Extend the Elasticsearch to support more ranking models

Jining Song (captain) – jinings2@illinois.edu

I've chosen to extend the Elasticsearch to support more ranking models and evaluate the performance of each model with the professor information dataset which we created at the beginning of the semester. Elasticsearch is one of the most popular and modern way to index large amount of document, enabling creating large scale searching applications. The service itself and the ranking models which I'll be working on are highly related to this class.

Currently, the available similarity modules (ranking models) on Elasticsearch are BM25, DFR, DFI, IB, LM Dirichlet, and LM Jelinek Mercer. In this proposed study, I'll be implementing at least two more state-of-the-art ranking models through the 'scripted similarity' functionality provided by Elasticsearch.

I'll use the professor information dataset to verify the integration and test the performance of the two new ranking models. I'll also compare the two new ranking models with the rest of pre-defined models.

This project will be implemented in python or JavaScript and the expected workload for this project is 20 hours (This will be a single person team).