

Team Members Contribution

Sameedha Bairagi (sbairagi)

1. Designing strategy for Task 2.
2. Forming hierarchical model for all businesses so that we can query things very easily.
3. Dividing whole data in two sets: Training Data and Test Data in ration 60:40
4. Indexing all reviews and tips using Lucene.
5. Extracting top n features from the reviews
6. Executing three similarities (Vector, BM25 and LMDirichlet) to extract matching attributes with respect to train and test data.
7. Managed data on MongoDB.
8. Generating restaurant business json (businesses with category as restaurants) and featureMap json from Word2Vec (used directly to get search query for each attribute).
9. Git Repo Version Control.

Deepak Bharanikana (dbharani)

1. Helped with algorithm and module design for task 1.
2. Helped with strategy and module design for task 2.
3. Helped with Attribute map creation.
4. Extracting most talked about features from the reviews database.
5. Researched & learned mongodb usage.
6. Worked on project presentation slides.
7. Worked on project report.

Milind Gokhale (mgokhale)

1. Prepared the design and strategy for task 1.
2. Implemented task 1 along with indexing and mongo database.
3. Created datatypes for business, review, tip and sub types for json files parsing and data handling.
4. Created indexes for the dataset – training and test set.
5. Extracted category features.
6. Also POS tagged the review text to prune category features.
7. Collated the outputs for task 1.
8. Carried out evaluation of task 1.
9. Helped prepare the project report for task 1.
10. Managed Git repo and version control to manage the project.

Namrata Jagasia (njagasia)

1. Designed Strategy for task 2.
2. Forming hierarchical model for all businesses so that we can query things very easily.
3. Dividing whole data in two sets: Training Data and Test Data in ration 60:40
4. Indexing all reviews and tips in lucene.
5. Extracting top n features from the reviews
6. Run BM25 on the extracted attributes as a query string to get score for that attribute for that city.
7. Managed data on mongodb.
8. Optimized data retrieval .Reduced the time from 12 hours to less than 30 minutes.

Siddharth Jayasankar (sidjayas)

1. Helped with the design and strategy for task 1.
2. Created the data set of tips, reviews and business files by filtering out only restaurant businesses.
3. Helped with the algorithm and strategy for task 2.
4. Explored the possibility of using Topic modeling (LDA) by using Mallet in task2
5. Explored the possibility of incorporating Word2Vec by using DL4J in task2
6. Performed the evaluation of task2
7. Prepared the Slides for the class presentation
8. Prepared the project report
9. Compiled and managed all project related documents.