**Paper Review for USA Admission**

1. Binu P k A worked for Cloud Based Data Analysis and Prediction System for University Admission as a first author. He is from India. So his team basically predicted that total number of candidates applying for admission and total number of students from each state in the upcoming year. For this prediction they collected previous 10-15 years data from university admission details. After data collection they used Hadoop for data processing and analyzing. Here they used ANN with Backpropagation algorithm and ARIMA model for predicting but precision , accuracy are slightly better in ANN than ARIMA model. That’s why most of the time they used ANN because that was the cloud based problem where every years data was stored and forecasting for upcoming years. They claims for future they can used a hybrid combination of different models to get more accuracy.
2. Franziska T. Fischer worked for Sex-Specific Differential Prediction of College Admission Tests: A Meta-Analysis as a first author. He is a student of University of Konstanz. So his team predicted differential prediction of undergraduate and graduate college admission tests for women and men by meta analysis. This paper has two part. So the previous part was where they showed differential prediction of admission tests by gender and restrictions of these reviews. Then the present part was potential underprediction of women's academic performance and the potential overprediction of men's academic performance by undergraduate and graduate admission tests. Basically their predictions based on tests and tests combined with grades of undergraduate and graduate students. Here they used some statistical term for making decision such as t test, standard deviation , confidence interval, credibility interval etc and data collected from university websites, google scholar using search terms and different types website. And Finally they suggested to focus on sex differences in noncognitive factors like study habits and motivational factors of undergraduate students rather than on test or criterion differences for future work.
3. Abdul Hamid M Ragab worked for HRSPCA: Hybrid Recommender System for Predicting College Admission. He is from Saudi Arab. So his team proposed a system for recommendations. Basically this paper based on it. The system which name is HRSPCA consists of two cascaded hybrid recommenders working together with the help of college predictor, for achieving high performance. The first recommender assigns students tracks for preparatory year students and the second recommender assigns the specialized college for students who passed the preparatory year exams successfully. They claimed their system is adaptive, tasks faster and fairly. They used prediction algorithm which name is Cascaded Hybrid Recommender & DM for predicting with high performance. They collected data from their university portal which name is KAU. So shortly I say that they proposed a new college admission prediction technique based on using two cascaded hybrid recommenders and use of data mining discovery knowledge rules for achieving university college admission with high performance fairly and accurately. And finally they proposed this system only based on KAU data but it can applicable to other Saudi Universities.
4. Donald W. Irvine worked for Multiple Prediction of College Graduation From Pre-Admission Data. He is stayed from University of Georgia as a first author. In this paper his purpose was to investigate the practicality of predicting graduation of college students on the basis of information available prior to acceptance and matriculation. He collected data on each subjects from permanent records and questionnaires. Data features like SAT scores, high school average, third of high school graduating class, number in high school graduating class, and the number of high school units earned in English, mathematics, science, social studies, and foreign languages. He used multiple regression model for predicting and used different types of statistical term such as intercorrelation, multiple correlation coefficient , prediction probability of graduation scores, f, pearsonian correlation etc. He said for future work ‘If justified on the basis of further development

and cross-validation, it is recommended that probability of graduation data for various institutions be utilized by secondary school counselors in assisting students in making college selections’.

1. Maarten Pinxten worked for At-risk at the gate: prediction of study success

of first-year science and engineering students in an open-admission university in Flanders—any incremental validity of study strategies? As a first author. His team predicted an open admission procedure based on first-year science and engineering students data in university of Flanders. In this paper they used different types of context for open admission risk analyzing and decision making. Such as Prior achievement and course-taking in secondary school, Teacher judgements of academic achievement, Self-efficacy beliefs and self-regulatory skills, Study skills and strategies, University admission in Flanders , Advice teacher board secondary school. For this analysis part they used statistical term like ANOVA test, Pearson test, t test, P value, F, R2 etc. After analyzing they concluded that from a student guidance perspective, the latter finding is of paramount importance in an open admission system given that students cannot be formally excluded from registration on the basis of weak prior achievement. Poor achievement like a low math level background, poor math and science GPAs, a negative advice from secondary school teachers etc. So focus on a malleable skill set provides valuable training opportunities for low performing students who decide to enroll in a science and engineering program. Finally this paper limitations has they noted and that was there are more variables at stake that directly or indirectly influence student achievement at the end of the first year but this paper only focused on malleable student characteristics at the start of the academic year.

1. Joseph E Gaugler is a first author of Predicting nursing home admission in the U.S: a meta-analysis. His team basically predict that Nursing Home admission based on U.S patients. For this prediction they collected data from different types of database resource such as MEDLINE, PSYCINFO, CINAHL, and Digital Dissertations by searching. Then they collected data with using survey, clinical rating, medical charts/records, etc. Basically they divided data in different portion for analyzing and predicting such as age, annual income, gender, education , disease type etc. They followed longitudinal designs and community-based samples. Then they used Information on number of nursing home admissions, length of follow-up, sample characteristics, analysis type, statistical adjustment, and potential risk factors were extracted with standardized protocols and predicted by Random effects models which used to separately pool the logistic and Cox regression model results from the individual data sources. For analyzed purpose they used statistical term which were odds ratio, confidence interval, cognitive impairment, P value , Hazard Ratio etc. For future works they thought they used systematic data , different types of model to improve their assumption.
2. Simon Fong is a faculty of Science and Technology University of Macau and he is the first author of Applying a Hybrid Model of Neural Network and Decision Tree Classifier for Predicting University Admission. This paper based on RASU (Recommender System of Admission to University) system. RSAU system prototype was implemented in Weka which an application software written in Java open source with a collection of machine learning algorithms. So the system analyzes various sources of secondary school students’ data, to predict their chances of admission to universities. It provides decision support about recommendations to university for secondary school administrators, teachers and senior secondary students. They choose the Macau education system as an application scenario here for testing the hybrid model of admission recommender because it is one of the most complicated around the world and they used mixed types education system which were British, Portuguese and Chinese. They collected raw data from samples of 2400 secondary school students in Macau. The records were merged from three databases: the students academic year study report data, the student personal information, and the school alumni data. They use backpropagation neural network algorithm to sort out the relative important input variables from all the available variables which called feature selection. For predictions they used Decision Tree Classifier algorithm. They did not say about their work limitations or future work.
3. Sashank Sridhar worked for A University Admission Prediction System using Stacked Ensemble Learning as a first author. His team tried to understand that every year huge number of candidates applied for higher studies but high ranking universities are rejecting the maximum applications because of students lacings. That’s why in this paper they predicted university admission based on students information who are chanced in high ranking university. Here they collected data from Edulix by using the python library ‘BeautifulSoup’. The scraped data had a total of 22 features and 50, 000 samples. The dataset features such as TOEFL, IELTS, GRE, GMAT, CGPA, etc. After data collected and processing they built model for prediction. Here they used Ensemble learning method which is a type of machine learning method where multiple weak models are trained and their results are merged and used to solve a particular problem. They also used backward propagation neural networks algorithm. Here Ensemble learning method the combination of Decision Tree, Random Forest, KNN (K-Nearest Neighbours), Naive Bayes Classifier, Logistic Regression, SVM (Support Vector Machine), Linear and Quadratic Discriminant Analysis . For result analysis they used evaluation metrics such as Precision, Recall, F1-Score, and Accuracy. Their proposed method provides the best performance with an accuracy of 91%. In the future, the model can be improved as they gain more data about students. Details regarding the applicant’s Statement of Purpose essay and Letters of Recommendation can be used to improve the prediction accuracy. An alternative is to use Natural Language Processing methods to evaluate the essays and letters.
4. Amal Al Ghamdi is a first author of A Machine Learning Approach for Graduate Admission Prediction. His team used a machine learning approach which is developed to automatically predict the possibility of postgraduate admission to help graduates recognizing and targeting the universities which are best suitable for their profile. For this prediction they used three machine learning algorithm such as linear regression, decision tree and logistic regression model and data collected from Kaggle. Data features like GRE Score, TOEFL Score, University Rating, Statement of Purpose, Letter of Recommendation, CGPA, Research Experience, Chance of Admit. After evaluating model by RMSE score they said logistic regression performs better than others model. The future goal of this paper is to create software by using machine learning, especially using Logistic Regression which helped students can know the how the possibility of postgraduate admission in universities with suited their requirements.
5. Dineshkumar B Vaghela is a PhD Scholar of Gujarat Technological University, Chandkheda. He is a first author of Students' Admission Prediction using GRBST with Distributed Data Mining paper. His has one team member. So his team tried to predict students admission in college by using Global Rule Binary Search Tree (GRBST). In this paper basically used data mining procedure and data structures rules. For this paper they proposed framework of the system where have two sites which called Local Site and Remote Site. Local Site consists of Application layer, Middleware and Meta Data directory. Remote Site consist Data warehouses at different servers. They did not collect any data here but they said their training data will done by admin or automatically once a year as there will be no major changes in the data and thus it will not affect the decision making and testing phase can be done by users and admin as well where they can enter data for single user or multiple data in the provided format at their site. After then they used Consolidation Technique and Binary Search Tree Construction for making decision . The time complexity of generating the Binary Search Tree from the Decision table is very less and also this BST has efficient time complexity to predict the result. In this paper they showed four columns as instance and there were parul, SVIT, BVN and BABARIA. So the prediction for first 3 instances are parul, SVIT and BVN respectively, and probability of predicted value is 100% and BABARIA’s predicted value is 80%. In future the security parameters can also be considered from them .
6. Null
7. T.Stanko is a first author of On possibility of prediction of academic performance and potential improvements of admission campaign at IT university paper. He has three team members for this paper. Their papers goal was evaluating the admission procedure, in terms of possibility to adequately predict a student’s future academic performance based on available data. They collected students admission and academic performance data. For checking this data statistically significant or not they used pearson correlation on existing data including admission and mid-semester grades. But there has low correlation between admission and academic data. After using PCA and correlation analysis Authors used two types of models linear regression and logistic regression model. Linear regression models provide a prediction of grade, based on admission data, whereas logistic regression forecasts the probability of a D grade in a given course for each student. For linear and logistic regressions respectively R2 metric and Accuracy, F1 score was used. After analysis they concluded students previous academic success obviously needed for admission as a prerequisite but it did not justify students will be successful in university academic performance. About future works they said their further research focus will be the possible influence of students expectations and intrinsic motivation on academic success at the university.
8. I think 13 numbers paper is not relatable with our works. That’s why I skip this.
9. Predicting the academic success of architecture students by pre-enrolment requirement: using machine-learning techniques paper written by Ralph Olusola Aluko as a first author. His three team members worked for this. They predicted two things here. First one is developed models for predicting academic success of undergraduate architecture students using prior academic performance as predictors and second is identify the most important predictors which have a significant impact on academic success. For predicting they used two types of algorithm which are k-NN algorithm and linear discriminant analysis and collected data from the Department of Design and Architecture, Olabisi Onabanjo University, Ogun State, Nigeria. The collected data contained information on 102 students that completed the undergraduate program between 2011 and 2014. This dataset carried 13 features and 1 target variable. After using two algorithms they saw KNN worked with better accuracy than linear discriminant analysis. Here linear discriminant analysis to predict the academic success of architecture undergraduate students (‘Pass’ or ‘Fail’) with 50% classified accuracy and KNN with 73.33% accuracy. Then they used sensitivity analysis technique to extract additional information on the importance of each independent variable in predicting the dependent variable in machine-learning models. In the end, they concluded prior academic achievement is a significant predictor of academic success in this undergraduate architecture program in Nigeria and prior academic performance in mathematics, physics, chemistry, and local language are significant determinants of academic success in the undergraduate architecture program. For future works they added another criteria like learning style, teaching methods in this area for better predictions.
10. A Graduate School Recommendation System Using the Multi-Class Support Vector Machine and KNN Approaches paper written by Alisha Baskota as first author. She has one team member for this work. In this paper they tried to proposed a recommendation system that eliminates the tedious application search process imposed on graduate school applicants by designing a single platform which can shortlist the universities/colleges appealing to the applicants. For this works they collected data from educational portals, such as Edulix.com and Yocket.com. Here they collected two types datasets. The first dataset on various universities/colleges, called Grad Sch, is generated by combining data and is used by the KNN algorithm. The second dataset, called Std Info, consists of former graduate school applicants data used by multi class SVM algorithm. After data processing and scaling they used KNN algorithm for feature selections under the forward wrapper method. Before selecting SVM algorithm they compared with other machine learning algorithm which are Decision tree, Multi-layer perceptron and Naïve bayes algorithm. In the end they used SVM with KNN algorithm together generate a number of graduate schools which are the actual ranked recommendations and they claimed their recommender is effective and advantageous as the suggestions generated by recommendation system are relevant, and the user study has also verised that they are accurate.
11. Null
12. Criteria evaluation and selection in non‐native language MBA students admission based on machine learning methods paper written by Xiaojun Wu as a first author. He has only one team member who was a second author here. Basically, this paper worked on data of Shanghai International MBA Program in China. From 2007 to 2014 enrolled students data used here as a sample. In this paper they claimed English is an important language to admit and introduce yourself. So who can speak better English, who got the better opportunity for MBA program. For this paper they collected 596 students data from SIMBA who successfully applied for admission and completed the MBA program. After the above data preprocessing, 549 valid sample data were obtained. Then they used three research methods here. So 1st one was two-sample test of a hypothesis, 2nd was multiple linear regression analysis and the last was machine learning algorithms (Ridge regression, SVM, Random forest, GBDT). Also they used different types of statistical formula such as Descriptive statistics analysis, T test, ANOVA test etc. They said two hypothesis like Mean GPA for students who have high math exam grades and Mean GPA for students who have low math exam grades. All of the hypothesis performed 95% level of confidence by T test. Second research method based on multiple linear regression which is provide Sum of squares, df, Mean Square, F, Sig. by ANOVA testing. And last research method in this paper is machine learning algorithms (Ridge regression, SVM, Random Forest, GBDT) which result based on Mean absolute error, Root mean squared error, The optimized parameters through parameters selection in Weka parameter. Lastly they said the shortcomings of this study are mainly refected in the following: the criteria used in this study are highly representative in the selection of MBA students but do not include all the widely used criteria and In the future, further consideration can be given to the predictive efects of these criteria.
13. A Comparison of Regression Models for Prediction of Graduate Admissions paper written by Mohan S Acharya as a first author. He works under Dept. of ECE, National Institute of Engineering, Mysuru in India. He has two team member who works under same institute. Basically, they predicted graduate students admission in MSc programs. Because after graduation some students can no idea what will be the best choice for applying from their past skill and experience. So this paper helped to them for making the right decision. For analyzing they collected data from Kaggle and prepared typically from an Indian student’s perspective. Datasets features was GRE, TOEFL and Undergraduate GPA. Statement of Purpose and Letter of Recommendation are two other important entities. A unique feature of this dataset is that it contains equal number of categorical and numerical features. Then they used four regression methods which were Linear Regression, Support Vector Regression, Decision Tree and Random Forest for error case calculating. They checked evaluation metrics like Mean Squared Error (MSE), Root Mean Squared, Error Mean Squared Log Error and R2 Score for all algorithms. After evaluation they declared that Linear Regression performs the best on their dataset, with a low MSE and high R2 score, closely followed by Random Forest. For future woks they said, they developed a plausible solution for the problem keeping in mind the various factors that affect the chances of admission and also used Deep Neural Networks as another plausible model to understand the subjective nature of admission.
14. Ioannis Hatzilygeroudis was a first author of this paper which name is PASS: An Expert System with Certainty Factors for Predicting Student Success. He has more than two members for this works. From this paper we saw a keyword which name is PASS (Predicting Ability of Students to Succeed). PASS is a rule based system that uses a type of certainty factors which is used to predict how certain is that a student of a specific type of high school in Greece will pass the national exams for entering a higher education institute. Here they used two methods, an expert system approach and a well-known statistical method, namely logistic regression, to achieve their objective. For this works they collected data from students database in technical and vocational secondary education in Greece which contained 201 records of students, who took the entry examinations during the last three years. Datasets features were sex, age, specialization, grade A (average mark of all first year courses), grade B (average mark of all second year courses), grade SC (average mark of the three courses to be examined in the national exams at the end of the first semester of the third year) etc.
15. Graduate Admission Prediction Using Machine Learning paper written by Sara Aljasmi as a first author. He has three member for this works. The goal of their paper was to assist students in knowing in advance if they have a chance to get accepted into the MSc program abroad. So for this prediction they used some machine learning algorithm such as multiple linear regression, k-nearest neighbor, random forest and Multilayer Perceptron and collected 500 raw data from admission dataset. Datasets features like GRE, TOFEL, University ratings, SOP, LOR, CGPA out of 10 and Research experience and the target variable was Chance of admission. After data division, they applied feature selection by linear regression equation because of all variable were numeric. Then they used some statistical analysis such as descriptive analysis for checking features importance, Independent variables importance by using Random Forest classifier, Histogram which presented the frequencies of continuous numbers and to show the distribution of the data selected, Shapiro-Wilk Normality Test checked P value, Multicollinearity Issue checked correlation between one or more features and lastly Regression model which explained 83% variation in their dataset and P value less than 0.05 which rejected null hypothesis and showed the model is statistically significant. After statistical analysis they used ML model through WEKA and decided which model performed best based on MAE value. According to result, multilayer perceptron has the smallest MAE equivalent to 3.37% which means that it was the best model. For the future works they said, more models can be conducted on more datasets to learn the model that gives the best performance.