|  |  |  |
| --- | --- | --- |
| **Task**  **Don’t forget module prediction**  **https://ai.stackexchange.com/questions/20193/which-classifier-should-i-use-for-a-dataset-with-one-feature**  [**https://machinelearningmastery.com/feature-selection-with-real-and-categorical-data/#:~:text=rank%20coefficient%20(nonlinear)-,Numerical%20Input%2C%20Categorical%20Output,the%20categorical%20target%20into%20account**](https://machinelearningmastery.com/feature-selection-with-real-and-categorical-data/#:~:text=rank%20coefficient%20(nonlinear)-,Numerical%20Input%2C%20Categorical%20Output,the%20categorical%20target%20into%20account)**.** | **Notes** | **Date** |
| ~~Extract and plot dropout distributions~~ | ~~Note: Do across courses.~~ | **~~Sat 13 Aug – 3:30~~** |
| ~~Work out a definition for dropouts based on the distribution~~ | ~~Note: Do across courses.~~ | **~~Sat 13 Aug – 3:30~~** |
| Create new interaction strings using only the important slides. | Note: Compare performance of new interaction strings with old, and visualise well.  Note: Do across courses. | **Sat 13 Aug – 4:30** |
| **WORKOUT** | **WORKOUT** | **WORKOUT** |
| ~~Create, run and evaluate simple classifiers for dropout~~ | ~~Note: Performance of classifiers might not be great if there are clusters. Explain why.~~  ~~Note: Look at the data and then explain which model might be best – e.g. if sequence, then maybe LSTM.~~  ~~Note: Do across courses.~~ | **~~Sat 13 Aug – 9:00~~** |
| Create, run and evaluate simple classifiers for performance | Note: Difference between performance measures (# of slides completed and final score considering submissions)  Note: Performance of classifiers might not be great if there are clusters. Explain why.  Note: Look at the data and then explain which model might be best – e.g. if sequence, then maybe LSTM.  Note: Do across courses. | **Sun 14 Aug – 9:00** |
| **Create a classifier with number of submissions made and performance** | Note: Again, perform an initial clustering analysis on the data – are there different types of students e.g. students who make high number of submissions but don’t do well (i.e. spammers), students who make low numbers of submissions but do well, people in the middle etc. | **Sun 14 Aug – 11:00** |
| **Create a classifier with number of submissions made and dropout** |  | **Sun 14 Aug – 3:00** |