SENTIMENT ANALAYSIS OF CUSTOMER REVIEW IN ZOMATO DATASET

INSIGHT STRATEGISTS

TEAM MEMBERS: INSIGHT STRATEGISTS

AARUSHI AGARWAL - PES2UG19CS004

ARJUN HARISH - PES2UG19CS062

ANWESHA KELKAR - PES2UG19CS058

MOHNISH SRIKANTH- PES2UG19CS241

PRESENTATION HIGHLIGHTS

FOCUS AREAS

- Problem statement and data set used to solve the problem.
- Importance of Problem statement.
- Approach chosen.
- Evaluation of solution/algorithm we implemented.
- Insights and interesting information gained from the project.

PROBLEM STATEMENT

TO EXTRACT WORDS FROM THE REVIEWS AND ANALYZE THEM, SO WE KNOW HOW USERS INTERACT IN THE ZOMATO REVIEW SECTION AS ZOMATO HAS NO SUCH FEATURE.

DATASET USED

ZOMATO DATASET FROM KAGGLE WAS CHOSEN.
IT CONSISTS OF TABULAR FORM CONTAINING SUMMARISED INFORMATION SUCH AS NAME, ADDRESS, PHONE NUMBER, WEBSITE URL, AND CUSTOMER FEEDBACK.
THE DATA IS IN A CSV FORMAT AND IS APPROXIMATELY 547MB IN SIZE. THE DATASET CONTAINS 17 VARIABLES ALL OF WHICH WERE SCRAPED FROM THE ZOMATO WEBSITE. THE DATASET CONTAINS DETAILS OF 51,717 RESTAURANTS IN BENGALURU.(17 COLUMNS AND 51717 ROWS)

IMPORTANCE

OF THE PROJECT

Increases efficiency and saves the cost of manual analysis

Understanding the customer's overall feedback

Aids with making better business decisions

APPROACH CHOSEN

SENTIMENT ANALYSIS USING LSTM MODEL

SENTIMENT ANALYSIS:

SENTIMENT ANALYSIS IS A PART OF NATURAL LANGUAGE PROCESSING (NLP) WHICH IS A SUBFIELD OF ARTIFICIAL INTELLIGENCE THAT DEALS WITH UNDERSTANDING AND DERIVING INSIGHTS FROM HUMAN LANGUAGES SUCH AS TEXT AND SPEECH.AS THE NAME SUGGESTS, IT MEANS TO IDENTIFY THE VIEW OR EMOTION BEHIND A SITUATION. IT BASICALLY MEANS TO ANALYZE AND FIND THE EMOTION OR INTENT BEHIND A PIECE OF TEXT OR SPEECH OR ANY MODE OF COMMUNICATION.

MODEL CHOSEN

LSTM:

LONG SHORT-TERM MEMORY (LSTM) IS A SPECIAL TYPE OF RECURRENT NEURAL NETWORK (RNN) ARCHITECTURE THAT WAS DESIGNED OVER SIMPLE RNNS FOR MODELING TEMPORAL SEQUENCES AND THEIR LONG-RANGE DEPENDENCIES MORE ACCURATELY.

APPROACH USED

1. PRE-PROCESSING: DATA PRE-PROCESSING, VISUALISATIONS & TEXT PRE-PROCESSING
2. BUILDING A MODEL
3. EVALUATION

IMPLEMENTATION AND EVALUATION

MODEL BUILDING METHODS:

-FIT THE KERAS TOKENIZER

-TOKENIZE AND PAD TRAINING DATA

-KERAS EMBEDDINGS LAYER

-TRAIN THE MODEL

EVAULUATION:

TO EVALUATE THE MODEL, WE USED KERAS'S MODEL.EVALUATE() METHOD AND PASSED THE TESTING DATASET AS PARAMETERS.

EXPERIMENTAL RESULTS

FOR THIS PROJECT, WE CHOOSE ACCURACY AS OUR MAIN EVALUATION CRITERIA.

THE MODEL SHOWS OVERALL ACCURACY AS 0.96 AND THE SCORE AS 0.12.

DATASET CHALLENGES

THAT COME WITH LARGE DATASET AND HOW TO HANDLE THEM

SENTIMENTAL ANALYSIS

HOW TO USE AN LSTM MODEL FOR SENTIMENTAL ANALYSIS

EFFECTS ON BUSINESSES

HOW BUSINESSES CAN BE PROFITTED USING SENTIMENT ANALYSIS

WHAT WE LEARNT & INTERESTING INSIGHTS

INDIVIDUAL CONTRIBUTION

ANWESHA KELKAR

1. Literature Review Report(Phase 1): 3 papers
2. Data Visualisation
3. Building a model
4. Final evaluation and conclusion of project
5. Final Report (Phase 2)

MOHNISH SRIKANTH

1. Literature Review Report (Phase 1): 3 papers
2. Data Visualisation
3. Pre-processing Text
4. Final evaluation and conclusion of project
5. Final Report (Phase 2)

AARUSHI AGARWAL

1. Literature Review Report (Phase 1): 3 papers
2. Data Pre-processing
3. Pre-processing Text
4. Final evaluation and conclusion of project
5. Final Report (Phase 2)

ARJUN HARISH

1. Literature Review Report (Phase 1): 3 papers
2. Data Pre-processing
3. Building a model
4. Final evaluation and conclusion of project
5. Final Report (Phase 2)