Students' Academic Performance Analyzing System using RFID and Raspberry-Pi

A. Hemalatha
Dept of ECE
Periyar Centenary Polytechnic College

G. Gandhimathi
Dept of ECE
National Institute of Technology

Abstract:- **The aptitude to timely predict the academic performance propensity of engineering graduate students is very important and useful for tutors and their placements. As big data is the hottest trend in the tech industry at the moment, machine learning is incredibly powerful to make predictions or calculated suggestions based on large amounts of data. The scope of this research is to explore which is the most competent rfid with machine learning, in predicting the final grade of the students. Consequently, five academic courses are chosen, each constituting an individual dataset, two well-known classification(k-nearest neighbourhood (knn), and decisionthen different prediction model such as linear regression, k-nearest neighbourhood (knn), decision tree are experimented with. Knn achieved the best prediction results, which are very satisfactory compared to those of similar approaches.**

Keywords: Machine learning algorithms, rfid. predictive analysis

I.INTRODUCTION

1.1 PREDICTIVE ANALYTICS:
It is about predicting future outcome based on analyzing data collected previously. It includes two phases:

- **TRAINING PHASE***: learn a model from training data
- **PREDICTING PHASE:** use the model to predict the unknown or future outcome

1.2 PREDICTIVE MODELS:
Many predictive models are there, each based on a set of different assumptions regarding the underlying distribution of data. Two general types of problems have discussed here.
- **CLASSIFICATION :** about predicting a category (a value that is discrete, finite with no ordering implied)
- **REGRESSION:** about predicting a numeric quantity (a value that's continuous and infinite with ordering).

1.3 MACHINE LEARNING ALGORITHMS*:*
ML can play a key role in a wide range of critical applications, such as data mining, natural language processing, image recognition, and expert systems. Ml provides potential solutions in all these domains and more, and is set to be a pillar of our future civilization. It can be divided into 3 broad categories.

- SUPERVISED LEARNING*:*
Where a property **(label)** is available for a certain dataset **(training set),** but is missing and needs to be predicted for other instances.

- UNSUPERVISED LEARNING:
Where the challenge is to discover implicit relationships in a given **unlabeled** dataset (items are not pre-assigned).

- REINFORCEMENT LEARNING:
Falls between these 2 extremes—there is some form of feedback available for each predictive step or action, but no precise label or error message.
Our research problem falls under supervised learning category.

1.4 K-NEAREST NEIGHBORS:
The training process involves memorizing all the training data. To predict a new data point, we found the closest k (a tunable parameter) neighbors from the training set and let them vote for the final prediction.
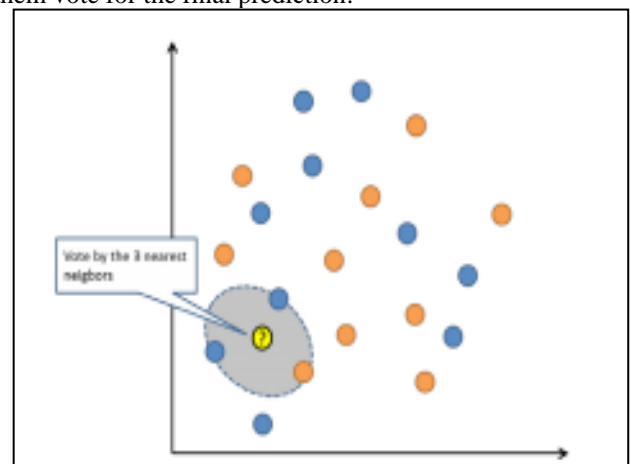
To determine the "nearest neighbors," a distance function needs to be defined (e.g., a euclidean distance function is a common one for numeric input variables). The voting can also be weighted among the k-neighbors based on their distance from the new data point. Here is the r code using k-nearest neighbor for classification.
The strength of k-nearest neighbor is its simplicity. No model needs to be trained. Incremental learning is automatic when more data arrives (and old data can be deleted as well).

Special Issue - 2019

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
CONFCALL - 2019 Conference Proceedings

The weakness of knn, however, is that it doesn't handle high numbers of dimensions well.

Choosing a suitable k is both an art and a science. Both low and high values of k have their own advantages. However, the optimal value of k largely depends on the data in the training setas shown in fig :1.

1.5 DECISION TREES:

Mostly used in classification problems, that is problems and works for both categorical and continuous input and output variables, with a predefined target variable. Decision trees learn and train themselves from given examples and predict for unseen circumstances by practical methods for inductive inference (inductive inference is the process of reaching a general conclusion from specific examples) as shown in fig:2 & fig :3 and it is widely used.
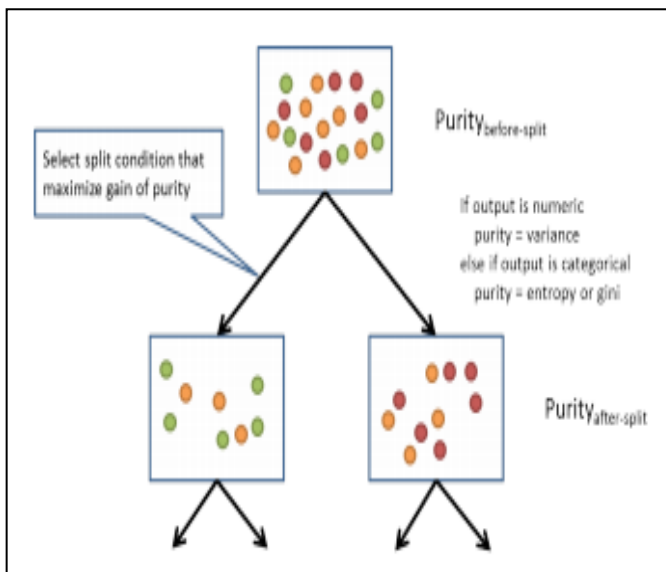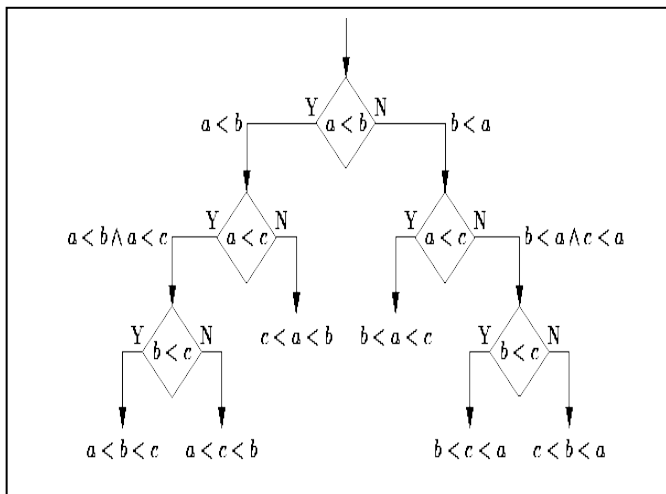

FIG:2   DECISION TREES

II MODEL DEVELOPMENT

In this paper, both authendication, classification and prediction problem has taken by the authors. The students examination data set of single academic course is used for this purpose. The structure of model development is shown in fig:4
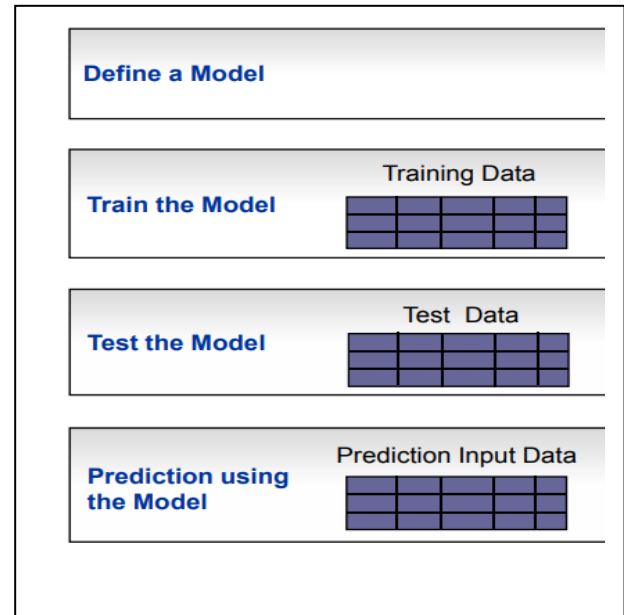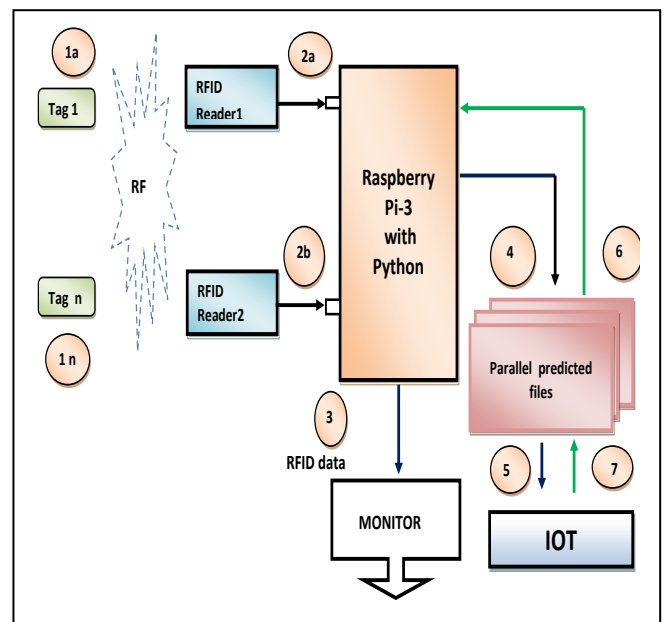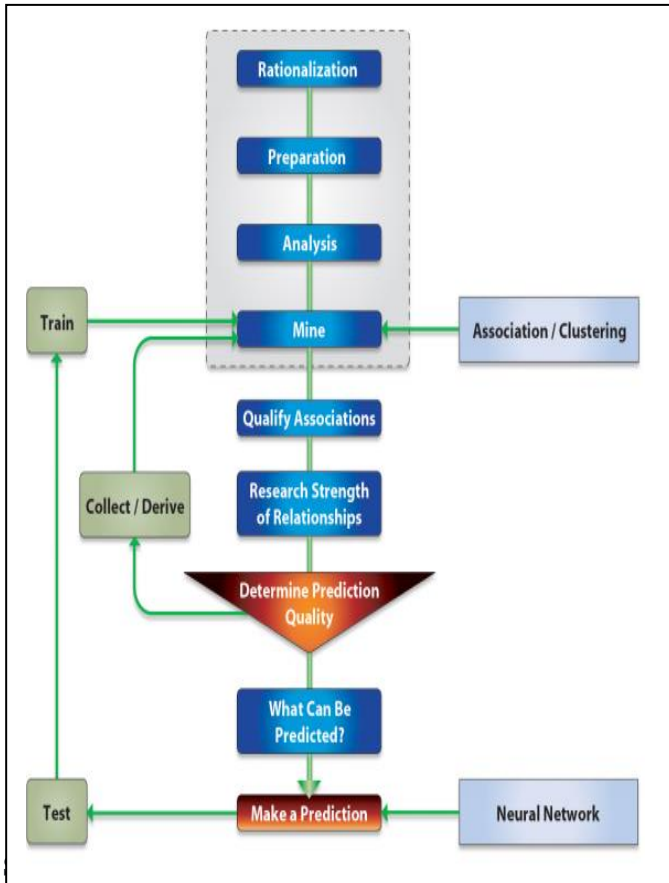

FIG: 4 MODEL DEVELOPMENT

For authentication, rfids are used, each student has a unique rfid, and their marks and other academic details are linked through raspberry –pi . This is shown in the fig:5

For classification and prediction problem machine learning algorithm such as knn, decision tree are written in the python language and loaded in raspberry –pi . The process flow is shown in fig :6

Special Issue - 2019

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
CONFCALL - 2019 Conference Proceedings

STEP: 2 develop predicting modeling approach to provide schedule and cost measure during program execution phase.

STEP: 3 develop preliminary models using appropriate algorithm and mining existing data set.

III   IMPLEMENTATION

3.1 MODULE: 1
Classification machine learning systems as 'student success rate' (ie. Whether the student has pass or fail in their final exam). Systems where we seek a yes-or-no prediction and the option will be pass or fail. Each instance can be assigned with multiple categories, so these types of problems are known as 'multi-label classification' problem, since we have a set of target labels as pass or fail.

3.2 MODULE:2- PROBLEM TRANSFORMATION:

*T*o transform our multi-label problem into single-label problem(s).this is the simplest technique, which basically treats each label as a separate single class classification problem.

The data set size of 395 samples are used , out of which 300 samples have taken for training and 95 samples have taken for testing, and success rates had predicted as pass/ fail. Here, where x is the independent feature and y's are the target variable. 'x' has three continuous internal assessment marks (t1-1,t-2,t-3) as x1,x2,x3 ,three assignments marks (a-1,a-2,a-3) as x4,x5,x6,one

attendance as x7, one regular /lateral as x8, one internal x9, one external x10 "represents the attributes or the independent variables and final decision (y) represents the target variables or the dependent variable. Its implementation in python, the multi-learn library provides the relation between x and y.

IV RESULTS& DISCUSSION

Regression machine learning systems as 'student final mark prediction', the input csv file(dataset.csv)of 500 students mark statements of multiple academic course. Out of which 385 samples have training data and 115 samples have testing data. Then the model is implemented and the predicted results are shown infig:7&   fig:8

V. CONCLUSION

Consequently, five academic courses are chosen, each constituting an individual dataset, two well-known classification (k-nearest neighbour
Hood (knn), then three different prediction model such as linear regression, k-nearest neighbourhood (knn), decision tree are experimented with. Furthermore, the datasets are enriched with demographic and in-term performance. The imbalance in the distribution of class values and less.

Special Issue - 2019

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
CONFCALL - 2019 Conference Proceedings

challenges in our present research work. Several techniques, like resampling, feature selection, and margin selection are employed to address these issues. Knn and decision tree are achieved the best prediction results, which are very satisfactory compared to those of similar approaches.

REFERENCES

[1] yuru, z., delong, c., & liping, t. (2013). The research and application of college student attendance system based on rfid technology. International journal of control and automation, 6(2), 273- 282.

[2] sunehra, d., & goud, v. S. (2016, october). Attendance recording and consolidation system using arduino and raspberry pi. In signal processing, communication, power and embedded system (scopes), 2016 international conference on (pp. 1240-1245). Ieee.

[3] sayanekar, p., rajiwate, a., qazi, l., & kulkarni, a. (2016). Customized nfc enabled id card for attendance and transaction using face recognition. International research journal of engineering and technology, 3(9), pp. 1366- 1368.

[4] noor, s. A. M., zaini, n., latip, m. F. A., & hamzah, n. (2015, december). Android-based attendance management system. In systems, process and control (icspc), 2015 ieee conference on (pp. 118-122). Ieee.

[5] kohalli, s. C., kulkarni, r., salimath, m., hegde, m., & hongal, r. (2016). Smart wireless attendance system. International journal of computer sciences and engineering, 4(10), pp. 131-137.

[6] phil simon (march 18, 2013). Too big to ignore: the business case for big data. Wiley. P. 89. Isbn 978-1-118-63817-0.

[7] ron kohavi; foster provost (1998). "glossary of terms". Machine learning. **30**: 271–274.

[8] http://www.learningtheory.org/

[9] settles, burr (2010), "active learning literature survey" (pdf), computer sciences technical report 1648. University of wisconsin–madison, retrieved 2014-11-18.

[10] rubens, neil; elahi, mehdi; sugiyama, masashi; kaplan, dain (2016). "active learning in recommender systems". In ricci, francesco; rokach, lior; shapira, bracha. Recommender systems handbook (2 ed.). Springer us. Doi:10.1007/978-1-4899-7637-6. Isbn 978-1-4899-7637-6.