James von Kaenel

December 19, 2016

Predicting Medical Procedure Times

I Problem Statement and Hypothesis

Medical procedures are very expensive and time consuming, both for the patient, the doctor and the hospital. However, despite this expense, procedures are often not booked precisely. The doctor and hospital only make educated guesses as to how long each procedure will take. This leads to patients waiting long hours for other procedures to end, or for procedures to reschedule. Sometimes the procedures will also finish more quickly than anticipated, causing operating rooms to be under-booked and losing opportunities for more procedures to take place on a given day.

My hypothesis is that by using past data concerning procedures, the length of procedure time in minutes can be predicted better the average procedure time as well as better than the educated guesses made by the doctors and hospital staff

II Description of Dataset

The dataset I used were the 2016 procedures from one Spine Intervention Medical Specialist, with the procedures being performed at Fresno Surgery Hospital. The data was given in a text format which was difficult to clean. No names were included in the data which protected the privacy of the patients.

The main categories of data included were:

Demographic Data: This included data specific to the patient, such as Patient Age, BMI, Gender, Insurance

Patient Medical History: How many procedures has the patient had, what allergies does the patient have,

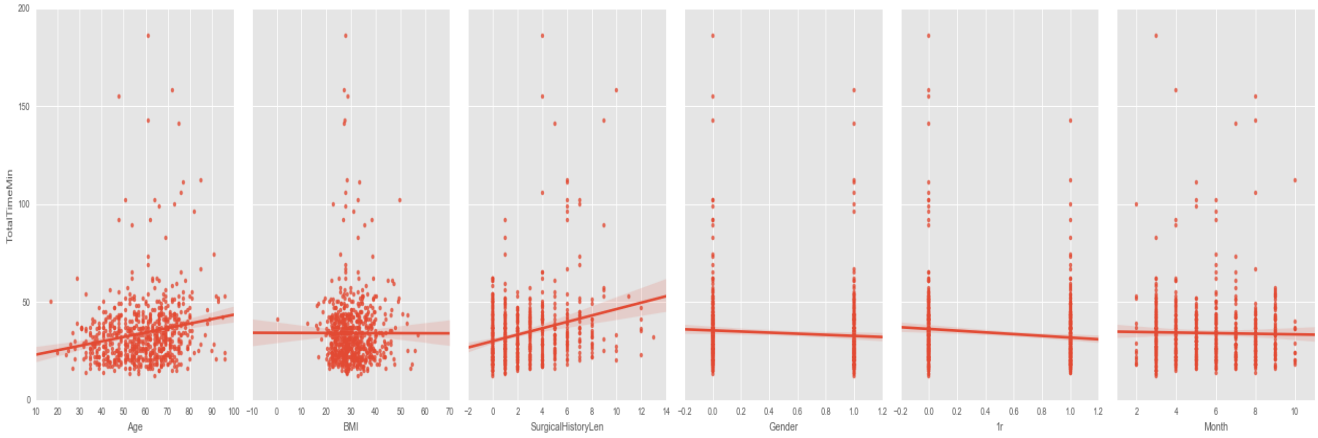
Procedure Operational Data: When the procedure took place, were there one or two rooms available, what were the CPT codes describing the procedure, what was the name of the procedure, how long did the procedure take.

Some of the data was missing values that were crucial to my predictions, namely total procedure time, so I was forced to throw this out. The complete data totaled 672 total procedures. The average length of a procedure is 34.18 Minutes

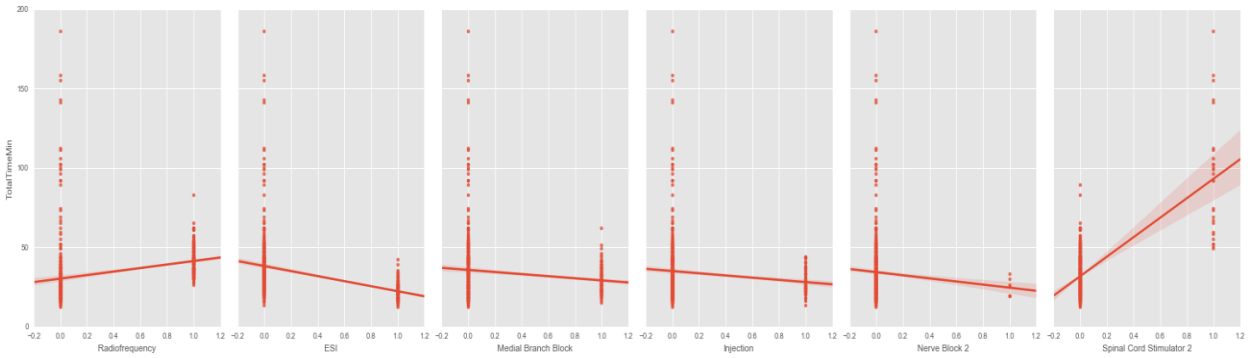
III Data Exploration

Through Data Exploration I could see that the most crucial elements in predicting the total time of each procedure are data which concerns the type of procedure being performed. For instance

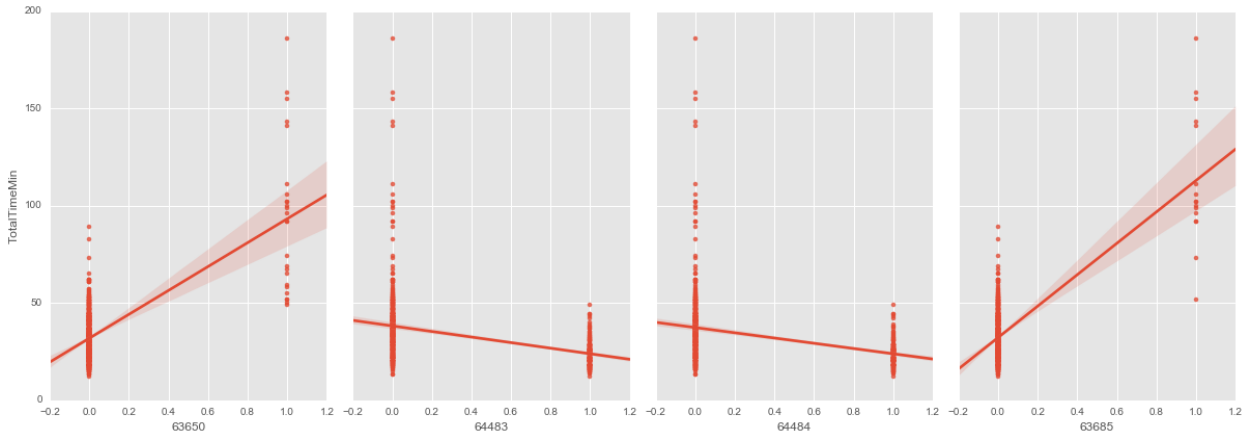
Demographic Data is not very predictive:



While predictions using Procedure Names are better:



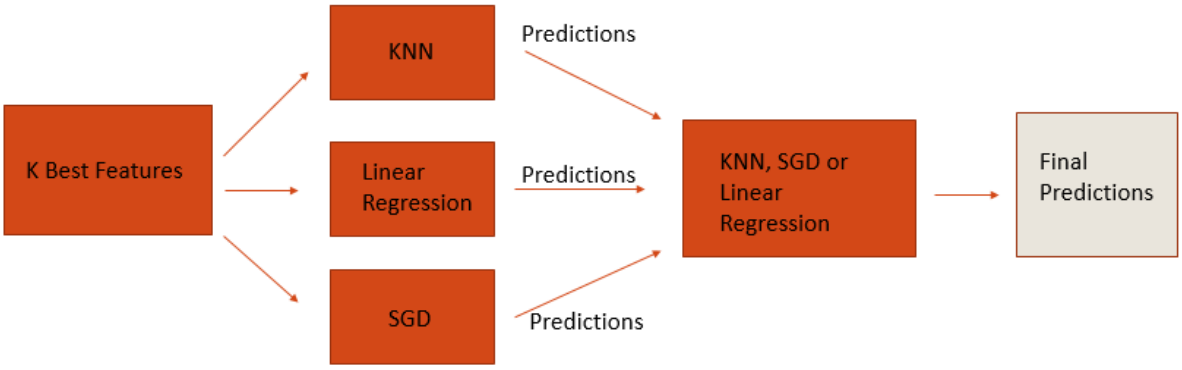
And using CPT Codes are best:



IV Feature Selection and Modeling

Algorithm Selection

I found that simpler algorithms tended to work better then more complicated ones. Algorithms like KNN and linear regression outperformed Random Forests. Also, these models were faster to run and were much less variable then KNN or Linear regression, making them easier to work with. I also created an ensemble of the predictions of KNN, Linear Regression and SGD, which I fed into another KNN, Linear Regression or SGD Algorithm.



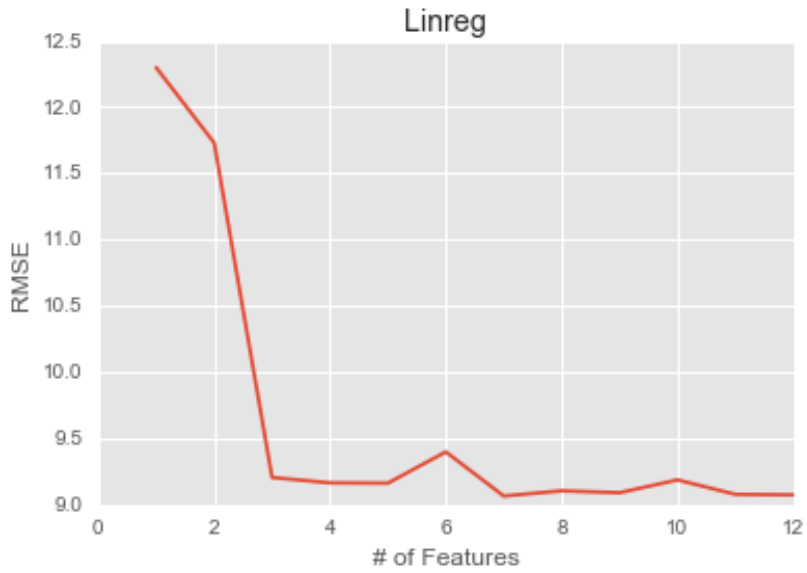
Feature Selection

I used stepwise feature selection where I selected the K best features at every feature length (i.e the best one feature, the best two feature etc.). Initially when I used this method I was getting very large results when I would try to replicate my algorithm in a cross-validation outside of my initial feature selection. For instance, in the initial selection code I would get a RMSE of 8.9 while in the Cross Validation outside of the code I would return a RMSE of 120,208. To combat this I looped my Select K best code and totaled the features that were selected most often. I then used these features as my "Best" features, with the feature selected the most, as the first feature to be used in my algorithms.

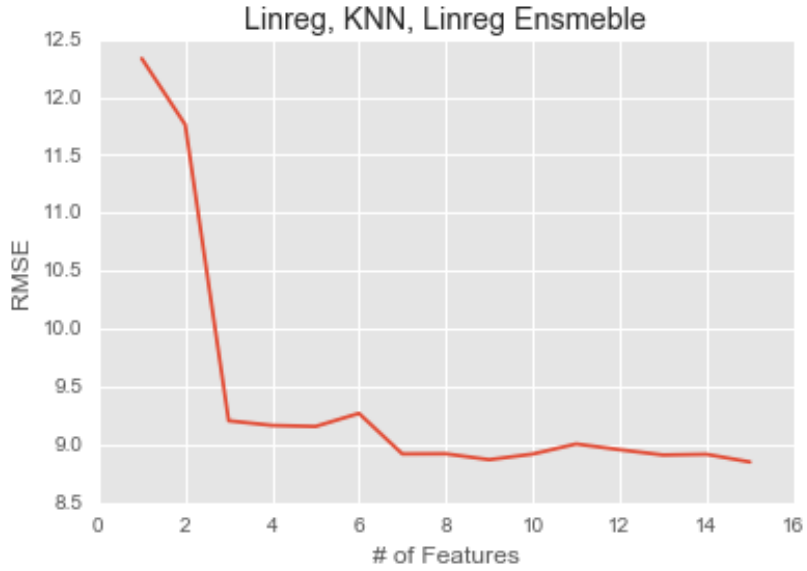
Results:

I found that Linear Regression algorithms or ensembles which were fed into linear regressions were the most effective. I also found that seven features was probably the correct number of features to choose. The RMSE could go lower after 7 features, but there was no obvious pattern showing that further features were useful. Below are three examples of such algorithms, with their outputted RMSE plotted with each additional feature.

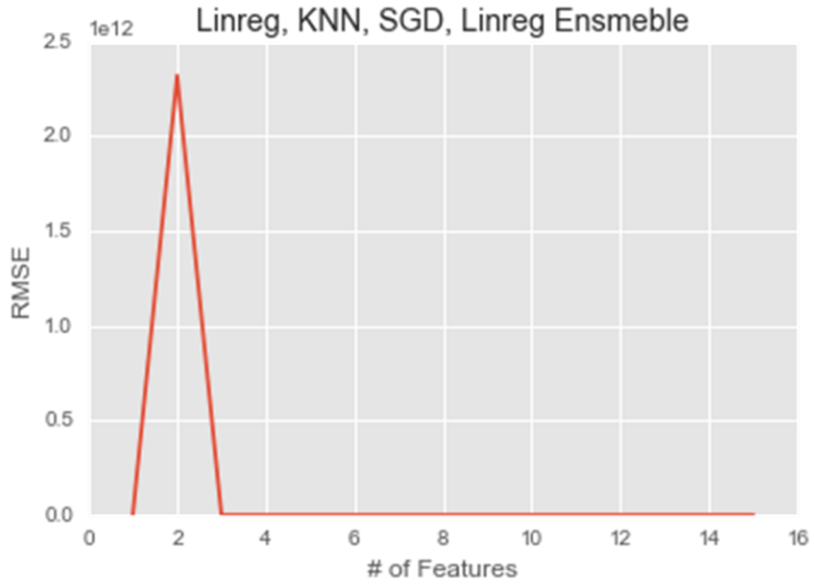
Standard Linear Regression:



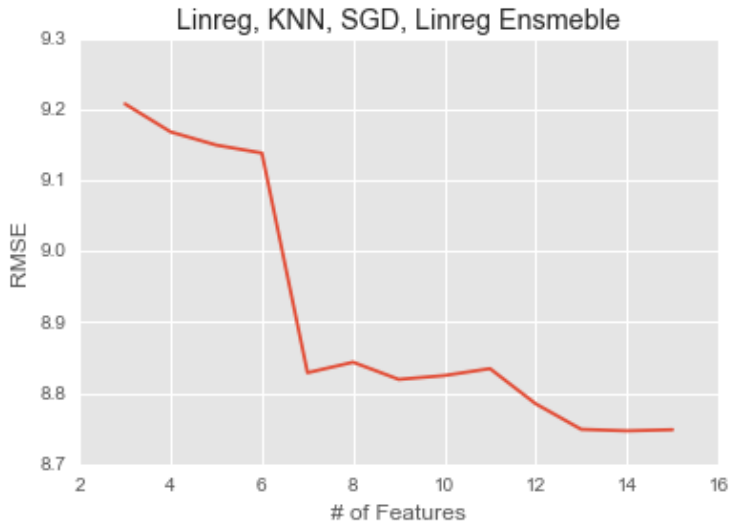
Linear Regression, KNN predictions ensembled into a Linear Regression:



Linear Regression, KNN and SGD predictions ensembled into a Linear Regression:



Same graph as previous, starting after the second feature.:



The best performing algorithm was the Linear Regression, KNN and SGD – Linear Regression algorithm with a final RMSE of 8.82 and these 7 features:

['63685', '63650', 'Radiofrequency', 'Lumbar Radiofrequency', '64636', '64635', 'ESI‘]

I will also add the feature for whether there is or is not one room, as I was told by the doctor that this is important. This feature brings the RMSE down to **8.8**

V Conclusion

In conclusion, I was able to beat the null score, which using the average procedure time was **RMSE of 17.04**. I also beat the approximated educated guess of the doctor staff, which I estimated to be the average procedure time, for each procedure type, this was a **RMSE of 10.23**. I believe that if the doctor does use my algorithm that he will have significantly improved wait times for the patients and that they will be able to fit more procedures into the schedule.