Performance evaluation of MLPNN and NB: A Comparative Study on Car Evaluation Dataset

Zia Ul Rehman[1], Hira Fayyaz[2], Asghar Ali Shah[3], Numan Aslam[4], Muhammad Hanif[5], Sagheer Abbas[6]

[1,3,4,5]Department of Computer Sciences, Bahria University Lahore Campus, Pakistan
[2]Department of Computer Science, University of Management and Technology, Pakistan
[6]School of Computer Science, National College of Business Administration & Economics, Pakistan

Summary
Cars are vital in everyday life. It plays an important role as it's a comfortable mean of transportations. Every car has a distinct flavor in term of price, feature, safety and the level of luxury it provides. People tend to make clear choices when they decide to buy car for themselves. They evaluate different cars on various parameters. Manufacturing and business are interested to know the popular features on which buyers make their choice as it can enhance their business value. Data mining algorithms can be employed in this respect. Various data mining algorithms perform differently. The purpose of this research work is to equating two influential algorithm evaluating dataset acquired from the University of California Irvine. This research focuses on comparing and contrasting speed, accuracy and performance of these algorithms.

Keywords:
Multi-layer perceptron (MLPNN), Naïve Bayesian (NB), Artificial Neural Network (ANN)

1. Introduction

Cars offer diverse characteristics in terms of model and manufacturer preferences. Cost, safety and luxury are three imperative factors which are considered when buyers make their choice. These factors significantly contribute towards the reduction of accidents occurring. Some standard equipment is also vital to consider when buying cars. Which includes performance enhancers, conveniences and safety tools in cars. Safety as already mentioned is one of the imperative factors for car buying decision. Same is the case for convenience which has attributes such as maintenance, door and luggage boot. Cost deliberation is also crucial to make sure that car which is bought is worth what it has cost to the owner. Financial responsibility also comes with owning a car as it need to be maintained for convenience. This particular research work utilizes attribute "buying" for assessing acceptability of car cost in comparison to the other attributes it is offering such as doors, lug boot, person and safety.

Data mining is a subdivision of Artificial Intelligence that is functional in a varied sphere extensively as it is analysis of data for relationships between parameters which previously have been not explored. This analysis of large dataset proves fruitful in determining future aspects in field of manufacturing, medical, business, education and many more. Data mining techniques employs algorithms such as Artificial Neural Network, Support Vector Machines, Naive Bayes, multi-layer perceptron and these all perform differently in varying circumstances. The focus of this research work is to compare two influential algorithms; Naive Bayes (NB) and multi-layer perceptron artificial neural network (MLPNN) in term of speed and accuracy they depict on the data set of cars.

This paper is structured as follows; it initiates with an introduction in Section 1, Section 2 proceeds with brief insight into the previous work done in this field, Section 3 provides an elaboration on data set utilized and various techniques applied on it such as data cleaning and normalization, Section 4 demonstrate experiments and result and finally section 5 presents conclusion the research work conducted.

2. Literature Review

One crucial step in data mining projects is to find an efficient classification algorithm so that results can be trusted upon. It also depends on the experimental design of the system. If the selection of algorithm is not made thoughtfully the outcomes of data mining task could be compromised, consequently resulting in invalid conclusions. Researchers have focused on this point and have compared various algorithms in terms of accuracy and speed. This section presents a brief overview of the essential work done in this domain.

In research work conducted by S. Makki [1] backpropagation neural network (BNN) and naïve Bayesian classifier (NB) has been employed for data mining classification for evaluation on car data set. These two algorithms are tested on data set and results show that BNN is much more accurate as compared to NB although it is difficult to implement and it runs slow. In [2] author presents a comparative study on multiple prediction

Manuscript received September 5, 2018
Manuscript revised September 20, 2018

algorithm for analyzing breast cancer survivability. In experiments a large data set with 10-fold cross validation has been used. Results demonstrate that decision tree is the most accurate, artificial neural networks takes third place and logistic regression method is the last in terms of accurate results. R. Russo [3] in his work has applied machine learning algorithm to dataset which describes movie. The basic aim is to create a movie recommender system for movie watchers. Neural networks, NB, simple rule classifiers and decision tree are compared. Results indicate that NB and neural networks perform better in terms of evaluating given dataset. In [4] author proposes a methodology to evaluate an adaptive tourist service of onboard cars. The system evaluated provides personalized information to tourist on cars. In the research work layered sampling strategy is employed and system suggestions to users are compared for accuracy. S. Singh [5] evaluates the performance of different classification methods. Three algorithm are studied in this research; K-Nearest Neighbors, Support Vector Machines (SVM) and Artificial Neural Networks (ANN). The results demonstrate that SVM and ANN are better predictors.

3. Data Set Description

This research focuses on evaluation of car data set which has been obtained from UCI dataset repository. This dataset records specific attributes of car and is denoted by Marco Bohance [6]. The car dataset is a derivative of simple hierarchical decision. Table 1 describes the dataset categorically.

Table 1: Car Evaluation Dataset

Dataset Characteristics:	Multivariate
Attribute Characteristics:	Categorical
Associated Tasks:	Classification
Number of Instances:	1728
Number of Attributes:	6
Missing Values?	No

The class attributes utilized in this dataset are described in table 2

Table 2: Class Attributes

Attributes	Denoted as
Acceptable	“acc”
Good	“good”
Unacceptable	“unacc”
Very Good	“vgood”

For checking the performance of algorithm under study we first need to perform standard data analysis so that some pattern in data can be recognized. The attribute range and their frequency are calculated and presented in tabular form for further comparison. Figure 1 describes distribution of the class attributes in the car dataset.

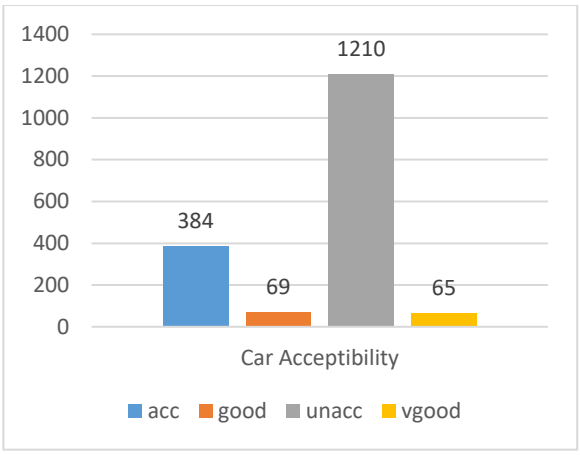

Figure 1: frequency of class output from the dataset

Figure 1 demonstrate the frequency of class attributes. 1728 cars are evaluated. The results obtained are elaborated below:

- Acceptable: 384 (22.28%)
- Good: 69 (4.05%)
- Unacceptable: 1210 (69.85%)
- Very good: 65 (3.82%)

4. Experiments and Results

This section aims at demonstrating the experimental setup and the results obtained. It describes the basic working of algorithm, how data is prepared for testing and what results claims about the performance of both data mining algorithms.

4.1 Classification Method

The classification methods employed in this research work are Naïve Bayesian (NB) and Multilayer perceptron artificial neural network (MLPNN). NB is named after Thomas Bayes (1702-1761), it is learning and statistical method for classifying data. The Naïve Bayesian is based on Bayes theorem which assumes that there are no dependencies between predictors. It is easy to implement particularly with satisfying outcomes. A multilayer perceptron works by mapping set of inputs data onto suitable outputs. It consists of multiple layers ensuring that each layer is fully connected with next one. Excluding the input layer, each layer in composed of neurons which are processing elements. Backpropagation techniques are employed for training networks and then testing data. The experiments are conducted using these two classifiers. The basic purpose was to find out which classifier demonstrate best results with the given car data set. The process starts by pre-processing data, training data, testing data and then making predictions.

4.2 Data Cleaning

In classification problems, data cleaning is employed for achieving better results. The basic purpose of data cleaning is to remove any kind of inconsistency in training data. This make us sufficient enough to use reliable data for developing efficient classification model as unclean data affect the accuracy of results obtained. The dataset utilized in this research is also cleaned to ensure quality for model creation. We have converted nominal attributes into numeric attributes. This conversion is required as it makes normalization of the data possible. Table 3 shows the conversion.

Table 3: nominal to numeric conversion

Attribute	Nominal	New Numeric Value
Buying	Vhigh	4
	High	3
	Med	2
	Low	1
Maintenance	Vhigh	4
	High	3
	Med	2
	Low	1
Luggage Boot	Small	3
	Med	2
	Big	1
Safety	Low	1
	Med	2
	High	3

4.3 Data Pre-processing

Once the dataset has been chosen, raw input data should be pre-processed, otherwise it will negatively affect the results obtained. It is extremely crucial to the performance of neural network. The two basic pre-processing techniques are data transformation and normalization. Transformation manipulates raw data inputs creating a single input to the network, while normalization tends to distribute data evenly scaling it onto an adequate range. This can help network in learning process enhancing its ability to understand the association between given inputs and generated outputs. For this particular research min-max normalization approach is being employed instead of z-score methodology. As a basic rule, min-max normalization will always generate results between the range 0 to 1.

4.4 Dataset Split

The pre-processed dataset is split into two shares of varying sizes for utilizing one half as training data and second half as testing or validating network. The methodology of data splitting can have considerable influence on the performance of model. Inappropriate data splitting can result in incorrect and extremely variable performance. Classifying algorithm uses training data for learning. Training model is built by comparing the attributes of dataset with class/label. After training, the model is tested on test data which is the other half of split dataset. In this research work four splits are being tested. 10-folds cross validation is also checked in experiment. 10-folds means that fitting procedure will be completed ten times, with each fit consisting of training set off 90% and testing set of 10 %. Table 4 demonstrate the data splitting.

Table 4: car dataset split for model creation

Training and Testing % split
90% 10%
66% 34%
50% 50%
10 Folds

4.5 Results

This section presents results of the experimentation setup. The process is as follows; it is supervised learning method. We have trained the model utilizing attributes inclusive of class attributes. As it is a supervised model, the model is built basing on the class values in correspondence to the values of attributes individually. Weka 3.8 is used for simulation purpose. The results achieved by various experimentation setup in Naïve Bayesian and multilayer perceptron artificial neural networks are elaborated in Table 5 and 6 respectively. The tables show the percentage splits employed which are; 90:10, 66:34, 50:50 and 10-folds cross validation. The time taken to build the model and test is also show in association with each split. The last column demonstrates the percentages of correct and incorrect results.

Table 5: Performance of Naive Bayesian

Percentage Split		Time in Seconds		Naïve Bayesian	
Training %	Testing %	Build	Test	Correct %	Incorrect %
90	10	0.02	0.02	79.76	20.23
66	34	0	0.05	83.84	16.15
50	50	0	0.02	83.68	16.31
10 Folds		0	-	82.34	17.65

Table 6: Performance of MLPNN

Percentage Split		Time in Seconds		MLPNN	
Training %	Testing %	Build	Test	Correct %	Incorrect %
90	10	2.16	0	94.79	5.20
66	34	2	0	93.19	6.80
50	50	1.98	0	92.70	7.29
10 Folds		1.97	-	94.09	5.90

5. Conclusion

The fundamental objective of this research work was to compare and contrast two data mining algorithms; Naïve Bayesian and Multilayer Perceptron Artificial Neural Network in terms of accuracy they offer. Paper initiates with an introduction of the domain and deliberates about the previous influential work conducted. Next it come towards a detailed elaboration of the experimentation setup describing dataset, data cleaning and pre-processing. Results with both the algorithm are presented.

The results demonstrate Multilayer Perceptron Artificial Neural Network (MLPNN) outperforms Naïve Bayesian (NB) in every sampling methodology employed in experiment. The comparative study shows that the accuracy of MLPNN is far much better than NB. In all the ratios tested; 90:10, 66:34, 50:50 and 10-folds cross validation MLPNN depicts accurate results in comparison to NB. Although MLPNN has limitation of time, that means it takes longer to build and test model but proves to best in presenting accuracy.

References

[1]. Makki, S., Mustapha, A., Kassim, J. M., Gharayebeh, E. H., & Alhazmi, M. (2011). Employing neural network and naive Bayesian classifier in mining data for car evaluation. In Proc. ICGST AIML-11 Conference (pp. 113-119).

[2]. Delen, D., Walker, G., & Kadam, A. (2005). Predicting breast cancer survivability: a comparison of three data mining methods. Artificial intelligence in medicine, 34(2), 113-127.

[3]. Russo, R. (2006). Bayesian and neural networks for motion picture recommendation. Technical Report, Boston COllege.

[4]. Console, L., Gena, C., & Torre, I. (2003). Evaluation of an on-vehicle adaptive tourist service. In Proceedings of the Second Workshop on Empirical Evaluation of Adaptive Systems, held at the 9th International Conference on User Modeling UM2003, Pittsburgh (pp. 51-60).

[5]. Singh, S. (2005). Modeling Performance of Different Classification Methods: Deviation from the Power Law. Project Report, Department of Computer Science, Vanderbilt University, USA (April 2005).

[6]. Bohanec, M., & Rajkovic, V. (1988). Knowledge acquisition and explanation for multi-attribute decision making. In 8th Intl Workshop on Expert Systems and their Applications (pp. 59-78).