**LITERATURE SURVEY**

Students’ social activities and background details were used by Ching-Chieh Kiu [1] in the year 2018 for predicting the student’s academic performance. This study used several data mining technique has been applied on predicting the accuracy of each dataset, of which decision tree J48 has given the best accuracy rate. The decision tree model has achieved 95% accuracy rate compared to other data mining technique. The dataset was divided into three subsets, students background details, student’s social activities and student course work. The algorithms were implemented in WEKA tool. A comparative study on decision tree and random forest was done by Fergie Joanda and Reymon in the year 2018. The dataset was collected using questionaries’ from the computer science students. It had 249 instances out of which 3 different classes of data are used. The study proves that decision tree gives more accurate result of 66.9% compared to random forest which gave 61.44%. The comparison was implemented using WEKA tool [2].

A comparative study was done on Naïve Bayes, Decision Tree, K-Nearest Neighbor and Discriminant Analysis by Samuel, Nor Bahiah and Siti Mariyam in the year 2019. The study was done to identify the best data mining technique for predicting the student’s academic performance. It used 10 datasets from the University of California Irvine Repository. The decision tree out performed with the accuracy of 81.94% compared to other data mining techniques. The accuracy of other techniques was Naïve Bayes-73.61 %, KNN-80.56 % and Discriminant Analysis -77.78% accuracy rate. The tool used in the study was WEKA. In future hybrid metaheuristics algorithms will be for feature selection on the student data [3]. According to the study of Nongnuch Ketui, Warawut Wisomka and Kanitha Homjun in year 2019, Gradient boosted trees has given the best accuracy compared to other classification data mining techniques like Decision Tree, Weighted Decision Tree, Iterative Dichotomiser 3 (ID3) and Random Tree. WEKA is the data mining tool which is used for implementing the data mining techniques. A raw dataset was collected from the Rajamangala University of Technology Lanna Nan. The gradient boosted tree and decision tree gave good accuracy rate of 92.31% and 91.03% compared to other techniques. The Weighted Decision Tree 84.14%, ID3-89.66% and Random Tree-84.14% accuracy rate. The classification technique is widely used in predicting the student’s performance [4].

According to Romero, Cristóbal, et al [5] SMO gives more accurate result compared to other techniques like BayesNet, Naïve Bayes Simple and EM. The paper used four different datasets and each dataset produced its own accuracy. The algorithms were implemented using WEKA tool. The SMO produced 82.4% accuracy result and BayesNet produced accuracy result of 81.5% accuracy. The Naïve Bayes Simple produced an accuracy of 82.4% and EM has 80.7% accuracy rate. All the algorithms performed equally good. In the year 2014 [6], Hu, Ya-Han, Chia-Lun Lo, and Sheng-Pao Shih made a study on predict students’ online learning using C4.5, CART and LGR. The WEKA tool was used for implementation. The dataset used was learning portfolio data and C4.5 produced more accurate result of 93.4%. Yu, Liang-Chih, et al [7] used Sentiment analysis in order to predict the student’s academic performance in the year 2018. The study used unstructured dataset. The WEKA tool was used for implementing the sentiment analysis which produced an 76% accuracy rate.

According to Deepika, K., and N. Sathvanaravana Support Vector Machine produces more accurate result compare to Linear Regression and Random forest. The study used student dataset of various academic disciplines of higher educational institutions in Kerala, India. The Linear Regression produced 89.96% accuracy and Random forest produced 89.98% and Support Vector Machine produced 91.43% accuracy result [8]. A comparative study was done in the year 2018 by Uzel and Vahide Nida on Multilayer Perceptron, Random Forest, Naïve Bayes, Decision Tree and Voting classifiers. The study used an educational dataset (xAPI) which is generated  
from an e-learning system includes 480 instances and 16 attributes. The Voting classifiers has highest accuracy of 80.6% [9]. The Artificial Neutral Network outperformed the decision tree and Naïve Bayes. The study used dataset from Kalboard 360 e–learning system with 500 instances and 17 attributes. The Decision Tree has 71.1% accuracy, Naïve Bayes has 67.5% accuracy and ANN has highest accuracy of 78.1% [10]. According to Rawat, Keshav Singh, and I. V. Malhan a hybrid classification gives more accurate result for predicting the student’s academic performance. The study used data set of Department of Computer Science with 27 instances and 11 attributes. The Decision Tree produced 86.7% and KNN produced 87.5%, ANN produced 81.3% and NB produced 87.5%, Hybrid produced highest accuracy rate of 93.3%[11]