Capstone Project Proposal Template

Notes:

- This should take no more than one hour to complete – the clearer you are about the business problem you're working to solve with your ML-driven solution, the easier your proposal will be to complete
- This will be uploaded to your repo, which will be a part of your final submission
- Due date for submission is 1/16

Instructions:

1. Download this document as a Word Doc
2. Answer each question using a few sentences, at most
3. Save your completed proposal as a PDF
4. Create a project GitHub repo (if you have yet to do so)
5. Add your instructor as a collaborator (username dodgy719) to your project repo
6. Add your mentor as a collaborator
7. Push your proposal PDF (created in Step 3) up to your repo
8. Copy the URL corresponding to the location of the PDF in your repo
9. Submit the copied URL using this link

Telecommunication Customer Churn Predictions

Business Understanding

- What problem are you trying to solve, or what question are you trying to answer?
  Trying to predict whether a customer will change telecommunications providers, which is also known as 'churning'
- What industry/realm/domain does this apply to?
  Telecommunications Industry
- What is the motivation behind your project? (Saying you needed to do a capstone project for flatiron is not an appropriate motivation)
  A new Telecommunications company is trying to make its footing in the telecommunications industry and wants to get data on if a customer would change providers and on what basis

Data Understanding

- What data will you collect?
  State, how many months the customer has been with the telecommunications provider, if the customer has an international plan, total number and charge of day calls, total mins of evening calls, customer churn, etc
- Is there a plan for how to get the data (API request, direct download, etc.)?

I will get the data from a dataset on Kaggle – direct download

- What are the features you'll be using in your model?
  Customer churn, area code,if customer has international plan, etc

Data Preparation

- What kind of preprocessing steps do you foresee (encoding, matrix transformations, etc.)?
  One hot encoding
- What are some of the cleaning/pre-processing challenges for this data?
  Data set is straightforward – so no cleaning needed

Modeling

- What modeling techniques are most appropriate for your problem?
  Classification modeling techniques → decision trees, random trees
- What is your target variable? (remember - we require that you answer/solve a supervised problem for the capstone, thus you will need a target)
  Customer Churn (yes/no)
- Is this a regression or classification problem?
  Classification

Evaluation

- What metrics will you use to determine success (MAE, RMSE, Accuracy, Precision etc.)?

Accuracy

Tools/Methodologies

- What modeling algorithms are you planning to use (i.e., decision trees, random forests, etc.)?

Decision Trees, Random Forests