2022 IEEE 7th International conference for Convergence in Technology (I2CT)
Pune, India. Apr 07-09, 2022

Personalized College Recommender and Cutoff Predictor for Direct Second Year Engineering

Abdul Majeed Inamdar
Department of Information Technology
Fr. C. Rodrigues Institute of Technology
Navi Mumbai, India
inamdarmajeed94@gmail.com

Tanmay Mhatre
Department of Information Technology
Fr. C. Rodrigues Institute of Technology
Navi Mumbai, India
tanmaymhatre827@gmail.com

Pravin Nadar
Department of Information Technology
Fr. C. Rodrigues Institute of Technology
Navi Mumbai, India
13pravinnadar@gmail.com

Supriya Joshi
Department of Information Technology
Fr. C. Rodrigues Institute of Technology
Navi Mumbai, India
supriya.joshi@fcrit.ac.in

2022 IEEE 7th International conference for Convergence in Technology (I2CT) | 978-1-6654-2168-3/22/$31.00 ©2022 IEEE | DOI: 10.1109/I2CT54291.2022.9825378

***Abstract*—In Maharashtra there are around 342 polytechnics providing 110 engineering courses wherein throughout every year approximately 90,000 to 1.5 lakhs students pass out. In addition, there are only 338 engineering colleges that provide Direct Second Year Engineering (DSE) admissions for B.E and B.Tech studies. However, the engineering admission process is hectic and more than that is to find the college according to student preference i.e stream, location, university and seat type, etc. During the admission process the students always need to check the previous years cutoff of each college by visiting numerous websites & pdf lists which certainly consumes a lot of time. In our web application we will provide cutoff prediction of each college by the data-analysis of previous years cutoff, recommendation system for colleges listing according to student preferences, furthermore providing detailed comparison between institutions of their choice. The application is developed to bestow a personalized system so as to reduce time and ease the student's college selection process.**

Keywords—ML(machine learning), Direct Second year Engineering (DSE), Prediction Analysis, Recommendation System, Prediction System, CAP (Centralized Admission Process).

I. INTRODUCTION

The engineering admission process is hectic and on top of that is to find the college according to student preferences i.e course, location, university and seat type, etc. of their choice. During the admission process the students always need to check the previous years cutoff of each college by visiting numerous websites & pdf lists which certainly is time consuming like it takes around 1 or 2 days to make a proper institutes list. In addition, there are existing systems(college recommendation) in the market for students who need to get admission in the first year of engineering, yet there is no such system for the direct second year students who pass out from Diploma. Furthermore, the existing system does not provide a prediction analysis for particular institute cutoffs for the upcoming admission process.

The Personalized College Recommender and Cutoff Predictor for Direct Second Year Engineering is a college recommender and prediction system based web application, that helps pass out diploma students to choose colleges according to their percentage, merit-list, course, seat type, location and university ,etc. in order to get admission in colleges for Direct Second Year of Engineering. The web application has two systems i.e. Prediction-analysis and Recommendation system. The prediction-analysis system with the help of ML(Machine Learning) is used to predict the upcoming year cutoffs for a particular college by analyzing the cutoffs of previous few years of that particular college. The second one i.e. Recommendation system helps students to get a list of colleges as per their preferences(i.e percentage, course, seat type and location ,etc.).Subsequently, the recommendation system also provides a comparison between the colleges based on college details,courses offered, and infrastructure, etc. .The main objective of the project is to reduce time and ease the student's college selection process.

II. REQUIREMENTS

The data i.e.cutoffs percentage of previous years and college details are the main requirement for proper working of our system. The college recommendation system requires data from the latest CAP round containing percentage, seat-type, college name, college location and course name, etc. for college listing, this data is extracted from pdf provided by the mahacet.org for every college and is stored into a database. The college details are extracted from college's and other educational websites. The cutoffs prediction system requires data from previous years cutoffs for proper prediction analysis which is hardly available on some websites over the internet, that is not in proper format. The data used while preparing dataset is quiet hard to collect as data need to be extracted from the pdf files i.e. downloaded through mobile application(DSECutoff) which is further converted into csv file and these csv files of different colleges are then merged into one dataset(for model training) manually. Dataset contains percentage(CAP round 1 of 2018 up till CAP round 2 of 2020), seat-type, course name and college name as main features of our dataset required for prediction of future cutoffs.

III. RELATED WORK

[1] In the paper "Prediction and Analysis for Students' Marks Based on Decision Tree Algorithm." by Zhiwu Liu and Xiuzhi Zhang(IEEE) proposes, using decision tree algorithm C4.5 to establish a classification rule and an analysis-forecasting model for student's marks.Describing how the analysis-forecasting result can be used to find out the factors which can affect students' marks, so some negative learning habits or behaviors of students can be revealed and corrected in time. The effectiveness and

correctness of analysis and forecasting model and classification for students' marks based on decision tree algorithm C4.5 has been examined by an example.This system only provided Students marks and the accuracy of forecasting model is less than 80%.

[2]The "College Admission Predictor" by Annam Mallikharjuna Roa and Et al. proposes a web based application system in which students can register their marks along with their personal information. The main advantage of the project is the computerization of the entrance seat allotment process. Using this Application, the entrance seat allotment became easier and can be implemented using system Whatever may be their scores , this application helps to find the best colleges.The main objective of this system was to make the right choice of colleges.The proposed system only performed the seat allotment process.

[3]The "StudieMe: College Recommendation System" by Vidish Sharma and Et al. proposed a novel web platform for a college selection process. Having a recommendation system as a helping hand to give them detailed information about the options they have and the best options to choose from according to their caliber is a huge requirement. In this paper, they worked on designing a recommendation system that could understand the skill set and interest of a user through the data from the User's Profile to suggest recommended options of colleges for the users to select. They have developed the college recommendation system as a web platform which gives the result as top matched colleges for a particular user. In this paper, recommendation system is not based on cutoffs percentage.

[4]"HRSPCA: Hybrid recommender system for predicting college admission" by Abdul Hamid M Ragab ,Abdul Fatah S. Mashat and Ahmed M Khedra(IEEE) proposed a new college admission system using hybrid recommender based on data mining techniques and knowledge discovery rules, for tackling college admissions prediction problems.The proposed HRSPCA system consists of two cascaded hybrid recommenders working together with the help of college predictor, for achieving high performance.The first recommender assigns student's tracks for preparatory year students while the second recommender assigns the specialized college for students who passed the preparatory year exams successfully.The system analyzes student academic merits, background, student records, and the college admission criteria. Then, it predicts the likelihood of university college that a student may enter. In this paper, the recommendation system is not based on cutoffs percentage.

IV. Proposed Approach

A web application is developed using Django along with Mysql for storage and retrieval of data. The application consists of two system's(Recommendation and Prediction) so the solution methodology for both the systems are explained below.

The college recommendation system needs to provide a list of colleges based on student percentage and preferences from a large amount of data. The entities of the table are B-tree indexed for fast retrieval of data. B-Tree is a self-balancing search tree.Search is similar to the search in Binary Search Tree. Let the key to be searched be k. We start from the root and recursively traverse down. For every visited non-leaf node, if the node has the key, we simply return the node. Otherwise, we recur down to the appropriate child (The child which is just before the first greater key) of the node. If we reach a leaf node and don't find k in the leaf node, we return NULL. Searching for a B-Tree is similar to searching a binary tree. The algorithm is similar and goes with recursion. At each level, the search is optimized as if the key value is not present in the range of the parent then the key is present in another branch. As these values limit the search they are also known as limiting value or separation value. If we reach a leaf node and don't find the desired key then it will display NULL.The time complexity for search operation using B-tree indexing is O(log n). Furthermore after the list is generated students can change the order of the list according to their needs. For this functionality we have used Jquery(Sortable library).Subsequently for downloading the list into pdf we have used Javascript (jsPDF library).

The cutoffs prediction system predicts cutoffs of particular colleges by performing prediction analysis from previous years cutoffs percentage.For the prediction analysis we have used Linear regression on the dataset. Linear regression is one of the most commonly used predictive modelling techniques.It is represented by an equation $Y = a + bX + e$, where a is the intercept, b is the slope of the line and e is the error term. This equation can be used to predict the value of a target variable based on a given predictor variable(s).

The dataset used for both the systems are collected from the DSECutoff mobile application. Fig 1 shows the entire process of data collection.

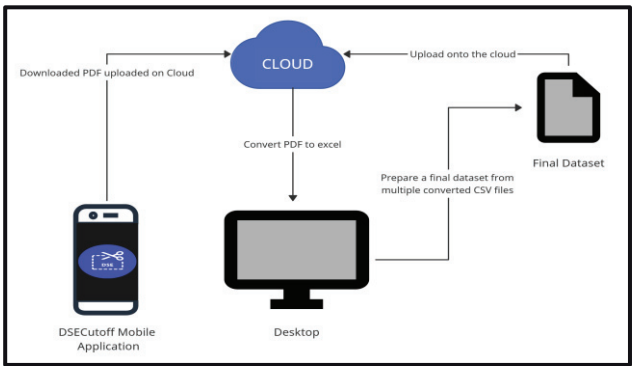


Fig. 1. Data-collection process for recommendation & prediction system.

V. System Architecture

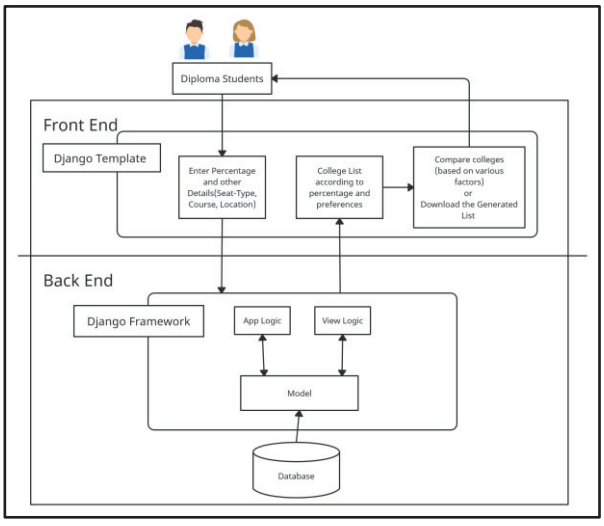


Fig. 2 Architecture Diagram of Recommendation System

The architecture diagram of both the system contains Django Template as their Frontend, Template is a presentation layer which handles User Interface part completely that takes students percentages & preferences,etc. as a input and these request is processed at the Backend by Django framework wherein, View is used to execute the business logic and interact with a model to carry data and renders a template. The working of each system is explained in the flowchart.

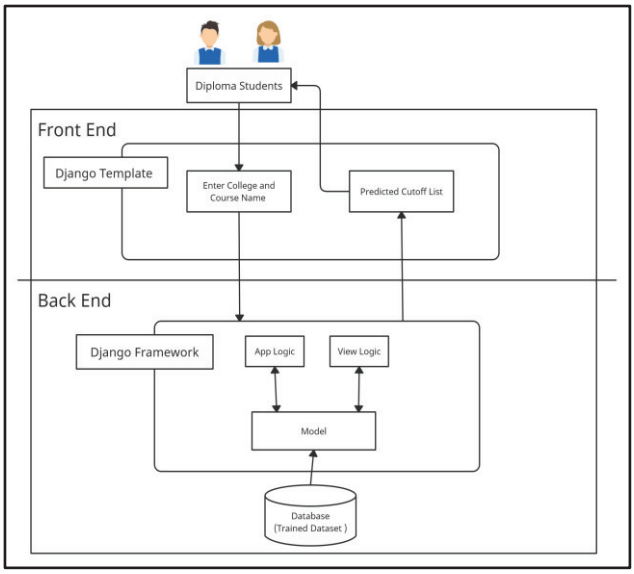


Fig. 3 Architecture Diagram of Prediction System

VI. EXPERIMENTAL RESULTS

The experimental result for the recommendation system is represented from (Fig. 4 to Fig. 7) given below. (Fig 4.) is the College recommender page wherein the student needs to provide the percentage range, details(seat type & minority status) and preferences(college location, courses name & college status). The form contains certain required fields and certain optional fields. Moreover, for efficient list generation, except percentage and location we have default values for the remaining fields. After the input and submission from student the list is generated as shown in (Fig. 5), student can also change the order of the list simply by dragging up and down the row with the help of jQuery (sortable library).Once the list is generated and student has rearranged the list, he/she can even download the list into a PDF format with the help of JavaScript(jsPDF library). The downloaded list is shown in the (Fig. 6). Subsequently, (Fig. 7) is the college comparison feature which comes under the recommendation system wherein the student needs to enter the college name for comparison.

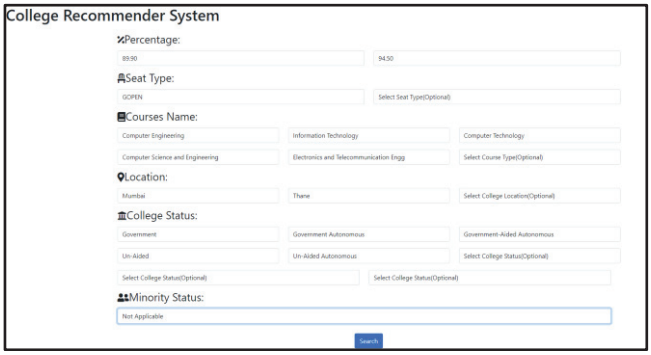
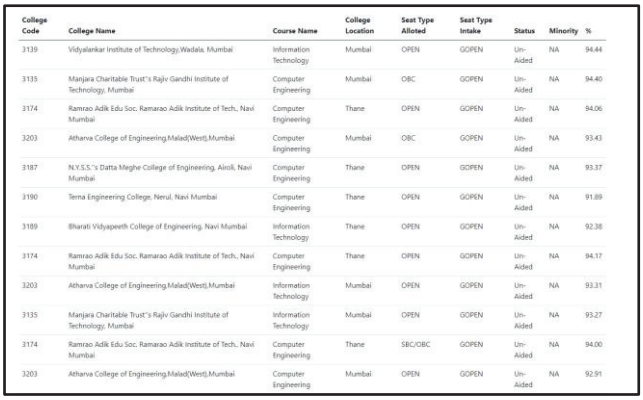


Fig. 4 Student Input Page(College Recommender)



College Code	College Name	Course Name	College Location	Seat Type Alloted	Seat Type Intake	Status	Minority	%
3139	Vidyalankar Institute of Technology,Wadala, Mumbai	Information Technology	Mumbai	OPEN	GOPEN	Un-Aided	NA	94.44
3135	Manjara Charitable Trust's Rajiv Gandhi Institute of Technology, Mumbai	Computer Engineering	Mumbai	OBC	GOPEN	Un-Aided	NA	94.40
3174	Ramrao Adik Edu Soc, Ramarao Adik Institute of Tech., Navi Mumbai	Computer Engineering	Thane	OPEN	GOPEN	Un-Aided	NA	94.06
3203	Atharva College of Engineering,Malad(West),Mumbai	Computer Engineering	Mumbai	OBC	GOPEN	Un-Aided	NA	93.43
3187	N.Y.S.S.'s Datta Meghe College of Engineering, Airoli, Navi Mumbai	Computer Engineering	Thane	OPEN	GOPEN	Un-Aided	NA	93.37
3190	Terna Engineering College, Nerul, Navi Mumbai	Computer Engineering	Thane	OPEN	GOPEN	Un-Aided	NA	91.89
3189	Bharati Vidyapeeth College of Engineering, Navi Mumbai	Information Technology	Thane	OPEN	GOPEN	Un-Aided	NA	92.38
3174	Ramrao Adik Edu Soc, Ramarao Adik Institute of Tech., Navi Mumbai	Computer Engineering	Thane	OPEN	GOPEN	Un-Aided	NA	94.17
3203	Atharva College of Engineering,Malad(West),Mumbai	Information Technology	Mumbai	OPEN	GOPEN	Un-Aided	NA	93.31
3135	Manjara Charitable Trust's Rajiv Gandhi Institute of Technology, Mumbai	Information Technology	Mumbai	OPEN	GOPEN	Un-Aided	NA	93.27
3174	Ramrao Adik Edu Soc, Ramarao Adik Institute of Tech., Navi Mumbai	Computer Engineering	Thane	SBC/OBC	GOPEN	Un-Aided	NA	94.00
3203	Atharva College of Engineering,Malad(West),Mumbai	Computer Engineering	Mumbai	OPEN	GOPEN	Un-Aided	NA	92.91

Fig. 5Generated List Page(College Recommender)

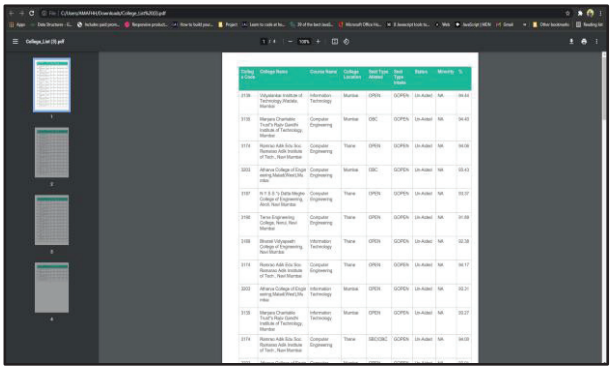


Fig. 6 Downloaded List(College Recommender)

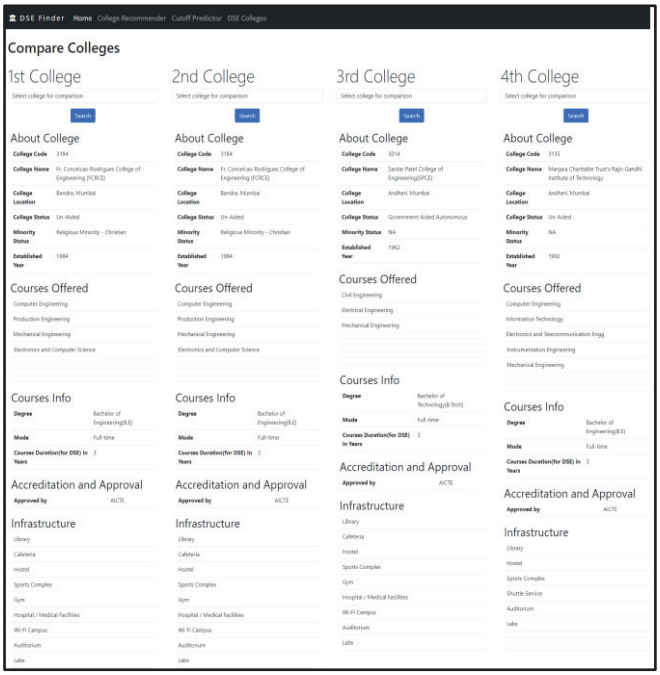


Fig. 7 College Comparison Page(College Recommender)

The experimental result for the prediction system is represented from (Fig. 8 to Fig. 11) given below. (Fig. 8) is the dataset prepared for Model training. We have used Linear regression for predicting the cutoffs. Before training the model we cleaned the data using Pandas library for eliminate any NULL values as shown in (Fig. 9). (Fig. 10 & Fig. 11) provides the accuracy for the trained model and predicted cutoff using linear regression respectively.

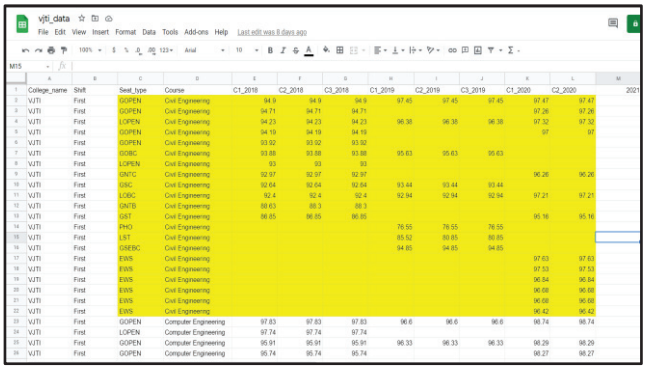


Fig. 8 Dataset prepared for Model Training(Cutoff Prediction)

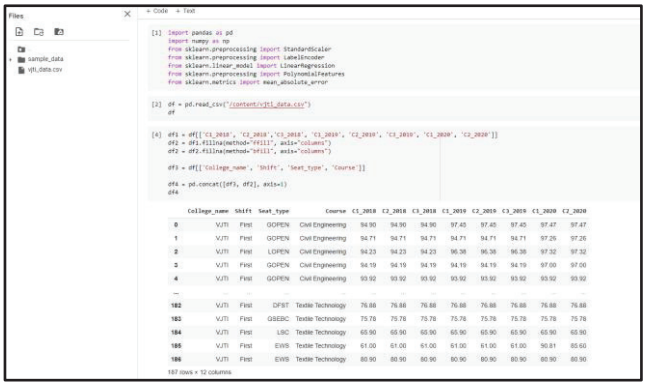
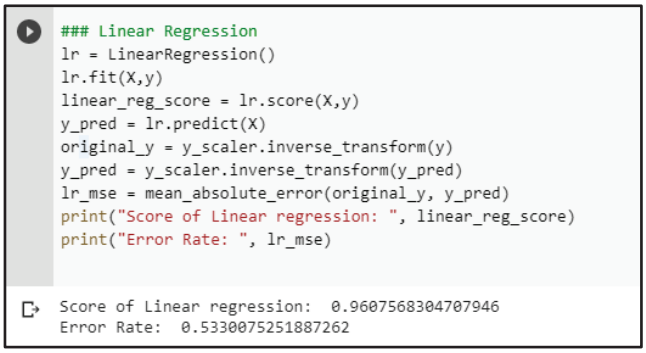


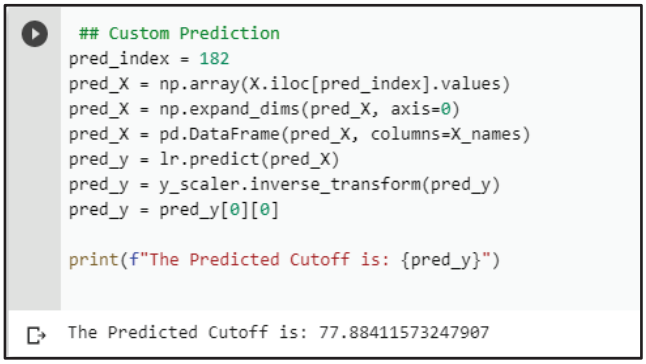
Fig. 9 Data cleaning using Pandas(Cutoff Prediction)



```
### Linear Regression
lr = LinearRegression()
lr.fit(X,y)
linear_reg_score = lr.score(X,y)
y_pred = lr.predict(X)
original_y = y_scaler.inverse_transform(y)
y_pred = y_scaler.inverse_transform(y_pred)
lr_mse = mean_absolute_error(original_y, y_pred)
print("Score of Linear regression: ", linear_reg_score)
print("Error Rate: ", lr_mse)
```

```
Score of Linear regression:  0.9607568304707946
Error Rate:  0.5330075251887262
```

Fig. 10 Accuracy of Trained Model(Cutoff Prediction)



```
 ## Custom Prediction
pred_index = 182
pred_X = np.array(X.iloc[pred_index].values)
pred_X = np.expand_dims(pred_X, axis=0)
pred_X = pd.DataFrame(pred_X, columns=X_names)
pred_y = lr.predict(pred_X)
pred_y = y_scaler.inverse_transform(pred_y)
pred_y = pred_y[0][0]

print(f"The Predicted Cutoff is: {pred_y}")
```

```
The Predicted Cutoff is: 77.88411573247907
```

Fig. 11 Predicted output using Linear Regression(Cutoff Prediction)

VII. CONCLUSION

Personalized College Recommender and Cutoff Prediction for Direct Second Year Engineering is built with an objective to bestow a personalized system so as to reduce time and ease the Direct Second Engineering(DSE) student's college selection process.The project overcomes the limitation of existing systems that doesn't provide a prediction analysis of particular college cutoffs for the upcoming admission process.Moreover, it is the first system that recommends colleges for diploma students who require admission in Direct Second Engineering(DSE). In recommendation system, the web application provides a list of colleges based on students preferences and generated list is downloadable at one click from the user.Furthermore students can also compare colleges base on college information(college location,degree type, mode, course duration), courses offered, accreditation and infrastructure,etc. . The prediction system provides the predicted cutoff for the upcoming years of a particular college.

REFERENCES

[1] Zhiwu Liu & Xiuzhi Zhang "Prediction and Analysis for Students' Marks Based on Decision Tree Algorithm." IEEE Third International Conference on Intelligent Networks and Intelligent Systems by Zhiwu Liu and Xiuzhi Zhang(IEEE), 2010

[2] Annam Mallikharjuna Roa , Nagineni Dharani , A. Satya Raghava , J. Buvanambigai , K. Sathish "College Admission Predictor" Journal of Network Communications and Emerging Technologies (JNCET), 2018 .

[3] Vidish Sharma, Tarun Trehan, Rahul Chanana and Suma Dawn"StudieMe: College Recommendation System" IEEE 3rd International Conference on Recent Developments in Control, Automation & Power Engineering (RDCAPE), 2019.

[4] Abdul Hamid M Ragab ,Abdul Fatah S. Mashat and Ahmed M Khedra "HRSPCA: Hybrid recommender system for predicting college admission." IEEE 12th International Conference on Intelligent Systems Design and Applications (ISDA), 2012 .

[5] Shaobo Huang & Ning Fang "Regression Models of Predicting Student Academic Performance in an Engineering Dynamics Course "

[6] Prediction using Supervised ML (Prediction of Marks)

[7] Zafar Iqbal, Junaid Qadir, Adnan Noor Mian & Faisal Kamiran "Machine Learning Based Student Grade Prediction: A Case Study"

[8] DSECutoff Mobile Application: https://play.google.com/store/apps/details?id=com.rohit.suthar.dsecutoff&hl=en_IN&gl=US