Rating prediction based on Disneyland reviews

JOËLLE BOSMAN, LARISA BULBAAI & WESSEL POELMAN

Introduction

Large amounts of reviews can provide companies with important information to market their product.

In our research, we will create an algorithm that can predict the rating of reviews on Disneyland.

Our system is useful for analyzing sources that are not yet rated. Making it possible to combine ratings from multiple sources in the same rating system, making further analysis easier.

Research Question

- "Is it possible to reliably assign the labels negative, indifferent and positive to textual reviews?"

Related work

KNN - Kalaivani and Shunmuganathan (2013) and Baid et al. (2017)

SVM - Basari et al (2013); Se et al (2016) and Kalaivani and Shunmuganathan (2013)

- Basari mentions that SVM achieves the best accuracy for text classification.
  - Uses: N-grams and feature weighting
- Kalaivani studied online movie reviews, using SVM, KNN & Naive Bayes
- Baid et al. uses KNN among other methods to identify the polarity of tweets.
- Stein et al. uses fastText for Word Embedding

Data

ID	Rating	Date	Country	Review text	Park
659366150	5	2019-3	United States	Relatively small compared to Disney World. Well maintained, clean, Admission fee the lowest is $61 US dollars.	Disneyland_Hongkong
253619202	1	2015-2	Canada	There were 9 attractions closed while I was there. Thus, all the other attractions were really lined up. Plus the park closed at 8 pm, making less time, less attractions for the same amount of money. The bathrooms were dirty.	Disneyland_California

Data

- Dates range from October 2010 until May 2019
- Some dates are missing

	Number of Reviews	Average rating
Disneyland California	19,406	4.41
Disneyland Paris	13,630	3.96
Disneyland HongKong	9,620	4.20
Total	42,656	4.22

Creating test, train and dev splits

- Train = 80%
- Test = 10%
- Dev = 10%

- Total positive: 33921
- Total indifferent: 5109
- Total negative: 3626

Somewhat skewed, test with more equal splits.

Implementation

Approach:

- Read data from CSV
- Clean review text (lowercase, tokenize, remove stop-words, lemmatize)
- Create word embeddings for review text (doc vector)
- Create train, test and dev splits
- Train
- Evaluate

Experimentation!

Evaluation

We evaluate our models based on:

- Accuracy
- Precision
- Recall
- F-score

	Accuracy	F-Score
Baseline	66 %	0.33
SVM	82 %	0.52
KNN	79 %	0.48
Decision Tree	70 %	0.44

Areas of improvement

- Overrepresentation of positive reviews
- Create more equal split?
- Confusion matrix SVM:

	Negative	Indifferent	Positive
Negative	34	66	423
Indifferent	16	177	195
Positive	20	36	3299

Division of labour

- Joëlle
  - Read in information from dataset
  - Split data into different sets
  - Analyzed new reviews
- Larisa
  - Evaluated the models
  - Interpreted the results
  - Compared SVM and KNN
- Wessel
  - Cleaned the data
  - Made word-embeddings
  - Trained models
  - Analyzed scores per period in the year