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

Presented by: Sadeep Gunathilaka

Overview

- Introduction
- Dataset
- Approach
- Experiments

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 : Why ABSA

- Identifying cases such as: is there a privacy concern? Are users satisfied with the new update? Which apps features matter most to users? Can help to gain a deeper understanding of specific apps features.
- To support such analysis, we can utilize Aspect-Based Sentiment Analysis (ABSA) [2], which identifies the sentiment with respect to a specific aspect.
- Having the aspect information leads to a fine-grained analysis that adds more accurate understanding of opinions [3].
- ABSA consists of three sub-tasks:

  (i)aspect category classification

  (ii) aspect term extraction

  (iii) aspect sentiment analysis.

[2] 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.
[3] 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.

Dataset : AWERE 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.

Dataset : AWERE Dataset

TABLE III
DEFINITIONS OF THE DERIVED ASPECT CATEGORIES.

Category	Definition
Aesthetics	Beauty of the appearance including graphics, audio, and video.
Cost	Price or value for money of the app or a feature
Compatibility	How the app can be launched on another environment or co-exist with another system.
Effectiveness	Usefulness and practicality of the app or a feature, including how the app offers its users the ability to achieve various objectives and reach the final goal.
Efficiency	Consumption of resources such speed, memory consumption, and battery consumption.
Enjoyability	Degree of pleasure derived from using the app or a feature such as fun, disappointment and engagement.
General	General positive or negative sentiment that does not fit any of the pre-defined categories.
Learnability	Ability to understand and master the app, such as difficulty in overcoming a particular challenge or understanding certain concepts.
Reliability	Crashing issues, failures, data loss, and connection errors.
Safety	Safety of users such as concerns about harassment, stalking and bullying.
Security	Personal information protection, such as encryption and authentication mechanisms, and how it is used and controlled.
Usability	Ease of use or convenience in using the app, including properties such as accessibility, look & feel, placement or existence of components, etc.

Dataset : AWERE Dataset

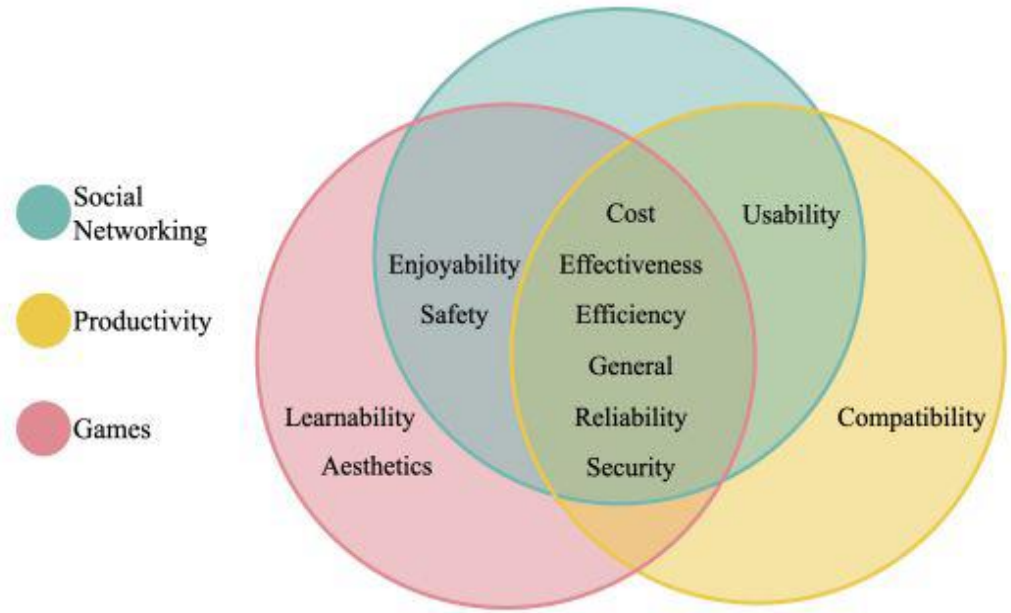
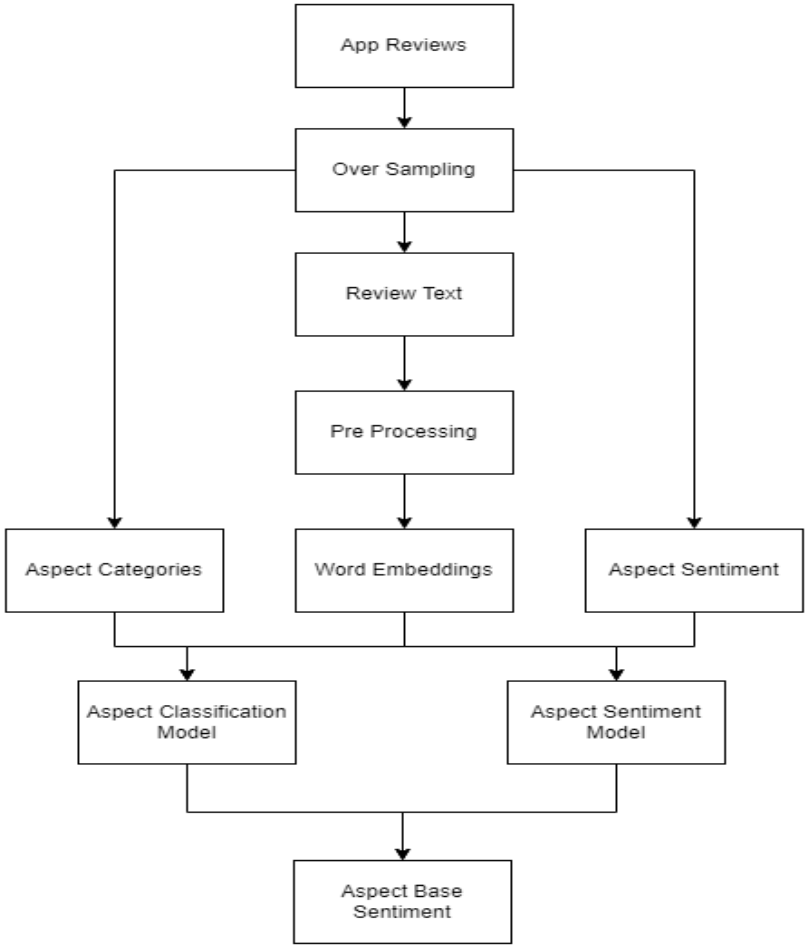


Fig. 2. Logical Relationships of the Aspect Categories.

Approach : Overview

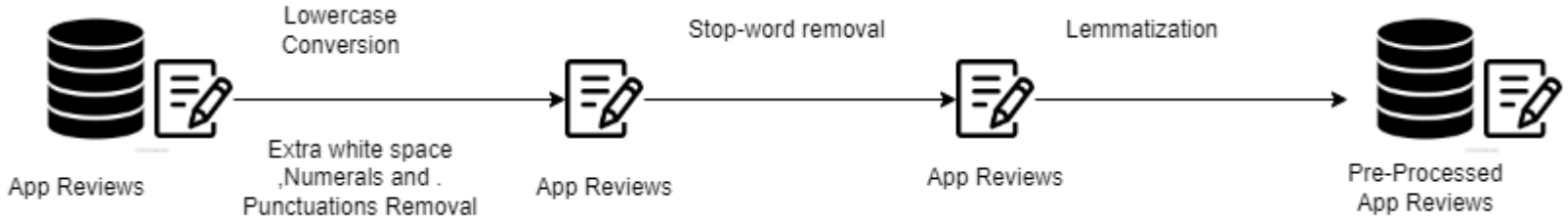


Approach : Over Sampling

- 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, lenght 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-and-forth, 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

Approach : Preprocessing

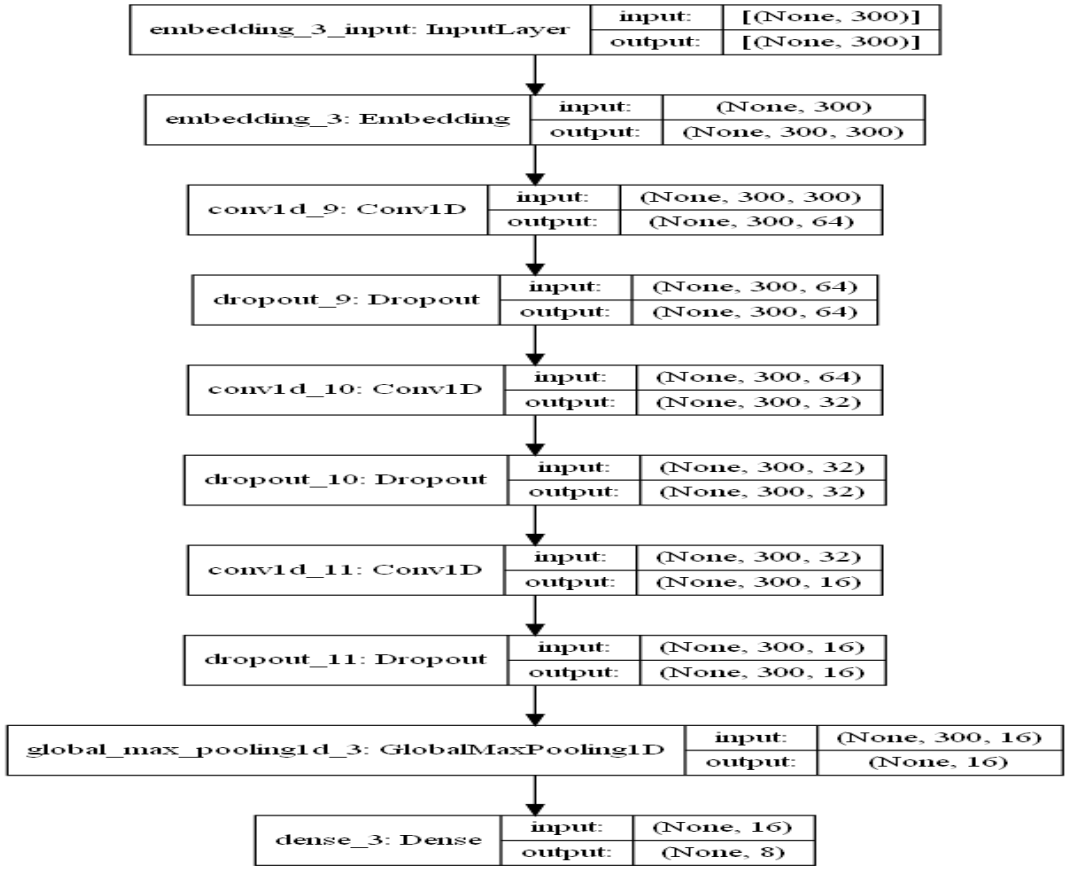


Approach : Embeddings

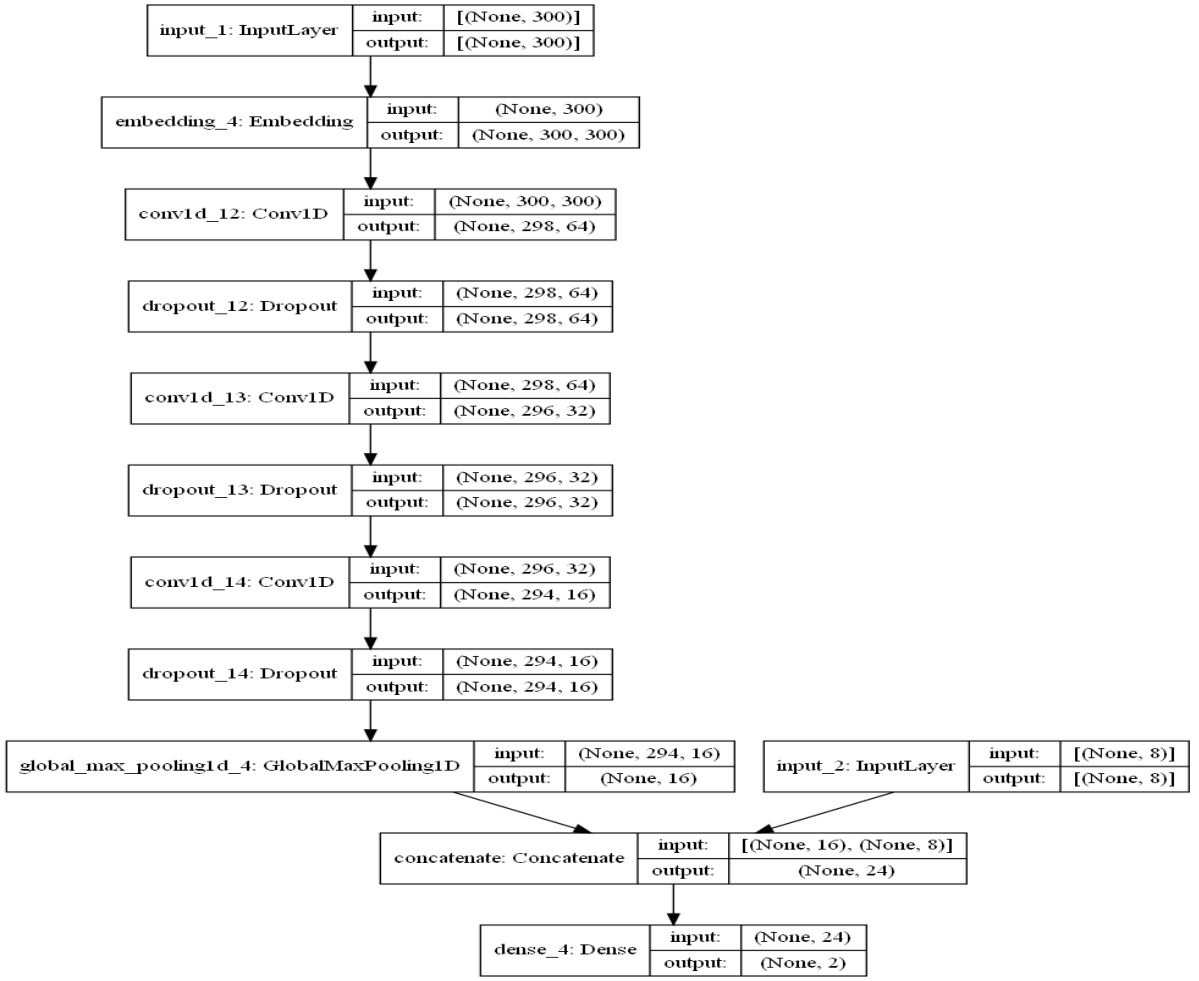
Pre-trained Models

- FastText
- Glove
- Word2Vec

Approach : Aspect Category Classification Model



Approach : Aspect Sentiment Classification Model

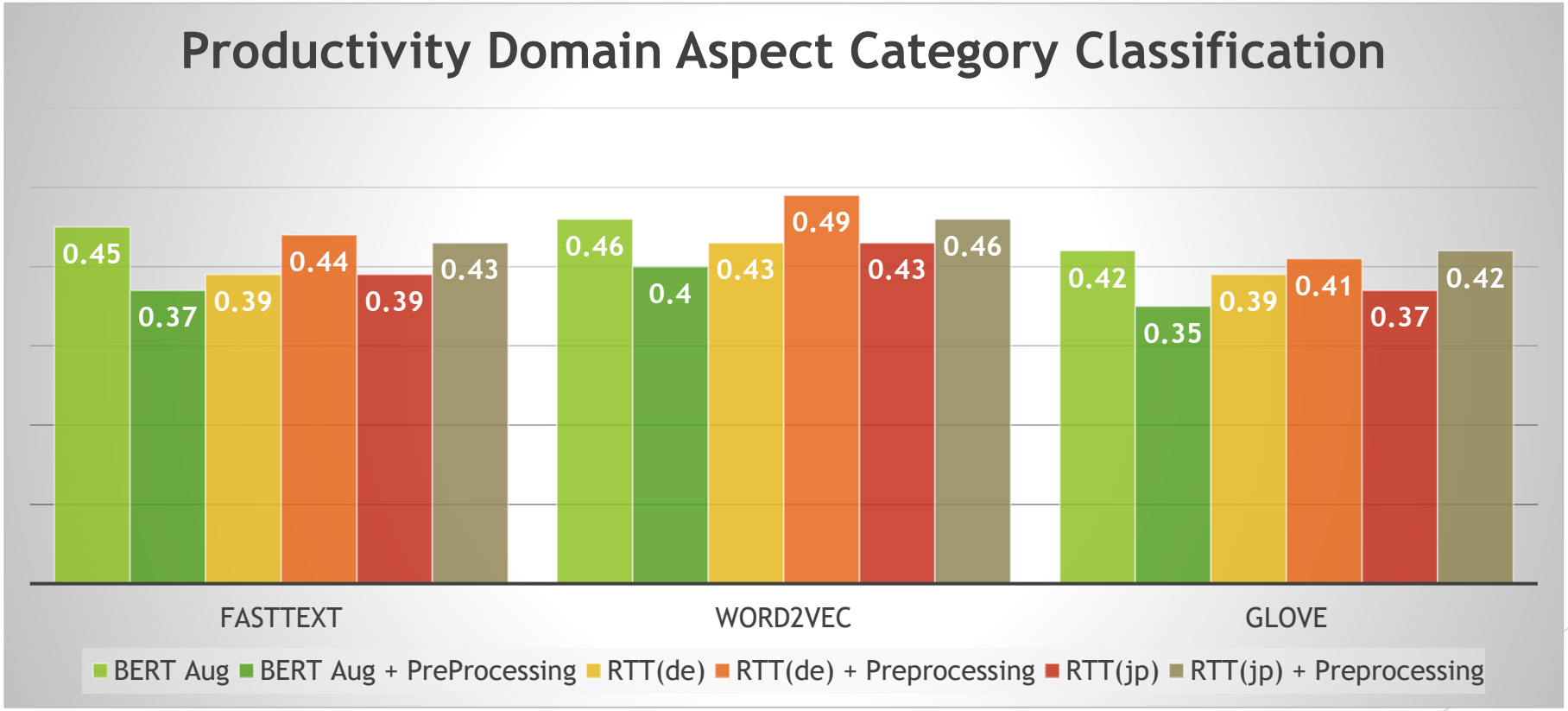


Experiments : Baseline

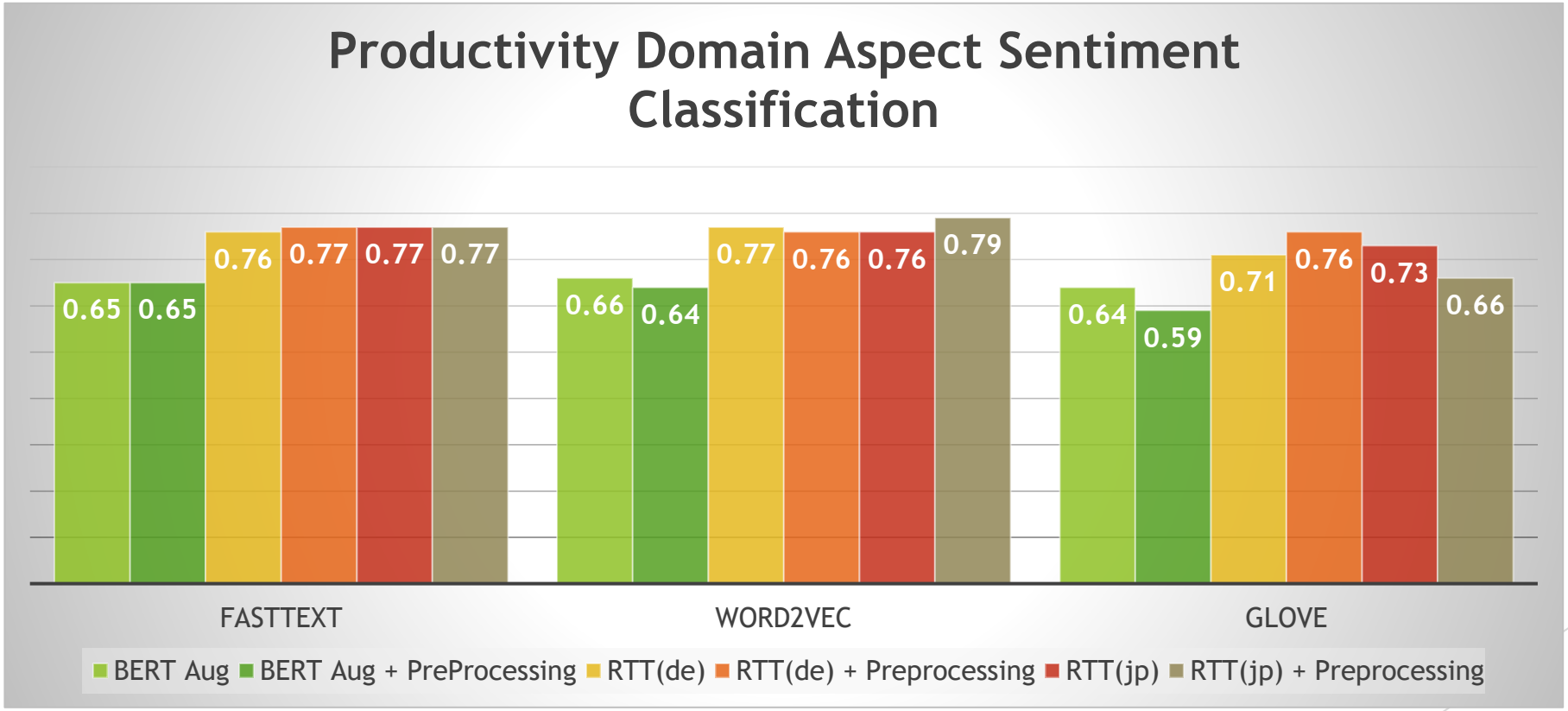
TABLE IV
BASELINE MODELS RESULTS.

Task		Model	Metric	Result
Aspect Category Classification	Productivity	SVM	F1	0.33
		MLP		0.32
	Social Networking	SVM		0.32
		MLP		0.31
	Games	SVM		0.32
		MLP		0.29
Aspect Sentiment Classification	Productivity	SVM	Acc.	68.71%
		MLP		66.11%
	Social Networking	SVM		69.72%
		MLP		67.32%
	Games	SVM		67.49%
		MLP		64.79%
Aspect Term Extraction		POS	F1	0.82

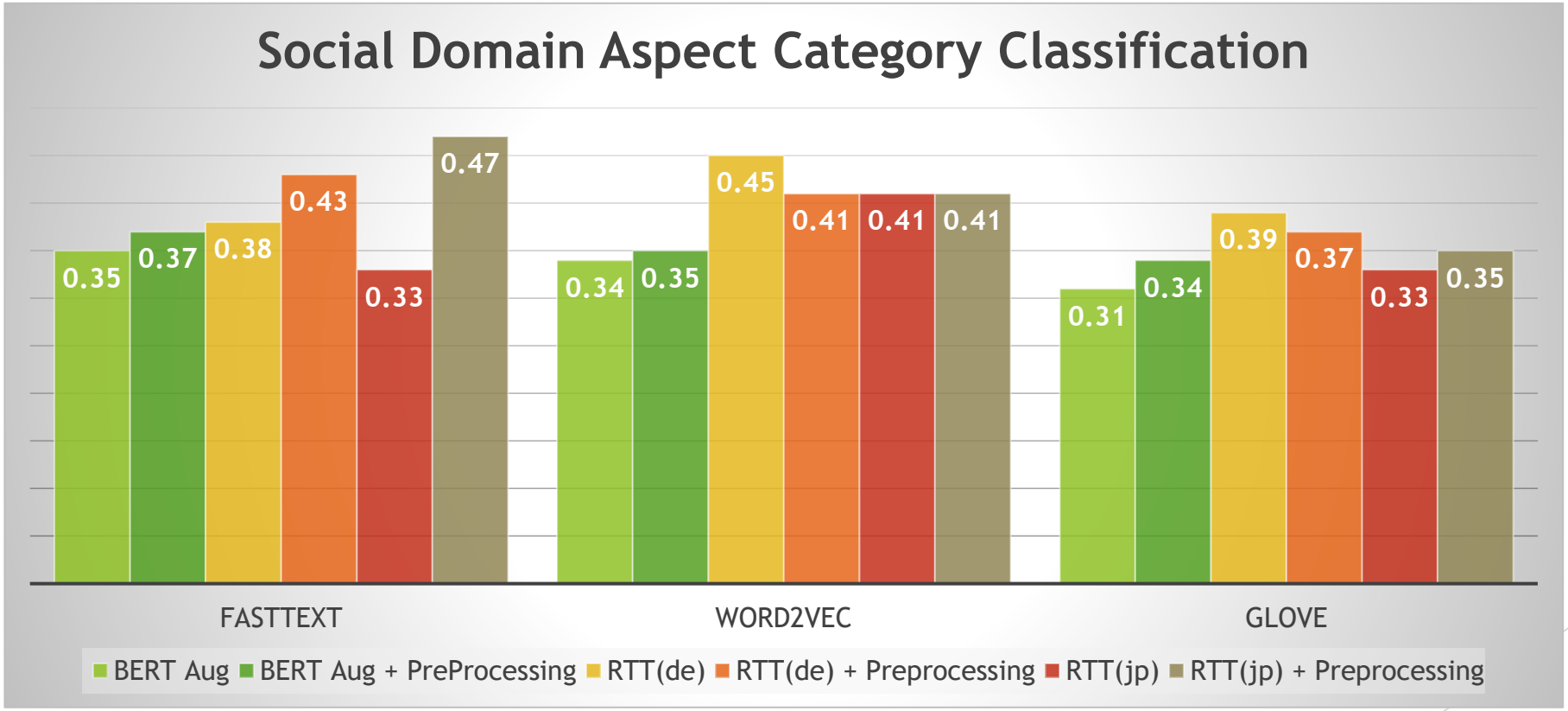
Experiments : Results



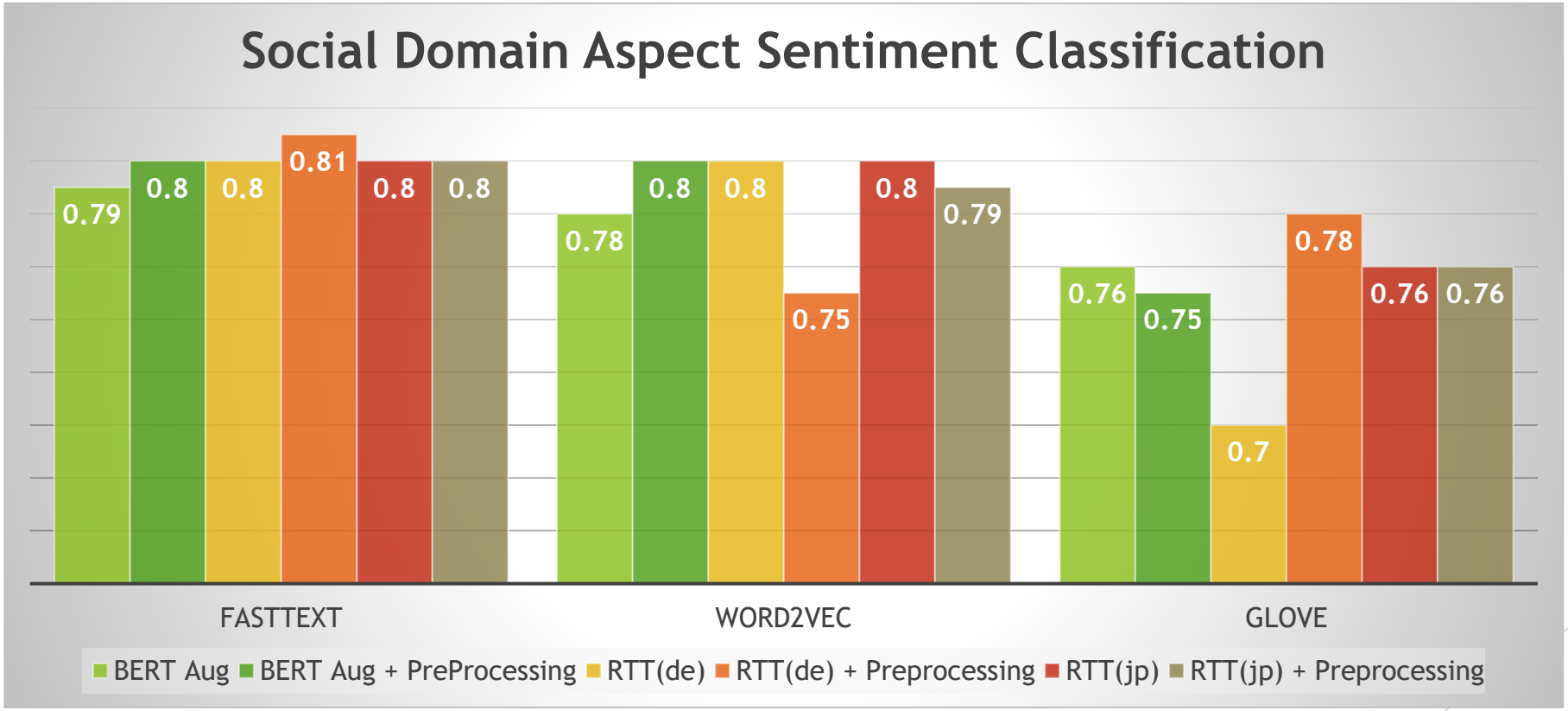
Experiments : Results



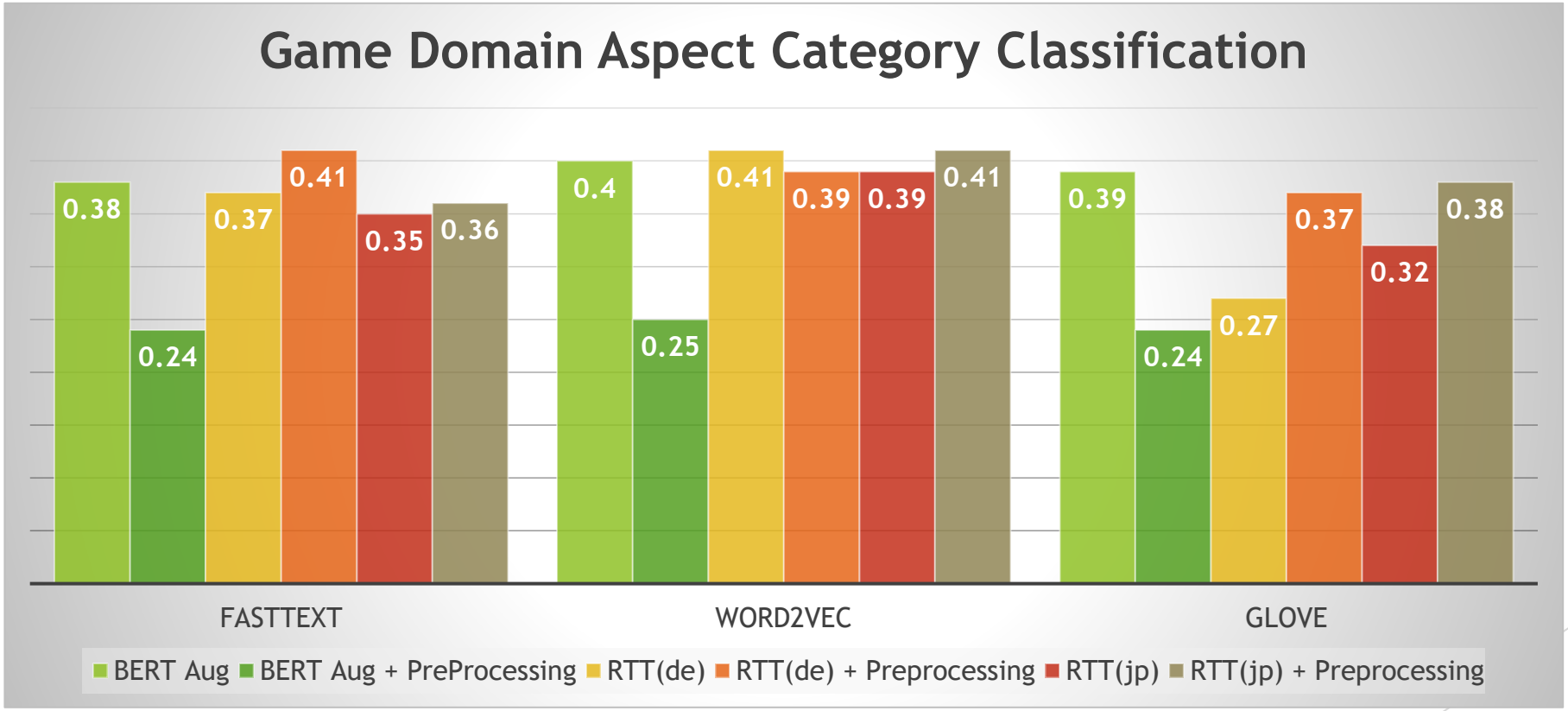
Experiments : Results



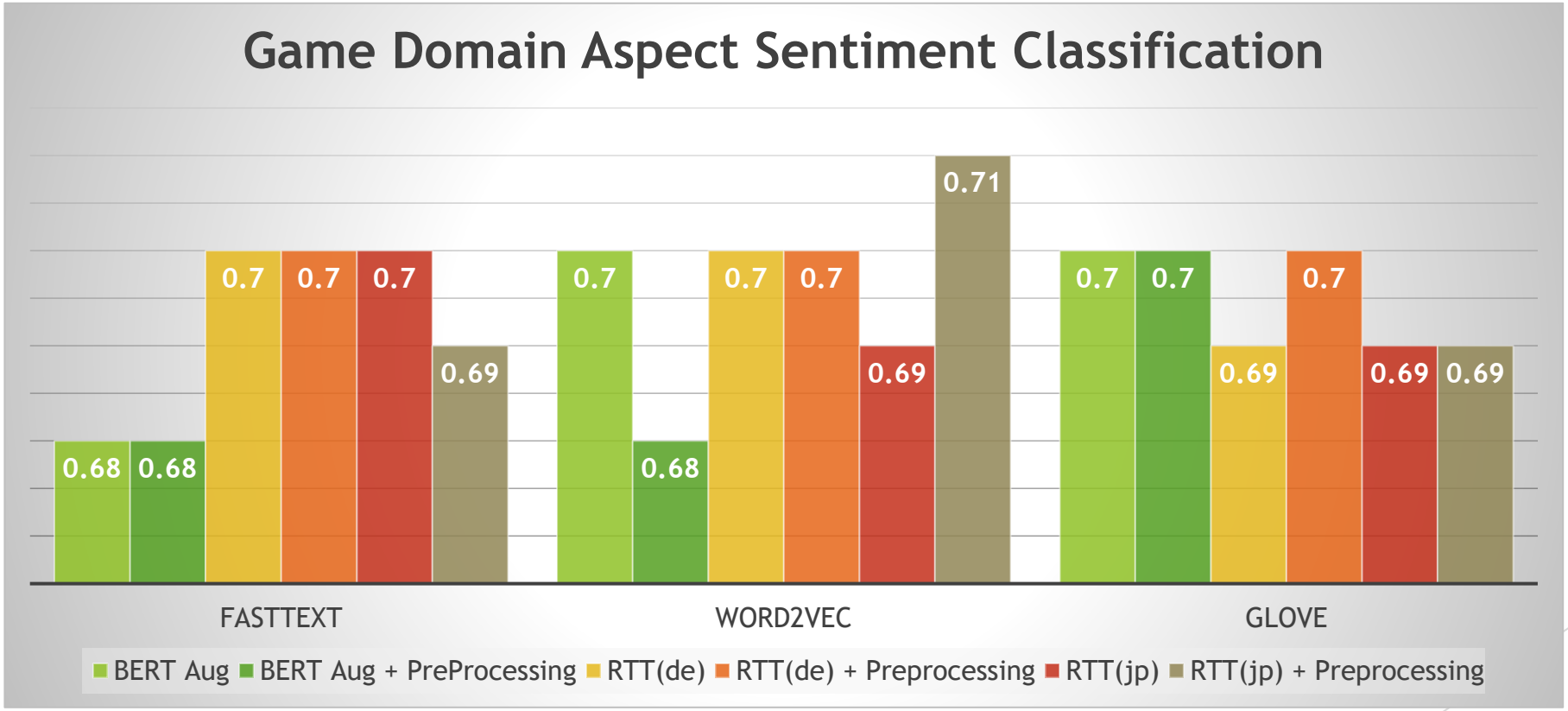
Experiments : Results



Experiments : Results



Experiments : Results



Thank You

Q n A