Classification Experiments for Number Plate Recognition Data Set Using Weka

Atul Kumar[1], Sunila Godara[2]

[1]Department of Computer Science and Engineering
Guru Jambheshwar University of Science and Technology

[2]Assistant Professor Department of Computer Science and Engineering
Guru Jambheshwar University of Science and Technology

***Abstract*— the Paper presents the comparison of different classification techniques for the task of classifying Number plate image data set. The comparison was conducted using WEKA (Waikato Environment for Knowledge Analysis) open source which mainly consists the collection of machine learning algorithms for data mining purpose. The main purpose of this paper is to investigate efficiency of different classification methods by applying on the Number Plate Image data set. Which will further used in Number Plate Recognition process. The methods or algorithms performed are J48 (Decision Tree), MLP (Multilayer Perceptron), Naïve Bayes, K-NN (K-Nearest Neighbors) and the result of MLP was better than others received in range of 78% - 96% as compared to other techniques.**

***Keywords*— J48(Decision Tree), MLP, Naïve Bayes, K-NN.**

I. INTRODUCTION

In all aspects of modern life, there is a demand for information systems for data processing in respect of vehicles. These systems needed the data to be archived or by a human or by a special team which is able to recognize vehicles by their license plates in real-time environment and reflect the facts of reality in the information system. In mostly cases vehicles are identified by their license plate numbers, which can be easily read by humans but not machines. For the machines, a number plate is just a dark spot that is within a region of an image with a certain intensity and luminosity. Due to this it is needy to design a robust mathematical system able to perceive and extract what we want from the captured image. The functions which implemented or mathematical patterns in what is called "NPR Systems" (Numbers Plate Recognition) and mean a transformation between the real environments is perceived and information systems need to store and manage all that information.

The design of these systems is one of the areas of research in areas such as Artificial Intelligence, process of Pattern Recognition and Neural Networks. Systems of automatic recognition of license plates are sets of hardware and software to process a signal that is converted into a graphical representation such as static images or sequences of them and recognize the characters in the plate.

Here for our system we need to classify the data in term of images. One of the crucial tasks among computer vision field in term of image processing is an object classification. Image classification is the process of labelling the images into one of a number of predefined categories. Here we have the Data set of Images which includes the Number Plate of the vehicles. The large data set was recorded from the various different location and then processing is provided. From the large data set of images experiment was performed on no of images. There are various algorithms or methods which were applied on the data set and result was recorded. Different techniques or methods available are like J48 (Decision Tree), MLP (Multilayer Perceptron), Naïve Bayes, K-NN (K-Nearest Neighbors) [1] [3]. Where this data mining task provide the method to best classify the image of number plate .

II. METHODOLOGY

Methodology is very simple. The Data for experiment is collected from various location like images of vehicle captured and then stored as data base. Here using Weka different classification algorithms or methods were applied and useful results were achieved which will be very helpful for our Number plate recognition System and also for the new researchers. Let is the Process in further steps discussed below.

III. DATA SOURCE AND TOOL USED

Data set of images of Number Plates is taken into consideration for the classification process. Images were captured manually by Nikon Digital camera and collected from different sites. All images carry the number plate of vehicle. An image captured from camera was in jpeg format. While our all process of classification take place on Weka and Weka support only CSV or ARFF format files. Then we need of conversion take place. For all further processing Images

ISSN: 2278 – 1323

from data set was taken and converted to CSV and uploaded. Then further process of classification take place. Here some of test images given below used for our test purpose and results of these has been discussed in result section.

Figure 1: Test Image1

Figure 2: Test Image2

Figure 3: Test Image3

WEKA

WEKA the open source software provided under the General Public License. Where the System was developed at University of Waikato(New Zealand). WEKA mainly stands for the Waikato Environment for the purpose of Knowledge Analysis. Weka is available freely at http://www.cs.waikato.ac.nz/ml/weka. The system was written using the object oriented language Java. While the many different levels are at which Weka can be used [1]. Which generally provides the implementations of state of art data mining and various machine learning algorithms .

Figure 4: Front View of WEKA Tool

Weka contains the modules for clustering, data preprocessing, and classification and association rule extraction. The format of input data supported by WEKA is mainly ARFF (Attribute Relation File Format) and CSV (Comma Separated Values). Where ARFF format consists of special tags which indicate the different things in the input data like attribute values, attribute types, attribute names and data etc. The actual functionality in Weka is provided by interface Explorer. While simulation is performed by clustering and classification methods of Weka for comparison purpose . To work on WEKA is not important to having deep knowledge of Data Mining that's why also Weka is popular. While it also provide the Graphical user interface to the user and many more facilities [1] [2].

GUI interface consists mainly four buttons each for major task. These four buttons shown in starting phase can be used to start the following applications:

- Explorer: Provides the environment for data exploring.
- Experimenter: Provides environment for conducting statistical tests and performing experiments.
- Knowledge Flow: Provides environment which supports essentially the same task as the Explorer but having drag and drop interface. Also provides incremental learning.

ISSN: 2278 – 1323

- Simple CLI: Provides the simple Command line interface which allows directly execution of Weka commands [2].

A. LOAD DATA INTO WEKA

As we know Weka only supports the two kind of format ARFF (Attribute Relation File Format) and CSV (Comma Separated Values). Then our input data is in jpg format so we need to convert it into CSV or ARFF. Here in our purposed system we converted the Jpg image to CSV format and then uploaded to Weka. Let's see how the process takes place

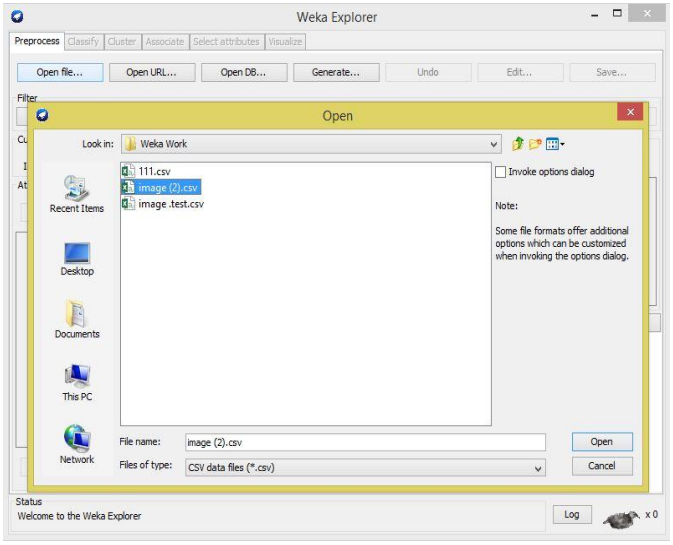


Figure 5: Loading Data to Weka

B. DATA PREPROCESSING

After the data uploaded it is also essential to be data in the format or type which support the further processing. Then before further process we need to take preprocessing of data then in this phase data is converted to Nominal type. Let's see the process of task with the help of image shown below.

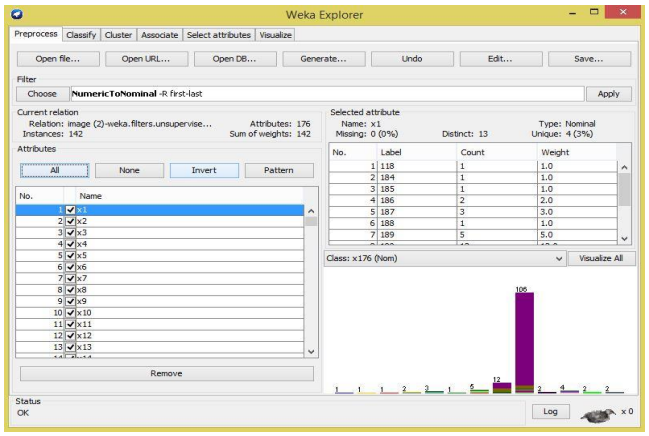


Figure 6: Conversion from Numerical to Nominal

C. APPLYING CLASSIFIER

In Data mining task provided by Weka there are different techniques or methods available for classification purpose like J48 (Decision Tree), MLP (Multilayer Perceptron) [4] [6], Naïve Bayes, K-NN (K-Nearest Neighbors). There are available various classifiers which we implemented some of which and the blow snapshot showing the interface

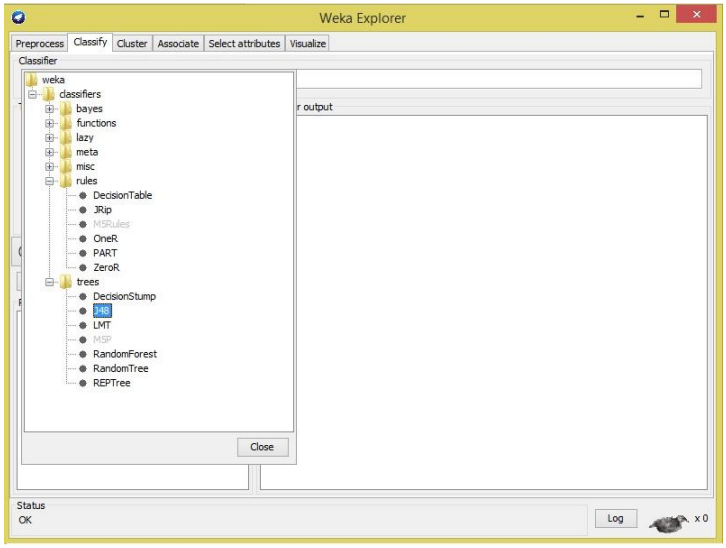


Figure 7: Applying Classifier into Weka

IV. EXPERIMENTAL RESULT AND PERFORMANCE EVALUATION

Here we have applied different classifier on No of images and achieved different results. Where the accuracy achieved on different image of number plate is achieved is different by applying different classifier. But mainly we achieve the good accuracy by applying the Neural Network approach. We take 10 images as sample for classification experiment purpose. So of which are shown above. Here you will see a table blow which shows the accuracy of image correctly achieved by using different classifier.

Testing Image	Recognition Accuracy
Test Image1	84.3
Test Image2	79.2
Test Image3	81.0
Test Image4	76.8
Test Image5	82.1
Test Image6	73.5
Test Image7	86.5
Test Image8	80.2
Test Image9	73.4
Test Image10	79.7

Table 1: Matching Percentage of 10 Sample Number Plate Images

ISSN: 2278 – 1323

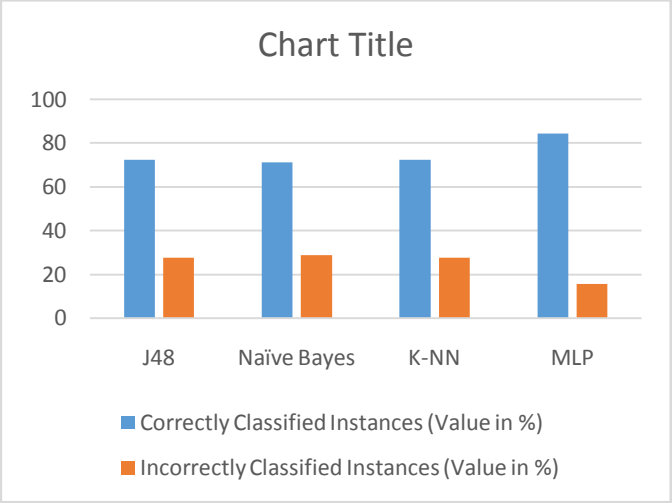
Here we have shown that the one test Image data set having 142 instances. Where different classifier which we have applied like Decision Tree, Naïve Bayes [8], K Nearest Neighbor, Neural Network (Multilayer Perceptron)[3] [4] [6] and achieved the results as shown in below table.

Result of Each Algorithm for NPR Image Dataset

Name of Algorithm	Correctly Classified Instances (Value in %)	Incorrectly Classified Instances (Value in %)	Time Taken (In seconds)	Kappa Statistic
J48	72.5352	27.4648	0.04	0.299
Naïve Bayes	71.1268	28.8732	0.02	0.4269
K-NN	72.5352	27.4648	0	0.3851
MLP	84.3333	15.6667	0.60	-0.174

Table 2: Result for NPR Data Set

Image Showing simulation for image Dataset (% age of correctly and incorrectly classified instances)

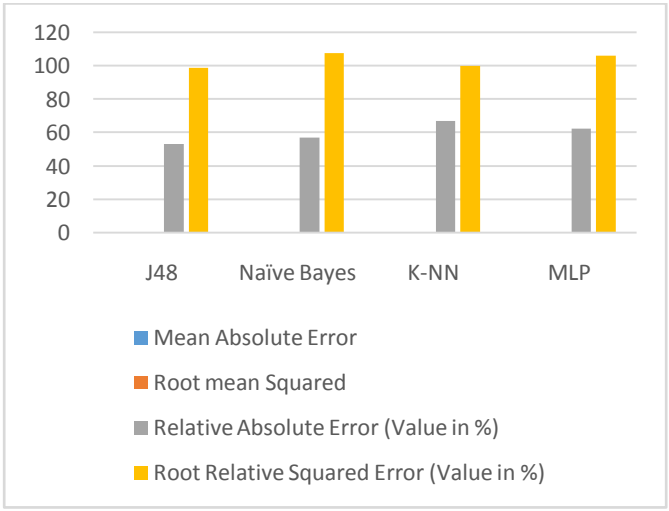


Graph 1: Showing Simulation for image Dataset

Simulation Error of Each Algorithm for Image Dataset

Name of Algorithm	Mean Absolute Error	Root mean Squared	Relative Absolute Error (Value in %)	Root Relative Squared Error (Value in %)
J48	0.0285	0.1532	53.3437	98.9567
Naïve Bayes	0.0304	0.1667	56.9436	107.66
K-NN	0.0358	0.1552	66.9509	100.2491
MLP	0.0332	0.1355	62.6582	106.258

Table 3: Simulation Error Image Dataset



Graph 2: Simulation Error of Each Algorithm

Image Showing simulation Error for image Dataset (% age of Mean Absolute Error, Root mean Squared, Relative Absolute Error, Root Relative Squared Error to build classifier)

While the same experiment was performed over another image having 2536 instances and result of different classifier like J48, Naïve Bayes [8], KNN, MLP was recorded. Like how many % of the classified correctly or how many were incorrectly, also the time taken. This is shown below in the form of table and also indicated into Graph.

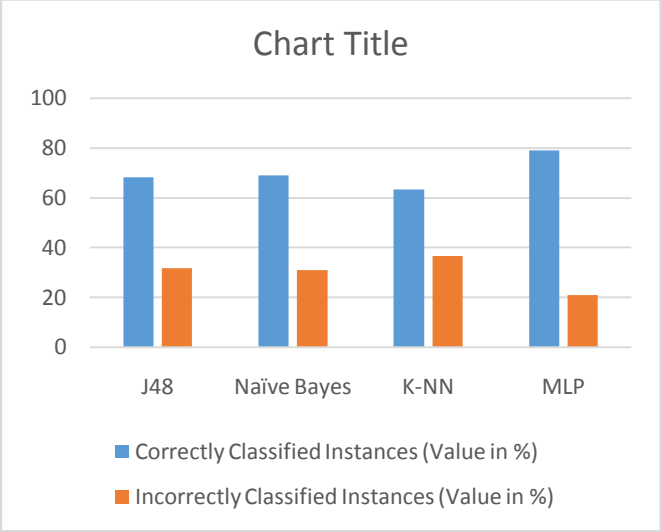
Result of Each Algorithm for Image Dataset

Name of Algorithm	Correctly Classified Instances (Value in %)	Incorrectly Classified Instances (Value in %)	Time Taken (In seconds)	Kappa Statistic
J48	68.3333	31.6667	0.02	-0.0105
Naïve Bayes	69.1268	30.8732	0	0.2484
K-NN	63.3333	36.6667	0	0.0418
MLP	79.2214	20.7786	0.40	-0.0183

Table 4: Result for Image Dataset

Image Showing simulation for image Dataset (% age of correctly and incorrectly classified instances) .

ISSN: 2278 – 1323



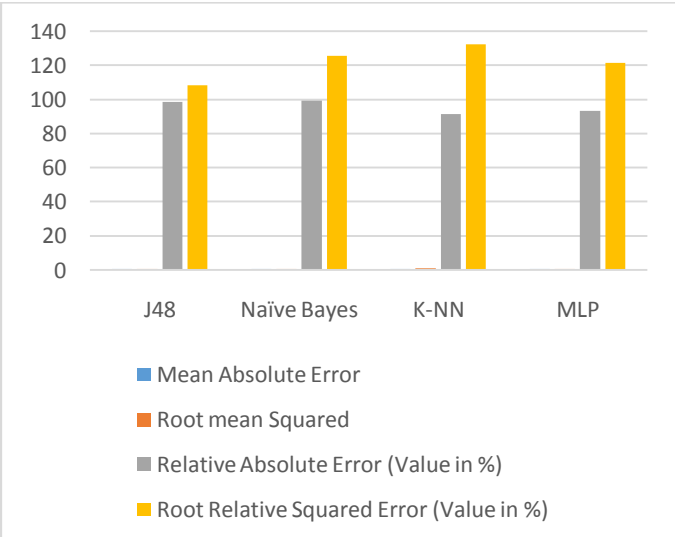
Graph 3: Showing Simulation for image Dataset

Simulation Error of Each Algorithm for Image Dataset

Name of Algorithm	Mean Absolute Error	Root mean Squared	Relative Absolute Error (Value in %)	Root Relative Squared Error (Value in %)
J48	0.4658	0.4586	98.4521	108.2211
Naïve Bayes	0.3568	0.5421	99.4125	125.8521
K-NN	0.3874	0.6582	91.532	132.5284
MLP	0.3502	0.559	93.255	121.5647

Table 5: Simulation Error for Image Dataset

Image Showing simulation Error for image Dataset (% age of Mean Absolute Error, Root mean Squared, Relative Absolute Error, Root Relative Squared Error to build classifier)



Graph 4: Simulation Error of Each Algorithm

V. CONCLUSION

Our assessment of different groups of classification suggests that Neural Network algorithm provides the best classification result for the data set of image of number plate. Algorithm provides the fast result in case simple data set. While the negative aspect of the approach is it is time consuming process when data set is of complex type. While K-NN can also be adopted when complex data set is taken into consideration.

REFERENCES

[1] Payal P.Dhakate, SuvarnaPatil, K. Rajeswari, DeepaAbin, "Preprocessing and Classification in WEKA Using Different Classifiers,"*Int. Journal of Engineering Research and Applications*, Volume 4, Issue 8, August 2014

[2] Sonam Narwal and Mr. Kamaldeep Mintwal, "Comparison the Various Clustering and Classification Algorithms of WEKA Tools,"*International Journal of Advanced Research in Computer Science and Software Engineering IJARCSSE*,Volume 3, Issue 12, December 2013.

[3]Rohit Arora, Suman, "Comparative Analysis of Classification Algorithms on Different Datasets using WEKA,"*International Journal of Computer Applications*,Volume 54, September 2012.

[4] Izzat AlsmadiandIkdam Alhami, "Clustering and classification of email contents," *Journal of King Saud University Computer and Information Sciences*,SCIENCEDIRECT, January 2015.

[5] Jaikee Kumar Singh, ShrutiPareek, "Comparison of Different Classification Techniques Using WEKA for Hepatitis," *Indian Journal of applied research*, Volume: 4, Issue: 10, October 2014.

[6] NurshazwaniKamarudin, MokhairiMakhtar, "Comparison of image classification techniques using CALTECH 101 data Set,"*Journal of Theoretical and Applied Information Technology*, Vol.71, No.1, January 2015.

[7]WalidMoudani, Abdel Rahman Sayed, "Efficient Image Classification using Data Mining," *International Journal of Combinatorial Optimization Problems and Informatics*, Volume 2, No. 1, January/April 2011.

[8] Tina R. Patil, Mrs. S. S. Sherekar, "Performance Analysis of Naive Bayes and J48 Classification Algorithm for Data Classification," *International Journal of Computer Science and Applications*, Volume 6, No.2, Apr 2013.

[9] Andreas Meier, Mark Gonter, Rudolf Kruse, "Precrash classification of car accidents for improved occupant safetySystems," *2nd International Conference on System-Integrated Intelligence*, SCIENCEDIRECT 2014 .

ISSN: 2278 – 1323

[10] Ranjitakumari Dash, "Selection of the Best Classifier from Different Datasets Using WEKA," *International Journal of Engineering Research & Technology*, Volume 2, Issue 3, March 2013 .

[11] Shilpa DhanjibhaiSerasiya, Neeraj Chaudhary, "Simulation of Various Classifications Results using WEKA," *International Journal of Recent Technology and Engineering,*Volume 1, Issue 3, August 2012.

[12] Meenakshi, Geetika, "Survey on Classification Methods using WEKA,"*International Journal of Computer Applications*, Volume 86, No 18, January 2014.

ISSN: 2278 – 1323