**DatasetReader.java (inside package ir.yelp.jsonparser)**

Given the file name, this reads the json file and returns json objects.

**BusinessCategories.java (inside package ir.yelp.jsonparser)**

Puts the categories of all businesses into a hashmap with businessId as the key

**LocationExtractor.java (inside package ir.yelp.jsonparser)**

Puts the relevant fields and location information of all businesses into a hashmap with businessId as the key

**PART 1:**

**IndexCreator.java (inside package ir.yelp.indexcreator)**

Creates index in batches. Gets a concatenation of review and tip of 10,000 businesses and indexes them. If the business is already in index, it would retrieve business from index if it is already there in the index, modify it and update the index.

**indexCreatorNonBatch.java (inside package ir.yelp.indexcreator)**

Creates indexes of all the business at once. Appends all the review and tip to a hashmap and then reads and hashmap and indexes each business.

**statsOnIndex.java (inside package ir.yelp.indexcreator)**

Prints statistics on the index like number of businesses in the index, vocabulary, etc.

**categoryListAndCount.java (inside package ir.yelp.jsonparser)**

Gets a list of unique categories and the counts of businesses with that category

**QueryFire.java (inside package ir.yelp.indexcreator)**

Loops over the list of categories and fires each of the as a query. Improves the query if there is improvement in precision for the business. Prints to a file, the category as initial query and the top documents.

**AccumulateCategories.java (inside package ir.yelp.filemanipulations)**

It read the file of the format

Business1 Category1

Business2 Category2 …..

And converts it to

Business1 Category1|Category2….

Business2 Category3|Category2….

**evaluatePredictedCategories.java** **(inside package evaluation)**

It computes the precision and recall for each category, and then computes the average precision and recall.

**Part 2:**

**dataWriter.java (inside package ir.yelp.jsonparser)**

Reads the json files for review, and extracts reviews for the two city Montreal and Edinburgh for businesses of category “Restaurants” and writes the two dataset to two files.