

Course: Big Data programming / CMPT 732 / Final Project

Team members: Bhuvan Chopra, Junaid Qazi

Project proposal:

“Business Intelligence based on product rating and customer reviews.”

In this project, we will focus on both entities, the customers and the company, to implement recommender systems to save their time, money and provide data driven solutions. The project will provide suggestion to the customer for a relevant product based on rating and their product search in departments with-in the store. The idea is also to provide business intelligence to the company so that they buy and stock the most relevant and popular product based on the customer reviews and rating. Another bigger aspect of this project is targeted marketing which is one of the big interests for all companies to improve their customer base and product sale. We have plans to work on the dataset from bigger stores such as Canadiantire, Walmart or BestBuy etc.

Strategies and possible technologies:

- Web scrapping for product reviews and star rating
- Cassandra database for ETL
- Python and spark libraries for Machine learning and natural language processing.
- Possible implementation of feature engineering
- Scalable project to run on the cluster using big data technologies.

We would be pleased to consider and discuss other possible technologies with the project / course mentor.