Github repository:

CSCE 5290: Natural Language Processing

Project Report

Project title and team members:

Sentiment analysis – Trip advisor reviews

Nghia Dang NghiaDang2@my.unt.edu

Goals and objectives:

Motivation: The fact that hiring specialists to read and summarize reviews from customers is time and finance consuming. Therefore, our team would like to conduct a sentiment analysis on trip advisor reviews to help reduce costs of hospitality businesses that they can better make use of their resources

Significance: As the size of data is fairly large (20491 x 2), we may have to utilize modules that specialize in large-scale data sets, such as Tensorflow or Pytorch. For the same reason, vectorization rather than apply will be used to process the original data.

Objectives:

After cleaning the data set, we will extract important information, especially features that customers of the businesses find the most valuable, among with the importance percentage of each. From this point, a future project after this when we are given enough resort is to design a model that automatically choose the best businesses for the customers after they have filled in a survey of their preferences – recommendation system.

Github repository:

Features:

Some techniques that are useful in this project may include text summarization, stop words, lemmatization, Bag of Words, TF-IDF, Part-Of-Speech Tagging, dimensionality reduction, Spacy, NLTK, and Pytorch library.

Github repository:

Works Cited

Agarwal, Basant, et al. “Sentiment Analysis Using Common-Sense and Context Information.” *Computational Intelligence and Neuroscience*, vol. 2015, 2015, pp. 1–9, www.ncbi.nlm.nih.gov/pmc/articles/PMC4381572/, 10.1155/2015/715730. Accessed 1 Apr. 2019.

Bhadane, Chetashri, et al. “Sentiment Analysis: Measuring Opinions.” *Procedia Computer Science*, vol. 45, 2015, pp. 808–814, www.sciencedirect.com/science/article/pii/S1877050915003956, 10.1016/j.procs.2015.03.159. Accessed 24 Nov. 2019.

Cyril, C Pretty Diana, et al. “An Automated Learning Model for Sentiment Analysis and Data Classification of Twitter Data Using Balanced CA-SVM.” *Concurrent Engineering*, 20 July 2021, p. 1063293X2110314, 10.1177/1063293x211031485. Accessed 26 Oct. 2021.

Devika, M.D., et al. “Sentiment Analysis: A Comparative Study on Different Approaches.” *Procedia Computer Science*, vol. 87, 2016, pp. 44–49, 10.1016/j.procs.2016.05.124. Accessed 14 June 2019.

Le, Bac, and Huy Nguyen. “Twitter Sentiment Analysis Using Machine Learning Techniques.” *Advanced Computational Methods for Knowledge Engineering*, 2015, pp. 279–289, 10.1007/978-3-319-17996-4_25. Accessed 15 Oct. 2020.

“Optimization of Sentiment Analysis Using Machine Learning Classifiers | Human-Centric Computing and Information Sciences.” *Rdcu.be*, rdcu.be/cAaGY. Accessed 26 Oct. 2021.

Puschmann, Cornelius, and Alison Powell. “Turning Words into Consumer Preferences: How Sentiment Analysis Is Framed in Research and the News Media.” *Social Media + Society*, vol. 4, no. 3, July 2018, p. 205630511879772, 10.1177/2056305118797724.