Machine Learning Model as College Recommendation System, Admission Prediction System and Career Counseling

By

Yash Suhas Mayur

4 August 2023

Abstract:

Educational organizations are one of the important parts of our society and playing a vital role for growth and development of any nation. For that getting appropriate college is of foremost importance. Many elements influence this choice for students, including financial, social, demographic, and cultural factors. If a student makes a poor choice, it will have implications for their academic life as well as their professional life. These implications may include having to change their major, which will cause a delay in their graduation, having a low grade-point average (GPA) in their chosen major, which will cause difficulties in finding a job, or even dropping out of university. The purpose of this project is to develop a system which is reliable, convenient, money saving for students and robust to any failures.

1) Problem Statement:

1.1) Problem for Student:

- In India, especially for 10th passed student have the problem of selecting field of interest.
    - If they secure higher marks their parents force them to take science field for more secure future or a commerce field to just acquire CA.
    - Even after secured higher marks, still unable to get good college with low fees or in locality for low income family.
- After 12th pass out, students normally gets confused on getting different streams like Engineering, Architecture, Bsc or BCA etc.
- After selecting there stream somehow, now there may be confusion on different colleges
    - Those colleges may be local or may be very far from their residence.
    - The Fees of the college may be expensive to some families.
    - At this time wrong selection may happen.

1.2) Problem for Colleges:

1.2.1) Improper Marketing:

- Currently colleges market themselves by presenting their ads on newspapers, paid ads on educational websites like (shiksha.com), printing brochures, hiring volunteers to give brochures to the student who visits campuses, hence this method is more expensive for the colleges.
- Now, students have to visit campus individually and get the brochure for college's full information.

1.2.2) Unavailability of Information:

- Some colleges may hide or may not have complete details on their placements, professor qualifications, quality of teaching, campus environment, student life etc.
- Sometimes wrong student may also get admission with the admission quota the college provides.

2) Market/Customer/Business needs assessment:

2.1) Students Needs:

- Currently, in India or more specifically in Maharashtra, There is no any convenient system which can address the above problem.
- The student has to visit some educational websites and collect some info, ask their friends and seniors for career guidance. They ask their school teacher or private tutors for stream and college selection. If any student comes in contact with good career counselor still their fees are expensive and still the result will be unsatisfactory.
- Currently, students need a system which saves time and addresses above problem.

2.2) College Needs:

- For colleges, In Maharashtra, they are using expensive methods to market themselves.
- Non-deserving candidate may get selected in admission quota which has the policy of "First Come –First Serve".
- Colleges are in search of a system which addresses above loopholes.

2.3) Business Needs:

- In India, There are many papers are written about Machine Learning, Deep Learning, AI, recommendation system etc but actual system is not produced yet.
- Career guidance is a big issue, as many students take guidance from other field experts.
- For example, a student wants to pursue a data science but he asks about college and stream to a counselor who was a former software engineer. And if a counselor wants to give guidance about software development he may not get appropriate candidate.
- There is a need of a system which will act as a bridge between service providers and the customers in the field of education in India.

3) Bench Marking Products:

- In the era of information, recommender systems play a main role in alleviating information overload, having been widely adopted by many online services, including E-commerce, online news and social media sites. The key to a personalized recommender system is in modeling users' preference on items based on their past interactions.
- Many research studies have been conducted to support students in their decision to select the appropriate major and colleges.
- As of now, whatever resources I found online has only research papers on this personalized recommender system for colleges and students, but any convenient app or a web portal is not available, or hasn't been developed.
- Present situation for a counselor, college management and students are present with struggle for their basic needs.
- ML-Based recommender system solves the following task:
    - Bridge the gaps between students and counselor.
    - Bridge the gaps between college management and students.
    - Reliable system, easy to maintain, scalable and updateable.

4) Business Model:

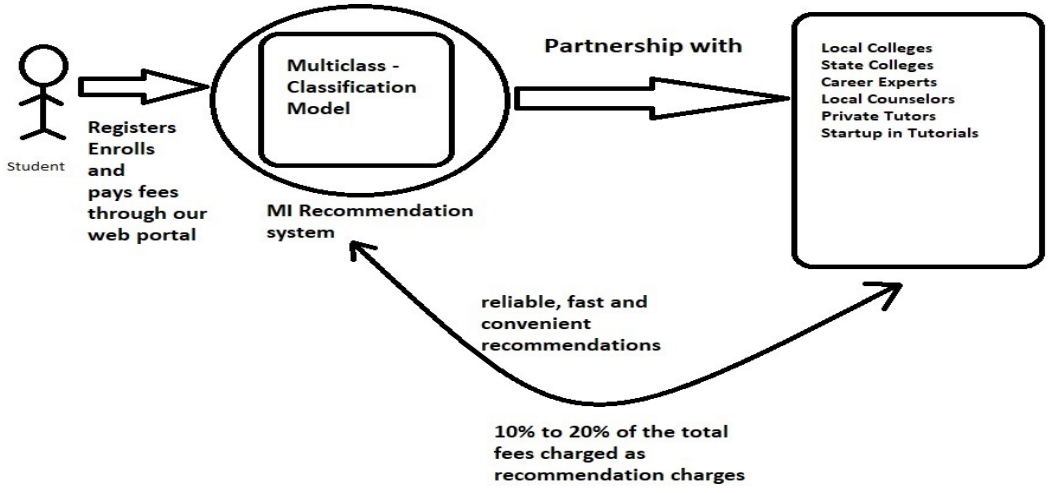

Fig: Business Model

- In India, the system which is close to college, counselors and students exists but in complicated way which has no direct communication with either at state level or at local level.
- The model shown in schematic diagram will have more effective way of communication, secured structure and can be optimized for future use.
- As this type of system hasn't developed yet, this will be a good opportunity to work on and develop some good business opportunity at the local level. Then we can further develop this system to higher level.

5) Concept Generation:

- There is a need of a system which can bridge the gaps between students, colleges and the counselor.
- This system should be user-friendly, reliable, fast and convenient.
- Software development should be scalable, and can be optimized in the future with updates.

6) Concept Development:

- A web portal with an artificial recommendation system will recommend:
  - Best colleges to students.
  - Best students to colleges.
  - Best counselor to students (if they wish for career guidance).
  - Or just a simple web portal which will provide free college and university information, brochures and reviews.
- Web portal will be user-friendly, and easy to all.

7) Final Product Prototype with Schematic Diagram:

7.1) Step 1:

- Student has to enter some information regards with marks, locality, budget, field of interest etc.
- Our recommender system will recommend some section of colleges to the student.
- Our recommender system will also recommend a reference of a student with the marks provided by the consent of student if he decides to enroll. At the same time the student data will be stored at the private highly secured database system (for future use of monitoring the machine learning model).

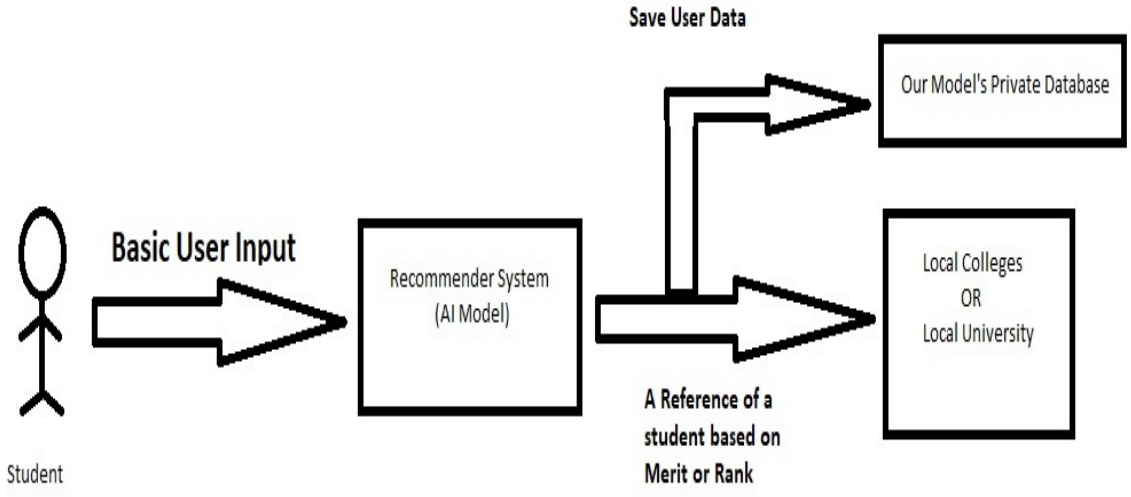

Fig: User Input

7.2) Step 2:

- If student decides to enroll to recommended college, an application fee will be charged. If he enroll through our web portal, that student may get 10% fee waiver on application fee.
- After enrollment the student will be recommended to the college's admission team.
- As the management admission quota of the college is "First Come-First Serve" depending on the availability, the offer will be provided by the college in few days after deadline.(if college rejects the student, that student will automatically considered for the other second recommended college.
- If 3 attempts of the recommendation is over and still student face the rejection then the application fee will be refunded back to student and the career expert will be allotted for counseling.

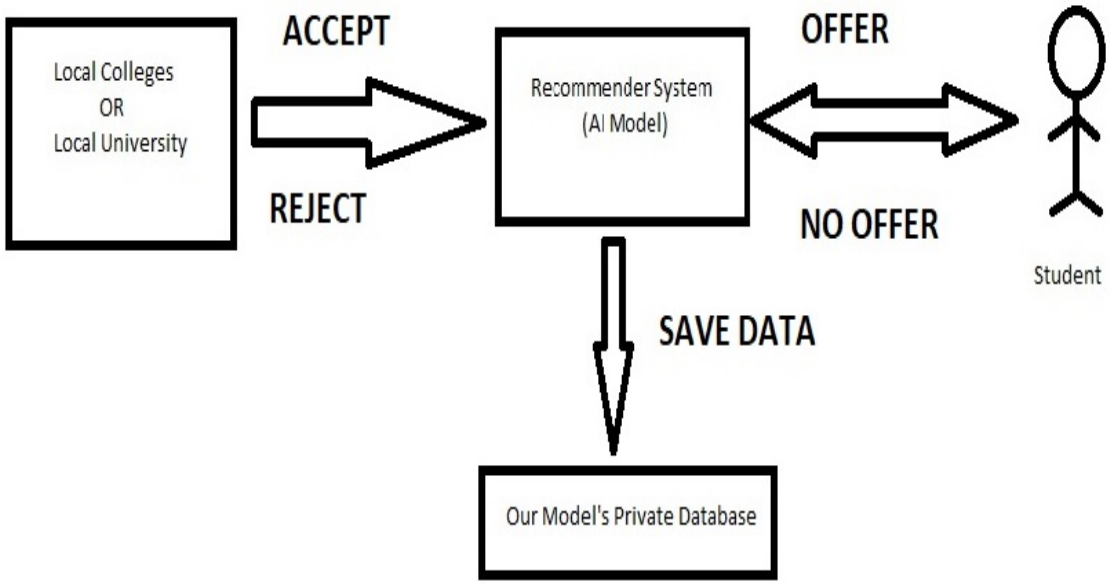

Fig: College decision and AI model

7.3) Step 3:

- If student accepts the offer from college and decides to enroll with 1 year admission fee for first year, then we will have our commission of about 10% - 20% of the first year admission fee from the college for simplifying the process.

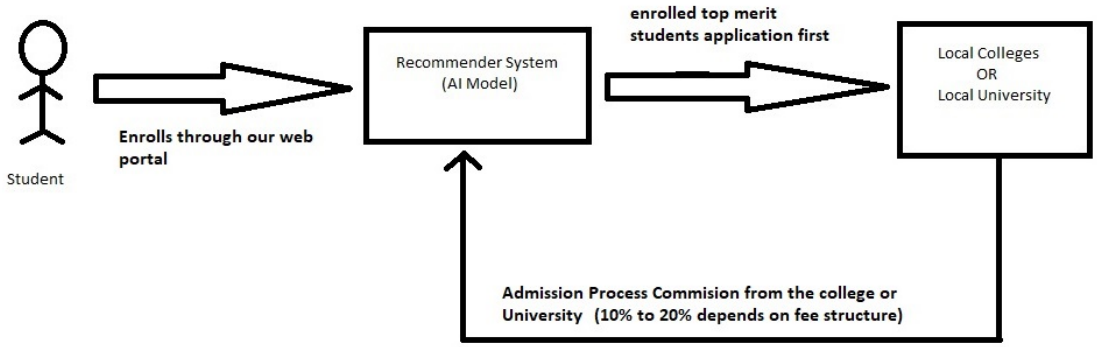

Fig: final stage

8) Code Implementation:

- ✓ **Github link** - https://github.com/ysmayur1992/College_Recommender

8.1) Data Analysis and Model Building:

8.1.1) Data Collection and Preparation:

- The idea of the project is somewhat unique and only some papers were available. So the dataset used in this project is created synthetically i.e. dataset information is random.
- Total of 2 datasets are used for 10^{th} passed student and 12^{th} passed student.

8.1.2) Data Processing and Visualization:

- Feature engineering techniques are used for data preprocessing.
- As the data is randomly set, so there is no missing values and other problems but for original dataset, we have to take care of missing values, outliers, convert to Gaussian normal distribution if not in Gaussian form.
- Scaling of the continuous numerical values is done with standard scaler.
- Below figures are some graphs of Gaussian distribution of marks and bar graph of previous stream of study.

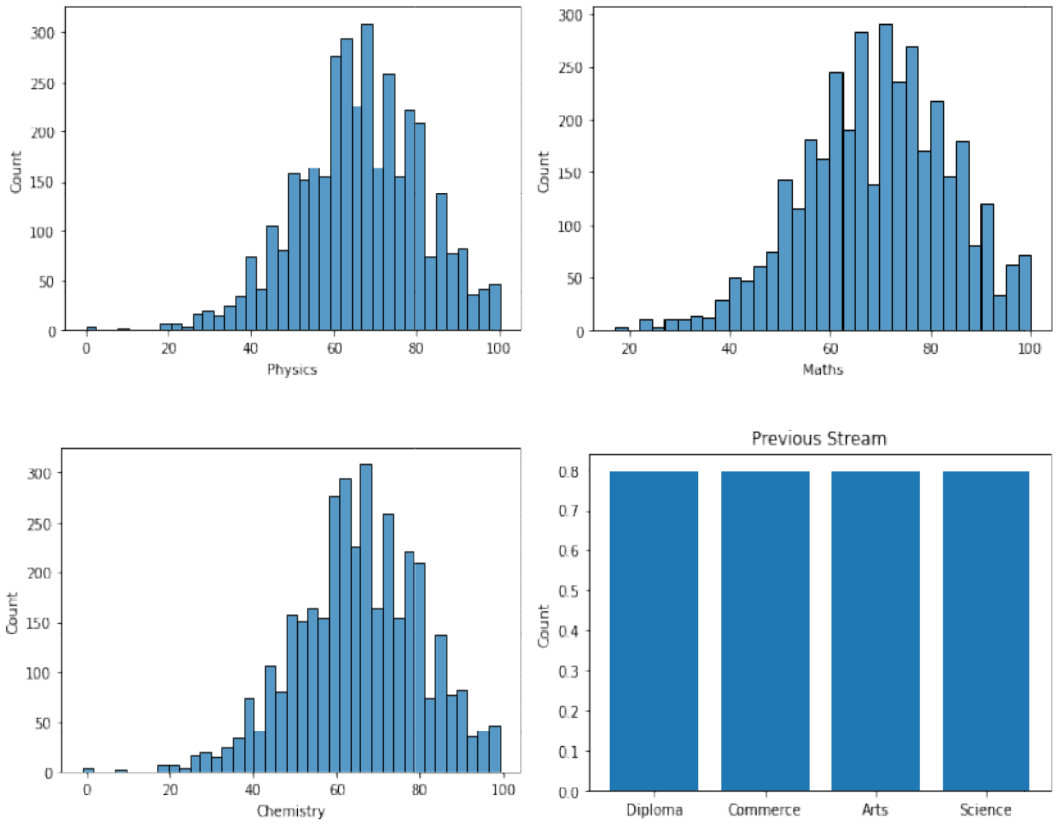

Fig: Histograms and Bar graph

- Below is the figure on correlation of the independent and dependent features.
- Note: Feature selection is not used as there is very less features and the data is inaccurate.

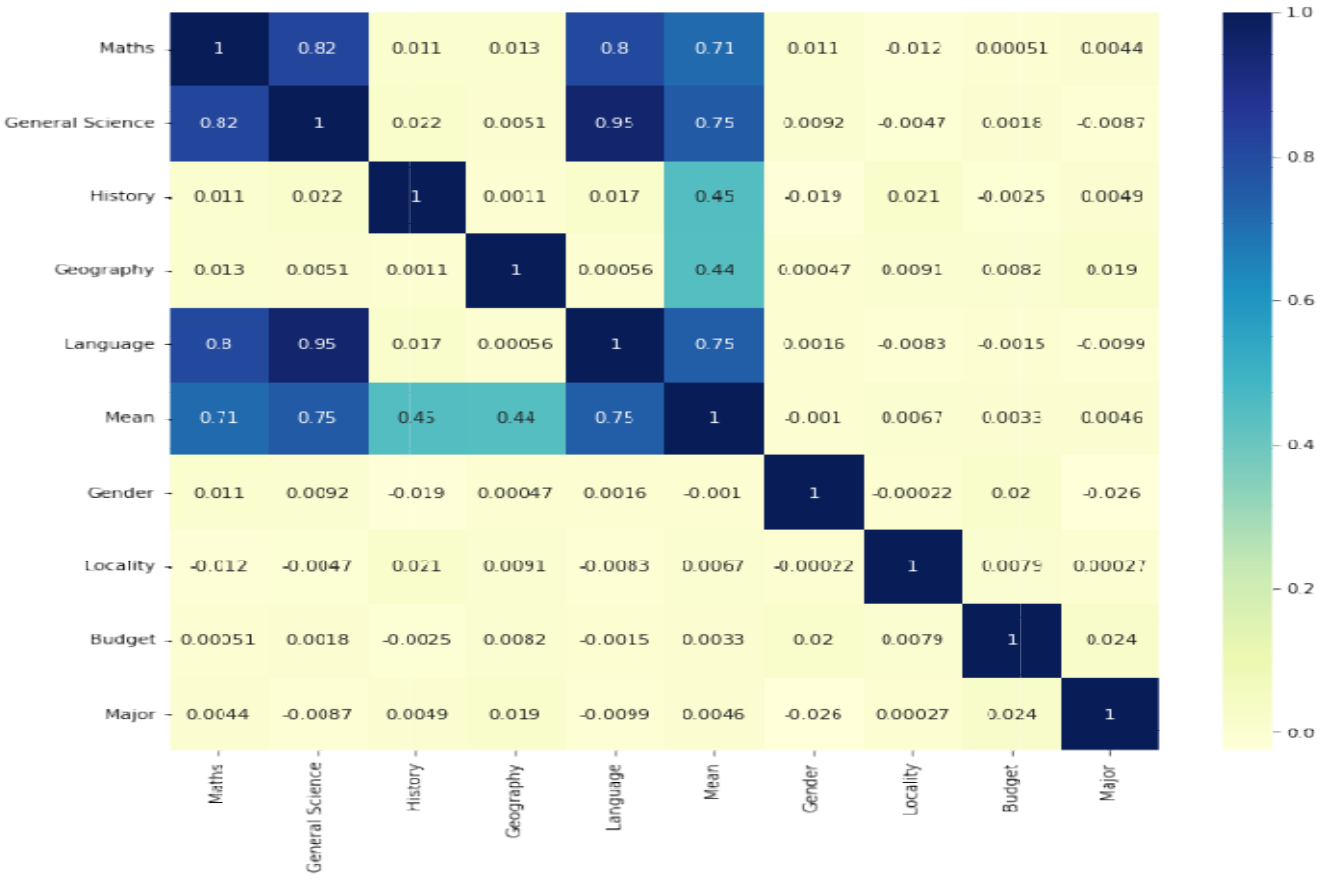

Fig: Correlation graph of dataset for 10th passed student

8.1.3) Model Classification:

- Total of 9 different classifier algorithms are used from sklearn.
- One classifier with best accuracy score is used for final model creation.
- Note: we can use classification report, confusion matrix, ROC curve, true negative and true positive rate to check the performance of the model. As the dataset is randomly picked this performance metrics is ignored for this project.

8.2) Modular Coding:

- Follow the above link for the code. Run app.py in your ide and follow the web link which will appear in Recent_Logs folder.
- Project has the following modules:

8.2)1. **Data Ingestion**: Responsible for reading original dataset, splitting into training and testing dataset.

8.2)2. **Data Transformation:** Responsible for creating preprocessor files in pickle (.pkl) format which is used for feature engineering, encoding of categories and scaling of data.

8.2)3. **Model Trainer:** Responsible for creating machine learning models in pickle (.pkl) format.

8.2)4. **Pipeline:** This module is responsible for processing of data received from html document, loading the machine learning models, and predicting the results.

- Below is the picture of flask app for prediction of college for 12th passed student.

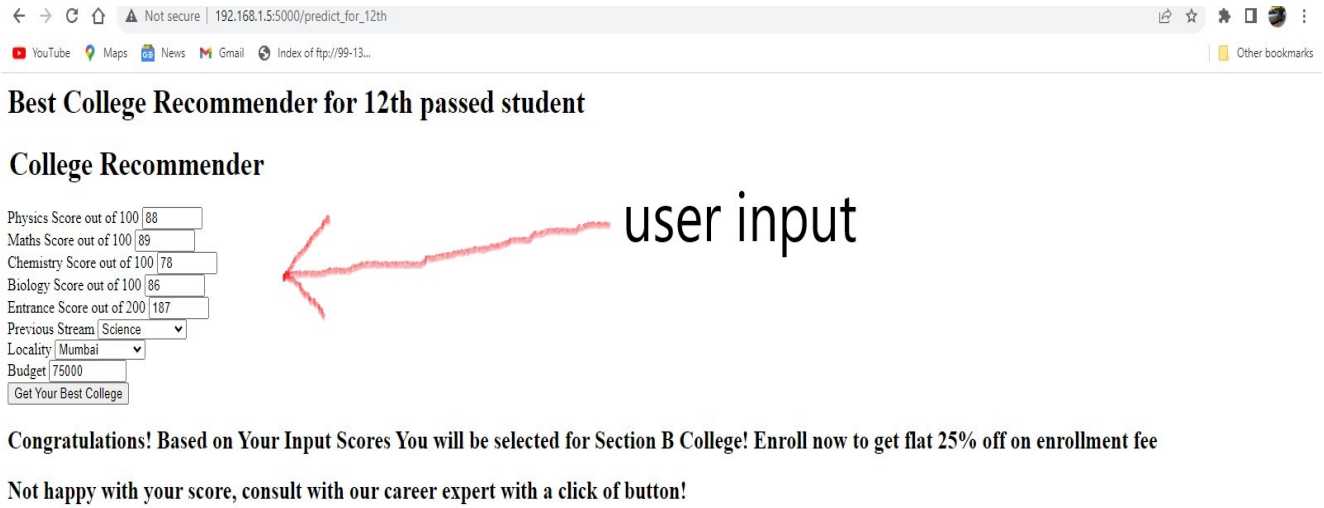

Fig: Web portal for 12th passed student

Github link - https://github.com/ysmayur1992/College_Recommender

Conclusion:

Data science and artificial intelligence is one of the powerful tool which can make ease of work for humans. Currently big tech companies are using this tool in their products and slow and steadily this tool will cover the small scale industry as well.

In this project, I have presented a general idea of model which can touch to the ease of admission process and can benefit student and middle class families at local level and the scalable software presented above can be updated to higher level as well which can be optimized in the future.

Reference:

[1].V. Sharma, T. Trehan, R. Chanana and S. Dawn, "StudieMe: College Recommendation System," 2019 3rd International Conference on Recent Developments in Control, Automation & Power Engineering (RDCAPE), 2019, pp. 227-232, doi: 10.1109/RDCAPE47089.2019.8979030.

[2] Prof. Sunil Sonawane, Miss. Madhavi Gosavi, Miss. Aparna Pawar, Mr. Omkar Pawar, Mr. Dhananjay Dighe: College recommendation system using artificial intelligence and machine learning International Research Journal of Modernization in Engineering Technology and Science doi: 03/2023

[3] Yara Zayed, Yasmeen Salman and Ahmad Hasasneh: A Recommendation System for Selecting the Appropriate Undergraduate Program at Higher Education Institutions Using Graduate Student Data doi:12/2022