***Credit card Approval detection using Machine Learning through financial data***

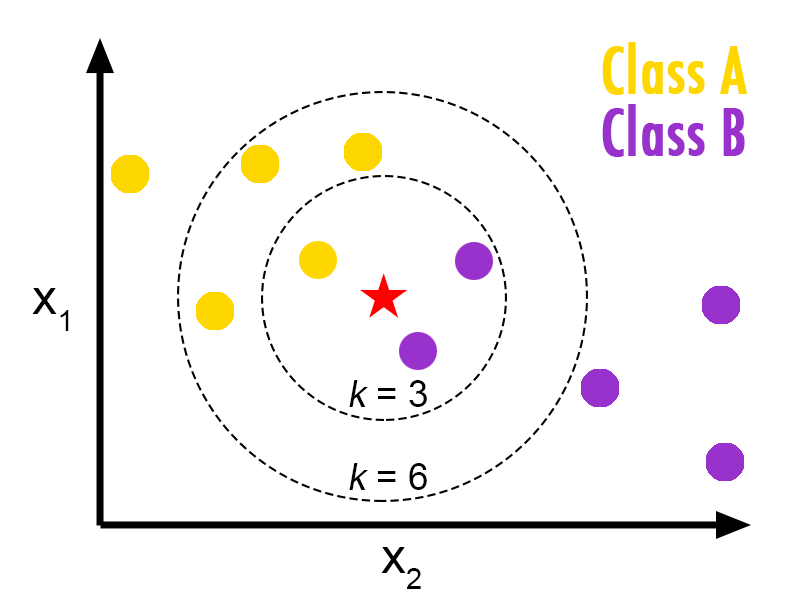
Credit card companies must be capable of detecting credit card fraud transactions in order to detect fraudulent transactions of products that the customer did not buy. This has driven a lot of financial institutions to put strict actions on their customers’ using credit cards. So, Data Science and machine learning are now assisting the financial institutions with more powerful and robust machine leaning models which can help the admins identify the fraudulent transactions involved with the user’s credit card.

***Base Paper***

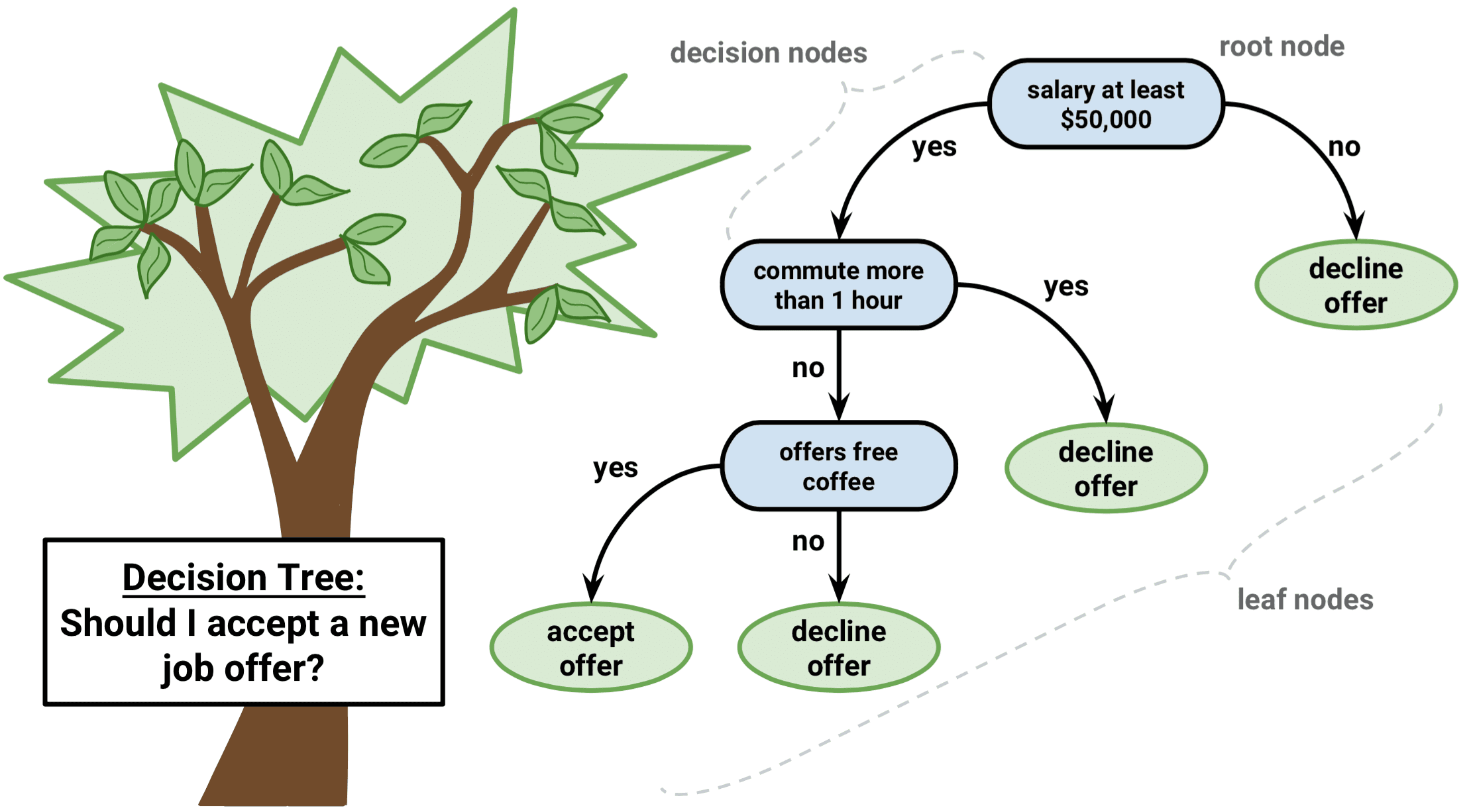
1. <https://www.researchgate.net/publication/40227011_Credit_card_fraud_and_detection_techniques_A_review>
2. <https://www.researchgate.net/publication/348755769_Application_of_Deep_Learning_for_Credit_Card_Approval_A_Comparison_with_Two_Machine_Learning_Techniques>

***Algorithm Description***

***Nearest Neighbour:*** KNN or K Nearest neighbours is a basic yet an efficient algorithm which is being used in most of the Machine learning application. Since it is a non-parametric i.e. This algorithm doesn’t make any underlying assumption like other algorithms do, such as having specify distribution of data to work with. So, this makes it very easy and understandable to all the users who are using it. The Technique KNN applies in predicting on new data is where it finds the nearest neighbours for the given point and takes a majority voting, whichever class is resided near to the new point, it will be considered as the new class for the new data point.



***Decision Tree:*** This algorithm itself has the answer and explanation around it i.e., Making the decisions by splitting the question in tree like structure. Decision trees are the most widely used algorithm along with Random Forest in any machine learning project, as we have discussed above decision tree is also a non-parametric algorithm which makes it easier to understand and implement due to its capability of handling any kind of data, such as Decision tree can be applied directly on data which is not normalized/standardized, since the output/target class is predicted by taking decision from root node to leaf node, there is no need of making the data normalized/standardized. Decision tree is used in application such as Operational research, specifically in any decision analysis or to help identify a strategy to reach a particular goal.



***Steps to Execute the Code!***

**Note:** Make sure you have added path while installing the software’s.

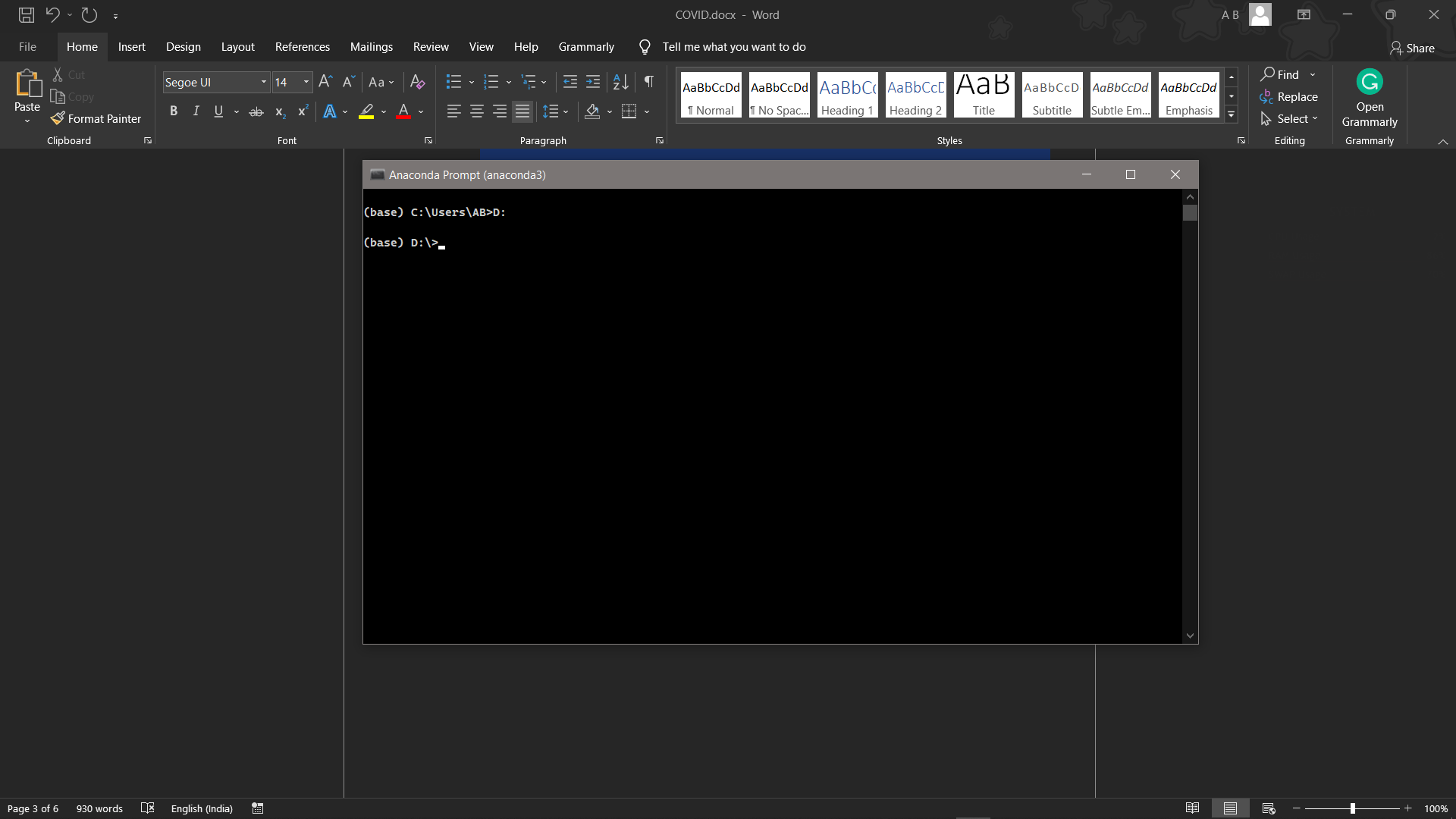
<https://techieyantechnologies.com/2022/07/how-to-install-anaconda/>

<https://techieyantechnologies.com/2022/06/get-started-with-creating-new-environment-in-anaconda-configuring-jupyter-notebook-and-installing-libraries-using-requirements-txt-2/>

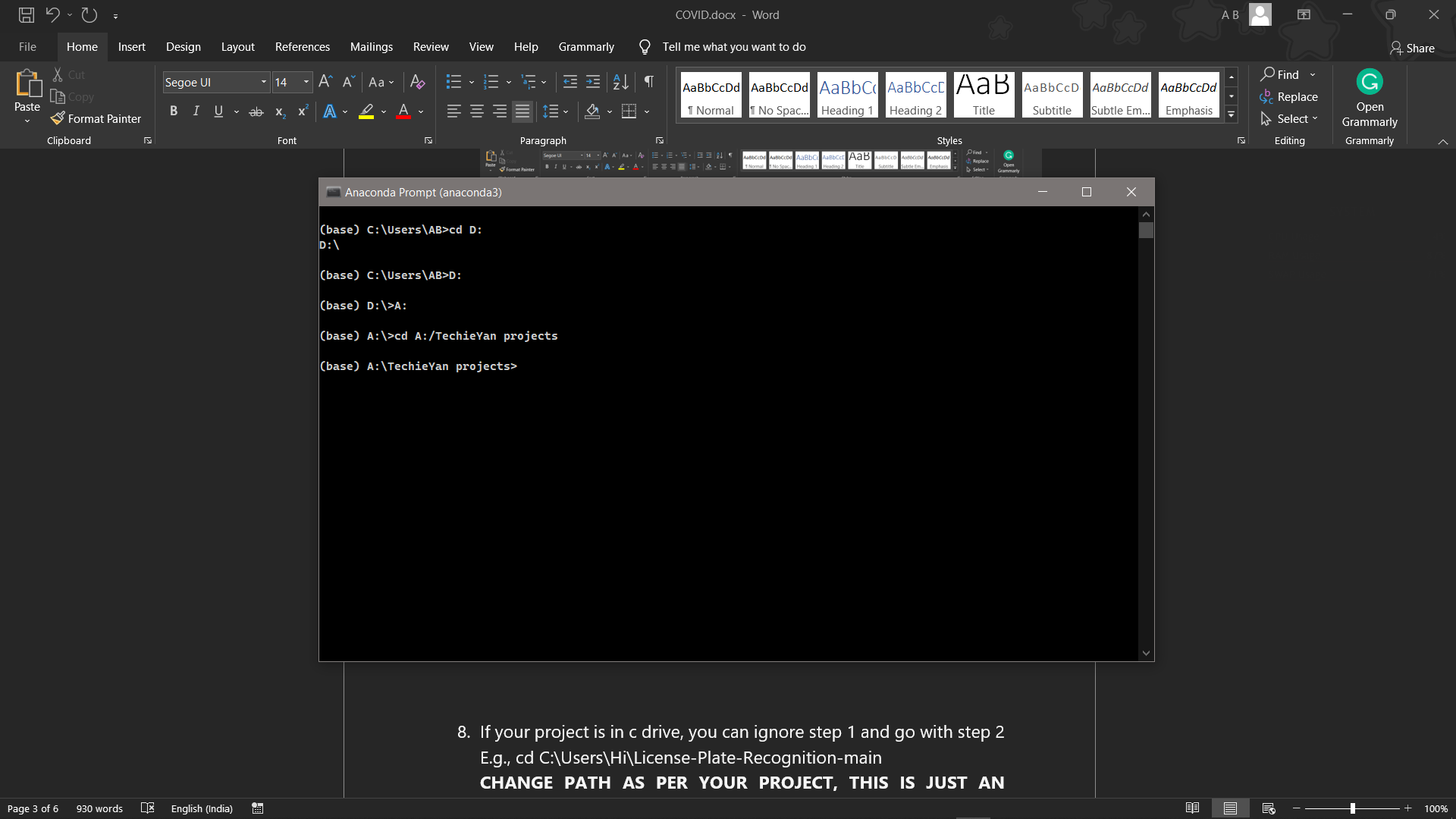
1. Install the prerequisites/software’s required to execute the code from reading the above blog which is provided in the link above.
2. Press windows key and type in anaconda prompt a terminal opens up.
3. Go to the directory where your requirement.txt file is present, not just requirement.txt, if you want to execute any .py or .ipynb files, you need to go to that specific folder or path, where they are saved.



1. <<directory of your file:>>. E.g., If my file is in d drive, then
2. Type d:



1. cd d:\License-Plate-Recognition-main #CHANGE PATH AS PER YOUR PROJECT, THIS IS JUST AN EXAMPLE

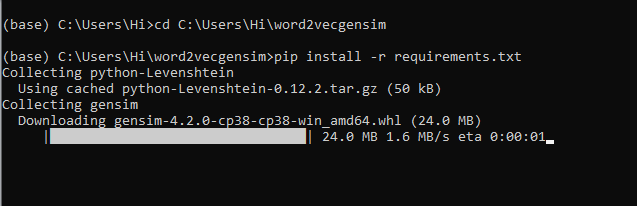


1. If your project is in c drive, you can ignore step 4 and go with step 5.

E.g., cd C:\Users\Hi\License-Plate-Recognition-main

**CHANGE PATH AS PER YOUR PROJECT, THIS IS JUST AN EXAMPLE**

1. Run pip install -r requirements.txt or conda install requirements.txt (Requirements.txt is a text file consisting of all the necessary libraries required for executing this python file. If it gives any error while installing libraries, you might need to install them individually.)



1. Run main.py in your anaconda terminal and make sure to change the path where your executable files are located.
2. If you would like to run .ipynb file, Please [follow the link to setup and open jupyter notebook](https://techieyantechnologies.com/2022/06/get-started-with-creating-new-environment-in-anaconda-configuring-jupyter-notebook-and-installing-libraries-using-requirements-txt-2/), You will be redirected to the local server there you can select which ever .ipyb file you’d like to run and click on it and execute each cell one by one by pressing shift+enter.

***Dataset Description***

The dataset was downloaded from a private data repository which might not be available now. The dataset is a .csv file splitted into train and test dataset. The train training dataset consists of around more than 2 lakh data entries of individuals from around the country. The test data consists of around 28000 data entries on individuals without the target class, which we need to predict by training a model. There are 13 columns associated with training file and 12 columns with test file. Some of the important columns which had a huge impact on the approval estimation are, profession, current\_house\_years, house\_owner, car\_owner and income.



***Issues Faced***

1. We might face an issue while installing specific libraries, in this case, you might need to install the libraires manually. Example: pip install “module\_name/library” i.e., pip install pandas

2. Make sure you have the latest or specific version of python, since sometimes it might cause version mismatch.

3. Adding path to environment variables in order to run python files and anaconda environment in code editor, specifically in any code editor.

4. Make sure to change the **path in the code** where your dataset/model is saved.

**Refer to the Below links to get more details on installing python and anaconda and how to configure it.**

<https://techieyantechnologies.com/2022/07/how-to-install-anaconda/>

<https://techieyantechnologies.com/2022/06/get-started-with-creating-new-environment-in-anaconda-configuring-jupyter-notebook-and-installing-libraries-using-requirements-txt-2/>

***Note:***

**All the required data has been provided over here. Please feel free to contact me for model weights and if you face any issues.**

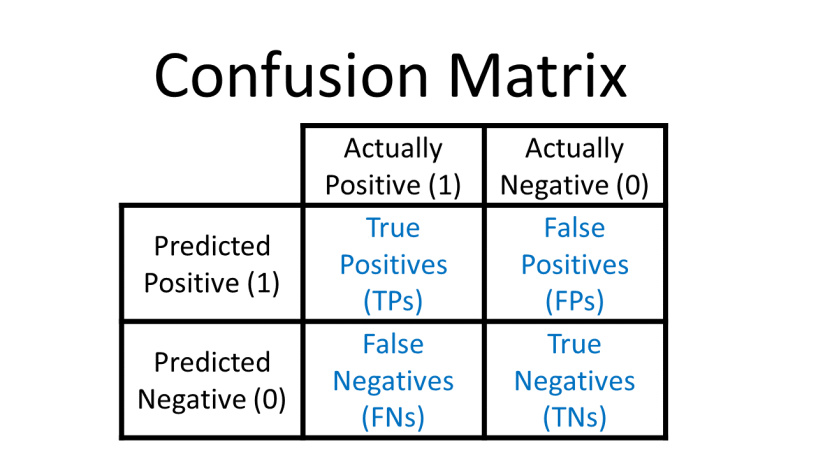
<https://www.linkedin.com/in/abhinay-lingala-5a3ab7205/>

***Yes, you now have more knowledge than yesterday, Keep Going.***

***Evaluation Metric***

Evaluation metrics are considered as one of the most important steps in any machine learning and deep learning projects, where it will allow us to evaluate how good our model is performing on the new data or on unseen data. There are a lot of evaluation metrics such as confusion matrix, roc\_auc\_curve, f1\_score, recall, precision and each of which work for specific problem we deal. So, for our project we have gone with confusion matrix the OG of every evaluation matric, where using it, we have calculated the accuracy and other metric, which has given a conclusion that the model is performing very well on new data.

***Confusion matrix:***



***Results***

