Aspect Based Sentiment Analysis

Data Mining S24: Project Proposal

Team: Sifael Sebastian Ndandala, Piumi Abeyrathne and Protogen Hahirwayabo, 12.03.2024

Traditional vs. Aspect Based Sentiment Analysis

Sample Hotel Review

The hotel service was fast and sufficient. We got checked in in under 5 minutes and had all our questions answered. However the food was terrible. Breakfast ran from 7-9 am and there was little variety. The wifi was also not very fast and we couldn't stream our favorite shows on TV. Overall the experience was ok.

Traditional Sentiment Analysis

Overall Sentiment: 😐

Aspect-Based Sentiment Analysis

Check In Service 😍 AMAZING

Food 👎😠 👎😠

Wifi 👎😠

Why this Topic?

Natural Language Processing is a relevant, active and consequential area of research in Academia and Industry

Build on Toolkits from Data Analytics Course in their Application to Textual Data

Opportunity to build and test a real-life NLP Product

Develop foundations for more advanced NLP Techniques i.e. GenAI

Implementation Steps

Aspect-Based Sentiment Classification: Implementation Stages

Data Acquisition	Data Preprocessing	Model Development	Model Evaluation	Deployment

Data Acquisition

Amazon Review Datasets
- Products and Services
- ~29 Categories

Data Preprocessing

3. From textfile to Dataset

Preprocessing and Feature Generation
- Stemming
- Lemmatization
- Stop Words Handling
- Vectorization
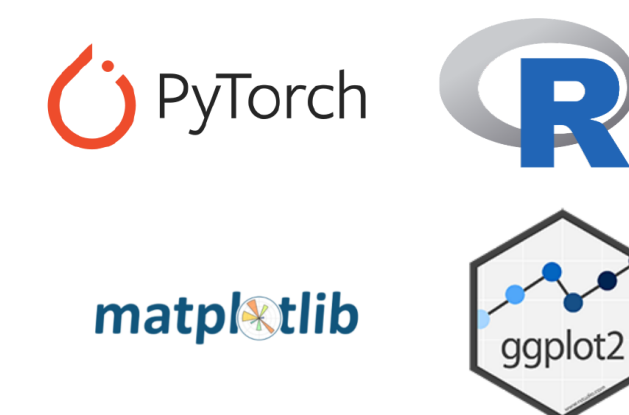- Word Embeddings (CBOW, GLOVE, Word2Vec)

Model Development

Stages of Model Development
- Traditional Sentiment Classification Model
- Topic Classification Modeling
- Aspect-Based Sentiment Classification using BERT

Model Evaluation

Model Evaluation
- Test Set Performance Evaluation and Visualization
- Saving final model file
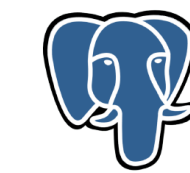
Deployment

Deployment of Model in a simple Web App
- R Shiny using R
- Django using Python
- PostgreSQL for Storage of User queries

Learning Outcome

Deepening Understanding of foundational NLP Concepts and Techniques

Developing Technical Competencies in NLP Technologies and Software

Integration of NLP into Real-World Applications (Django, Rshiny)

Collaboration and Version Control through Git