Individual Contribution and Reflection

David Sharp ds16797

Student no. 1618766

I would describe my role as a part of this project as primarily a researcher and secondarily project management to a degree.

During the initial stage of the project I did a lot of work carrying out preliminary data exploration which involved calculating correlation coefficients for pairs of variables. This helped inform future exploration by other group members.

Once the project had gotten fully underway, I spent time researching related work in the field to see what methods are commonly used for similar problems and what prerequisites must be met to allow some methods to be used. Time was also spent researching the medical side of the dataset to try and fully understand what some details meant e.g what tests were used to get metrics and what outcomes from those tests meant, this knowledge was then passed on to the rest of the group.

The non-project outcomes from this period was that I gained appreciation for the value in making your results simple to understand and more importantly as simple as possible for others to reproduce. It was surprisingly difficult for me reading some papers to try and comprehend what had been done with what data. Additionally it made me realise how necessary it is to have somebody in the group who comes into the project from the data background, as understanding of the medical data was often difficult to obtain while being entirely required to completely understand and use the dataset.

Throughout the project I was one of the members trying to handle the administrative side of meeting up digitally both as a group and with the supervisor as well as trying to handle the software needed to work as an effective group.

In terms of technical knowledge gained, I have learnt a lot about data science as a pipeline rather than my previous understanding of it as component parts. I now know that it is essential to treat data science akin to product development for a client where the problem has to be explored, a solution prototyped, and then evaluated, before the process begins again in order to iteratively reach a more complete solution. Where the lectures provided knowledge of the component parts, cleaning, modelling, etc. the project provided the framework to string the components together into a functioning system.