Shelton (2000), outlines compliance with soft law through the analysis of a wide variety of non-binding legal instruments within a variety of domains. Abbott & Snidal (2000), contend that soft law exists by virtue of the weakening of a legal arrangement with respect to the obligations delegation and precision of said arrangement. Having accepted the existence of soft law, it is pertinent to understand what factors motivate the enactment of international soft law.

Guzman & Meyer (2010), offer a worthwhile summation of why soft law is employed by states. The authors advance for complementary explanations for why states employ soft law that seek to describe a much broader range of state behaviour than has been previously explained.

First, and least significantly, states may use soft law to solve straightforward coordination games in which the existence of a focal point is enough to generate compliance.

Second, under what we term the loss avoidance theory, moving from soft law to hard law generates higher sanctions that both deter more violations and, because sanctions in the international system are negative sum, increase the net loss to the parties. States will choose soft law when the marginal costs in terms of the expected loss from violations exceed the marginal benefits in terms of deterred violations.

Third, under the delegation theory, states choose soft law when they are uncertain about whether the rules they adopt today will be desirable tomorrow and when it is advantageous to allow a particular state or group of states to adjust expectations in the event of changed circumstances. Moving from hard law to soft law makes it easier for such states to renounce existing rules or interpretations of rules and drive the evolution of soft law rules in a way that may be more efficient than formal renegotiation.

Fourth, we introduce the concept of international common law (ICL), which we define as a nonbinding gloss that international institutions, such as international tribunals, put on binding legal rules. The theory of ICL is based on the observation that, except occasionally with respect to the facts and parties to the dispute before it, the decisions of international tribunals are nonbinding interpretations of binding legal rules. States grant institutions the authority to make ICL as a way around the requirement that states must consent in order to be bound by legal rules. ICL affects all states subject to the underlying rule, regardless of whether they have consented to the creation of the ICL. As such, ICL provides cooperation minded states with the opportunity to deepen cooperation in exchange for surrendering some measure of control over legal rules.

Ho (2002), offers a worthwhile account of the economic and institutional determinants of the implementation of the first Basle accord employing traditional empirical method. The study of 107 countries explored the importance of factors such as banking sector concentration, democracy, macro-economic conditions and savings rates on the likelihood of implementation of the accord. As noted,

the study identified both a positive and significant effect associated with the nature of the jurisdictions degree of democracy, as measured by the author using an ordinal scale variable.

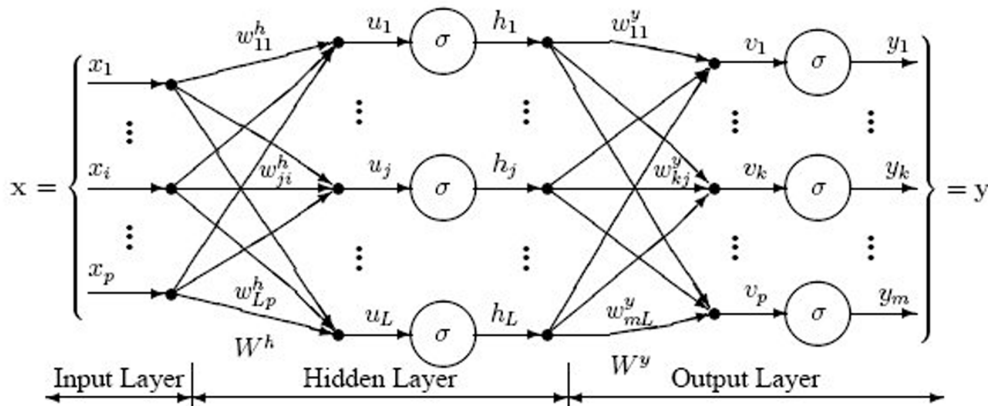
The noted study employs logistic regression modelling techniques. This is consistent with much of the extant research when exploring relationships involving dichotomous dependent variables. The most commonly method employed in empirical legal research is logistic regression, largely because it is well suited to the dichotomous dependent variables that serve as the subject of many studies within the legal sciences. While the contributions to empirical legal research have been apt, many complex non-linear problems have not been considered empirically. The present article proffers an alternative method asserted to be parsimonious for research in the legal domain, representing a viable alternative to the commonly employed logit model.

An Introduction to the Multilayer Perceptron

While there remains the need for further acknowledgement of the benefits associated with parametric estimation methodology within the law and economics, the so called second wave of law and economics (Richardson, 1989), it is apparent that novel non-parametric methodologies made possible by modern computing technologies represent an equally important sphere of inquiry. In particular, the advantages of multi-layer perceptron models in comparison to logit models when modelling complex nonlinear processes are pertinent to consider a number of recent studies have sought to advance the role of artificial intelligence methods within legal research and legal practice. Hanke (2017), considers the role of artificial methods intelligence in response to dispute resolution. The author contends that artificial intelligence applications will augment the capabilities of lawyers and as a consequence should make them more productive and reduce the cost associated with legal advice provision. Nissan (2017), offers a precis on recent advancements within law and A.I. as it relates to the legal profession. The author explains how machine learning technologies and non-machine learning based applications are supplementing and augmenting legal service provision and legal practice. Artificial intelligence methods and the role of such technologies in advancing justice outcomes has been the focus of much recent research (see inter alia Bench-Capon & Modgil, 2017; Branting, 2017; Verheij, 2017; D’Rosario, 2017a, 2017b). This notable interest extends upon the substantial interest in the law and A.I. field of research, deriving from recent advancements in autonomous vehicle technologies and their associated law and artificial intelligence ethics questions. There remains however, only a small number of studies employing artificial intelligence methods when addressing empirical legal questions pertaining to economic soft law. The multilayer perceptron model is simply a form of 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. More plainly it is a series of weighted, aggregative, non-linear values. Neural network models emerged out of the Artificial Intelligence body of research that should to 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, it is therefore well suited to the present research questions (D’Rosario, 2017). Rosenblatt (1958) posited the perceptron model, framing the concept of the retina layer, distributing input values. The work of Minsky & Papert (1969) asserted a number of the limitations of the perceptron model, the 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 that made problematic to train (D’Rosario, 2017). 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 or a multilayer neural network. The backpropagation training method and viable algorithm was first asserted in Rumelhart and McClelland (1986); it was the first practical method for training neural networks (D’Rosario, 2017).

Figure 1 illustrates conceptually the study’s posited three-layer perceptron network.

Figure 1. Three-layer perception network



This visual depiction is of a three-layer perceptron model, though more layers are possible, notwithstanding the significant increase in resources required for multiple hidden layers being employed. The three-layer perceptron is constituted by an input layer, hidden layer and an output layer. The first layer is the input layer. The input layer processes and standardises a vector of the predictor variables (x_1, \dots, x_p). 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, for example standardised vectors of amici curie and/or attitudinal variables, reach the hidden layer. At the hidden layer, the standardised values are multiplied by a weight (w_{ji}), with the weighted values aggregated (u_j) and passed forward to the transfer function. The values from the transfer function are passed to the output layer (h_j).

The output layer accepts the values from the hidden layer, multiplying the values by weights (w_{ki}), with the weighted values then aggregated resulting in a composite value v_j . The v_j value is then passed forward to the transfer function. The values from the transfer function are the model outputs y_k . Should the target variable be dichotomous, then there will be 2 neurons in the output layer producing 2 values, one for each of the categories of the target variable. Where the output value takes greater than to values ($y_k = n$), then the output layer will be constituted by n values.

Neural Models as an Alternative to Logistic Regression in Legal Science Research

The use of such methods is not uncommon in the social sciences but remains relatively underemployed within interdisciplinary legal research. Buchanan & Headrick (1970), offers an early speculative account in regard to the role of artificial intelligence methods within the legal profession asserting the role of artificial intelligence in modelling legal reasoning and argumentation. Rissland (1989), provided an account of the potential role of artificial intelligence methods in modelling legal process. The author provided practical insight into how reasoning could be captured and implemented within an artificial intelligence framework. This study extended upon the earlier work of McCarthy & Haze (1981), who addressed many of the issues associated with subsuming complex logic into an artificial intelligence framework. Bench-Capon (1997), provided a formal explication on the manner with which results-based frameworks and rule based legal information could be incorporated into artificial intelligence models of legal argumentation. The study outlined four areas where legal argument has been applied. Specifically, legal reasoning based on cases, explaining results based on rules of law and rule based legal information in resolution of normative conflict and problems on non-monotonicity

and the basis of dialogue games to support the modelling of legal argument. Walton (2005), offer some refinement to these methods in the authors text, Walton (2005) outlines methods of argumentation and their application in artificial intelligence methods research relating to law incorporating practical and legal profession relevant guidance. However, it is noteworthy to acknowledge that artificial intelligence methods remain underutilised within legal research in comparison with other disciplines.

Importantly, there remains a dearth of research considering key empirical questions, such as the impact of legal experience of lawyers and the role of *amicus curie* on legal outcomes; and an even greater dearth of literature positing practically framed deterministic models of legal outcomes utilising the posited 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. 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.

The first significant practical use of artificial neural networks in law was first developed by Stranieri et al (1999). They used a combination of rules and artificial neural networks and rules to determine what percentage of the common pool of assets a husband would receive following divorce in Australia. The argument structure of Toulmin (1958) was used to justify conclusions made by their Split-Up system. In Stranieri and Zeleznikow (2005) the authors provided a detailed discussion on how knowledge discovery from legal databases could be gainfully used to assist legal decision making.

While some of these assertions may possibly overstate the benefits of neural networks within the legal domain to an extent, 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 Multi-layer perceptron model is asserted to be a viable alternative to the logit 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. MP-ANN overcomes this challenge through the application of weighted, aggregative non-linear values based estimation. Neural networks have been advocated as an alternative modelling technique to logistic regression within the applied sciences, however there is a dearth of research considering the viability of such methods within law and economics research.

Within law and economics research, logistic regression is by far the most commonly employed method, given its suitability to dichotomous dependent variables. Within the applied sciences many more complex alternatives are being adopted. These methods present as viable alternatives to logit and probit modelling.

As asserted by Tu (1996) neural networks offer a number of advantages over logistic regression, including the ability to implicitly detect complex nonlinear relationships between dependent and independent variables, ability to detect all possible interactions between predictor variables, and the availability of multiple training algorithms. Driesietl and Ohno-Machado (2002) offer a formal methodological review of each modelling technique, noting the formal association between the techniques. The multi-layer perceptron model can be seen as an abstraction of the logistic model, employing a 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. ¹ Indeed, there are a number of benefits associated with neural networks that render the approaches advantageous in comparison to traditional logistic estimation. Firstly, the methods appear to evidence greater consistency with legal decision making. If it is assumed that all decision making is linear or sigmoidal, and therein all decisions can be estimated with a linear or sigmoidal function then neural networks and logistic regression offer similar estimates.

Neural models appear more consistent with legal decision making (see Kannai et al 2007, D’Rosario, 2017). Members of the legal 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 legal decision making.

Moreover, the perceptron may be more accommodating of non-linear decision processes. The perceptron model and back propagation method were developed to respond to complex non-linear processes and problems (n.b. Rumelhart, Hinton and Williams). This is because of the existence of the hidden layer that accepts input values from vectors of the predictors. This allows for the combination of vector inputs, rather than the direct pass through of the predictor weights directly to the output/dependent variable. The logistic regression method involves the direct pass through of the predictor weights. This is largely because the logistic regression method can be seen as a simple form of single layer feed forward neural network. It is this assertion that may result in greater consideration being afforded to non-parametric models of estimation.

Neural models have been shown to provide more accurate prediction rates in legal studies such as those of Stranieri et al. (1999), Kannai et al (2007), D’Rosario (2017) and D’Rosario (2017). As such where outcome prediction is assigned the highest priority it is a superior estimation methodology. Where there is complex interaction between predictor variables the neural model is likely to provide a better rate of prediction. Many legal decision processes involve complex interactions between predictor variables. As such these legal decision processes are not well suited to conventional regression based estimation, and in particular linear regression. Perceptron models appear to be well suited to modelling legal decision process because of the formal consistency of the modelling technique with the process of legal decision making.

The method is not without its challenges. The most obvious challenge associated with the use of neural network models are the computing resources required for the employment of such modelling techniques. Multilayer perceptron models are more resource intensive than traditional parametric techniques, though this is of little consequence given the ubiquity of modern computing technology.

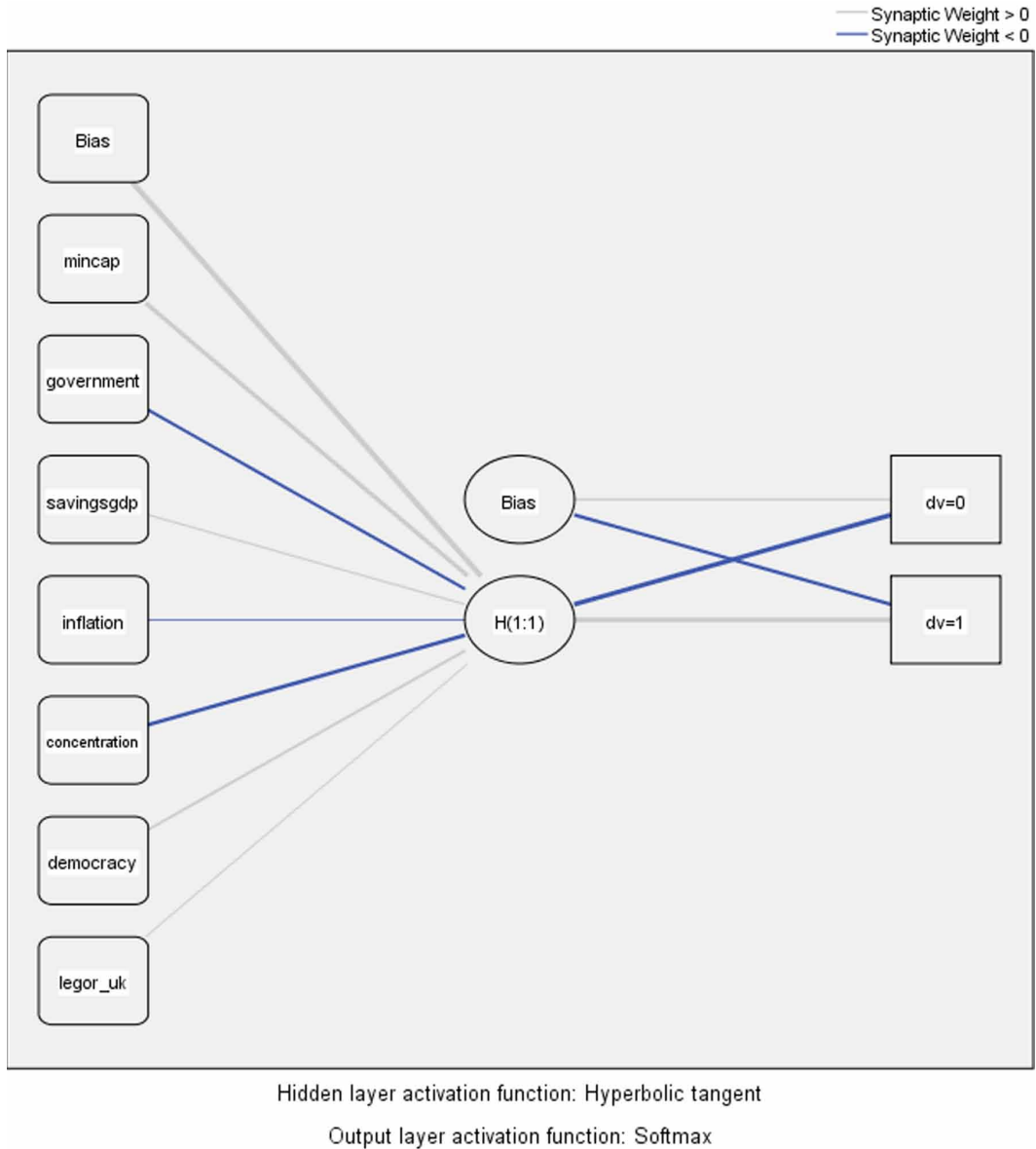
The most significant challenge of neural network-based modelling and machine learning generally is the perception of such methods as black boxes that are difficult to understand. This is a misnomer as there are now many techniques available to understand the relative importance of model variables and model parameter impacts. Given the techniques developed by Garson (1991) and Goh (1995), as well as more contemporary methods, this is certainly no longer an accurate representation of the methodology. Normalised importance analysis allows for the interpretation of neural network connection weights. Thus, providing a set of compelling metrics for the interpretation and analysis of predictor variables, and determinations about the relative importance of predictor variables.

The Logit model is the obvious choice when a dependent variable is non-continuous. It is also useful when addressing non-linear associations. However, the technique fails to properly measure threshold events and specification effects. This results in model specifications incapable of analysing and measuring complex interactions between regressors. A number of non-parametric techniques are more capable of accommodating complex relationships with threshold and interaction effects.

MODEL STRUCTURE AND PARAMETERS

The model seeks to determine the likelihood of soft law (Basle Accord) implementation given a series of macroeconomic, legal and demographic variables. Specifically, the specification models the implementation of the Basel accord based on a number of viable estimators. Figure 2 conceptually

Figure 2. Conceptual model diagram



illustrates the study's posited three-layer perceptron network. The model is a three-layer network where Basle Accord implementation, the dependant variable is a binary variable taking the values of implement or no implementation. (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 Ho (2002). Accord implementation is the variable denoting the implementation of the first Basle accord denoted as a binary dummy variable. Government is a scale variable denoting the effectiveness of government. SavingsGDP is a continuous variable denoting the level of savings relative gross domestic product measured in U.S. dollars (U.S. constant dollars base

year 2000). Inflation denotes the rate of inflation within each jurisdiction in average terms during the preceding year. Concentration is a measure of the level of banking sector concentration. Democracy is a proxy variable quantifying strength of democratic practices within the jurisdiction. Corruption (denoted in the model as *corrupti80-89*, as per Ho (2002)) is a measure of corruption deriving from World Bank corruption statistics. RL is a measure of the rule of law within each jurisdiction. *lngdpcap* is the log value of GDP per capita. *signatory id* is a variable denoting whether the nation is a signatory to the Basle accord. *oppositionul* is a variable denoting division within government. *debtwldperc* is a variable denoting the current level of debt. *lndebt* is the log value of the current level of debt. *lninswld* is a measure of international insurance and financial exposure. *Supervise* is a measure of banking sector supervision (see Table 1).

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 macroeconomic, governance and stability 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, government, *savingsgdp*, inflation, concentration, democracy, corruption, *rl*, *lngdpcap*, *signatory*, *oppositionul*, *debtwldperc*, *lndebt*, *lninswld*, *Supervise*, 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

Table 1. Variable List

Variable	Obs	Mean	Std. Dev.	Min	Max
concentration	95	68.04189	22.77217	12	100
democracy	153	4.241486	3.840977	0	10
corrupti80_89	140	0.085714	0.487037	0	4
RL	155	-0.0008	0.906071	-1.56681	2.129017
lngdpcap	177	7.572594	1.551932	4.61494	10.70118
signatory	217	0.059908	0.237865	0	1
oppositionul	153	0.281261	0.458103	0	1.571429
debtwldperc	140	0.714286	1.696409	0	11.58676
lndebt	135	23.58345	2.310281	18.00163	28.44928
lninswld	155	19.05335	2.258084	13.72721	24.46665
supervise	107	10.25234	2.588412	4	14
transfers	175	-0.28275	4.030983	-32.3727	6.9032
reglnorm	87	0.574702	0.060534	0.4772	0.6667
capacctsm	206	0.083658	0.592008	-3.17104	5.841527
insfinwldsm	160	1.827934	5.58895	0	42.2405
totradesm	161	5.35267	18.14126	0.002116	124.5998

Note: This table summarises the data employed within the study. All data was sourced from Ho, (2001).

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 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. The total number of units employed in the model was 16, and determined through optimisation, again indicating interaction effects.

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 Ho (2002). This is pertinent as it ensures that each predictive modelling framework can be compared on equal terms employing the same parameter set (see Table 2).

MODEL FINDINGS

The MP-ANN is able to robustly predict the decision to implement the Basel Accord 100 per cent of the time (with no converging minima) with accuracy in prediction and goodness of fit far greater than the traditional logistic regression. This suggests that the underlying variables are useful in the determination of the outcome of a soft law implementation process. 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 in Table 3.

The goodness-of-fit measures suggest that the MP-ANN is superior to the logistic model in the Ho study (Ho, 2002). While the Ho 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 a minimum of 14 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 economic soft law research and the benefit of MP-ANN modelling in the development of predictive models in the economics law.

Employing the methods of Garson (1991) and Goh (1995), the normalised importance analysis suggests that the democracy (the proxy of the strength of democratic practices) ranks as the most important determinant of litigant success, followed by the government effectiveness measure and the strength of the rule of law measure. The data is presented graphically in Figure 3. The variation in the hidden layer structure yields comparable results and robustly achieves a 100 percent predictive accuracy. Thought that it is notable that the relative ranked importance of government effectiveness and democracy was reversed in some model specifications, lending credence to the assertion that government effectiveness may be of equal importance to the strength of democratic practices when determining the likelihood of soft law implementation. These findings are broadly consistent with the findings of the Ho study, with a few notable points of difference (Ho, 2002).

Firstly, unlike the Ho study, the current study indicates that government effectiveness and the rule of law may rank as equally important to the strength of democracy in determining the likelihood of soft law implementation. Additionally, the network results suggest that threshold and interaction effects may exist between key independent variables democracy rule of law government effectiveness and savings GDP. Again, noting that substantial improvement in pseudo R-Squared values ranging between 0.65 – 0.81 in contrast of the present study of 0.99 to 1.0.

Table 2. Network Information

Input Layer	Factors	1	government
		2	savingsgdp
		3	inflation
		4	concentration
		5	democracy
		6	corrupti80_89
		7	rl
		8	lngdpcap
		9	signatory
		10	oppositional
		11	debtwldperc
		12	lndebt
		13	lninswld
		14	supervise
	Number of Units ^a		563
Hidden Layer(s)	Number of Hidden Layers		1
	Number of Units in Hidden Layer 1 ^a		16
	Activation Function		Hyperbolic tangent
Output Layer	Dependent Variables	1	Dv
	Number of Units		2
	Activation Function		Softmax
	Error Function		Cross-entropy

a. Excluding the bias unit

Note: This table summarises the structure of the model presented within this study. The model is a multi-layer perceptron employing hyperbolic tangent and soft max activation functions. The dependant variable is the implementation of the Basel Accord.

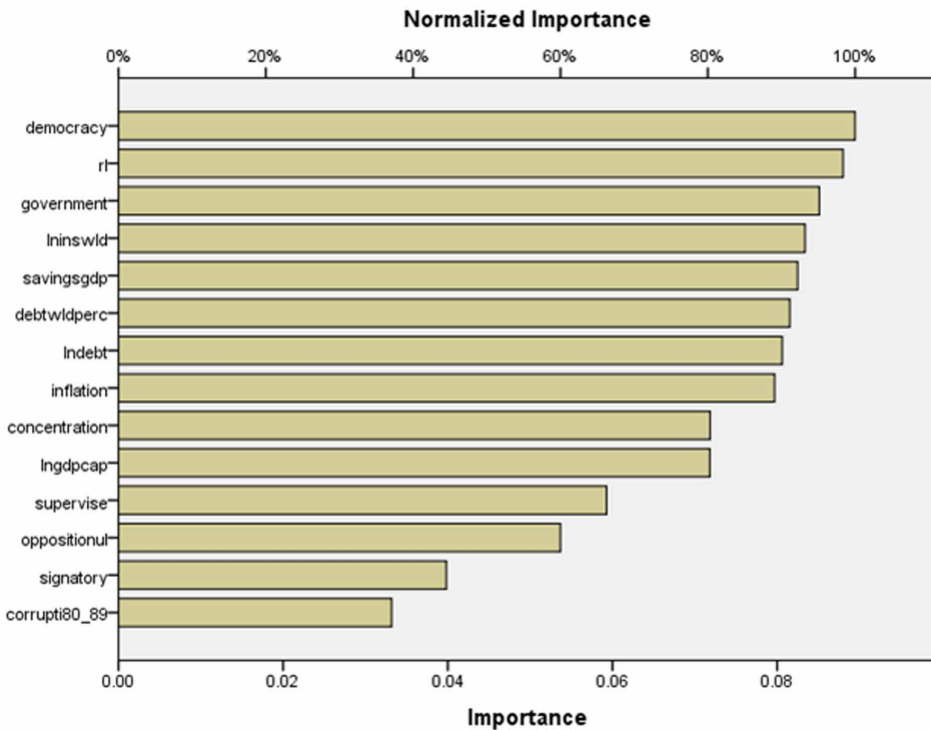
CONCLUSION

The study has evidenced the benefits associated with non-parametric predictive modelling and specifically the benefits of artificial neural networks in legal research. Unlike a number of recent studies employing time series and panel data in non-parametric modelling in legal research, specifically employing artificial neural networks, the present study makes evident the efficacy and efficiency of neural networks in legal research employing cross sectional data. The findings indicate the employed predictive methods may be of greater utility than logistic models when emphasis is placed on outcome prediction rather than casual association, which is particularly important when seeking to determine the likelihood of economic soft-law adoption. The study indicates that while logit and probit models are worthwhile exploratory tools when engaging in empirical work in law and economics and can accommodate non-linear problems; they are less capable of accommodating non-linear problems with complex interactions as efficiently as multilayer perceptron models, at least when the desire is to engage in a predictive exercise. Logit and probit models cannot capture threshold and interaction effects unless such effects are known and understood and can be specified in model prior to estimation. Additional research is nonetheless necessary to determine the potential generalizability of this claim but the claim appears to be highly plausible in light of recent studies.

Table 3. Normalized Variable Importance

Independent Variable Importance		
	Importance	Normalized Importance
Government	.085	95.1%
Savingsgdp	.082	92.2%
Inflation	.080	89.1%
concentration	.072	80.3%
Democracy	.089	100.0%
corrupti80_89	.033	37.1%
Rl	.088	98.4%
Lngdpcap	.072	80.3%
Signatory	.040	44.5%
oppositionul	.054	60.0%
Debtwldperc	.082	91.1%
Lndebt	.081	90.1%
Lninswld	.083	93.2%
Supervise	.059	66.3%

Figure 3. Independent variable importance analysis



The notion that perceptron modelling may be beneficial and potentially superior in soft law and social 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 pertaining to economic soft-law, and in particular methods beyond logistic regression.

Importantly the current research displays the potential benefits associated with legal system and legal structure data collection efforts when employed with modern parametric estimation techniques. The MP-ANN models may be useful tools in determining the relative importance of independent variables and the model of estimation to employ when engaging in casual analysis. The modelling methodology employed herein evidences the benefit of the presented empirical technique in predicting soft law 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. It is hoped that the present study shall inform future legal researchers engaging in empirical legal research employing non-parametric estimation technique such as those employed herein. 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 soft law variables.

ACKNOWLEDGMENT

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

REFERENCES

- Abbott, K. W., & Snidal, D. (2000). Hard and soft law in international governance. *International Organization*, 54(3), 421–456. doi:10.1162/002081800551280
- Ashley, K. D. (2017). *Artificial Intelligence and Legal Analytics: New Tools for Law Practice in the Digital Age*. Cambridge University Press. doi:10.1017/9781316761380
- 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.
- Baxter, R. R. (1980). International Law in “Her Infinite Variety.” *The International and Comparative Law Quarterly*, 29(4), 549–566. doi:10.1093/iclqaj/29.4.549
- Bench-Capon, T., & Modgil, S. (2017). Norms and value based reasoning: Justifying compliance and violation. *Artificial Intelligence and Law*, 25(1), 29–64. doi:10.1007/s10506-017-9194-9
- Bex, F., Prakken, H., van Engers, T., & Verheij, B. (2017). Introduction to the special issue on Artificial Intelligence for Justice (AI4J). *Artificial Intelligence and Law*, 25(1), 1–3. doi:10.1007/s10506-017-9198-5
- Bex, F., Prakken, H., van Engers, T., & Verheij, B. (2017). Introduction to the special issue on Artificial Intelligence for Justice (AI4J). *Artificial Intelligence and Law*, 25(1), 1–3. doi:10.1007/s10506-017-9198-5
- Branting, L. K. (2017). Data-centric and logic-based models for automated legal problem solving. *Artificial Intelligence and Law*, 25(1), 5–27. doi:10.1007/s10506-017-9193-x
- D’Rosario, M. (2017). The Impact of Legal Advocacy Experience Within the US Supreme Court on Trial Decision Outcomes. *International Journal of Strategic Decision Sciences*, 8(1), 65–76. doi:10.4018/IJSDS.2017010105
- D’Rosario, M. (2017). Fair Use Defences During Copyright Litigation: Is the Success of a Fair Use Defence Strategy Predictable? *International Journal of Strategic Decision Sciences*, 8(2), 31–51. doi:10.4018/IJSDS.2017040103
- Enns, K. M. (1998). Can a California Litigant Prevail in an Action for Legal Malpractice Based on an Attorney’s Oral Argument before the United States Supreme Court? *Duke Law Journal*, 48(1), 111–146.
- Epstein, L., Ho, D. E., King, G., & Segal, J. A. (2005). The Supreme Court during crisis: How war affects only non-war cases. *NYUL review*, 80, 1.
- Epstein, L., & Knight, J. (1997). *The choices justices make*. Sage (Atlanta, Ga.).
- Garson, G. D. (1991). Interpreting neural-network connection weights. *AI 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.
- Gressman, E., & Stern, R. (2002). *Supreme Court Practice: For Practice in the Supreme Court of the United States* (8th ed.). Bureau of National Affairs.
- Guzman, A. T., & Meyer, T. L. (2010). International soft law. *The Journal of Legal Analysis*, 2(1), 171–225. doi:10.1093/jla/2.1.171
- Hanke, P. C. (2017). Computers with Law Degrees? The Role of Artificial Intelligence in Transnational Dispute Resolution, and its Implications for the Legal Profession. *Transnational Dispute Management*, 14(2).
- Ho, D. E. (2002). Compliance and International Soft Law: Why Do Countries Implement the Basle Accord? *Journal of International Economic Law*, 5(3), 647–688. doi:10.1093/jiel/5.3.647
- Hunter, D. (1994). Looking for law in all the wrong places: Legal theory and legal neural networks. In *Legal knowledge based systems: The relation with legal theory* (pp. 55-64).
- Hunter, D. (1999). Out of their minds: Legal theory in neural networks. *Artificial Intelligence and Law*, 7(2-3), 129–151.
- Kannai, R., Schild, U., & Zeleznikow, J. (2007). Modeling the evolution of legal discretion – an Artificial Intelligence Approach. *Ratio Juris*, 20(4), 530–558. doi:10.1111/j.1467-9337.2007.00374.x

- Karnin, E. D. (1990). A simple procedure for pruning back-propagation trained neural networks. *IEEE Transactions on Neural Networks*, 1(2), 239–242.
- Makridakis, S. (2017). The forthcoming artificial intelligence (AI) revolution: Its impact on society and firms. *Futures*, 90, 46–60. doi:10.1016/j.futures.2017.03.006
- 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.
- Martin, A. D., Quinn, K. M., & Epstein, L. (2004). The median justice on the united states supreme court. *NCL review*, 83, 1275.
- Maskus, K. (2002). Intellectual Property Issues for the United States and Japan: Disputes and Common Interests. In *Issues and Options for US–Japan Trade Policies* (pp. 165–190).
- McAtee, A., & McGuire, K. T. (2007). Lawyers, justices, and issue salience: When and how do legal arguments affect the US 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. *The Journal of Politics*, 57(1), 187–196.
- Minsky, M., & Papert, S. (1969). *Perceptrons: An Introduction to Computational Geometry*. Cambridge, Mass.: MIT Press.
- Nissan, E. (2017). Digital technologies and artificial intelligence’s present and foreseeable impact on lawyering, judging, policing and law enforcement. *AI & Society*, 32(3), 441–464. doi:10.1007/s00146-015-0596-5
- 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(1-2), 135–150.
- Patterson, D. W. (1998). *Artificial neural networks: theory and applications*. Prentice Hall PTR.
- Prakken, H. (2017). On the problem of making autonomous vehicles conform to traffic law. *Artificial Intelligence and Law*, 25(3), 341–363. doi:10.1007/s10506-017-9210-0
- Rosenblatt, F. (1958). The perceptron: A probabilistic model for information storage and organization in the brain. *Psychological Review*, 65(6), 386.
- Rumelhart, D. E., Hinton, G. E., & Williams, R. J. (1988). Learning Internal Representations by Error Propagation. In *Readings in Cognitive Science* (pp. 399–421). Science Direct.
- Sankar, A., & Mammone, R. J. (1991, July). Optimal pruning of neural tree networks for improved generalization. In *Seattle International Joint Conference on Neural Networks IJCNN-91* (Vol. 2, pp. 219–224). IEEE.
- Sheehan, R. S., Mishler, W., & Songer, D. R. (1992). Ideology, status, and the differential success of direct parties before the Supreme Court. *The American Political Science Review*, 86(2), 464–471.
- Shelton, D. (Ed.). (2000). *Commitment and Compliance: The Role of Non-binding Norms in the International Legal System*. Oxford: Oxford University Press.
- Songer, D. R., & Sheehan, R. S. (1992). Who wins on appeal? Uppercourts and underdogs in the United States Courts of Appeals. *American Journal of Political Science*, 235–258.
- Stranieri, A., & Zeleznikow, J. (2005). *Knowledge Discovery from Legal Databases* (Vol. 69). Springer Science & Business Media.
- Stranieri, A., Zeleznikow, J., Gawler, M., & Lewis, B. (1999). A hybrid—neural approach to the automation of legal reasoning in the discretionary domain of family law in Australia. *Artificial Intelligence and Law*, 7(2-3), 153–183. doi:10.1023/A:1008325826599
- Toulmin, S. (1958). *The Uses of Argument*. Cambridge: Cambridge University Press.
- Verheij, B. (2017). Proof with and without probabilities. *Artificial Intelligence and Law*, 25(1), 127–154. doi:10.1007/s10506-017-9199-4

Weigend, A. S., Rumelhart, D. E., & Huberman, B. A. (1991). Back-propagation, weight-elimination and time series prediction. In *Connectionist models* (pp. 105-116).

Weil, P. (1983). Towards relative normativity in international law? *The American Journal of International Law*, 77(3), 413–442. doi:10.2307/2201073

ENDNOTES

- ¹ For a full methodological review see Dreiseitl, S. & Ohno-Machado, L. (2002). Logistic regression and artificial neural network classification models: a methodology review. *Journal of Biomedical Informatics*, 35(5–6), 352-359. ISSN 1532-0464

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 Price Waterhouse Coopers, Ford Motor Company, Korda Mentha, 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 artificial intelligence, law, econometrics, decision sciences, theoretical economics, and information systems. Michael’s work focuses predominantly on the application of artificial intelligence within the legal domain. Presently, his research is exploring the role of artificial intelligence in predicting judicial outcomes and decisions within the legal domain. Michael is developing hybrid models, employing traditional parametric and contemporary non-parametric techniques.

John Zeleznikow has conducted research and taught in Australian, US, French, Dutch, Israeli, Belgian, German, UK and Polish universities for forty years. He is the author of 3 research monographs, 65 refereed journal articles and over 200 refereed conference articles. John has also won over \$A 7 million in competitive research grants. Over the past twenty years, Professor Zeleznikow has focused on how Artificial Intelligence can be used to enhance decision-making. Specific examples have been created in the domains of law, negotiation and sport. His research findings have been utilised by law and mediation firms, Victoria Legal aid, Relationships Australia Queensland, Victorian Institute of Sport, Australian Institute of Sport and Relationships Australia Victoria. Professor Zeleznikow’s current teaching focuses upon research methods, research evaluation, negotiation, ethics and legal issues related to information technology.