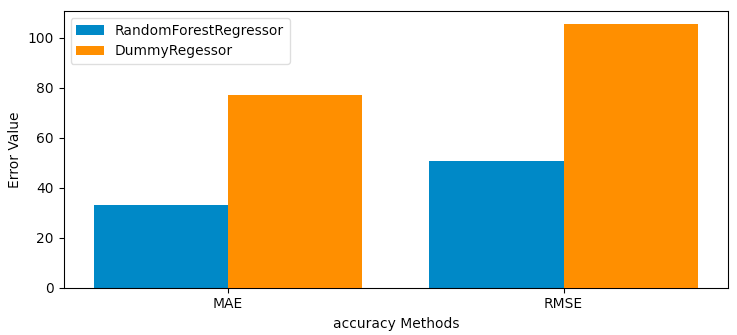
My Model – Random Forest Regression with sk-learn in python

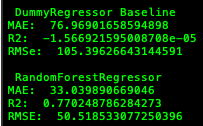
# Model Explanation

Originally, I was going to use decision trees as they are an exceptionally useful machine learning method to know how the prediction is made as each step in the decisions are easy to understand. However, my first few tests showed a flaw; these trees overfit very easily.

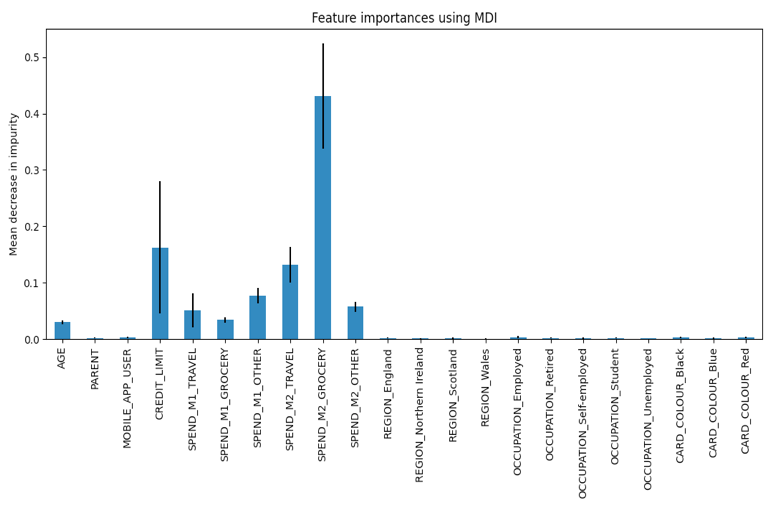
After a bit a research, I came across the idea of random forests. Essentially the idea behind it, is that we have multiple trees, each training on a subset of the training set. We then take all the trees we made and combine them all into on tree. Since they all only worked portions of the training set, any effect of overfitting is reduced as well as an increase in accuracy.

.

Below is a comparison between a regressor baseline and my model, by comparing the mean absolute error and the root mean squared error, it is evident that my model is a lot more accurate: (on the right you can see r2 scores – closer to 1 is more accurate)



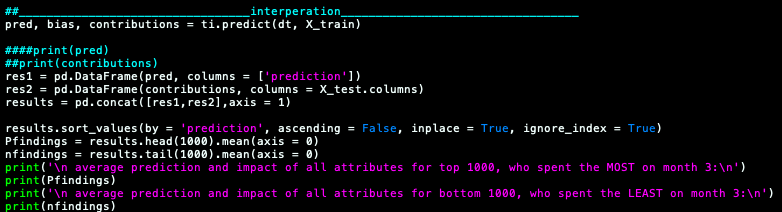
## Getting importance

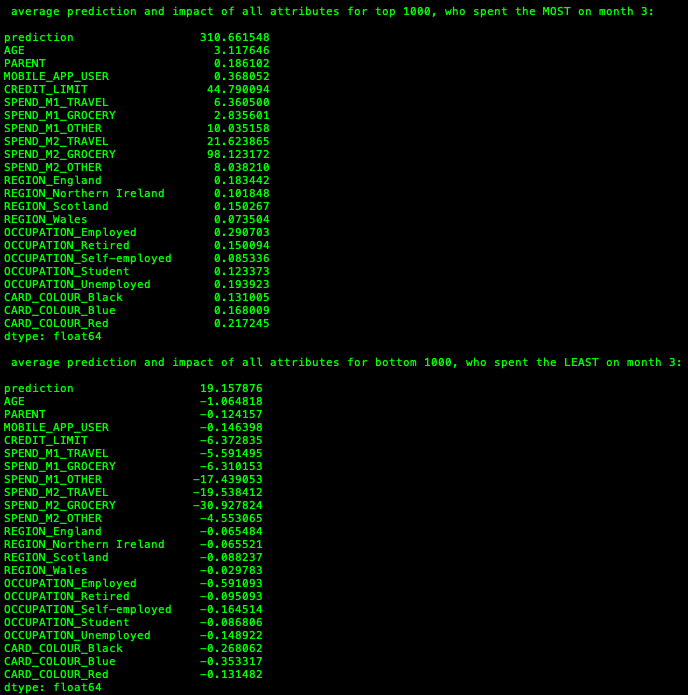
tree.feature\_importances\_ allows us to see which features effect the trees, prediction on the amount that a person would spend in month 3. We can use this to get the average importance of every feature in the random forest model. You can see this on the right →

You can see very clearly, according to the graph, that the region, occupation, and card colour, whether they are a parent, whether they use the mobile app user, are insignificant for the prediction of my model.

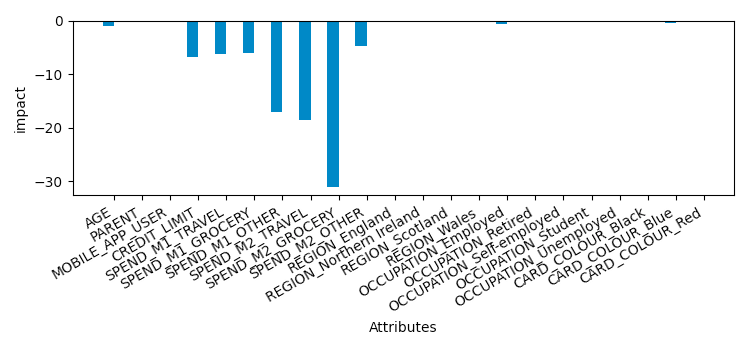
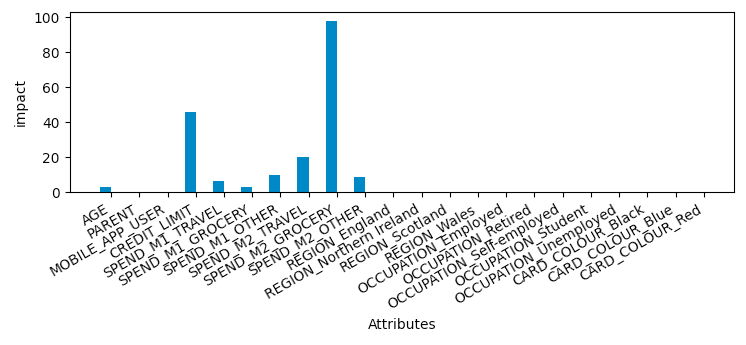
The remaining features impact the model’s prediction the most, with the largest impact of 0.42 being from the amount the person has on spent for grocery in month 2.

Now that we know which features are affecting the prediction of my model, let us look at how they impact it. To do this we will be using the treeinterpreter library.

Running the following section of code:

Gives us the output shown on the right→

Now you can see how each feature impacts the prediction of the randomForest model. An impact with a negative value suggests the model is lowering its prediction due to that feature, whilst a positive value suggest the opposite. The magnitude of these values allows us to see how much the prediction lowers/highers based on the feature.

However, this isn’t not easily readable, so let’s make it a bit more presentable.

Now you that we can see the effects of each feature and we know which features are most important we can explore and find patterns within the database provided to suggest advice to help keep customers involved with your credit card.

Evaluation and Advice:

# Credit\_limit

When the limit is 250 the model tends to give it a negative contribution (mostly between -25 and -5) 🡪 it is predicting that the person will spend less. No positive contribution.

When it is 500, the model gives a scores contribution within the range of -15 to 20 (excluding outliers below -40). This means it has mostly positive contribution 🡪 more likely to be spending more money. Some negative and positive contribution.

When it is 1000 the model gives huge positive contributions of range 20 to 100 🡪 much more likely to spend more money with the credit card. No negative contribution

From this we can conclude that the larger credit limits lead to more spending of money

So, in order to get customers to use their credit card more, you should consider increasing their credit limit.

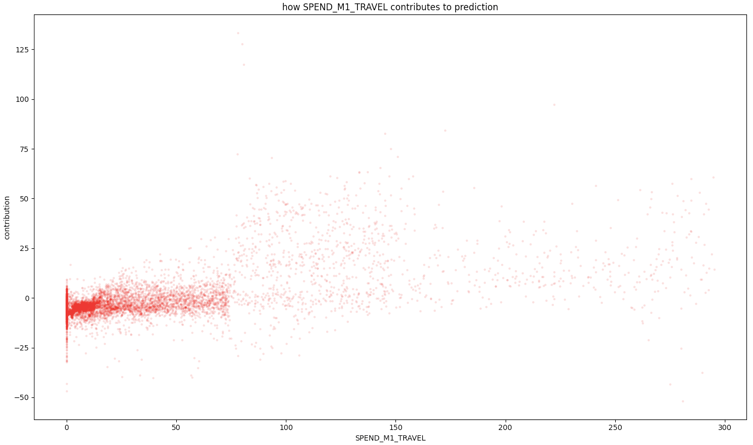
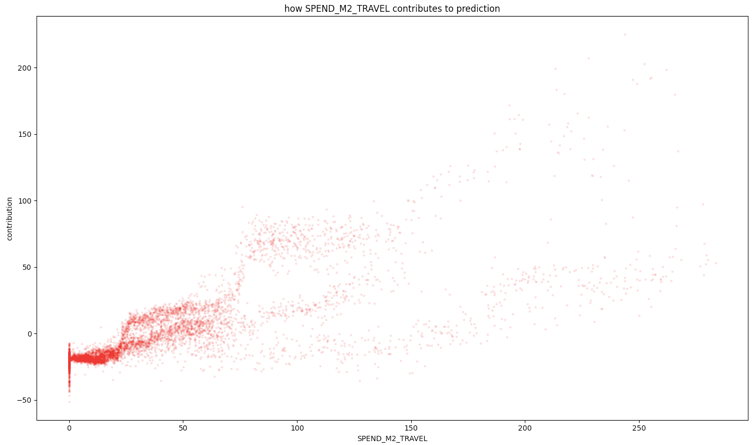
# Amount spent on groceries on month2 –

Here there is a very clear trend; Customers who spend less than £50 on their groceries are less likely to spend more money with their credit card.

However, the contribution of this feature curves up exponentially, showing us that, if the customers spend more than ≈ £70, they are a lot more likely to be spending more money using their credit card.

Due to the fact that people that there a large amount of people spending less money on groceries and that groceries can be considered essential items, it is fair to assume that the majority of these people are struggling financially.

Therefore, a way to increase their willingness to spend more would be to give special offers or rewards to them. One such offer could be a lowered interest rate for a certain period of time. This could allow those struggling to use more of their credit, to get themselves into a better place, from which they can start using money/credits more freely. This method would also make customers already spending more, to stay with your credit card.

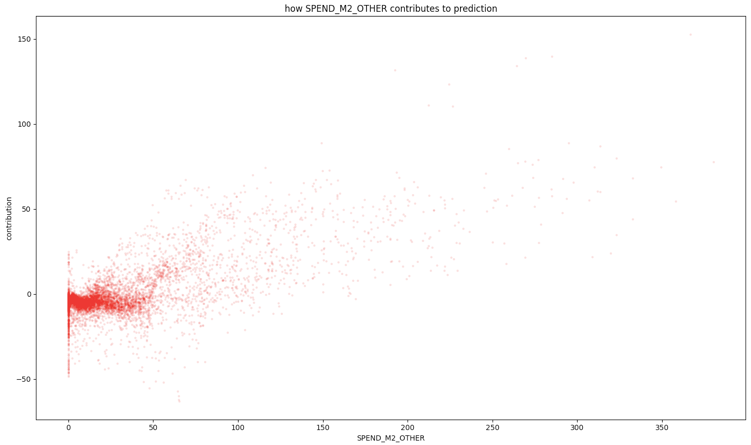
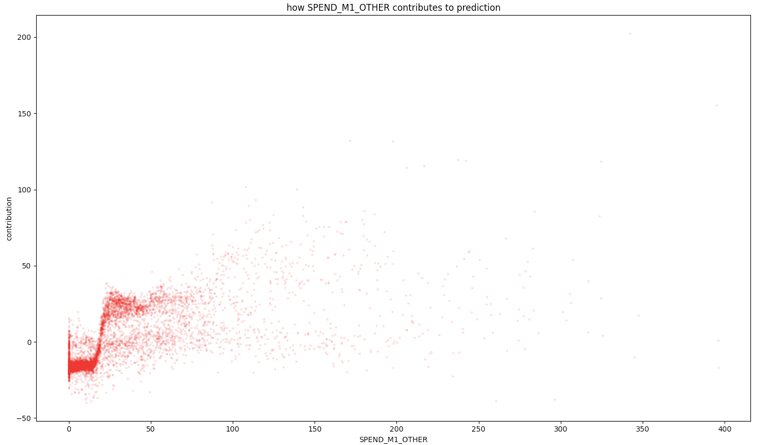


# Amount spent travelling both months-

Both of these graphs show that, those who spend very little (≈< 20) are more likely to be spending less in total. However, in month 2, we see a rapid rise in likelihood of them spending more with their card, where the customers spend ≈ >20. This is not scene for month 1.

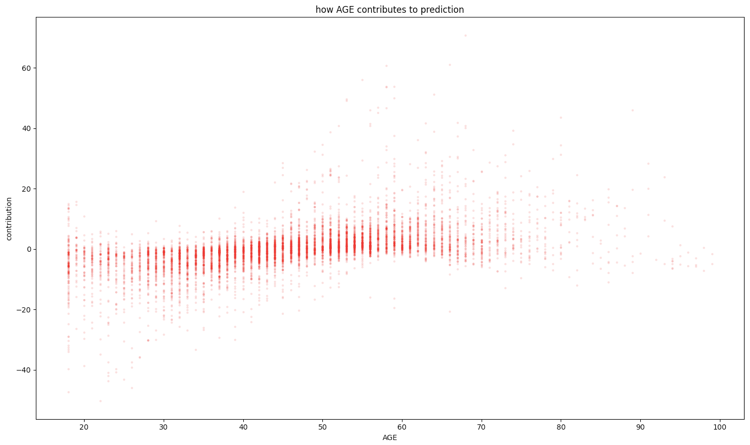
For both graphs, it is clear that, initially, there is a strong negative contribution, but as the amount spent increases, the contribution of this feature spreads, with slight positive correlation.

In this case, I suggest only focusing on those that are less likely to be spending more. This seems very similar to the contribution shown for those spending money on groceries, so I suggest the same approach of reward systems and offers.

Amount spent on other for both months-

While month 2 portrays higher contributions for those spending less, month 1 shows a rapid climb in contribution after ≈25. Both graphs show a lot more spread than previous graphs, but follow similar trends to the travel and groceries graphs. For that, I believe the same approach of reward system and offers should work.

Due to the spread, I would also recommend reminding customers the advantages of a credit card, such as deferred payment.

Their age –

The graph shows a peak in negative contribution around 35 years of age and peak in positive contribution around 60 years age. It is also clear that the majority of 19-year olds tend to be more unlikely to spend more.

I believe this links back to the credit Limit point, as younger people tend to have lower credit limit and credit limit is increased with age, this is likely due to the fact that younger people are less responsible. In this situation, it is best to increase credit limit, for those that are responsible and have better credit scores. I would also advice, giving offers and guidance to those with lower credit scores, so that they can increase theirs.

# summary

In simple terms, the best way to encourage customers to use your credit card more would be to increase their credit limit, give offers like increased cashback for purchases and final reminding them the advantage of a credit card over other methods.

In order to keep customers using your card, it would be best to implement a reward system.