The third entry for my portfolio will be based on the week 7 lecture on Random Forest and feature engineering. Following the week 6 lecture where we were first introduced to classification we learnt about classifiers such as Decision Trees and Random Forests or better known as ensemble models.

According to Eremenko ‘A decision tree functions similarly to a flow chart. It runs tests on individual attributes in your dataset in order to determine the possible outcomes and continues to add results as further tests are run, only stopping when all outcomes have been exhausted. ‘

Eremenko explains builds upon the principles of decisions trees through ensemble learning. Instead of there being only one tree, a random forest will use many different trees to make the same prediction, taking the average of the individual trees’ results.

With this dataset I am going to encode the lablel’s and hopefully predict an entirely new column. Our model will hopefully help us predict some features that it hasn’t yet seen.

For this assignment I used a dataset which has information of credit cards users

familiar dataset museums.csv this dataset was used in our first semester for the STEM assignment. I did notices that as rich as the dataset was with data there were a number of columns that missed values.

The Office of national statistics is a government departments which collates information on statistics on sales, population, birth death rates etc. It provides this data for free here. During my exploration of the dataset I came across a number of intresting data set from how many bananas are sold on a daily basis to how much immigration has grown in the last year. In the end I decided to use data from the ONS website and combine it with some general knowledge I have of a subject all ready and use that to conduct a finding.

There’s been alot of focus on the performance of the economy in the last few weeks which is unexpected as its an election year and the financial year is soon to end. The bank of England announced that the economy grew by … which was welcome news to the chancellor. They announced that this was mostly down to retail. This is an indication as to how a signle market can effect the entire economy.

This very rarely happens by chance, the performance of the retail industry is seasonal. Christmas is normally its peak time, while the preceeding period is the opposite

During World cup 2022 an estimated 14 millon people tuned in to watch England loose to Scotland. Other nationas such as Republic of Ireland and Wales were also present in the competion. This resulted in a rise in retail sales.

I am going to take the information from the performance of the economy in the quarter of a footballing competion compare the data with other factors like how many home nations are participating and how far in the competiotn they went and to guess whath these factors will be in the next tournament.

After searching through the office of national statistics site for