

University Admit Eligibility Predictor

Using Ibm Watson

INDEX:

- 1.INTRODUCTION
- 2.LITERATURE SURVEY
- 3.THEORITICAL ANALYSIS
- 4.EXPERIMENTAL INVESTIGATIONS
- 5.FLOW CHART
- 6.RESULT
- 7.ADVANTAGES AND DISADVANTAGES
- 8.APPLICATIONS
- 9.CONCLUSION
- 10.FUTURE SCOPE
- 11.BIBILOGRAPHY
- 12.APPENDIX
 - a.Source Code
 - b.UI output Screenshot

1.INTRODUCTION:

Students are often worried about their chances of admission to University. The aim of this project is to help students in shortlisting universities with their profiles. The predicted output gives them a fair idea about their admission chances in a particular university. This analysis should also help students who are currently preparing or will be preparing to get a better idea.

Overview:

1. Defining our classification categories
2. Importing the dataset
3. Train the model
4. Test our model

2. LITERATURE SURVEY :

To complete this project, we need to install the following packages:

- 1.Numpy: This package is used to perform numerical computations. This package is pre-installed in anaconda.
- 2.Pandas: Pandas is one of the most widely used python libraries in data science. It provides high-performance, easy to use structures, and data analysis tools. This package is pre-installed in anaconda.
- 3.Matplotlib: Matplotlib is a comprehensive library for creating static, animated, and interactive visualizations in Python. This package is pre-installed in anaconda.
- 4.Scikit-learn: This is a machine learning library for the Python programming language. This package is pre-installed in anaconda.
- 5.Flask: Flask is a lightweight WSGI web application framework.

If not installed use the following

Open anaconda prompt as administrator.

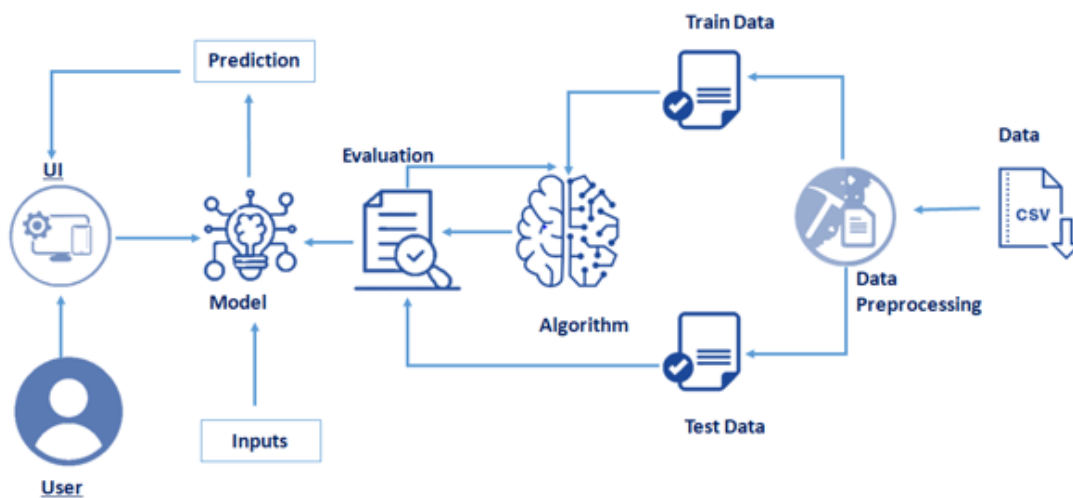
- Type “pip install numpy” and click enter.

- Type “pip install pandas” and click enter.
- Type “pip install matplotlib” and click enter.
- Type “pip install scikit-learn” and click enter.
- Type “pip install Flask” and click enter.

The above steps allow you to install the packages in the anaconda environment.

3.THEORITICAL ANALYSIS:

BLOCK DIAGRAM:



4 .EXPERIMENTAL INVESTIGATIONS:

- **Numpy**- It is an open-source numerical Python library. It contains a multi-dimensional array and matrix data structures. It can be used to perform mathematical operations on arrays such as trigonometric, statistical, and algebraic routines.
- **Pandas**- It is a fast, powerful, flexible and easy to use open-source data analysis and manipulation tool, built on top of the Python programming language.
- **Seaborn**- Seaborn is a Python data visualization library based on matplotlib. It provides a high-level interface for drawing attractive and informative statistical graphics.
- **Matplotlib**- Visualisation with python. It is a comprehensive library for creating static, animated, and interactive visualizations in Python.

5.FLOWCHART:

Project Flow:

You will go through all the steps mentioned below to complete the project.

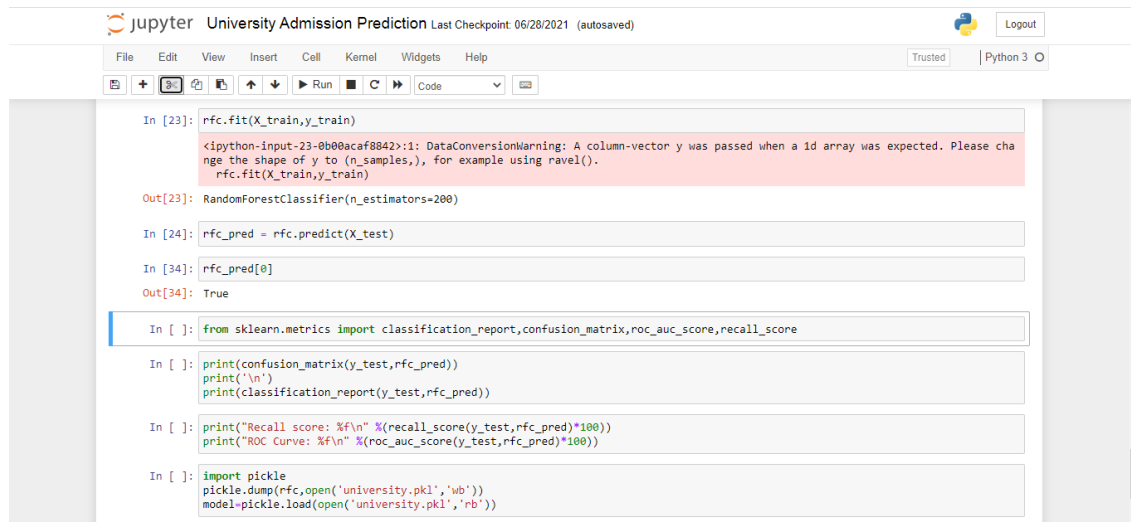
- User interacts with the UI (User Interface) to enter Data
- The entered data is analyzed by the model which is integrated
- Once model analyses the input the prediction is showcased on the UI

To accomplish this, we have to complete all the activities and tasks listed below

- Data Collection.
 - Collect the dataset or Create the dataset
- Data Preprocessing.
 - Import the Libraries.
 - Importing the dataset.
 - Checking for Null Values.
 - Data Visualization.
 - Taking care of Missing Data.
 - Label encoding.
 - One Hot Encoding.
 - Feature Scaling.
 - Splitting Data into Train and Test.
- Model Building
 - Training and testing the model
 - Evaluation of Model
- Application Building
 - Create an HTML file
 - Build a Python Code

6. RESULT:

- User interacts with the UI (User Interface) to enter Data.
- The entered data is analyzed by the model which is integrated.
- Once model analyses the input the prediction is showcased on the UI.



```

In [23]: rfc.fit(X_train,y_train)

<ipython-input-23-0b00acaf8842>:1: DataConversionWarning: A column-vector y was passed when a 1d array was expected. Please change the shape of y to (n_samples,), for example using ravel().
  rfc.fit(X_train,y_train)

Out[23]: RandomForestClassifier(n_estimators=200)

In [24]: rfc_pred = rfc.predict(X_test)

In [34]: rfc_pred[0]

Out[34]: True

In [ ]: from sklearn.metrics import classification_report,confusion_matrix,roc_auc_score,recall_score

In [ ]: print(confusion_matrix(y_test,rfc_pred))
        print('\n')
        print(classification_report(y_test,rfc_pred))

In [ ]: print("Recall score: %f\n" %(recall_score(y_test,rfc_pred)*100))
        print("ROC Curve: %f\n" %(roc_auc_score(y_test,rfc_pred)*100))

In [ ]: import pickle
        pickle.dump(rfc,open('university.pkl','wb'))
        model=pickle.load(open('university.pkl','rb'))

```

7. ADVANTAGES AND DISADVANTAGES :

ADVANTAGES:

- It reduces overfitting in **decision** trees and helps to improve the accuracy.
- It is flexible to both classification and regression problems.
- It works well with both categorical and continuous values.
- It automates missing values present in the data.

DISADVANTAGES:

- If there are enough trees in the forest, the classifier won't overfit the model.
- The main limitation of random forest is that a large number of trees can make the algorithm too slow and ineffective for real-time predictions.

8. APPLICATIONS:

- ✓ Banking Industry
- ✓ Health Care and Medicine

✓ E-Commerce

✓ Stock Market

9. CONCLUSION:

By the end of this project you'll be able to understand:

- Regression and Classification Problems
- To grab insights from data through visualization.
- Applying different algorithms according
- Evaluation metrics
- how to build a web application using the Flask framework.

10. FUTURE SCOPE:

Random forest algorithm generates many classification trees and generation of each tree is independent of each other. Thus, RandomForest is by nature a suitable candidate for parallel processing. Additionally, data mining is usually performed on very large datasets, and Random Forest can work well on datasets with large number of predictors. As mentioned, each parallel implementation of Random Forest is specific to some platform or language. Thus, there is scope for generalized Parallel Algorithm for Random Forest. With the geographical spread of business and the world getting connected with the Internet; business data is distributed at different locations. Hence, design of Distributed Random Forest algorithm is another important future research direction.

11. BIBLIOGRAPHY:

We used dataset from kaggle website to train the model.

<https://www.kaggle.com/rishal005/admission-predict>

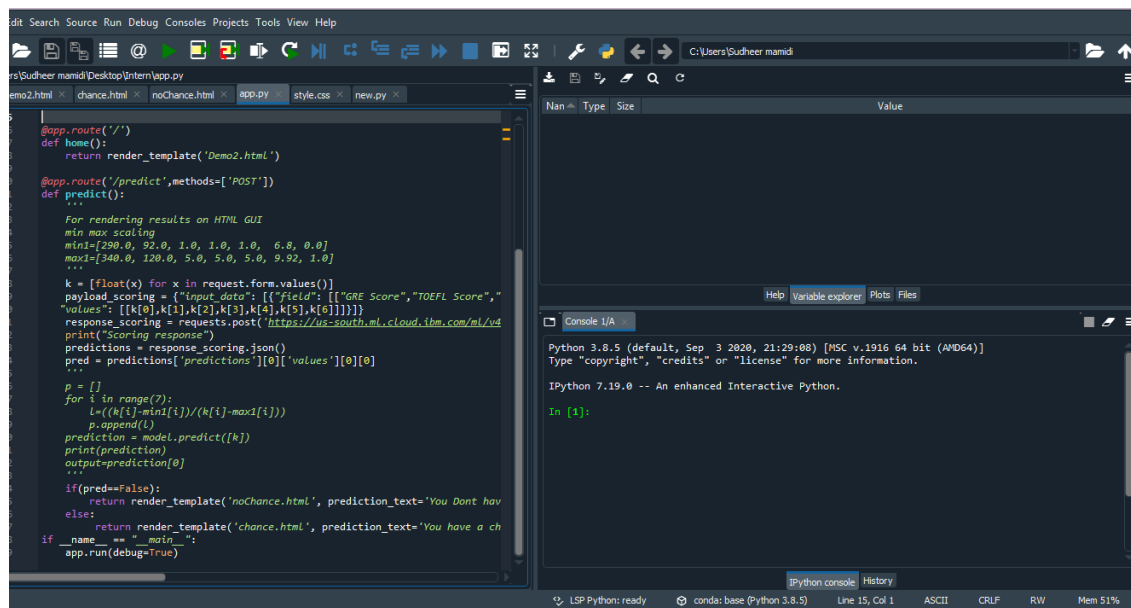
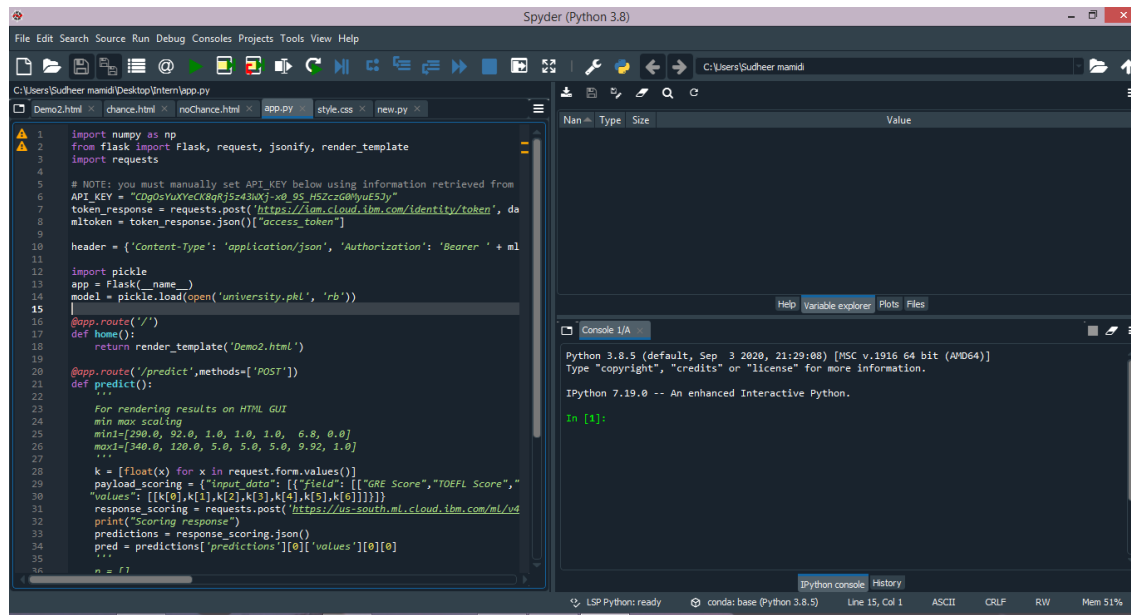
We saw some reference videos in you tube.

<https://www.youtube.com/watch?v=TysuP3KgSzc&t=4s>

<https://www.youtube.com/watch?v=ST1ZYLmYw2U>

12. APPENDIX:

a. Source Code



b. UI Output Screenshots

127.0.0.1:5000

127.0.0.1:5000

Apps Virtual lab litr - You... Online C Compiler ~... HTML Color Codes Contact Reading list

UNIVERSITY ADMISSION PREDICTION SYSTEM

Enter your details to get probability of your admission

Enter GRE score

Enter TOEFL score

Select University no

☐ 1

☒ 2

☐ 3

☐ 4

☐ 5

Enter SOP


Enter LOR

Enter CGPA

Research

☐ Research

☒ No Research



127.0.0.1:5000/predict

127.0.0.1:5000/predict

Apps Virtual lab litr - You... Online C Compiler ~... HTML Color Codes Contact Reading list

Predicting chance of Admission

A machine learning web app using flask

Prediction : You have a chance



127.0.0.1:5000

127.0.0.1:5000

Apps Virtual lab litr - You... Online C Compiler ~... HTML Color Codes Contact Reading list

UNIVERSITY ADMISSION PREDICTION SYSTEM

Enter your details to get probability of your admission

Enter GRE score

Enter TOEFL score

Select University no

☐ 1

☒ 2

☐ 3

☐ 4

☐ 5

Enter SOP


Enter LOR

Enter CGPA

Research

☐ Research

☒ No Research



127.0.0.1:5000/predict

127.0.0.1:5000/predict

Apps Virtual lab litr - You... Online C Compiler ~... HTML Color Codes Contact Reading list

Predicting chance of Admission

A machine learning web app using flask

Prediction : You Dont have a chance

