Dataset: You can download the dataset from:
Yelp Dataset Challenge: https://www.yelp.com/dataset/download
Kaggle: https://www.kaggle.com/yelp-dataset/yelp-dataset

Phase 1: Data Manipulation

- Remove Outliers
- Merge two data files

Data set: business.json,
review.json
Code: IE7275ProjectCode1.py
Language Used: Python
Libraries/packages Used: json
Results: trimmedData_2.json,
trimmedReviewData.json,
mergedData.json

Phase 2: JSON Manipulation

- Drop bad Attributes

Data set: trimmedData_2.json
Code: json_cleaning.ipynb
Language Used: Python
Libraries/packages Used: boltons and json
Environment Used: Jupyter Notebook
Results: trimmedData_2.json

Phase 3: JSON to CSV Conversion

- Scripting json to csv

Data set: trimmedData_2.json,
trimmedReviewData.json
Code: JSONtoCSVScript.R
Language Used: R
Libraries/packages Used: jsonlite
Environment Used: RStudio
Results: yelp_data_business.csv
yelp_data_review.csv

Phase 4: Open yelp_data_business.csv in Excel and save it as a workbook as yelp_data_business.xlsx

Phase 5: Sentiment Analysis and Data Preprocessing

- Sentiment Analysis
- Data preprocessing for model

Data set: yelp_data_business.xlsx,
yelp_data_review.csv
Code: Yelp_Sentiment_And_DataPreprocessing.ipynb
Language Used: Python
Libraries/packages Used: pandas, numpy, sklearn, re, matplotlib, sys, os, and nltk
Environment Used: Jupyter Notebook
Results: model_data.csv

Phase 6: FeedForward Neural Network

- Model
- Performance Evaluation

Data set: model_data.csv
Code: Yelp_FNN.Rmd
Language Used: R
Libraries/packages Used: readr, keras, RSNNS, tensorflow, ggplot2, tidyquant, and plotly
Environment Used: RStudio

Phase 7: LSTM Neural Network

- Model
- Performance Evaluation

Data set: model_data.csv
Code: Yelp_LSTM.Rmd
Language Used: R
Libraries/packages Used: readr, keras, RSNNS, tensorflow, ggplot2, tidyquant, and plotly
Environment Used: RStudio

Phase 8: Polynomial

- Model
- Diagnostics

Data set: model_data.csv
Code: Yelp_polynomial.Rmd
Language Used: R
Libraries/packages Used: mice, dplyr, corrplot, RColorBrewer, data.table, ggplot2, ISLR, polynom, locfit, multipol, car, MASS, leaps, bootstrap, orcutt, hydroGOF, modelr, mlbench, and caret
Environment Used: RStudio