

Comparison and Evaluation of Intelligence Methods for Distance Education Platform

Maysam Hedayati

Islamic Azad University-Ghaemshahr Branch, Ghaemshahr, Iran
Hedayati_maysam@yahoo.com

Seyed Hossein Kamali

Islamic Azad University of Qazvin Branch, Qazvin, Iran
prjkamali@live.com

Reza Shakerian

Payame Noor University, PO BOX 19395-3697, Tehran, Iran
r.shakerian@gmail.com

Abstract— In this paper two favorite artificial intelligence methods: ANN and SVM are proposed as a means to achieve accurate question level diagnosis, intelligent question classification and updates of the question model in intelligent learning environments such as E-Learning or distance education platforms. This paper reports the investigation of the effectiveness and performances of two favorite artificial intelligence methods: ANN and SVM within a web-based environment (E-Learning) in the testing part of an undergraduate course that is “History of Human Civilizations” to observe their question classification abilities depending on the item responses of students, item difficulties of questions and question levels that are determined by putting the item difficulties to Gaussian Normal Curve.

The effective nesses of ANN and SVM methods were evaluated by comparing the performances and class correct nesses of the sample questions using the same 3 inputs as: item responses, item difficulties, question levels to 5018 rows of data that are the item responses of students in a test composed of 13 questions. The comparative test performance analysis conducted using the classification correctness revealed yielded better performances than the Artificial Neural Network (ANN) and Support Vector Machine (SVM).

Index Terms— E-Learning, Intelligence Methods, ANN, SVM, Comparison

I. INTRODUCTION

By the development of technology, in today's world; e-learning gained a great deal of importance in education. Internet is being used almost in anywhere in life, Education as well. Since education is consist of instruction and learning, instructors are trying to reveal new strategies and learning methods to enhance learning in learners' side. By constructive approach of teaching, that is education has to be learner centered and learning

occurs in a cognitive manner in learners' mind by means of past experiences gained and active learning “learning by doing in nature.” Each student has its own learning style, and this yields a result that each student's performance in learning can not be assessed and evaluated in a unique and simple way. Therefore, teachers can have the best technology, they can do the best applications, best schools, best facilities, but they can not transfer their information to a learner explicitly even if they use best technology because learning occurs in learner's mind with cognitive steps. As a consequence of this fact, instructional technologies must be developed by considering this regard. If it is conducted with E-Learning project that is the research issue of this paper, a student's performance can not be evaluated and assessed in a web based application only by measuring the test results depending on the number of right and wrong answers. Hence, testing must be intelligent. It must behave as if a teacher asks questions to a student in real class environment. If student can not answer the question, teacher must ask easier question about similar subject and if student answer this time, then, again more difficult question must be asked. In [1] pointed out that one of the advantages of computer assisted assessment (CAA) is “the networked nature of approach, which provides for distribution of formative or summative assessment directly to multiple client computers with little or no additional hardware or software installation needed apart from standard access to the Web.” This is the behavior of a real teacher in class and hence system must simulate this real environment in regarding E-Learning platform. The regard is the testing part and the concern is to develop an intelligent testing application that will produce intelligent questions depending on the student's responses and performance during testing session. This kind of testing is called as Computer Adaptive Testing (CAT).

In this regard, the question levels must be determined somehow. In research project, item

difficulties of questions were estimated using item responses (Answers of students to each question) and then they were put to Gaussian Normal Curve in order to grade their level of difficulty. Initial levels were found in order to teach these levels to computer software at first time only, and then make software run classification automatically after learning. In order to train the software for classification; ANN and SVM methodologies were used and found the nearest class values.

II. RELATED WORK

At the past, many scientists made researches on the topic of question classification by artificial intelligence methods. List of literature is listed below as follows:

A. Question Classification for E-Learning by Artificial Neural Network

In [2], they made a research on question classification by using artificial neural network (ANN) methodology and reported the regarding results as a journal in IEEE. They developed a project that is related to an E-Learning web site has many questions in a question bank and they tried to classify them into classes with respect to the text of questions. However, each text has many properties in its own such as difficulty level, area of curriculum, type of skill being tested such as vocabulary, comprehension, analysis and application. For priority, they found the item difficulties of each question by using some inputs for the system. They tried to classify the questions into two categories as: hard, medium and easy. Inputs were the length of question, the passing number of answers in question text, the number of questions answered correctly by students. And at the end, the system they developed achieved 78% performance in classification.

B. Question Classification with Support Vector Machines and Error Correcting Codes

In [3], they considered a machine learning technique for question classification. They used Support Vector Machines (SVM) methodology for classification and they used question and answer inputs for classification.

C. Question Classification by Structure Induction

In [4], they introduced a new approach to the task of question classification. The approach extracted structural information using machine learning techniques and the patterns found were used to classify the questions. The approach made appropriate in between the machine learning and handcrafting of regular expressions (as it was done in the past) and combined the best of both: classifiers can be generated automatically and the output can be investigated and manually optimized if needed. In their research, they used two different classifier as: Alignment-Based Learning Classifier and Trie Classifier. The results showed that the POS (Part of Speech) information helps the performance of the ABL

(Alignment Based Learning) method but not of the trie-based method. Overall, the trie-based approach outperformed the ABL implementations. They obtained similar results with the fine-grained data as well. Their best results fall close to the best-performing system in the bag-of-words and bag-of-grams versions of [5]. Their results ranged from 75.6% (Nearest Neighbors on words) to 87.4% (SVM on-grams). Given the simplicity of their methods and their potential for further improvement, this is encouraging.

III. METHODOLOGIES

What this research is trying to solve as a problem is that E-Learning System has thousands of questions from different topics in a question pool but the difficulty levels are not determined. The aim is to classify them into 5 classes as: Very Easy, Easy, Middle, Hard, and Very Hard. And these questions are composed of text formatted and they are multiple-choice questions. They are multiple question questions because these types of questions are commonly used in electronic testing environments.

By intelligent question classification:

- Instructors will be able to assign different questions to different students. This kind of instructional methodology can develop education quality and efficiency.
- Intelligent question classification will be operated in Run-Time of our educational testing software (Web Based Application), hence; this will provide instantaneous automated classification during testing session.
- A good database integrity with question classification engine will provide us synchronized classification.

The intelligent question classification involves 5 parts as: Data Pre-Processing, Item Difficulty, Initial Leveling according to Gaussian Normal Curve, Classification using ANN and SVM filters.

A. Classification by ANN

The Two techniques were used to classify 13 questions into 4 categories. First technique that has been used to classify was ANN (Artificial Neural Network).

ANN is mostly used technique that is used in artificial intelligence related works. It uses such a neuron model that is the most basic component of a neural network. It is so important in classification as well. It is used to distinguish multiple broad classes of models. Fig. 1 shows an example of a simplified neuron model. The neuron takes a set of input values, x_i each of which is multiplied by a weighting factor, w_i . All of the weighted input signals are added up to produce n , the net input to the neuron [6]. An activation function, F , transforms the net input of the neuron into an output signal which is transmitted to other neurons.

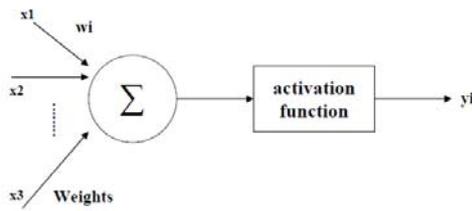


Figure 1. Simplified Neuron Model

Three inputs were used as follows:

- Item Difficulty
- Item Response
- Initial Gaussian Item Level

And it's been tried to compute what question belongs to what class of Gaussian Normal Curve. Sigmoid function (*tanh*) was used as an activation function in regarding research because it works in the domain range between -1 and 1 that is the domain borders of our inputs [6]. The sigmoid function is described by the following equation:

$$F(n) = \frac{1}{1 + e^{-n}} \quad (1)$$

Fig. 2 shows the sigmoid activation function.

There are two important properties of sigmoid function which should be noted. The first one is that the function is highly non-linear. For large values of activation, the output of the neuron is restricted by the activation function. The second important feature is that the sigmoid function is continuous and continuously differentiable. Both this feature and the non-linearity feature have important implications in the network learning procedure.

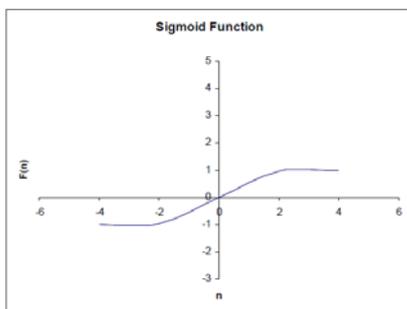


Figure 2. Sigmoid Function

This neural network type that is multilayer feed forward neural network distinguishes itself by the presence of one or more hidden layers, whose computation nodes are correspondingly called as hidden neurons or hidden units. The function of hidden neurons is to intervene between the external input and the network output in some useful manner [6]. The basic

structure of a feed forward neural network model is as follows (Fig. 3).

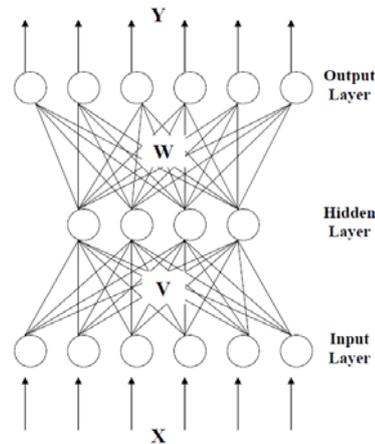


Figure 3. Basic Structure of Feed Forward Network

By adding one or more hidden layers, the network is enabled to extract higher order statistics. In a rather loose sense the network acquires a global perspective despite its local connectivity due to the extra set of synaptic connections and the extra dimension of neural interactions [7]. The ability of hidden neurons to extract higher-order statistics is particularly valuable when the size of the input layer is large. Learning happens depending on the activation function as well.

And here are the parameters that ANN used in regarding research as follows:

- Percentage of Data in Training Part: 50%
- Percentage of Data in Validation Part: 25%
- Percentage of Data in Testing Part: 25%
- Number of Inputs: 3
- Number of Patterns: 13
- Num of Hidden Layers: 4
- Number of Epochs: 500

After executing the ANN filter, 4 classes of questions were obtained. In other words, ANN found the correct classes of Gaussian Normal Curve. Because Gaussian Normal Curve classified questions into 4 classes, ANN yielded 4 classes. This means that ANN learned Gaussian Normal Curve classification logic correctly.

B. Classification by SVM

The second technique that was used for classification was SVM (Support Vector Machine). Same inputs of ANN and same percentages of data for training, validation and testing were used, however; SVM was not better than ANN in classification.

SUPPORT VECTOR machines (SVMs) [8, 9] offer a principled approach to machine learning combining many of the advantages of artificial intelligence

and neural-network approaches. Underlying the success of SVMs are mathematical foundations of statistical learning theory. Rather than minimizing training error (empirical risk), SVMs minimize structural risk which expresses an upper bound on the generalization error, i.e., the probability of erroneous classification on yet-to-be-seen examples.

From a machine learning theoretical perspective [8, 10], the appealing characteristics of SVMs are as follows:

- The learning technique generalizes well even with relatively few data points in the training set, and bounds on the generalization error can be directly estimated from the training data.
- The only parameter that needs tuning is a penalty term for misclassification which acts as a regularizer [8, 11] and determines a tradeoff between resolution and generalization performance.
- The algorithm finds, under general conditions, a unique separating decision surface that maximizes the margin of the classified training data for best out-of-sample performance.

SVMs express the classification or regression output in terms of a linear combination of examples in the training data, in which only a fraction of the data points, called “support vectors,” have nonzero coefficients. The support vectors thus capture all the relevant data contained in the training set. In its basic form, a SVM classifies a pattern vector X into class $y \in \{-1, 1\}$ based on the support vectors X^m and corresponding classes y_m as:

$$y = \text{sign}\left(\sum_{m=1}^M \alpha_m y_m k(X^m, X) - b\right) \quad (2)$$

Where $K(\cdot, \cdot)$ is a symmetric positive-definite kernel function which can be freely chosen subject to fairly mild constraints. The parameters α_m and b are determined by a linearly constrained quadratic programming (QP) problem [8, 10], which can be efficiently implemented by means of a sequence of smaller scale, sub problem optimizations, or an incremental scheme that adjusts the solution one training point at a time. Most of the training data X^m have zero coefficients α_m ; the nonzero coefficients returned by the constrained QP optimization define the support vector set. In what follows we assume that the set of support vectors and coefficients α_m are given, and we concentrate on efficient run-time implementation of the classifier.

Several widely used classifier architectures reduce to special valid forms of kernels $K(\cdot, \cdot)$, like polynomial classifiers, multilayer perceptrons,¹ and radial basis functions. The following forms are frequently used:

- inner-product based kernels (e.g., polynomial; sigmoid connectionist):

$$K(X^m, X) = f(X^m, X) = f\left(\sum_{n=1}^N X_{mn} X_n\right) \quad (3)$$

- radial basis functions (L_2 norm distance based)

$$K(X^m, X) = f(\|X^m - X\|) = f\left(\left(\sum_{n=1}^N |X_{mn} - X_n|^2\right)^{1/2}\right) \quad (4)$$

Where $f(\cdot)$ is a monotonically non decreasing scalar function subject to the Mercer condition on $K(\cdot, \cdot)$.

With no loss of generality, we concentrate on kernels of the inner product type (3), and devise an efficient scheme of computing a large number of high-dimensional inner-products in parallel.

Computationally, the inner-products comprise the most intensive part in evaluating kernels of both types (3) and (4). Indeed, radial basis functions (4) can be expressed in inner-product form as:

$$f(\|X^m - X\|) = f\left((-2X^m \cdot X + \|X^m\|^2 + \|X\|^2)^{1/2}\right) \quad (5)$$

Radial basis function was used in SVM and here are the statistical parameters and outputs of SVM in regarding research as follows:

- Percentage of Data in Training Part: 50%
- Percentage of Data in Validation Part: 25%
- Percentage of Data in Testing Part: 25%
- Average loss: 0.022227167
- Avg. loss positive: 0.0070014833 (965 occurrences)
- Avg. loss negative: 0.073067253 (289 occurrences)
- Mean absolute error: 0.022227167
- Mean squared error: 0.0033937631

IV. RESULTS

Be So far, two favorites’ artificial intelligence methods: ANN and SVM are proposed as a means to achieve accurate question level diagnosis, intelligent question for E-Learning or distance education platforms.

In order to observe their question classification abilities depending on the item responses of students, item difficulties of questions and question levels that are determined by putting the item difficulties to Gaussian Normal Curve, the effectiveness’s of ANN and SVM methods were evaluated by comparing the performances and class correctness’s of the sample questions (n=13) using the same 3 inputs as: item responses, item difficulties, question levels to 5018 rows of data that are the item responses of students in a test composed of 13 questions.

What has been done up to here is that we put the inputs to filters. At first, inputs were as follows: Inputs that entered to ANN and SVM commonly:

- Item Difficulty
- Item Response
- Initial Gaussian Item Level

Those inputs were put into two filters one by one and outputs were yielded as follows shown below:

TABLE I.
OUTPUTS OF ANN, SVM

Item ID	Actual Class	Class Predicated ANN	Class Predicated SVM
1	0	0,0031	0.0006
2	0	0,0003	0.0029
3	-0.5	-0,4999	-0.5100
4	0	-0,0033	0.0058
5	-0.5	-0,5090	-0.4954
6	-0.5	-0,5059	-0.5004
7	0.5	0,4998	0.5023
8	0.5	0,4890	0.5196
9	-0.5	-0,5087	-0.4959
10	0	-0,0091	0.0105
11	-0.5	-0,5120	-0.4905
12	0.5	0,4950	0.5010
13	1	0,7477	0.7926

If actual class values and predicted class values are benchmarked with respect to correctness parameter, following results are obtained. However, while calculating correctness values, all 13 questions were taken into account and average correctness values of 13 questions were obtained as in the table below.

TABLE II.
CORRECTNESS'S OF ANN, SVM

Classification Filter	Correctness
ANN	0,987702073954344
SVM	0,977920646004789

As seen above, ANN found the best results and predicted the nearest values with respect to actual class values. In order to understand how ANN and SVM classified 13 number of questions with respect to what rule, the Gaussian Normal Curve logic have to be known and it must be seen that system learned Gaussian Normal Distribution at the training parts of 2 methodologies. What the Gaussian Normal Distribution is that the normal distribution, also called the Gaussian distribution, is an important family of continuous probability distributions, applicable in many fields. Each member of the family may be defined by two parameters, location and scale: the mean ("average", μ) and variance ("variability", σ^2), respectively. It is often called as bell curve because the graph of its probability density resembles a bell. In addition, the normal distribution maximizes information entropy among all distributions with known mean and variance, which makes it the natural choice of underlying distribution for data summarized in terms of sample mean and variance. The normal distribution is the most widely used family of distributions in statistics and many statistical tests are based on the assumption of normality. In probability theory, normal distributions arise as the limiting distributions of several continuous and discrete families of distributions. This means that the area under the Gaussian curve is 1; that is the integration of Gaussian Curve function. From this probabilistic approach, normal distribution is used at classification as well.

A. Error Rates of Classification Methods

As it is known that supervised learning is a machine learning technique for creating a function from training data and in artificial intelligence (AI) concept, an error rating table is an indicator that is generally used in supervised learning. Each column of the table represents the classifier filter instance, while each row represents the error rates in a predicted class. It is beneficial for benchmarking multiple classes in a system. In fact, an error rating table contains information about actual and predicted class values done by a classification system.

TABLE III.
ERROR RATES OF CLASSIFICATION METHODS

	ANN	SVM
	Error Rates	Error Rates
-1	0	0
-0.5	0,0142	0,00312
0	0,00395	0,00495
0.5	0,0108	0,01526667
1	0,2523	0,2074

As seen above in the error rating table, the class named -1 has no error rates in all two ANN and SVM classifier because all two filters couldn't find any class of data in the domain of -1. At first, according to Gaussian Normal Curve, system learned the classifying system and there were no class as -1. Therefore, all error rates in class -1 are zero above.

B. C Relative Imaginary Classification Results Depending on the Error Rates

In previous section, the error rates of classifiers were found and now it is the time to find how much data were classified correctly and wrongly. It will be more accurate to handle true and false classification results at each class done by two classifiers. Why this methodology is needed is that confusion matrices are difficult or meaningless to use because classification results are so close to each other, therefore, it has been planned to show the classification accuracies by giving relative values depending on the error rates handled in previous section and the number of rows of data the system includes at first. The relative imaginary classification results depending on the error rates are as follows shown below:

TABLE IV.
ERROR RATES IN PREDICTION

	ANN		SVM	
	T	F	T	F
-1	0	0	0	0
-0.5	1902.594	27.406	1923.978	6.0216
0	1537.901	6.0988	1536.357	7.6428
0.5	1145.494	12.506	1140.321	17.678
1	288.6122	97.387	305.9436	80.056

Since the classification results are so close to each other, confusion matrices were not used to check the classifiers. Instead, the table above relatively shows the Accuracy, True-Positive, False-Positive, True-Negative, False-Negative and Precision values by relative imaginary true and false classification results.

V. CONCLUSION

In this paper, a classification model based on the two popular filter: ANN and SVM were proposed for the multiple choice question classification. Experiments were done in order to classify questions into 2 levels that are defined before entering filters by Gaussian Normal Curve. The effectiveness's of ANN and SVM methods were evaluated by comparing the performances and class correctness's of the sample questions (n=13) using the

same 3 inputs as: item responses, item difficulties, question levels to 5018 rows of data that are the item responses of students in a test composed of 13 questions. This paper provides an impetus for research on machine learning and artificial intelligence of question classification.

Future research directions are as follows:

- One of the shortcomings of the statistical method used in this project is that it lacks semantic analysis. To strengthen the "semantic understanding" ability, the supplemental method, natural language processing such as expression normalization, and sentence understanding should be used during the text preprocessing.
- To further test the effectiveness of the proposed model and to increase the generality of the empirical study, more extensive experiments should be conducted by using larger training and test sets.
- To find the most efficient classifier, additional machine learning or data mining methodologies can be used in latter works.

REFERENCES

- [1] Dalziel, J. "Integrating Computer Assisted Assessment with Textbooks and Question Banks: Options for Enhancing Learning", *Fourth Annual Computer Assisted Assessment Conference*, Loughborough, UK, 2000.
- [2] Fei T., Heng W.J., Toh K.C., Qi T. "Question classification for elearning by artificial neural network," *Information, Communications and Signal Processing, 2003 and the Fourth Pacific Rim Conference on Multimedia. Proceedings of the 2003 Joint Conference of the Fourth International Conference*, 3, pp.1757 – 1761, 2003.
- [3] Hacıoglu K., Ward W. "Question Classification with Support Vector Machines and Error Correcting Codes," in the *Proceedings of HLT-NACCL 2003*, Edmonton, Alberta, Canada, pp. 28-30, 2003.
- [4] Zaanen M., Pizzato L. A., Mollá D. "Question classification by Structure Induction". *Proceedings of the International Joint Conferences on Artificial Intelligence*, Edinburgh, U.K., pp. 1638-1639, 2005.
- [5] Zhang D., Lee W.S. "Question Classification using Support Vector Machines." *In Proceedings of the 26th ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR)*, Toronto, Canada, 2003.

- [6] Haykin, S. *Neural Networks: A Comprehensive Foundation*, 2nd edn., Prentice Hall International, Inc, 1999.
- [7] Churchland P.S. and Sejnowski T.J. *The computational Brain*, Cambridge, MA: MIT Press, 1992.
- [8] Genov R., Member, IEEE, and Gert Cauwenberghs. "Kerneltron: Support Vector "Machine" in Silicon", IEEE TRANSACTIONS ON NEURAL NETWORKS, 14(5), 2003.
- [9] Boser B., Guyon I., and Vapnik V. "A training algorithm for optimal margin classifier," in *Proc. 5th Annu. ACM Workshop Computational Learning Theory*, pp. 144–52, 1992.
- [10] Vapnik V. *The Nature of Statistical Learning Theory*, New York: Springer-Verlag, 1995. Girosi F., Jones M., and Poggio T. "Regularization theory and neural networks architectures," *Neural Comput.*, **7**, pp. 219–269, 1995.