Changelog

* Added abstract
* Expanded introduction
* Rewritten problem description with more background information and information overall
* Explained stratified sampling strata choices
* Added information on finding relations in the data
* Extended threats to validity and evaluation slightly
* Updated results;
  + removed most of the right skewed figures as they all looked the same;
  + added a table instead of the figures
  + added explanation of testing the model
  + added discussion of feature importance
  + Overall; added a lot to the results section
* Updated future work with other aspects
* Added conclusion

General

* Made text more formal
* Removed some more spelling mistakes
* Added references where needed