Convolutional Neural Network Based Aspect-Based Sentiment Analysis of Apps Reviews for Requirements Elicitation

A.D.S.R.Gunathilaka

Presentation Structure

- Introduction
- Methodology
- Experiments & Results

- Introduction
- Methodology
- Experiments & Results

Introduction

- What is Requirement Elicitation.
- Importance of user feedback in the context of mobile app development.
- Why ABSA?

[30] J. Verma and A. Patel, “Evaluation of unsupervised learning based extractive text summarization technique for large scale review and feedback data,” Indian Journal of Science and Technology, vol. 10, pp. 1–6, 05 2017.

Introduction : What is Requirement Elicitation

- Requirement Elicitation is the practice of understanding and capturing the business domain knowledge, stakeholder goals, and user needs.
- It is a critical activity in the Requirement Engineering (RE) process, and it plays a significant role in the overall quality of the RE outcome [1].
- Crowd-generated content (e.g. apps reviews) is an essential source of knowledge that can be utilized to create a customer-centric experience.
- Utilizing apps reviews instead of traditional approaches (e.g. surveys or interviews) brings big benefits and enhancement to the requirement elicitation activity.

[1] D. Zowghi and C. Coulin, “Requirements elicitation: A survey of techniques, approaches, and tools,” in Engineering and managing software requirements. Springer, 2005, pp. 19–46.

Introduction:Importance of user feedback in the context of mobile app development

- User involvement is a major contributor to success of software projects [1].
- Feedback typically contains multiple topics related to the application such as user experience issues, bug reports, and feature requests[2][3].
- Most of the feedback given by the users after a new release and the frequency of feedback submitted decreases over the time[3].
- Feedback content has an impact on download numbers of the application.
- According to a study by W. Maalej [3] majority of low star rating feedback usually contains shortcomings and bug reports of the application where four to five star ratings mainly consist of praise.It was noted that the feature requests are mostly coming from three to five star rating feedback.
- User comments can be used to improve user satisfaction of software products[4].

[1] M. Bano and D. Zowghi, “A systematic review on the relationship between user involvement and system success,” Information and Software Technology, vol. 58, 06 2014.
[2] D. Pagano and B. Bruegge, “User involvement in software evolution practice: A case study,” 05 2013.
[3] D. Pagano and W. Maalej, “User feedback in the appstore: An empirical study,” 07 2013.
[4] H. Li, L. Zhang, L. Zhang, and J. Shen, “A user satisfaction analysis approach for software evolution,” 2010 IEEE International Conference on Progress in Informatics and Computing, vol. 2, pp. 1093–1097, 2010.

Introduction : Why ABSA

- Consider following example: “UI is awesome and easy to use but applications drains the battery faster”. Here “ui” is an aspect with a positive sentiment and “battery” is another aspect with a negative sentiment.
- Having the aspect information along with their respective sentiment leads to a fine-grained analysis [2].
- To support such analysis, we can utilize Aspect-Based Sentiment Analysis (ABSA) [3], which identifies the sentiment with respect to a specific aspect.
- ABSA consists of three sub-tasks:

  (i)aspect category classification

  (ii) aspect term extraction

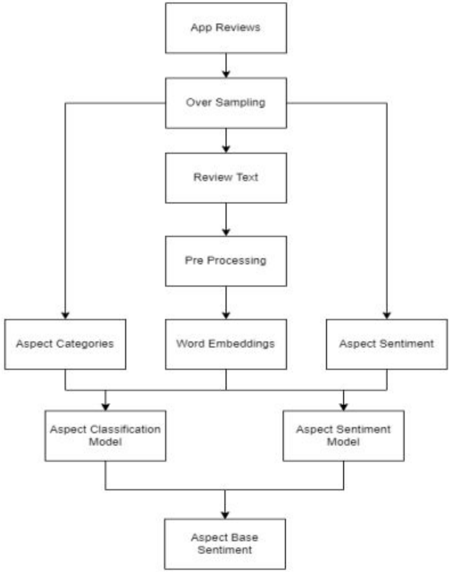
  (iii) aspect sentiment analysis.

[2] Y. Li, B. Jia, Y. Guo, and X. Chen, “Mining user reviews for mobile app comparisons,” Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies, vol. 1, no. 3, pp. 1–15, 017.
[3] M. Hu and B. Liu, “Mining and summarizing customer reviews,” in Proceedings of the tenth ACM SIGKDD international conference on Knowledge discovery and data mining, 2004, pp. 168–177.

Methodology

- Proposed Approach
- Dataset
- Over Sampling the Data
- Preprocessing
- Embeddings
- Feature extraction and classification

Methodology : Proposed Approach Overview



Methodology : Dataset[4]

- AWARE is benchmark dataset of 11323 apps reviews that are annotated with aspect terms, categories, and sentiment.
- It contains reviews that were collected from three domains: productivity, social networking, and games.
- The aspect categories for each domain were derived using content analysis and they validated them with domain experts in terms of importance, comprehensiveness, overlapping, and granularity level.
- The data set contains two aspect definitions
  - **Aspect Term**: A term describing an aspect of an app that is expressed by the sentiment and that exists in the sentence.
  - **Aspect Category**: A predefined set of domain-specific categories.

[4] N. Alturaief, H. Aljamaan and M. Baslyman, "AWARE: Aspect-Based Sentiment Analysis Dataset of Apps Reviews for Requirements Elicitation," 2021 36th IEEE/ACM International Conference on Automated Software Engineering Workshops (ASEW), 2021, pp. 211-218, doi: 10.1109/ASEW52652.2021.00049.

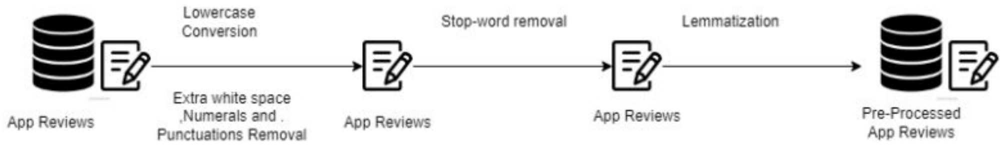
Methodology : Over Sampling the Data

- Contextual augmentation by Google Bert [5].
  - Contextual words embeddings assigns each words a representation based on its context. We used substitute actions for augmenting data. In substitute, length of sentence is same but some words are replaced. We utilized the NLPAug [7] open source python package for data augmentation.

- Data Augmentation by RTT.
  - Round-trip translation (RTT) is additionally referred to as re-cursive, back-andforth, and bi-directional translation. it's the method of translating a word, phrase or text into another language (forward translation), then translating the results back to the first language (back translation) .RTT is used as augmentation technique to extend the training data. We used Roundtrip translation python package[6] to augment data.

[5] Kobayashi, Sosuke. (2018). Contextual Augmentation: Data Augmentation by Words with Paradigmatic Relations. 452-457. 10.18653/v1/N18-2072.
[6] https://github.com/samhavens/roundtrip.
[7] https://github.com/makcedward/nlpaug

Methodology : Preprocessing

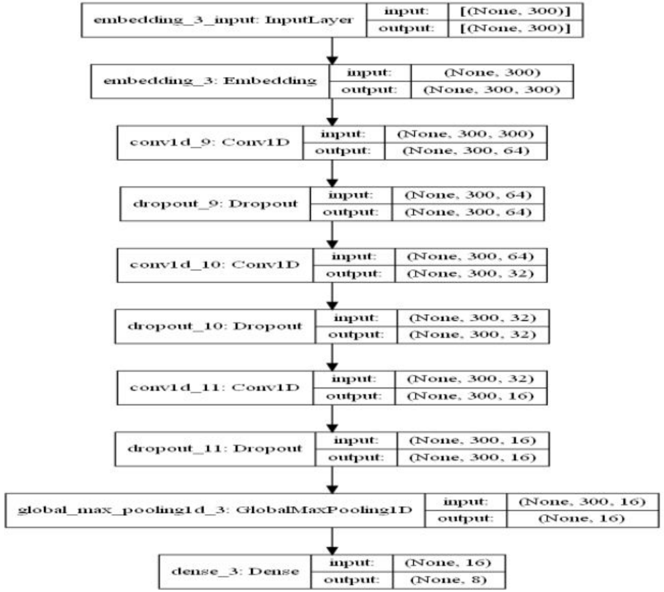


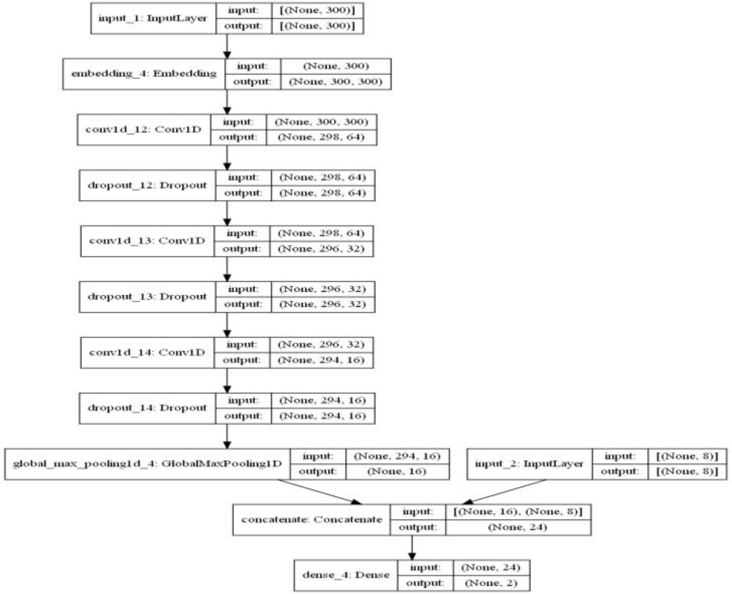
Methodology : Embeddings

Pre-trained Models

- FastText
- Glove
- Word2Vec

Methodology : Feature extraction and classification





- Introduction
- Methodology
- Experiments & Results

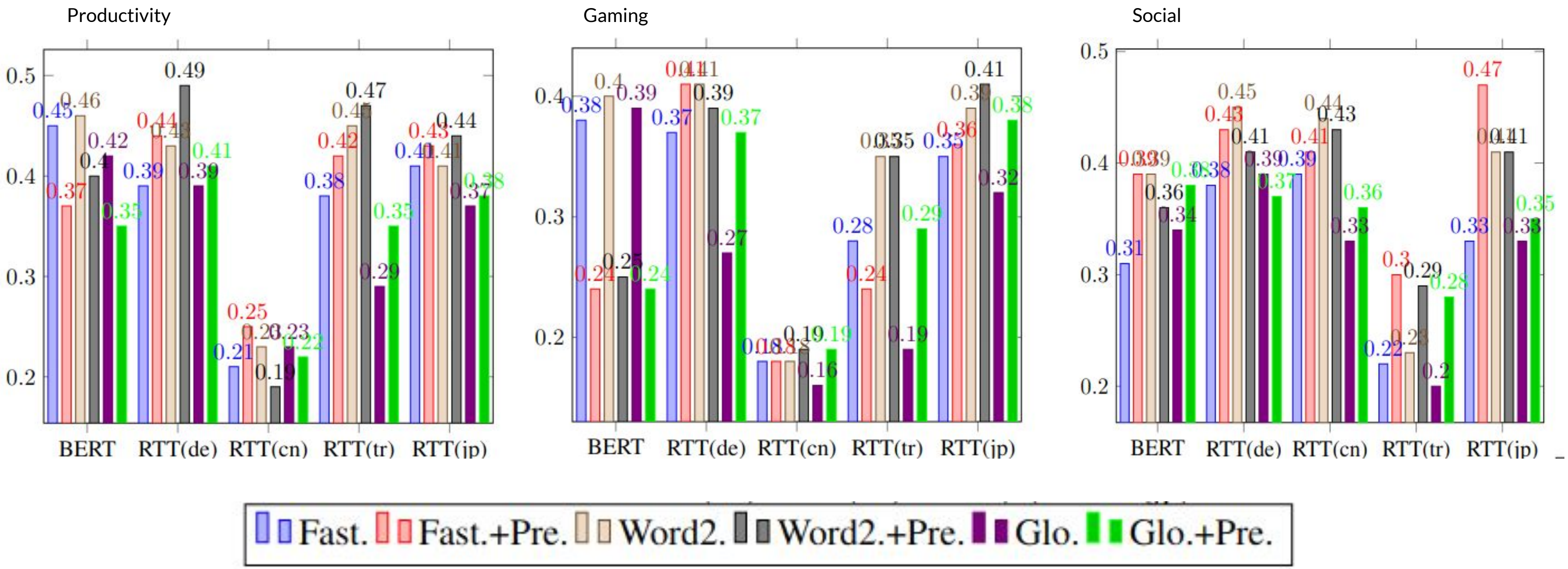
Experiments & Results

Baseline

TABLE IV
BASELINE MODELS RESULTS.

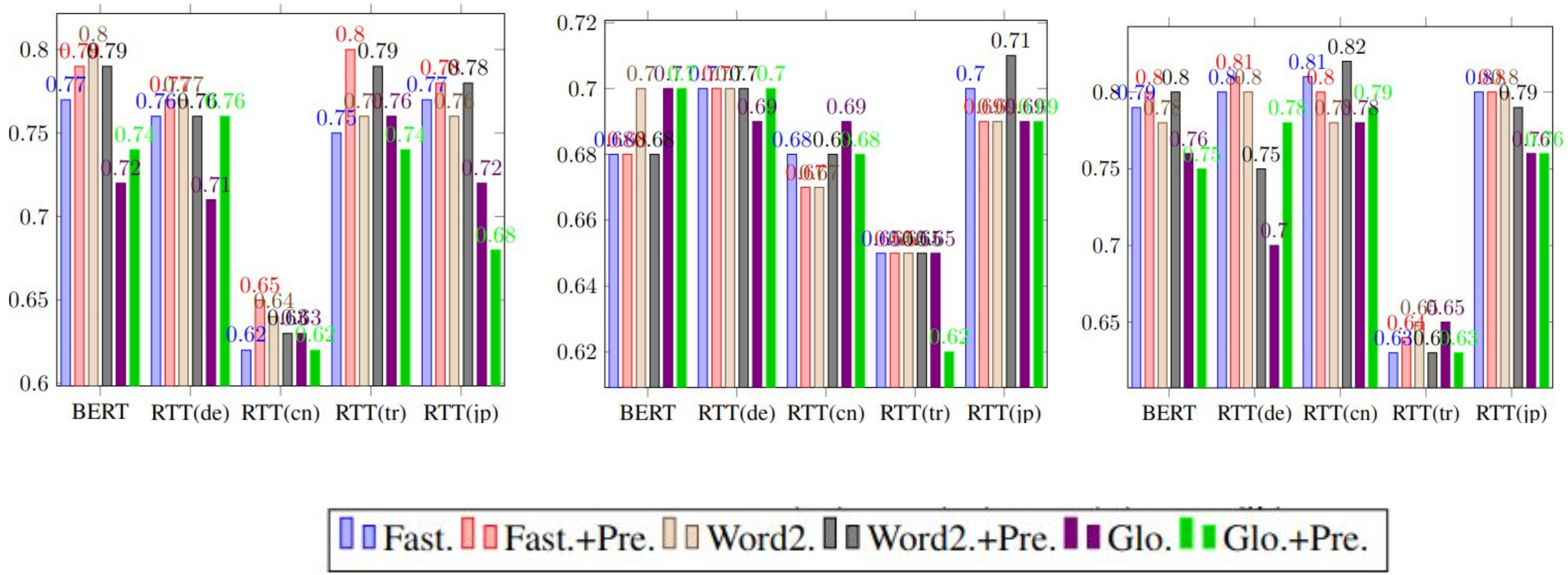
<table>
<tr><th colspan="2">Task</th><th>Model</th><th>Metric</th><th>Result</th></tr>
<tr><td rowspan="6">Aspect Category Classification</td><td rowspan="2">Productivity</td><td>SVM</td><td rowspan="6">F1</td><td>0.33</td></tr>
<tr><td>MLP</td><td>0.32</td></tr>
<tr><td rowspan="2">Social Networking</td><td>SVM</td><td>0.32</td></tr>
<tr><td>MLP</td><td>0.31</td></tr>
<tr><td rowspan="2">Games</td><td>SVM</td><td>0.32</td></tr>
<tr><td>MLP</td><td>0.29</td></tr>
<tr><td rowspan="6">Aspect Sentiment Classification</td><td rowspan="2">Productivity</td><td>SVM</td><td rowspan="6">Acc.</td><td>68.71%</td></tr>
<tr><td>MLP</td><td>66.11%</td></tr>
<tr><td rowspan="2">Social Networking</td><td>SVM</td><td>69.72%</td></tr>
<tr><td>MLP</td><td>67.32%</td></tr>
<tr><td rowspan="2">Games</td><td>SVM</td><td>67.49%</td></tr>
<tr><td>MLP</td><td>64.79%</td></tr>
<tr><td colspan="2">Aspect Term Extraction</td><td>POS</td><td>F1</td><td>0.82</td></tr>
</table>

Aspect Class Classification



Aspect sentiment Classification

Productivity Gaming Social



Thank You

Q & A