

Anya Rajan

4/28/23

CSE 144

Homework 1 Report

The screenshot shows the Kaggle competition page for 'Customer Churn Prediction 2020'. The header includes the competition title, a description 'Predict whether a customer will change telco provider', and statistics '27 teams · 3 years ago'. A navigation bar at the top has links for Overview, Data, Code, Discussion, Leaderboard, Rules, Team, Submissions, and a Late Submission button. The Submissions section is active, showing a summary: 'You selected 0 of 2 submissions to be evaluated for your final leaderboard score. Since you selected less than 2 submission, Kaggle auto-selected up to 2 submissions from among your public best-scoring unselected submissions for evaluation. The evaluated submission with the best Private Score is used for your final score.' Below this, there are filters for 'All', 'Successful', 'Selected', and 'Errors', and a 'Recent' dropdown. A table lists the submissions with columns for 'Submission and Description', 'Private Score', 'Public Score', and 'Selected'. One submission, 'submission_nn.csv', is listed as 'Complete (after deadline) · now' with a Private Score of 0.8819 and a Public Score of 0.91555. It is not selected.

| Submission and Description | Private Score | Public Score | Selected |
|---|---------------|--------------|--------------------------|
| submission_nn.csv Complete (after deadline) · now | 0.8819 | 0.91555 | <input type="checkbox"/> |

This assignment was very interesting! First, I learned how to prepare data for a neural network (`train_test_split`). In this process, I learned how this function properly splits given data into the training dataset and the validation dataset. Further, I learned how to implement one-hot encoding, which was pretty intriguing. I now understand what features can and cannot be one-hot encoded— Numerical data/yes and no features can be converted, but strings and large numbers cannot be. I learned about how to use the `LogisticRegression()` function, how to create a neural network, how to train that neural network, and, finally, how to evaluate it.