Recommender systems are increasingly popular for popular services like Amazon, Netflix and Spotify. One can hardly think of a widely adopted product that doesn’t include recommendations as part of the core value proposition. On the other hand, very few smaller services with inherently small data sets are even able to provide these valuable recommendation services. Through this project, I intend to explore the literature, implement a model that works on a small dataset generated from under 10k users, deploy it to a set of users and assess its impact and quality of recommendations. Progress permitting, I hope to implement a data pipeline that would enable continuous learning and deployment of the underlying recommender system.

The area of focus of this dataset and recommender system is generally food recommendations — both restaurants and specific items at a given restaurant. While preferences for songs, movies, articles, books and other emotion driven, visceral preferences have been decoded by popular products powered by recommender systems, food is a domain that still has not seen the application of a successful recommendation service. Certainly services like UberEats, GrubHub, Yelp and even Google Maps have the data and strong incentive to answer the daily question of “where should I go to eat?”, but no one service has a good enough recommendation service that makes it top of mind for most people.
Leveraging my combined interest in machine learning and software engineering, this project attempts to provide the deployment of an enjoyable recommender system. The main technical tasks that this entails, include:

1. Gather, process and structure raw data
2. Explore the literature for recommender services that work well with comparatively few data points
3. Implement a few different models from the most suitable discovered from the literature
4. Perform an offline evaluation of the different models and determine which should be shipped live
5. Deploy one model to product in a fast and scalable manner
6. Measure and test the efficacy of the model in deployment
7. Implement a data pipeline that enables the continuous training and re-deployment of the model

While machine learning is a new, challenging and exciting field of computer science in the past few years, industry and academia are realizing that so to is the development, deployment and maintenance of systems that support these machine learning models in production. I’m keen to learn software engineering best practices while working under Professor Piskac’s guidance. The time she’s spending advising this project is most appreciated!