Project Proposal IST707

I am looking at a data set I found on Kaggle with a bit of interesting backstory. It is a large feature set about various mobile phones and their attributes such as hardware, software, phone size, etc and the price bands for the phone.

In addition, the data provided with the labels there is a test set for a company looking to figure out how to price various new mobile phones they are looking to release.

The data set contains various data types that will be interesting to see how they affect the overall algorithm and learn what really drives what we pay for in the mobile phone market. Is it because a phone has LTE capability, the battery size, brand name or is there more under the hood that is being taken into consideration?

The data set also had enough variable that feature selection and preprocessing of the data will be important but also allows for working with different algorithms we have been studying to see which allows additional insights. Along with seeing which tuning measure are the best.