

Designing Optimal Knowledge Base for Neural Expert Systems

Dat-Dao Nguyen¹

¹Department of Accounting & Information Systems, California State University – Northridge, USA

Correspondence: Dat-Dao Nguyen, Department of Accounting & Information Systems, California State University – Northridge, 18111 Nordhoff Street, Northridge, CA 91330-8372, USA.

Received: April 5, 2017

Accepted: May 16, 2017

Online Published: May 18, 2017

doi:10.5539/ibr.v10n6p137

URL: <https://doi.org/10.5539/ibr.v10n6p137>

Abstract

One of the limitations of conventional expert systems and traditional machine induction methods in capturing human expertise is in their requirement of a large pool of structured samples from a multi-criteria decision problem domain. Then the experts may have difficulty in expressing explicitly the rules on how each decision was reached. To overcome these shortcomings, this paper reports on the design of an optimal knowledge base for machine induction with the integration of Artificial Neural Network (ANN) and Expert Systems (ES). In this framework, an orthogonal plan is used to define an optimal set of examples to be taken. Then holistic judgments of experts on these examples will provide a training set for an ANN to serve as an initial knowledge base for the integrated system. Any counter-examples in generalization over new cases will be added to the training set to retrain the network to enlarge its knowledge base.

Keywords: decision rule, expert systems, neural networks, knowledge acquisition, knowledge representation, knowledge-based systems

1. Introduction

Since expertise is scarce and not always available in some domain, Expert Systems (ES) seek to transfer the knowledge on problem solving from human experts to computer-based systems and then on to other human non-experts. To map the expert knowledge into the current state of ES technology, the expert's "know how" / production rule is required to be explained in a detailed manner. However, literature on cognitive psychology points out that most experts cannot express explicitly how the decisions were reached (Nisbett & Wilson, 1977; Ericsson & Simon, 1980). Therefore, the knowledge acquisition for conventional ES has to deal with seemingly incomplete and/or inconsistent information provided by the experts. Then a machine induction technique has been used to deduce new knowledge from exemplary solutions from problem domain provided by the experts. However, most machine induction techniques are hindered by the inadequacy of traditional statistical methods and/or the requirement of a large sample set to represent a problem domain.

To overcome these limitations, this paper reports on the design of an optimal knowledge base for an integrated system of Artificial Neural Network (ANN) and ES. In this framework, an orthogonal plan is used to define an optimal set of examples to be taken from a problem domain. Then holistic judgments of experts on these examples will provide a training set for an ANN. With its ability in pattern recognition and function approximation, the ANN will learn the decision patterns and production rules in the sample set to build the system's initial knowledge base and then generalize this rule base to new cases. Any counter-examples in generalization will be added to the training set to retrain the network in order to enlarge its knowledge base. The knowledge base of the integrated system will grow over time as the system learns new patterns in the same way human beings broaden their knowledge.

The paper is organized as follows. In the next section, limitations of conventional ES and machine induction technique are reviewed. Then the ability of ANN as a machine induction technique is discussed. After that, the design of an initial knowledge base for an integrated system of ANN and ES is presented. It is followed by an illustration on the capacity of such initial knowledge base in an integrated system. The paper concludes its findings and directions for future research.

2. Limitations of Conventional Expert Systems and Machine Induction of Expertise

The goal of knowledge acquisition for ES is to obtain necessary information to represent the knowledge by production rules in the form of cause/effect, situation/action, if-then-else (Chorafas, 1990). But discovering the

expert's mental models could be very difficult. Verbal protocols may be informative but incomplete. The technique may even yield erroneous information since people may state their beliefs and actually believe in one thing, but frequently act in quite a different manner. Furthermore, it has been noted that personal belief structures are not available for inspection (Norman, 1983).

Anderson (1983) proposes that the skill acquisition of human experts goes through three stages. In the *declarative stage*, new information is acquired in a declarative form containing domain-independent facts. Next comes the *knowledge compilation stage* in which the declarative knowledge is transformed into a domain-specific procedural form by recording the conditions under which a piece of declarative knowledge fits into production rules. The final stage is the *procedural stage* in which procedures become more automated. As the declarative knowledge becomes procedural, the ability to verbalize knowledge is lost. The cognitive changes in acquisition of knowledge reduce the ability to report on cognitive processes and increase the inaccessibility of proceduralized task knowledge.

Due to the automatization of cognitive skills and the proceduralization of domain knowledge, there is evidence that people generally have little or no introspective access to higher order cognitive processes. People are rarely able to give retrospective reports on what they were thinking about in solving a problem (Nisbett & Wilson, 1977; Ericsson & Simon, 1980). In fact, there is indication that the more people have about task knowledge the less ability they have to verbalize that knowledge (Berry & Broadbent, 1984).

Numerous literatures have reported on the limitations of human cognition (Best, 1998). Human memory is prone to forgetting and distortion of stored information (Baddeley, 1976). The rationality of human thinking in the formal logical sense can be questioned and some aspects of deductive reasoning prove to be problematic for human beings (Johnson-Laird & Wason, 1977). People have difficulty in handling combinations of uncertain evidence in multi-dimensional decision problems (Kahneman, Slovic & Tversky, 1982).

Within these cognitive limitations of human beings, one faces difficulties in building conventional ES since these systems require structured rule sets, but experts cannot always express explicitly the rules and procedures used in arriving at a solution. As the domain gets larger and more complex, the experts become unable to explain how they operate (Giarratano & Riley, 2004). However, they can supply suitable examples of problems and solutions reflecting their conceptualization of a domain. Hidden behind the holistic assessment of problem alternatives, there always exists implicitly a logic on the association between input-output patterns. The machine induction could be a suitable means to recognize and reveal these patterns.

It has been found that human beings in general do have the ability to make holistic judgments (Slovic & Lichtenstein, 1971), and expertise is based mostly on the recognition of patterns in the problem space (Smith, Adams & Schorr, 1978; Chase & Ericsson, 1982). These characteristics make machine induction a more appropriate technique to discover the decision patterns from a set of holistic assessments provided by experts.

Machine induction offers the possibility of deducing new knowledge (Hassoun, 1995; Haykin, 2009). It can list all the factors that influence the decision, without understanding their impacts, and induce a rule that works successfully (Turban, Sharda & Delen, 2011). The method needs only pre-classified examples and consideration of all samples in the domain. Apparently, one of its major disadvantages is in the requirement of a database containing sufficiently documented cases structured around human knowledge on a problem domain. In addition, the induced rules can be too large or too complex leading to unintelligibility.

Machine induction method has used various algorithms to convert knowledge on related attributes, values and relationships in problem solving, into rules. Such algorithms vary from statistical methods (e.g., decision tree, discriminant analysis, step-wise regression, principal component analysis, factor analysis...) to neural computing (Flores, 2011). Most machine induction techniques rely on traditional statistical methods. It has been pointed out that these traditional methods cannot handle the nonlinear relationship without imposing strong assumptions on behavior of the data.

3. Machine Induction of Expertise with Integrated ANN and ES

To overcome the inadequacy of traditional machine techniques in handling the nonlinear relationship and to avoid imposing *a priori* restrictions on the data, Artificial Neural Networks – ANN could be an appropriate technique in machine learning and ES building. This technique is particularly useful in recognizing the input-output patterns in heuristics, which cannot be expressed explicitly by the experts. An ANN does not require *a priori* elaborate models or pre-specified probability distribution function of the data in order to learn the underlying discriminant patterns and association.

An ANN contains processing/computing units called neurons (or nodes). These nodes are arranged into layers, in

which a node in one layer has a weighted connection to each node of the next layer in a particular configuration. A node, as a processing unit, receives inputs from other nodes or from an external stimulus. A weighted sum of these inputs constitutes the argument to an activation or transfer function.

Most applications have used 3-layer networks consisting of one input, one hidden and one output layer. The hidden nodes are needed to introduce nonlinearity into the network. In some cases, more hidden layers are necessary to approximate a higher order function. An input node provides an external signal to the network. An output node produces an output of the network as a whole. A hidden node that is necessary for the computation of complex functions. Node inputs and activations can be discrete, taking on values $\{0, 1\}$ or $\{-1, 0, 1\}$, or be continuous, taking on values in the interval $[0,1]$ or $[-1,1]$. Each node u_i computes a single numerical node output or activation. Output of a node can be the output of the network as a whole and/or it can be the input to other nodes. Every node, other than input nodes, computes its new activation u_i as a function of the weighted sum of inputs directed to it from other nodes:

$$S_i = \sum_{j=0}^n w_{i,j} u_j \quad (1)$$

$$u_i = f(S_i) \quad (2)$$

where the activation function, $f(\cdot)$, is usually a nonlinear, bounded and piecewise differentiable function such as the sigmoid function,

$$f(x) = 1/(1 + e^{-x}) \quad (3)$$

Such an ANN produces a response, which is the superposition of n sigmoid functions, where n is the number of hidden nodes, to map a complex function. As one adds more hidden layers, ANN will be able to map higher order functions. Therefore, the function mapping with ANN is more general than the regression of traditional methods.

In theory, an ANN can learn from past data and generalize over new cases. With its ability as a universal function approximator, an ANN can discover the associations between groups, elements of the domain space and those of the problem space. It has been shown that an ANN can approximate a functional relationship using a sample of input-output patterns, and learn probabilities and statistical distributions from the data (Cybenko, 1989; Hornik, Stinchcombe & White, 1989).

In practice, an ANN can offer the advantage of computer execution speed. The ability to learn from cases and train the system with data rather than to write programs may be more cost effective and even more convenient when frequent updates are needed (Medsker, 1995). In applications where rules are unknown, an ANN may be able to represent those rules implicitly in its stored connection weights.

However, previous integration of ANN and ES (Medsker, 1995) is hindered by the availability of examples and the cognitive effort of experts in providing judgment on these examples. The cognitive limitation of human experts made the assessment of a large sample impractical if not erroneous. Most importantly, these systems did not consider the trade-off in multi-criteria decision making which transcends all business problems. The representation of trade-offs in an appropriate rule set for conventional ES is extremely difficult. Taking into account the tradeoffs in their assessment, experts might provide judgments that seem to be contradictory and/or inconsistent with the formal logic. These limitations could be overcome by building an optimal training set for the ANN containing holistic assessments of experts on a set of examples defined by an orthogonal plan. Knowledge acquired from this optimal training set will serve as the initial knowledge base for an integrated system of ANN and ES.

4. Orthogonal Plan as Initial Knowledge Base for Integrated ANN and ES

To alleviate the burden of information processing on human experts in knowledge engineering, one can implement an orthogonal plan to define the minimum size of sample to be assessed. In the experimental design and analysis of variance, an orthogonal main-effect plan (Addelman, 1962) permits the study of several factor effects without going into every combination of factor levels. A factorial experiment is called *symmetrical* when each of its factors has the same number of levels. The experiment with factors having different number of levels is called *asymmetrical* factorial.

Consider a symmetrical factorial experiment involving $(s^n - 1)/(s - 1)$ factors, each of which has s levels with s^n treatment combinations, where s is a prime or a power of a prime number. These $(s^n - 1)/(s - 1)$ factors can be presented by n factors, each having s levels and their generalized interactions. Therefore, the treatment combinations of a main-effect plan for $(s^n - 1)/(s - 1)$ factors in s^n trials may be obtained by choosing the

treatment combinations of a complete s^n factorial plan and generating the remaining factors by the interactions of the s^n experiment (Addelman, 1962). For an asymmetrical factorial experiment, the orthogonal main-effect plan can be constructed by collapsing factors occurring at s_1 levels to factors occurring at s_i levels by using a many-to-one correspondence of the set of s_1 levels to the set of s_i levels.

It has been estimated that, by using an orthogonal design, a set of 49 holistic assessments of production rules can cover the dimensionality of decision problems having up to eight criteria, with seven levels per criterion (Barron & Person, 1979). For example, in a four-criterion problem with five levels per criterion, instead of making 4^5 or 1024 assessments to define a set of production rules for a conventional ES, with an orthogonal plan one needs only 25 assessments to capture the main effects of problem factors. Orthogonal plans not only help to reduce the information processing burden on the experts, but also serve to define initial knowledge bases for integrated systems of ANN and ES.

This paper investigates the implementation of an optimal knowledge base designed by an orthogonal plan in an integrated system of ANN and ES. Such integration will help to overcome the difficulties often encountered by the conventional ES technology. Within the proposed framework, one starts with an appropriate orthogonal plan to define a set of basic examples from the problem domain. The holistic judgments of an expert on these basic examples constitute a training set for the ANN. The holistic assessments intend to overcome difficulty of the expert in explaining explicitly production rules of his/her heuristics. Once the decision patterns of this set are learned, the network acquires an initial knowledge base on the problem domain.

After being trained, the network will be used to generalize over new cases. The expert will then be asked to evaluate the network performance on these cases. Any wrong prediction will constitute a counter-example to be added to the training set. Consequently, the network will be retrained to learn new decision patterns emerging from the addition of counter-examples to the training set. One notes that the integrated system can be put in production, even if it has learned only partial information on the problem domain. Since the system is trained with basic patterns of problem domain, it can generalize well over new cases.

Table 1. Knowledge Acquisition Process of the Integrated ANN and ES

<ol style="list-style-type: none"> 1- Select an appropriate orthogonal plan to define the optimal training set. 2- Acquire holistic judgments of expert on the training examples to build the initial knowledge base. 3- Train the network to learn the decision patterns from the initial knowledge base. 4- Generalize the trained network over new examples. 5- On the opinion of the expert, add counter-examples to the training set. 6- Retrain the network to acquire a new knowledge base.
--

The process of “learning by doing” continue as needed to acquire a robust knowledge base for the integrated system. Over time, the system will enrich its knowledge base with additional decision patterns as it learns from its own production and the expert’s opinion. This knowledge acquisition process is intuitive since it is similar to the one used by human beings. The knowledge base of the integrated system will grow over time as the system learns new patterns in the same way human beings broaden their knowledge.

5. An Illustration: Project Evaluation and Economic Appraisal

The following illustrates the implementation of an optimal knowledge base in an integrated system of ANN and ES. The problem relates to project evaluation and economic appraisal of proposals for new products in a manufacturing company. The proposal is evaluated on five criteria: (i) Net Present Value (NPV) of cash flows generated over the next five years, (ii) initial capital investment requirement, (iii) market growth rate in the next five years, (iv) capability to market the new product, and (v) prospect of technical success. Before submitting any proposal, the related department estimates the values of each project criterion. Some criteria are quantitative, others are qualitative. In the appraisal process, the quantitative data are converted to categories to avoid the difficulty in dealing with data on a continuous scale. For example, the flow of NPV is classified into five categories ranging from one to five million dollars. Similarly, the amount of initial investment is classified into five categories ranging from half a million to two and half million dollars. In the appraisal, there exist trade-off among criteria and no single criterion absolutely dominates any others. The experts will rate each proposal on a scale ranging from 0 to 100. In this study, the appraisal was taken by a senior manager, a senior engineer and a chartered accountant indicated as Expert A, Expert B, and Expert C respectively.

Table 2. Project Evaluation and Economic Appraisal: The Problem Domain

NPV of Cash Flow	[\$1.0	2.0	3.0	4.0	5.0] millions
Initial Investment	[\$2.5	2.0	1.5	1.0	0.5] millions
Market Growth Rate	[fair	good	very-good]		
Capability to Market	[fair	good	very-good]		
Prospect of Technical Success	[fair	good	very-good]		

For this problem domain, a complete set of 864 production rules would be needed to build a knowledge base for a conventional ES and make such ES functional. To alleviate this cognitive burden in this study, an orthogonal plan was used to define a sample of only 24 examples of production rules from the problem domain to be assessed. To validate the performance of the integrated system, two test sets, each containing five out-of-sample examples, are set up. Each expert was asked to provide his/her overall preference judgment for each example of the training set and test sets. For instance, the preference of Expert A for the first training example (1 2 2 3 3 50) in Appendix 1 is interpreted in the following production rule.

If NPV is 1 million dollars,
 AND Initial Investment is 2 million dollars,
 AND Market Growth Rate is good,
 AND Capability to Market is very good,
 AND Prospect of Technical Success is very good,
 then the Proposal score is 50/100.

In this manner, the sample constitutes a set of production rules implicitly expressed by the expert. Using this set, one trains an ANN to acquire an initial knowledge base for the integrated system. To test the performance of the system, one uses the acquired knowledge base to generalize over two test sets. The expert is asked to evaluate the network prediction of his/her preference on the first test set. If there is any disagreement between a network prediction and an expert judgment, this counter-example will be added to initial training set. Then the network is retrained to learn the new decision pattern and make generalization on the second set. The knowledge base of the system will grow over time with the learning of new patterns emerging from the counter-examples.

In this study, an ANN is configured with five input nodes, one for each criterion, five hidden nodes and one output node. This network uses a backpropagation algorithm with sigmoid transfer function, a learning rate of 1 and a momentum of .9. The training and testing tolerance are set at .1 and .3 respectively.

For Expert A, the network converged after 119 training epochs with a Root Mean Square of Errors (RMSE) of .0304. In generalization over the two test sets, it made RMSE of .0170 and .02275 respectively. For Expert B, the network converged after 104 training epochs with an RMSE of .0333. In generalization over the two test sets, it made RMSE of .0381 and .0473 respectively. In these cases, the experts judged that the system had provided satisfactory generalization on both test sets.

The case of Expert C is illustrated in Appendix 2. In this case, the network converged after 2425 training epochs, with an RMSE of .0334. On the first test set, it made two wrong predictions with an estimate of 81.70 for the example having a preference score of 70 and an estimate of 79.42 for the example having a preference score of 65. These two counter-examples were added to the training set and the network was retrained to learn new patterns in these examples. In the second run, the network converged at 1769 epochs with an RMSE of .0369. On the second test set, it made only one wrong prediction with an estimate of 93.34 for the example having a preference score of 60. The network training for this case takes longer time to learn the decision patterns and production rules of Expert C. However, the system's knowledge base and predictability for preference of Expert C will increase over time as it learns more about the decision patterns of this expert.

6. Conclusion

The limitation of conventional ES and traditional machine induction techniques in knowledge engineering is in their requirement of explicit production rules from an expert on a large structured sample taken from a problem domain. This study has shown that an integrated system of ANN and ES can perform well with an initial knowledge base designed with the orthogonal plan to capture expertise in a problem domain. This optimal

sample contains only a subset of all possible production rules in a problem domain. Consequently, much less cognitive effort is required from the experts since they make judgments only on a set of fewer examples. In addition, holistic assessment does help experts to overcome the difficulty in expressing explicitly the production rules of their heuristics. Any counter-examples in generalization will be added to retrain the network in order to enlarge its knowledge base. This approach is particularly attractive, as the knowledge base of the integrated system will grow over time as the system learns new patterns in the same way human beings broaden their knowledge.

To acquire a robust knowledge base for the integrated ANN-ES system, one may consider aggregating judgments of multiple experts in a composite training set or in a composite individual ANNs. The result from this aggregation may provide a starting point for negotiation in a group decision support system. In a future work, we shall report on the implementation of such system and show that once the system aggregates group preference from individual judgments, it will provide a consensus solution for the group in a decision problem.

References

- Addelman, S. (1962). Orthogonal Main-Effect Plans for Asymmetrical Factorial Experiments. *Technometrics*, 4(1), 21-46. <https://doi.org/10.1080/00401706.1962.10489985>
- Anderson, J. R. (1983). *The Architecture of Cognition*. Cambridge, MA: Harvard University Press.
- Baddeley, A. D. (1976). *The Psychology of Memory*. New York: Harper and Row.
- Barron, F. H., & Person, H. B. (1979). Assessment of Multiplicative Utility Functions via Holistic Judgments. *Organizational Behavior and Human Preference*, 24, 147-166. [https://doi.org/10.1016/0030-5073\(79\)90023-0](https://doi.org/10.1016/0030-5073(79)90023-0)
- Berry, D. C., & Broadbent, D. E. (1984). On The Relationship Between Task Performance and Associated Verbalizable Knowledge. *Quarterly Journal of Experimental Psychology*, 36A, 209-231. <https://doi.org/10.1080/14640748408402156>
- Best, J. (1998). *Cognitive Psychology (5th Edition)*. New York: Wadsworth Pub Co.
- Chase, W. G., & Ericsson, K. A. (1982). Skill and Working Memory. In G.H. Bower (Ed.). *The Psychology of Learning and Motivation*, 16, New York: Academic Press. [https://doi.org/10.1016/S0079-7421\(08\)60546-0](https://doi.org/10.1016/S0079-7421(08)60546-0)
- Chorafas, D. (1990). *Knowledge Engineering*. New York: Van Nostrand Reinhold.
- Cybenko, G. (1989). Approximation by Superpositions of a Sigmoid Function. *Mathematics of Control, Signals, and Systems*, 2, 303-314. <https://doi.org/10.1007/BF02551274>
- Ericsson, K., & Simon, H. A. (1980). Verbal Reports as Data. *Psychological Review*, 87, 215-251. <https://doi.org/10.1037/0033-295X.87.3.215>
- Flores, J.A. (ed.) (2011). *Focus on Artificial Neural Networks*. New York: Nova Science Publishers.
- Giarratano, J. C., & Riley, G. D. (2004). *Expert Systems: Principles and Programming (4th Edition)*. London: Thompson Learning.
- Hassoun, M. H. (1995). *Fundamental of Artificial Neural Networks*. Cambridge, MA: The MIT Press.
- Haykin, S. (2009). *Neural Networks and Machine Learning (3rd. Edition)*. New Jersey: Pearson
- Hornik, K., Stinchcombe, M., & White, H. (1989). Multi-Layer Feedforward Networks Are Universal Approximators. *Neural Networks*, 2, 359-366. [https://doi.org/10.1016/0893-6080\(89\)90020-8](https://doi.org/10.1016/0893-6080(89)90020-8)
- Johnson-Laird, P. N., & Wason, P. C. (Eds.) (1977). *Thinking: Readings in Cognitive Science*. Cambridge: Cambridge University Press.
- Kahneman, D., Slovic, P., & Tversky, A. (Eds.) (1982). *Judgment Under Uncertainty: Heuristics and Biases*. Cambridge: Cambridge University Press. <https://doi.org/10.1017/CBO9780511809477>
- Medsker, L. R. (1995). *Hybrid Intelligent Systems*. Boston: Kluwer Academic Publishers. <https://doi.org/10.1007/978-1-4615-2353-6>
- Nisbett, R. E., & Wilson, T. D. (1977). Telling More than We Can Know: Verbal Report on Mental Processes. *Psychological Review*, 84, 231-259. <https://doi.org/10.1037/0033-295X.84.3.231>
- Norman, D. A. (1983). Some Observations on Mental Models. In D. Gentner & A.S. Stevens (Eds.), *Mental Models*. Hillsdale, NJ: Erlbaum.
- Slovic, P., & Lichtenstein, S. (1971). Comparison of Bayesian and Regression Approaches to the Study on

Information Processing in Judgment. *Organizational Behavior and Human Performance*, 6, 649-774. [https://doi.org/10.1016/0030-5073\(71\)90033-X](https://doi.org/10.1016/0030-5073(71)90033-X)

Smith, E. E., Adams, N., & Schorr, D. (1978). Fact Retrieval and the Paradox of Interference. *Cognitive Psychology*, 10, 438-464. [https://doi.org/10.1016/0010-0285\(78\)90007-5](https://doi.org/10.1016/0010-0285(78)90007-5)

Turban, E., Sharda, M., & Delen, D. (2011). *Decision Support and Business Intelligent Systems (9th Edition)*. Boston: Prentice Hall.

Appendix 1. Results in Learning and Predicting Preference of Experts A and B

Training Set	Expert A		Expert B	
	Score	Prediction	Score	Prediction
1 2 2 3 3	50	54.63	45	47.64
1 3 3 1 2	40	39.50	46	48.07
1 4 3 2 1	45	44.90	54	53.86
1 5 1 3 3	60	60.49	65	62.25
2 1 2 2 2	35	33.43	42	44.89
2 2 3 3 1	50	49.79	52	51.02
2 3 3 1 3	55	55.08	57	56.56
2 4 1 3 1	40	38.26	58	57.88
2 5 1 1 3	45	44.46	63	64.50
3 1 3 3 3	70	69.89	58	56.85
3 2 3 1 1	35	34.78	52	51.85
3 3 1 2 3	50	49.48	62	61.35
3 4 1 3 2	55	55.81	69	69.81
3 5 2 1 1	40	40.39	70	70.78
4 1 3 3 3	75	73.78	65	65.46
4 2 1 1 2	30	29.72	55	53.62
4 3 1 3 1	45	43.61	66	66.33
4 4 2 1 3	60	60.88	76	76.05
4 5 3 2 1	65	68.24	83	82.35
5 1 1 1 1	20	24.21	52	49.95
5 2 1 2 3	55	54.31	68	69.03
5 3 2 3 1	60	61.66	75	77.07
5 4 3 1 3	75	73.99	85	82.91
5 5 3 3 2	90	82.35	93	87.42
<u>Test Set 1</u>				
3 3 1 2 1	30	29.83	55	54.75
5 1 3 1 1	40	38.39	58	58.26
1 5 2 3 3	70	69.51	68	65.09
4 4 2 2 2	60	62.35	76	77.24
1 4 3 3 2	65	65.45	62	58.44
<u>Test Set 2</u>				
3 1 2 2 2	40	38.16	48	48.65
3 4 1 2 2	45	44.49	63	66.28
1 4 3 3 3	75	72.76	65	60.48
2 2 2 2 3	50	49.44	52	50.61
4 1 1 2 3	45	42.47	55	53.71

Appendix 2. Results in Learning and Predicting Preference of Expert C

First Run			Second Run		
<u>Training Set</u>	<u>Score</u>	<u>Prediction</u>	<u>Training Set</u>	<u>Score</u>	<u>Prediction</u>
1 2 2 3 3	30	26.95729	1 2 2 3 3	30	27.86282
1 3 3 1 2	40	38.94524	1 3 3 1 2	40	37.15381
1 4 3 2 1	60	59.76158	1 4 3 2 1	60	60.95514
1 5 1 3 3	50	50.36586	1 5 1 3 3	50	50.32877
2 1 2 2 2	30	32.01082	2 1 2 2 2	30	34.2081
2 2 3 3 1	60	60.53619	2 2 3 3 1	60	60.18925
2 3 3 1 3	50	51.25394	2 3 3 1 3	50	52.09837
2 4 1 3 1	60	59.45174	2 4 1 3 1	60	59.10262
2 5 1 1 3	70	69.31223	2 5 1 1 3	70	69.52824
3 1 3 3 3	95	87.93129	3 1 3 3 3	95	88.05566
3 2 3 1 1	95	94.61041	3 2 3 1 1	95	92.54405
3 3 1 2 3	80	81.79112	3 3 1 2 3	80	80.6783
3 4 1 3 2	90	89.84055	3 4 1 3 2	90	92.66843
3 5 2 1 1	90	96.23819	3 5 2 1 1	90	96.90807
4 1 3 3 3	80	84.30698	4 1 3 3 3	80	83.20506
4 2 1 1 2	70	68.95001	4 2 1 1 2	70	67.50552
4 3 1 3 1	70	70.09775	4 3 1 3 1	70	69.33841
4 4 2 1 3	90	88.72336	4 4 2 1 3	90	88.78882
4 5 3 2 1	95	96.38002	4 5 3 2 1	95	97.11317
5 1 1 1 1	60	60.36818	5 1 1 1 1	60	60.69984
5 2 1 2 3	60	59.5041	5 2 1 2 3	60	61.01624
5 3 2 3 1	80	79.36691	5 3 2 3 1	80	79.62003
5 4 3 1 3	90	89.9671	5 4 3 1 3	90	90.27477
5 5 3 3 2	95	95.8607	5 5 3 3 2	95	97.06953
			3 3 1 2 1	70	70.63234
			5 1 3 1 1	65	65.18605
<u>Test Set 1</u>			<u>Test Set 2</u>		
3 3 1 2 1	70	81.70384	3 1 2 2 2	50	63.24843
5 1 3 1 1	65	79.42801	3 4 1 2 2	60	93.34921
1 5 2 3 3	70	63.60191	1 4 3 3 3	60	59.30336
4 4 2 2 2	95	94.17838	2 2 2 2 3	40	49.93819
1 4 3 3 2	60	59.23135	4 1 1 2 3	60	55.38448

Copyrights

Copyright for this article is retained by the author(s), with first publication rights granted to the journal.

This is an open-access article distributed under the terms and conditions of the Creative Commons Attribution license (<http://creativecommons.org/licenses/by/4.0/>).