

D&D Combat Length Predictor

A Dungeons and Dragons (D&D) data science project using the large FIREBALL dataset to predict the length of a given combat in number of rounds.

FIREBALL Dataset

The dataset we are using has a GitHub repo here: <https://github.com/zhudotexe/FIREBALL>. We have gotten permission from the managers of this dataset to use it for this project. The data consists of a collection of 24,748 JSONL files, each one containing a script for one full combat encounter run on Discord with the help of the Avrae bot. Each line contains an utterance by a player, and then is followed by an action, with periodic updates on the current game state (each character and their remaining health and resources). We extracted one observation per combat file, with a total of 4 seemingly independent features for predicting the target variable of combat length.

Modeling Approach

We began with some baseline models. These were predicting the mean, median or mode of the combat lengths of the training set. We compared this against linear regression to start. For our more sophisticated models, we used ensembles of trees: Random Forest Regressor, Gradient Boosting Regressor, and XGBoost Regressor, tuned via a gridsearch. To compare the models, we used cross-validation and calculated their average mean squared error, which is an easily interpretable metric which penalizes larger errors more than, say, using mean absolute error.

Results

When the tuned models were compared using 5-fold cross-validation, every model did better than the three baseline models, and XGBoost Regressor did the best. This is the one we then trained using the full training set, and we achieved an MSE of approximately 3.627.

We also created a simple app where a user can plug in the details about their encounter and see the predicted number of rounds. Playing around with this app, it quickly becomes clear that our model struggles with large party sizes (> 7). This is probably because such samples make up a very small portion of our data, and it is uncommon to use this many players in a game of D&D.

Future Directions

To further improve the MSE of our final model, the most impactful change we could make is engineering more features. Four features is relatively small, and the model did quite well considering this. Some possible features to consider which could be extracted from the FIREBALL dataset are various categorical features about the player party, perhaps capturing the party composition beyond just how many characters and their average level.