Effects of Named Entities on Amazon Sentiment Analysis

Anne Turmel
University of Southern Maine
anne.turmel@maine.edi

Owen Bean
University of Southern Maine
owen.bean@maine.edu

Abstract

Sentiment analysis is an important task for understanding how a group feels about a product, brand, or service. Currently, there are a lot of competitive models out there to handle Amazon sentiment analysis. Our paper experimented with masking name entities from the Amazon product review to see if there is an improvement to models compared to keeping the Amazon review text from too much noise in name entities. Results show name entities have small sentiment value, but does not improve the sentiment analysis of a model. One small outlier did reduce noise for Amazon product review sentiment analysis. Code used: https://github.com/owenbean400/AmazonSentimentAnalysis

1 Introduction

Sentiment analysis looks into extracting subjective information from a source. Some words have more weight in sentiment analysis compared to other words. With reviews from Amazon, name entities are mentioned in the review text.

There have been plenty of models for predicting sentiment analysis of Amazon reviews. The top 3 models; accuracies unsupervised data augmentation for consistency training [8], deep pyramid convolutional neural networks for text categorization [2], and disconnected recurrent neural networks for text categorization [7]; all have been tested with Amazon product reviews from users. These models' accuracies are between 60%-70% percent. These models were trained on the raw Amazon product review full.

A name entity is a phrase that contains the names of persons, organizations, locations, times, and quantities [1]. Does simplifying the name entities into one mask word effect Amazon sentiment analysis in positive or negative accuracies? The possibility of removing the context of name entity effects the training and tests of sentiment analysis could reduce noise for better model performance. Moreover, this will determine if name entities hold any sentiment value.

2 Related Work

Polarity on Sentiment Analsys Specific words have neutral, positive, or negative polarity when it comes to sentiment analysis. Recognizing Contextual Polarity in Phrase-Level Sentiment Analysis research determined a new stragety of figuring out polarity of words in senitment analysis. The research looked into how specific parts of a speech effect the polarity of a word. More specifically, nouns were classified as neutral polar. Name entities are nouns, so name entities is considered a neutral polarity. Therefore, name entities can be proposed as not having sentiment value for sentiment analysis.

Part of Speech Subjectivity Research into subjectivity compared with the objectivity of a word looked into parts of speech effects of sentiment analysis from Bing Liu [4]. It has been noted from Bing Liu sentiment analysis on subjectivity that adjectives, adverbs, nouns, and verbs can obtain sentiment detail. Furthermore, an object in object extraction from sentiment is a noun, proper noun, and sometimes a verb. The paper didn't look into named entities and give the impacts of training within the context of the named entities.

Name Entity Polysemy Ming Zhoa-Yan and Chua Tat Seng experimented with polysemy in entity linking for context modeling [5]. Polysemy occurs when a word or phrase has multiple different meanings. Some name entities are polysemy. For example, the article mentions that ABC may stand for American Broadcasting Company, Alphabet Network, or ABC. This shows there are multiple entities linking between names. The research did not look into the impact of multiple entity linkings on sentiment context for sentiment analysis.

3 Task

The task is to see if named entities contain information needed for sentiment analysis in Amazon product review models. One model is trained on unchanged Amazon product review, while the other model is trained on the same Amazon product reviews where named entities are replaced with a mask. The goal is to compare the two models to see if removing named entities has any impact on sentiment analysis.

4 Method

Amazon product review data comes with a rating and review that can be used for training. The product must parse out named entities with a replacement mask to represent them. This process removes the context of the named entity. An example of removing named entities' context can be shown.

```
"It's just like having $100. Except I
couldn't get Walmart to accept it.
Apparently you can only use it on the
interweb at Amazon."

Detect Amazon and Walmart as named entity,
so replaced with <ENT>

"It's just like having $100. Except I
couldn't get <ENT> to accept it.
Apparently you can only use it on
the interweb at <ENT>."
```

4.1 LUKE

LUKE is an entity extraction model that can be used for extracting entities in an Amazon product review. LUKE-500K pre-trained data had an F1 score of 94.3 on named entity recognition from the CoNLL-2003 dataset [9]. This gives the capability for a machine to detect the named entity in a passage. Once the entity is found in the passage, LUKE replaces it with one mask, so the context of the named entity is not known. To conclude, LUKE will parse out the named entity before sending the review text to the model.

4.2 albert-base-v2

Albert is a pre-trained BERT model on BookCorpus [3]. The pre-trained model is used to understand the English language. The purpose of Albert is to have a pre-trained model ready for fine-tuning. The reason for Albert is the fine-tuned task of making a decision based on whole sentences. The fine-tuning process will learn from Amazon product review text to classify the rating of the review. The rating of the review will show the sentiment analysis of the fine-tuned model.

5 Experiments

The Amazon product review data came from the University of California, San Diego Amazon product review dataset [6]. The data consists of the rating and the review text that came with the rating for training and testing. A rating from 1-5 will be used for the sentiment feeling of the Amazon review. The experiment trained and tested on the following 5 categories: toys and games, office products, appliances, all beauty, and video games.

5.1 Masking Named Entities

For each category of Amazon product review, there are two copies of each product review. One copy has the original Amazon product review text. The second copy has the same text, but named entities are replaced with a mask (ENT). There are at least 10,000 Amazon product reviews for each category that the models were trained on. This leaves over 1,000 review texts being tested for evaluation of the model.

5.2 Training the Models

The albert-base-v2 is used for fine-tuning on detecting sentiment analysis. Each dataset is split as follows: 80% training, 10% validation, and 10% testing. The models are tested on accuracy between the untouched reviews and the review that has the named entity replaced with a mask. The only change with the models is the named entity masking. The results show how the model compares with or without the context of the proper noun.

When training with Albert pre-trained model, the PyTorch seed is set to 1324224321 to eliminate randomness when training the model. Moreover, the learning rate, batch size, and number of epochs stay the same. Keeping these values the same will ensure the only change between the two models is whether the review text is masked or not. These factors make the results of training and testing reproducible for others.

5.3 Results

As shown below, the left column is the category the Amazon product review data was in. The middle column is the original unaltered product review text accuracy for testing evaluation after training. The right column is the alternative product review with name entity masked review text.

In the table, 1e Stands for one epoch and 5e is five epochs used during training. Ori. is the Amazon product review text that was not touched. Accur. is the accuracy results of 10% of the Amazon product review data. The batch sizes were both set to 16. Lastly, the learning rate was 2e-5. All of these values were set the same for all training, validation, and testing.

albert-base-v2 fine-tuned accuracy

Category	Ori. Accur.	LUKE Accur.
Video Games (1e)	0.789	0.779
Video Games (5e)	0.784	0.775
Appliances (1e)	0.740	0.712
Appliances (5e)	0.712	0.714
Toys + Games (1e)	0.853	0.850
Toys + Games (5e)	0.841	0.846
All Beauty (1e)	0.809	0.801
All Beauty (5e)	0.781	0.801
Office (1e)	0.820	0.809
Office (5e)	0.807	0.811

For epochs level 5, there is no significant difference between training, validating, and testing on the review text compared to masked name entities from LUKE, showing that masking name entity has no effects on its data. For 1 epoch, the original review text data accuracy was slightly better than masked name entity text data. The fine-tuning did not have enough to feed the data itself back through the model, so it did not understand the masking.

We experimented on the saved different Amazon categories models from previous training and applied them to test 10,000 records from other Amazon categories. Both the original review text model and LUKE masked review text model were applied to the original review text and LUKE masked review text. The model in the table is from the video games category of Amazon product review with epochs of 5.

Video Games Original Model Testing

Category	Ori. Accur.	LUKE Accur.
Appliances	0.713	0.712
Toys and Games	0.832	0.837
All Beauty	0.792	0.791
Office	0.797	0.797

Video Games LUKE Model Testing

Category	Ori. Accur.	LUKE Accur.
Appliances	0.921	0.925
Toys and Games	0.806	0.807
All Beauty	0.768	0.768
Office	0.777	0.777

Videos Games trained on the original data did slightly better than video games trained on the LUKE mask model except for Appliances. Doing sentiment analysis from video games LUKE masked model with appliances Amazon product reviews was significantly better than video games original data model and any appliances models. However, every other video games model with other categories besides appliances did slightly worse or significantly worse than the category model tested on its 10% unseen data. This shows that the context of masking name entities on video games was able to detect more accurately on sentiment classification for appliance reviews.

6 Conclusion

From this experiment, we conclude that masking named entities in review text does not improve sentiment analysis for Amazon product reviews. There is only one outlier where the LUKE masking review model was significantly better for another category compared to the original review text data. There is a small percentage of name entities that do hold sentiment context for sentiment analysis, where the original Amazon product review models are slightly more effective than name entities masked models. Because name entity masking did not have more effective results on Amazon sentiment analysis compared to not masking, name entities are not necessary as it is inefficient to do an extra step of masking data.

References

[1] Language-independent named entity recognition. 2003.

[2] Rie Johnson and Tong Zhang. Deep pyramid convolutional neural networks for text categorization. 2017.

[3] Zhenzhong Lan, Mingda Chen, Sebastian Goodman, Kevin Gimpel, Piyush Sharma, and Radu Soricut. ALBERT: A lite BERT for self-supervised learning of language representations. *CoRR*, abs/1909.11942, 2019.

[4] Bing Liu. Sentiment analysis and subjectivity.

[5] Zhao-Yan Ming and Tat Seng Chua. Resolving polysemy and pseudonymity in entity linking with comprehensive name and context modeling. *Information Sciences*, 307:18–38, 2015.

[6] Jianmo Ni, Jiacheng Li, and Julian McAuley. Justifying recommendations using distantly-labeled reviews and fine-grained aspects. 2019.

[7] Baoxin Wang. Disconnected recurrent neural networks for text categorization. 2018.

[8] Qizhe Xie, Zihang Dai, Eduard Hovy, Minh-Thang Luong, and Quoc V. Le. Unsupervised data augmentation for consistency training. 2020.

[9] Ikuya Yamada, Akari Asai, Hiroyuki Shindo, Hideaki Takeda, and Yuji Matsumoto. LUKE: Deep contextualized entity representations with entity-aware self-attention. In *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP)*. Association for Computational Linguistics, 2020.