Mini-Project 2: Yelp Business Rating Prediction using Tensorflow

Version 1.0

CSC215, Fall 2018

Oct 8th, 2018

Prepared by: Chandini Nagendra

Siddharth Chittora

Contents

[1. Problem Statement 2](#_Toc526176076)

[2. Methodology 3](#_Toc526176077)

[3. Experimental Results and Analysis 3](#_Toc526176078)

[4. Task Division 3](#_Toc526176079)

[4.1. Chandini Nagendra: 3](#_Toc526176080)

[4.2. Siddharth Chittora 3](#_Toc526176081)

[5. Project Reflection 3](#_Toc526176082)

[6. Additional Features 3](#_Toc526176083)

# Problem Statement

In this project, we aim to predict a business's stars rating using the reviews of that business and review count based on neural network implementation in Tensorflow. This project is twofold:

Task 1: Consider this problem as a regression problem. Compare the RMSE of the BEST **Tensorflow regression neural network model** you obtained with that of **regression model** you achieved in the last project.

Task 2: Consider this problem as a classification problem. Compare the accuracy of the BEST **Tensorflow classification neural network model** you obtained with that of **each classification model** you achieved in the last project.

# Methodology

## Regression Problem

### Linear Regression

show predicted and actual for 10 business

show RMSE

### Logistic Regression

show predicted and actual for 10 business

show RMSE

### Simple Regression

show predicted and actual for 10 business

show RMSE

**Activation**

### Regression with early stopping and modelcheckpoint with relu

show predicted and actual for 10 business

show RMSE

### Regression with early stopping and modelcheckpoint with sigmoid

show predicted and actual for 10 business

show RMSE

### Regression with early stopping and modelcheckpoint with tanh

show predicted and actual for 10 business

show RMSE

**Optimizer**

### Stochastic gradient descent optimizer(SGD)

show predicted and actual for 10 business

show RMSE

RMSProp

Adagrad

Adadelta

Adam

Adamax

Nadam

## Classification Problem

# Experimental Results and Analysis

# Task Division

## Chandini Nagendra:

## Siddharth Chittora

Discussed together on how to improve the model and came up with the following solution

# Project Reflection

label encoding does not work for tensor flow, when using classification. do one hot coding

with 25 neurons in the 1st layer the r2 score was 60 and with 60 neurons in the 1st and 30 in the second the R2 was 72.

used zscore for tanh

# Additional Features