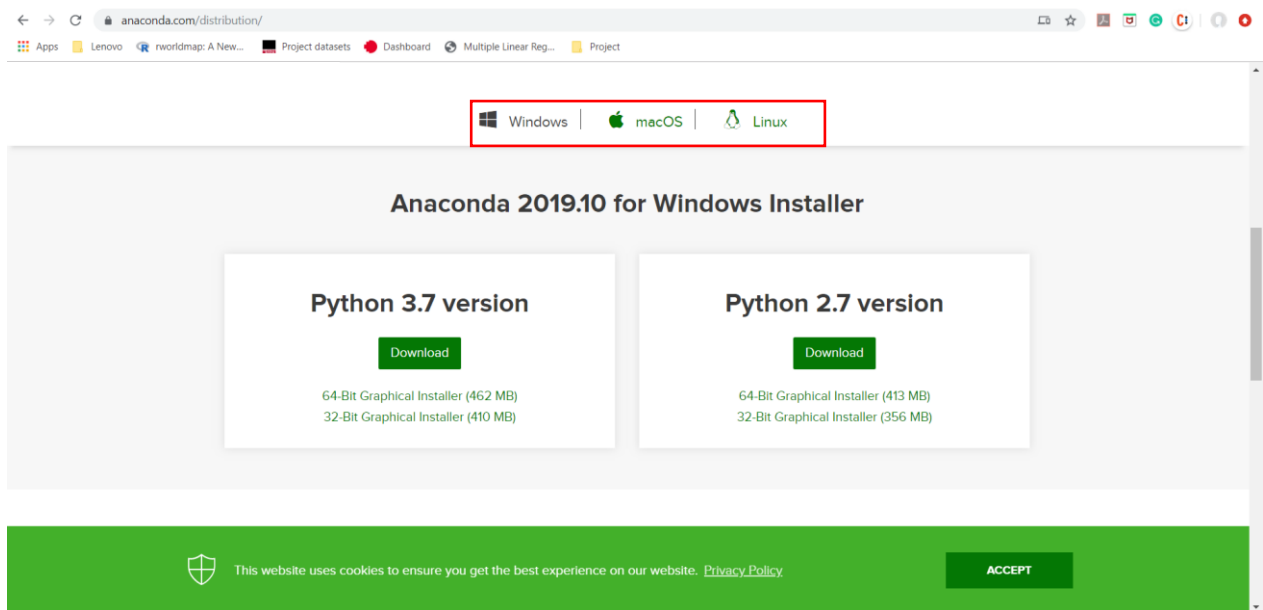


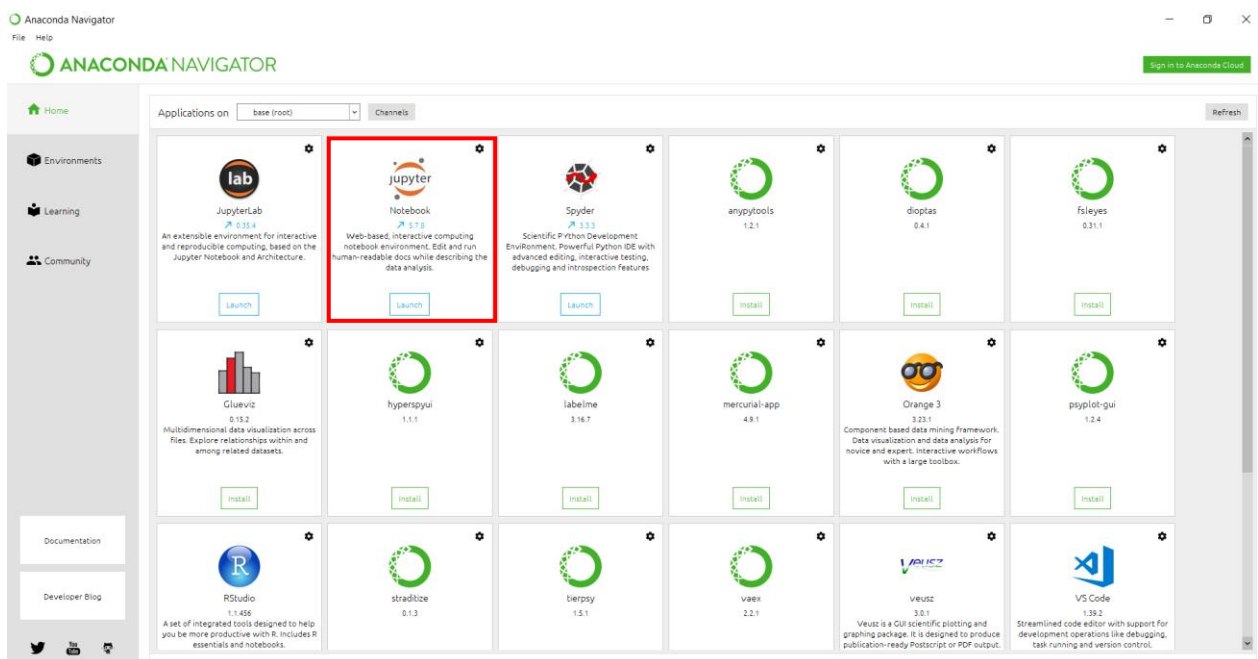
- Install Anaconda 3 using - <https://www.anaconda.com/distribution/>

Select Windows, macOS or Linux as per your operating system and then click on download button for Python 3.7 version.



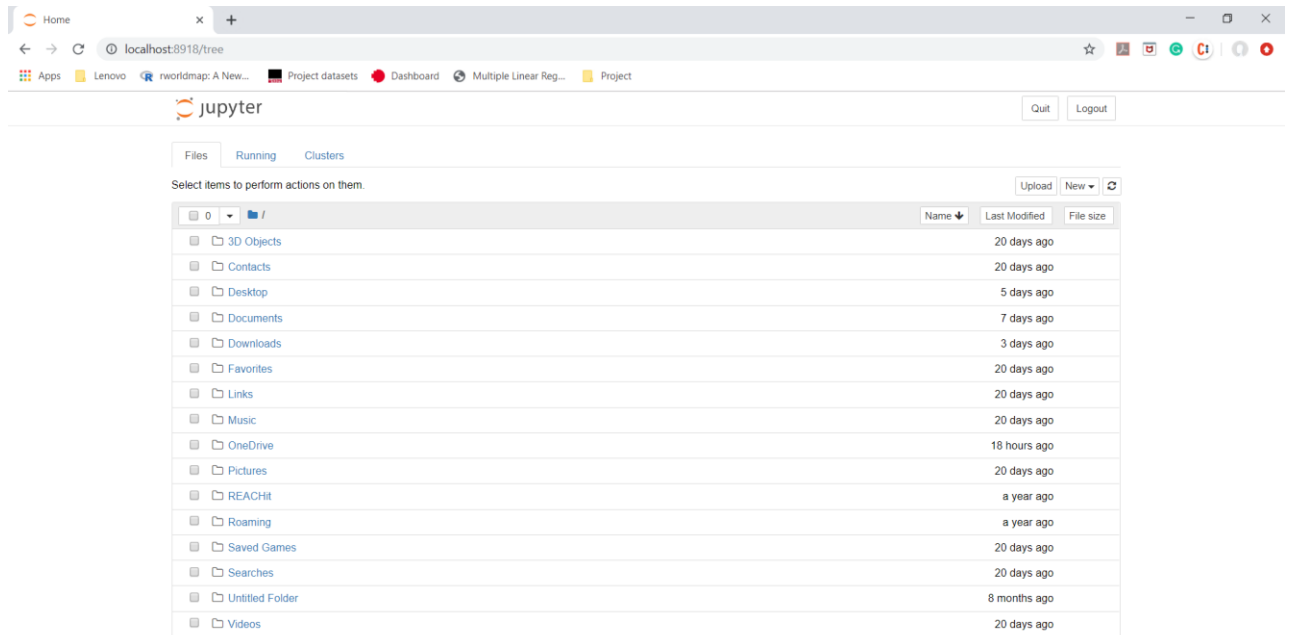
Install the downloaded file using recommended settings.

- Open Anaconda Navigator and launch - *Jupyter Notebook*



It will open in your Default Browser (preferably Chrome).

- This is how the screen looks -

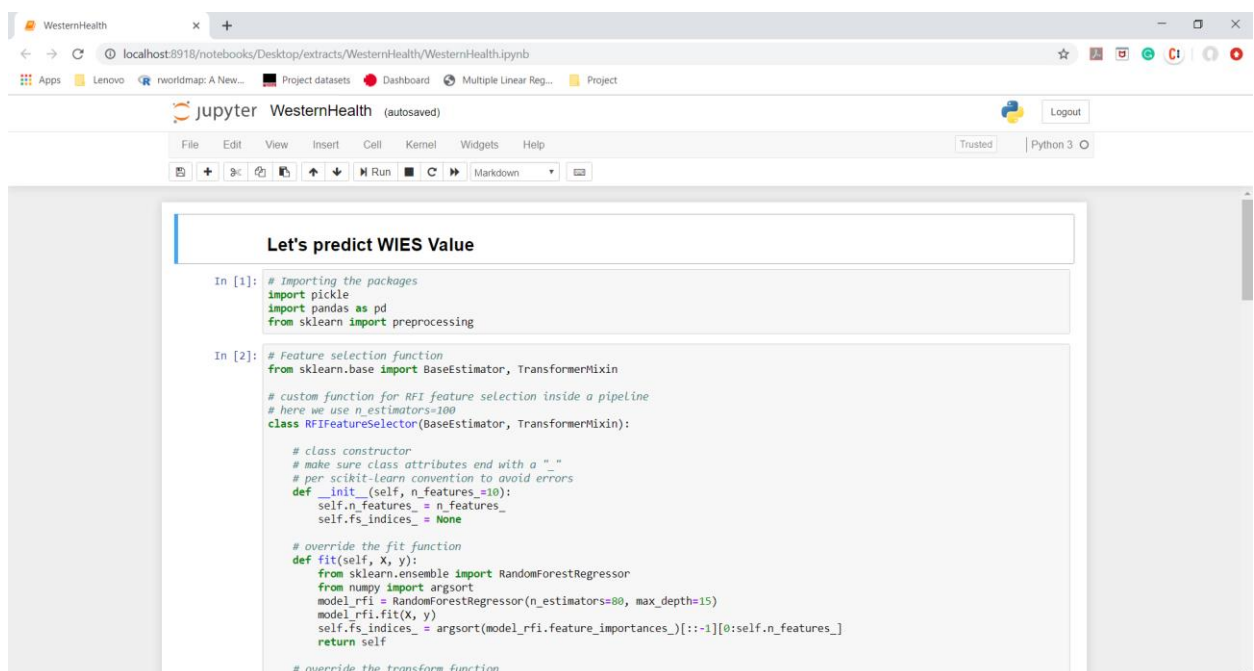


- Select the location where all the files are saved. Example – Make a folder on Desktop with all the supplied files in it and open it from here.

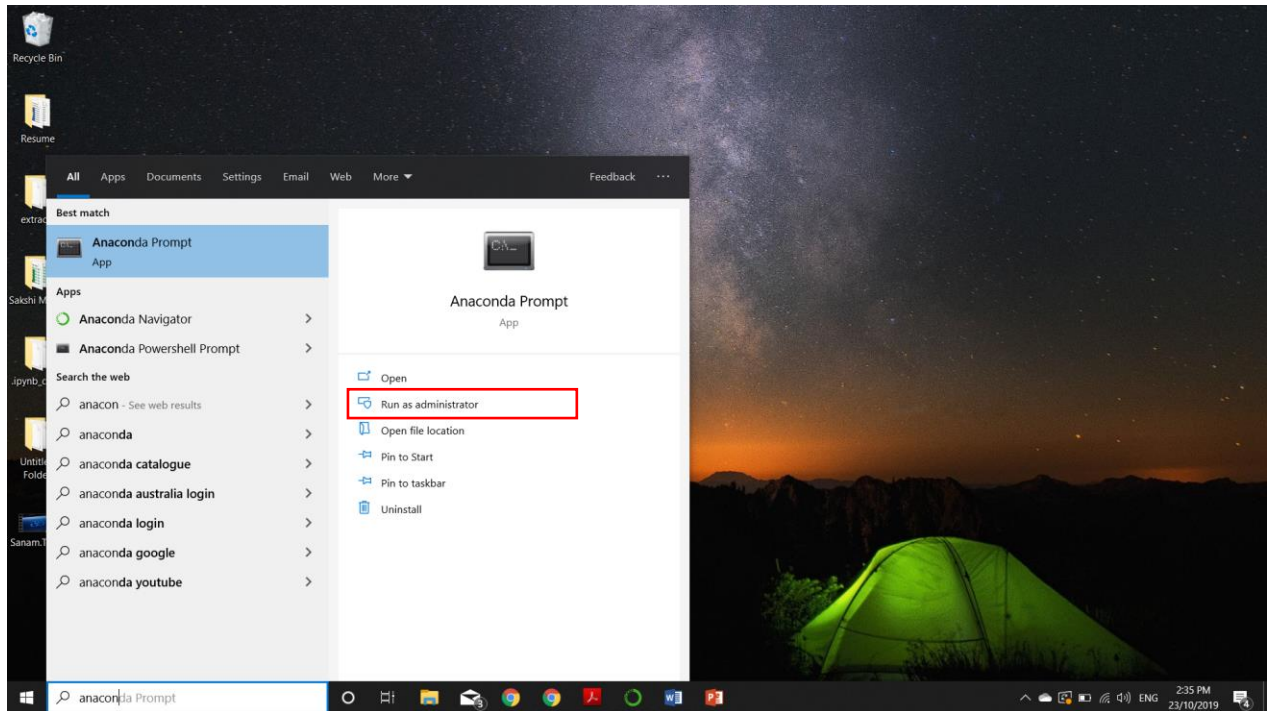
Required files-

1. WesternHealth.ipynb
2. PredictWIES.sav
3. TheRelevant.csv
4. Test.csv (The data on which prediction is to be made)

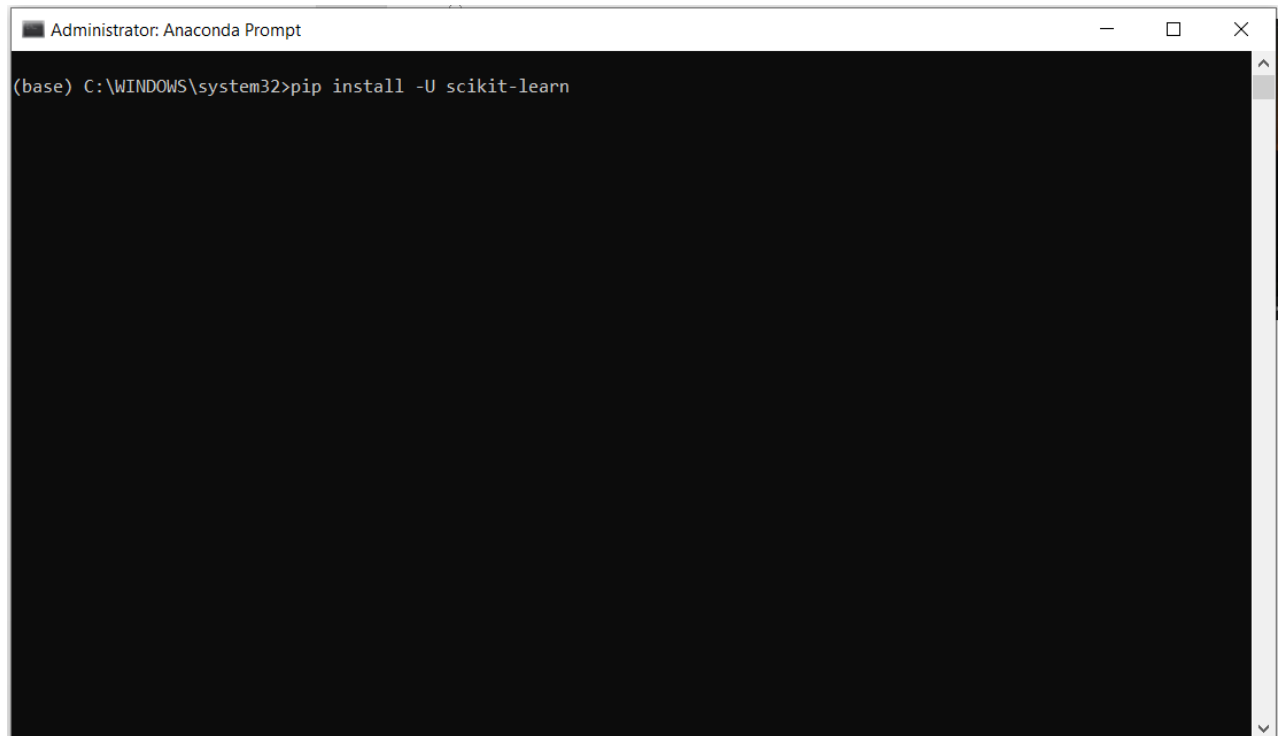
Open WesternHealth.ipynb



- You might have to install the required packages-
Search Anaconda Prompt – Run as Administrator – Select ‘Yes’



- In the command prompt type the following one by one and press enter-
 1. `pip install -U scikit-learn`
 2. `pip install pickle`



- If this is successfully installed you can start running the code.

- Press the run button on each cell.

WesternHealth (autosaved)

File Edit View Insert Cell Kernel Widgets Help Trusted Python 3

Let's predict WIES Value

```
In [1]: # Importing the packages
import pickle
import pandas as pd
from sklearn import preprocessing

In [2]: # Feature selection function
from sklearn.base import BaseEstimator, TransformerMixin

# custom function for RFI feature selection inside a pipeline
# here we use n_estimators=100
class RFIFeatureSelector(BaseEstimator, TransformerMixin):

    # class constructor
    # make sure class attributes end with a "_"
    # per scikit-learn convention to avoid errors
    def __init__(self, n_features=10):
        self.n_features_ = n_features_
        self.fs_indices_ = None

    # override the fit function
    def fit(self, X, y):
        from sklearn.ensemble import RandomForestRegressor
        from numpy import argsort
        model_rfi = RandomForestRegressor(n_estimators=80, max_depth=15)
        model_rfi.fit(X, y)
        self.fs_indices_ = argsort(model_rfi.feature_importances_)[::-1][:self.n_features_]
        return self

    # override the transform function
```

- While reading the data TheRelevant file is standard every time and is given, for Test data please make sure it has only the required columns nothing extra or less and is a csv file.

WesternHealth (autosaved)

File Edit View Insert Cell Kernel Widgets Help Trusted Python 3

```
In [4]: # Read Main data
DataAll = pd.read_csv('TheRelevant.csv', header = 0, sep=',', names = ['PatientID','IPEpisodeID','Gender Code',
'Age On Admission','Birth Weight','ATSI Flag',
'Admission Source Code','Admission Type Code',
'Admission Specialty','Discharge Specialty',
'Admission Ward Code','Discharge Ward Code',
'Separation Method Code','DRG Code',
'ICU LOS Hours','Total LOS Days','HITH Days',
'LOS Type','WIES Value'])

DataAll.head()

Out[4]:
```

	PatientID	IPEpisodeID	Gender Code	Age On Admission	Birth Weight	ATSI Flag	Admission Source Code	Admission Type Code	Admission Specialty	Discharge Specialty	Admission Ward Code	Discharge Ward Code	Separation Method Code	DRG Code	ICU LOS Hours
0	973538	3320175	M	56	NaN	No	T	P	HIHC	HIHC	F2B	FHITH	H	K01C	NaN
1	1198129	3327705	F	81	NaN	No	S	S	GBS	GBS	SHITH	SHITH	H	G70A	NaN
2	319442	3333120	M	82	NaN	No	H	C	45	26	F2C	FHITH	H	K01C	NaN
3	1210448	3353681	M	88	NaN	No	H	C	CRGS	UGIG	F2E	F3E	H	G02A	146.62
4	387487	3355364	M	72	NaN	No	H	C	IMSD	IMSD	S2A	S3E	S	F83A	NaN

```
In [5]: # Read the data you want to test on. Rename Test.csv to your file name.
# Please make sure your test data has columns as specified here
DataTest = pd.read_csv('Test.csv', header = 0, sep=',', names = ['PatientID','IPEpisodeID','Gender Code',
'Age On Admission','Birth Weight','ATSI Flag',
'Admission Source Code','Admission Type Code',
'Admission Specialty','Discharge Specialty',
'Admission Ward Code','Discharge Ward Code',
'Separation Method Code','DRG Code',
'ICU LOS Hours','Total LOS Days','HITH Days',
'LOS Type'])

DataTest.head()
```

- Hope the whole code runs successfully without any error and predicted value is printed at the end.

```
In [8]: # Prediction
prediction = loaded_model.predict(Test).sum()

print("Predicted WIES Value - ", prediction)

Predicted WIES Value - 1883.3838907122718
```

If you get stuck anywhere please feel free to contact us.

Best regards,
RMIT Team.