Supervisor Notes

26 June

What have I done:

Implemented positional encoding for the coordinate variables – not started working on a transformer model yet.

Started to work with CatBoost as an additional decision tree method.

Implemented randomisation of numbering of players for position and distance data – had little impact on results so far.

Working on project plan.

Questions:

Some papers talk about an ensemble method combining XGBoost with a neural net is best – any advice on how to implement this?

Ethics requirement.

Feedback:

Permutation invariance – network which has sets as inputs.

Possibly using TabnNet or NN architecture designed for tabular data.

10 July

What have I done:

Finished project plan.

Done benchmarking for my different methods at the moment.

Improved XGBoost method slightly.

Performed hyperparameter optimisation for neural network method – observed longer optimisation time for neural network than XGBoost.

Gone back through data gathering/prep and EDA.

Questions:

EDA – 2D histograms vs kde plot?