PREDICTING STUDENT PERFORMANCE USING MACHINE LEARNING APPROACH

1Dhanashree Mane, BE-computer, VIIT-pune
2Pranali Namdas, BE-computer, VIIT-pune
3Pooja Gargade, BE-computer, VIIT-pune
4Dnyaneshwari Jagtap, BE-computer, VIIT-pune
5S. S. Rathi Computer Engineering, VIIT, Pune.

Abstract- In this paper we are using machine learning for predicting student performance. Machine learning is the ability of system to automatically learn from past experience and improve performance. While predicting student performance we faces many challenges such as 1) Students from different backgrounds differs in course. 2) Previous techniques are not that much accurate to predict student performance 3) Only analysis is done in previous system. In our system we are using college student data like attendance and marks of particular subject. Using these two parameters we are analyzing and predicting student performance. In this system we are using algorithms like C4.5, ANN and Reinforcement algorithms. C4.5 algorithm is used for classification purpose. ANN is used for teacher approval in our system. Reinforcement is a supervised learning. We are using Q learning which is subtype of reinforcement algorithm for assigning task to student.

Keywords — machine learning; supervised learning; reinforcement learning; decision tree; C4.5; ANN

I. INTRODUCTION

Machine learning is the ability of system to automatically learn from past experience and improve performance. Now a days machine learning for education gains more attention. We are going to use different algorithms to predict student performance such as C4.5, reinforcement, ANN, etc. Also we are going to use software tool called Weka which checks algorithm complexity and analysis prediction result. Machine learning is used for analyzing data based on past experience and predicting future performance. Reinforcement machine learning algorithms is a branch of artificial intelligence. It automatically determines the behaviour of environment and maximizes its performance. In this algorithm agent is used for taking decision. With the help of agent best action is selected. It gives feedback by adapting the environment. It is an automated learning system. In this system we give input in the terms of marks and attendance. The intention of our system is to analyse and predict performance of student that will help to improve the marks of that student. In this paper this system is implemented for helping student to improve their academic performance.

II LITERATURE SURVEY

1. A Machine Learning Approach for Tracking and Predicting Student Performance in Degree Programs.[3] Novel machine learning method is used in this paper. Bilayered structure and Ensemble predictors algorithm is used for predicting students performance in particular state. Data driven approach is used to predicting students academic performance. Input parameters given to this system is students subjects and marks which are predicting future marks using the past results of those parameters.

2. Student performance prediction using machine learning[4]: In this paper they used Neural Network algorithm to predict students performance. They used MATLAB software to develop their system which based on machine learning approach. It takes parameters such as Marks, living location, medium of teaching, background details and provide student result set Poor, average and excellent performance.
3. Predicting and Analysis of Student Performance Using Decision Tree Technique[5]: In this paper they have used only decision tree for predicting and analysing students performance by taking tests on various academic subjects. Paper needs to select subject for giving test and output calculated is Results of given test and teacher guidance.

4. Predicting postgraduate student performance using machine learning techniques: In this paper Decision tree C4.5, K-NN, Naïve bayes classifier, RIPPER algorithm, SVM, Sequential minimal optimization. The input taken by system was Students marks, result and background details predicted result of student is Poor or good performance.

5. Early Prediction of Students Performance using Machine Learning Techniques[6]: In this paper Decision tree, C4.5, Naïve Bayes, 1 Nearest Neighbourhood. Previous years marks and background details are taken as input early prediction of results.

III PROPOSED WORK

A machine learning Approach for tracking and predicting student performance in degree programs[3] in this paper we were only tracking the degree programs and not predicting the overall performance of students and also they were using a data driven approach and clustering method based on probabilistic matrix factorization this is quite difficult to predict the student performance and ensemble learning algorithms were used which is hard to implement and complex data set is allowed. The proposed system automatically handles students data to predict their performance. System predict the defaulter and non-performing students. The proposed system will give student chance to improve themselves by giving them rewards as per assigned tasks. The system automatically assigns task and provide rewards using Q-learning algorithm hence teachers time is saved and also get the detained and non-performing student list.

There are few ways to resolve this issue which are as follows: 1. using ANN machine learning algorithm reduce the time needed for execution. 2. To use the machine learning based Reinforcement Q-learning algorithm.

3. To implement the algorithms training dataset is well calculated. In this report we are proposing and implementing the various types of machine learning algorithm, that is best learning algorithms to predict the student performance

IV. SYSTEM DESIGN
Artificial Neural Network (ANN):
ANN or neural networks are a computational algorithms. It is capable of machine as well as pattern recognition. It is a structure like neurons which can compute values from input. A neural network is a machine learning algorithm based on human neuron. We can apply neural network for classification and also for regression. The neural network performs well with a linear and non-linear data. Neural network works efficiently for large and complex data and even if one or few units fail to respond to the network. For assigning tasks we used ANN.

Q-Learning:
A model used to identify action-selection policies to optimally control for any given Markov Decision Process (MDP). It learns from action-selection function, often denoted by $Q(s, a)$. A policy denoted by $\pi$, is a rule that agent follows in selection actions, given the state it is in. Q learning finds an optimal policy, it is able to compare the expected utility of the available actions without requiring a model environment.
Fig 4.2 Q-learning

Q is initialized to be possibly arbitrary fixed value at each time t action selected is ‘at’ and the reward observed is ‘rt’. Q(St, at) is old value, St+1 is new state depends upon the previous state ‘st’ and the action selected. α is learning rate (0<α≤1). If the factor is 0, it will make agent not learn. While the factor is 1, agent consider only recent information. Often, we use learning rate αt=0.1 for all t. γ is discount factor numbered between 0 and 1 (0<γ≤1). Max(Q(St, a)) is a estimate optimal future value.

Algorithm:
step 1. Initialize value table, Q(st, at) and Observe the current state as st.
step 2. Choose an action according to above values.
step 3. Take the action, observe the reward, rt, and new state st.
step 4. Update the q-value for the state using the observe reward.
step 5. Maximum reward that can possible for the new state.
step 6. Set the state to new state and repeat the process till reach to end state.

C4.5 Algorithm
C4.5 is a computer program that accepts both continuous and discrete features, it handles the incomplete data. We are using C4.5 algorithm in classification section where attendance and marks being classified further C4.5 algorithm implements decision trees which helps data get properly classified.

Algorithm Description
1. Select one instance from dataset.
2. Select initial instance from dataset and use the subset of dataset.
3. Check the correctness of dataset.
4. If all data is correctly classified then stop.
5. If dataset is not correctly classified, then go to step 1.
6. Iterate
   ➢ Till data is not get classified properly
IV. CONCLUSION

Machine learning techniques can be useful in the field of students performance prediction considering that they helps to identify from the beginning of academic year. The aim of this paper is to apply machine learning algorithms for prediction of student performance. An early analysis of student having poor performance helps the management take timely action to improve their performance through predicting their academic details. Accurately predicting student performance based on their ongoing academic records is predicted. Also we conclude that proposed system is helping us to make the student performance better. In this paper machine learning can prove to be powerful tool and all algorithms we used increases with increase in dataset size. However students’ presence in class and their attendance, marks of bachelor degree in very important for classifiers.

REFERENCES


