

COMP6237 Coursework 1 "Predictive Data Mining Group Project"

Feedback

Team: Llanfair PG

Experimentation and Analysis	Analyse the problem and results obtained	30 35 /35
Good amount of experimentation and analysis. I do wonder however if you might have got better results by training on a data set that was less biased - ie by resampling the training data so that all classes were approximately equally likely. I also wonder if you should have spent some time thinking more about the types of user in the dataset, and if by first broadly categorising them you could build better models.		
Application of techniques	Show ability to apply predictive data mining techniques and preprocessing operations	28 /35
Overall approach to pre-processing & feature engineering seems sound. Range of predictive techniques used, although I wonder if you might have been better starting with a simple baseline and building incrementally (e.g. naive bayes, knn, single decision tree, ...) - that would help better quantify and understand the performance seen. AUC but seems weak.		
Reflection	Reflect on what the experimental results tell us about the problem and the techniques used	14 /20
Quite a lot of reflection around the techniques you tried, but a lack of context with respect to the other teams in the competition and also more general literature on the prediction of events from web log data. Also, more review and comparison to existing techniques for tackling cold-start would have been welcome.		
Reporting	Clear and professional reporting	8 /10
Abstract? Overall needs a good proof read - lots of small errors making it hard to read. Not much in the way of references.		
Total Mark		80 85 /100
Days Late	0	Adjusted Mark 80 85 /100