Overview

Which tasks have been completed?

- We have cleaned three datasets (CISI, NPL, Cranfield) by processing their documents, queries, and relevance files. We have the data stored in the needed format that enables us to use it for parameter tuning and evaluating.
- We gained a better understanding of all of the MeTA scoring functions and the appropriate ranges of their parameters.
- We researched and implemented grid search and random search tuning methods.
- We investigated assessment techniques such as k-fold cross-validation, as well as training-test separation and validation.
- We added a config table for the parameter tuning layer, and added a generic wrapper to search through the scoring functions we support to give an overall best model based on test score.

Which tasks are pending?

- Potentially implementing more training dataset collection and pre-processing.
- Implement more Hyperparameter tuning methods (currently implemented methods are grid_search and random_search using sklearn).
- Potentially adding more scoring methods (other than NDCG@k).
- Testing and optimization
    - Implementing functionality to determine the most optimal parameter configurations across every MeTA scoring function and testing the functionality on all datasets.
    - Plot graphs of metrics (i.e. NDCG@k) vs parameter value.
- Improve model evaluation methods.
- Documentation and presentation.

Are you facing any challenges?

- Limited documentation of data, resulting in difficulty in figuring out how to process and translate each dataset into a consistent format.
- Ensuring that the hyperparameters chosen through the tuning layer are not merely optimal for a particular scoring function or dataset but exhibit effective generalization across a range of datasets or functions.
- With limited size of training data it’s generally hard to evaluate on unseen/test data.