

---

# LITERATURE SURVEY

## UNIVERSITY ADMIT PREDICTOR

### DOMAIN-APPLIED DATA SCIENCE

---

SEPTEMBER 24

---

Authored by: G.HARSHINI NAIDU

B.SAI SRAVANYA

CH.HIMABINDU

B.EEKSHITHA

---

## Paper 1 - University Admissions Predictor

Publisher-Research Gate.

Reference-

[https://www.researchgate.net/publication/345391208 University Admissions Predictor](https://www.researchgate.net/publication/345391208_University_Admissions_Predictor)

### ADVANTAGES-

- Provide customers access to the prediction model.
- Provide answers to most common FAQs regarding PG Admissions abroad.
- Provide administrator access to all records.

Provide analysis of how the various academic factors affect university admission.

### DISADVANTAGES-

- The system is built on a limited data set, this could affect the accuracy of the predictions as a whole.
- The system cannot guarantee that our predictions will be a 100% guarantee of admissions because a lot other factors such as the Personal Interview also plays a major role in the admissions procedure.
- Other factors such as changes in policies by the university or by the country can also affect chances of admissions in a way that is beyond the scope of this project. Admissions also depend on the individual university's policy regarding the intake of foreign students and is not modelled by our system.

---

## Paper 2 - U Graduate Admission Prediction Using Machine Learning Techniques

Publisher-International Journal of Advanced Research in Science, Engineering and Technology.

Reference-

<http://www.ijarset.com/upload/2021/july/07-mail2vkk-07.PDF>

### ADVANTAGES-

- A comparative approach by developing four machine learning regression models: linear regression, support vector machine, decision tree and random forest for predictive analytics of graduate admission chances.
- A user interface is provided through which an actor can interact with the system. The algorithm with improved accuracy will act as a backend for the user interface.

### DISADVANTAGES-

- Many machine algorithms are used to the prediction of Graduate Admission. The existing system compares the four machine learning algorithms on the basis of accuracy.
- The algorithms are Linear Regression, Support Vector Regression, Random forest Regression, Decision Tree Regression. In this system Linear Regression performs the best on the dataset with low MSE and high R2 score.

---

## **Paper 3- Product Recommendation using Machine Learning Model.**

Publisher-Research Gate.

Reference-

[https://www.researchgate.net/publication/317399986\\_Product\\_Recommendation\\_using\\_Machine\\_Learning\\_Model](https://www.researchgate.net/publication/317399986_Product_Recommendation_using_Machine_Learning_Model)

### **ADVANTAGES-**

- Using the Linear Regression Methods we are identifying which one is optimized one from the dataset of the laptop of different types processor (i3, i5, i7) and it depends on the customer rating and Price. From the linear regression method of above Table V, the machine takes the optimized cost value by using machine learning method.
- The variable we are foreseeing is known as the foundation variable and is alluded to as Y. The variable we are constructing our forecasts in light of is known as the indicator variable and is alluded to as X. At the point when there is just a single indicator variable, the expectation strategy is called basic relapse.

---

## **Paper 4 - A Review on Data Mining and Machine Learning Methods for Student Scholarship Prediction.**

Publisher- IEEE.

Reference-

<https://ieeexplore.ieee.org/document/9418376>

### **ADVANTAGES-**

- This review paper consists of literature survey to prediction of scholarship by using Machine Learning and Data Mining technique.
- Along with this it contains a small description of ML/DM which are used by the researchers. It also describes data sets as very important in ML/DM methods. Machine Learning becomes most popular in the field of IT industry.
- Nowadays Machine Learning and Data Mining turn as a powerful technique which applicable for various fields such as IT, Education sector and also in business sector too.
- The different types of ML/DM algorithms are addressed by using all this technique. The algorithms which give more accuracy results in detection of continuity of every student's scholarship such as Naïve Bayes, Decision Tree and k-NN.
- Finally, the proposed model will provide a list of candidates, who deserve to have a scholarship and also discussion has been made on accuracy of each techniques which was used to get a result.

---

## Paper 5- Survey of Pre-processing Techniques for Mining Big Data.

Publisher- IEEE International Conference on Computer, Communication, and Signal Processing (ICCCSP-2017).

Reference-

<https://ieeexplore.ieee.org/document/7944072>

### ADVANTAGES-

- Data cleaning addresses a variety of data quality problems, that includes noise, outliers, inconsistent data, duplicate data, and missing values. In this survey, a brief analysis of the existing work in data cleaning techniques are performed. In summary, data cleaning techniques developed at the data collection stage are focused on detecting and removing low level errors and inconsistencies due to an imperfect data collection process and type of data whether homogeneous or heterogeneous.

### DISADVANTAGES-

- Data cleaning is more challenging in data relevant to fields of forensics, medicine, educational models, oceanographic data, weather forecasting, electricity consumption, transportation, cyberspace, air traffic control, banking and e commerce data models the rate of error finding is very low.
- Wrapper method is cannot be employed for a large and diverse dataset due to exhaustive search involved in the process. Due to large variation in data sets and the data types the error correction must be enhanced, in other words our data cleaning must be done on different levels after taking the appropriate sampling rates. Moreover it is a very costly process for large data sets and also for each evaluation, it is required to learn and validate a model. It is prohibitive to complex classifiers.

---

## Paper 6- Data Science – Cosmic InfoSet Mining, Modeling and Visualization.

Publisher- 2018 International Conference on Computational and Characterization Techniques in Engineering & Sciences (CCTES) Integral University, Lucknow, India, Sep 14-15, 2018.

Reference-

<https://ieeexplore.ieee.org/document/8674138>

### ADVANTAGES-

- Amazon prime and Netflix data mines movie interest patterns to analyze on what movie cards a user is interested in, and then uses all the information to predict and generate the movie lists.
- Targets features i.e., what are major range of customer within its base and the unique shopping interests within those group range. This helps them in guiding to message to different market group of audiences.
- Gamble and proctor utilizes time series models to more lucidly and intelligibly stats future need, which helps in planning for optimum production levels. Amazon and Flipkart uses recommendation engines for spotting the products, so that it can put the product to remain in the user's vision, using algorithms. Spotify uses algorithms to recommend songs to the user.
- Spam filter of Gmail works with the algorithm for the junk mails and put accordingly the spam, junk and not junk mails in the distinct folders.

### DISADVANTAGES-

- Technology and Hacking  
At the beginning, it should be cleared that we are not talking about hacking as in making the information as key by getting into computers. We're referring to the technical coder subculture meaning of hacking – i.e., creativity and inventive in using technical skills to create or generate things and finding tactical solutions to problems as expressed.

---

## **Paper 7- A University Admission Prediction System using Stacked Ensemble Learning.**

Publisher-IEEE

Reference-

<https://ieeexplore.ieee.org/document/9213205>

### **ADVANTAGES-**

- With an increasing number of students looking to pursue a degree, the domain of university admission prediction gains more relevance. In this paper, an effective method has been proposed to predict the chances of a student being admitted to a specific university. In addition, we have compared the performance of various machine learning algorithms to the proposed method in predicting admits. It is observed that the proposed method provides the best performance with an accuracy of 91%.

### **DISADVANTAGES-**

- The model can be improved as we gain more data about students. Details regarding the applicant's Statement of Purpose essay and Letters of Recommendation (a score can be assigned to rate these) can be used to improve the prediction accuracy. An alternative is to use Natural Language Processing methods to evaluate the essays and letters. In addition to this, the weightage given to the features can be varied according to past trends.



---

## Paper 8- Multi Disease Prediction Model by using Machine Learning and Flask API.

Publisher-IEEE

Reference-

<https://ieeexplore.ieee.org/document/9137896>

### ADVANTAGES-

- In multi disease model prediction, it is possible to predict more than one disease at a time. So user no need to traverse many models to predict the diseases. It will reduce time and also due to predicting multiple diseases at a time there is a chance of reducing mortality rate.
- The advantage of multi disease prediction model in advance can predict the probability of occurrence of various disease and also can reduce mortality ratio.

### DISADVANTAGES-

- Many of existing analysis involved analyzing particular disease. When a user wants to analyze diabetes needs to use one analysis and same user wants to analyze heart disease then user has to use one more model. This is a time taking process. And also if any user having more than one disease but in existing system if it is able to predict only one disease then there is a chance of mortality rate increase due to not able to predict the other disease in advance.