



# Sentiment Analysis on Amazon Reviews of Mobile Electronics

Samruddhi Dhumane, Aashna Shroff, Dharita Desai, Varshin Bhaskaran

Under Prof. Lu Xiao

## OVERVIEW

In the realm of e-commerce, sentiment analysis plays a crucial role in understanding customer feedback and gauging product satisfaction. By analyzing product reviews, businesses can identify areas for improvement, enhance customer satisfaction, and ultimately boost sales. This project delves into the realm of sentiment analysis, employing both spaCy's traditional NLP methods and advanced BERT transformers and compare them. It focuses on analyzing Amazon customer reviews for mobile electronics, incorporating Exploratory Data Analysis (EDA) for deeper insights into the dataset.

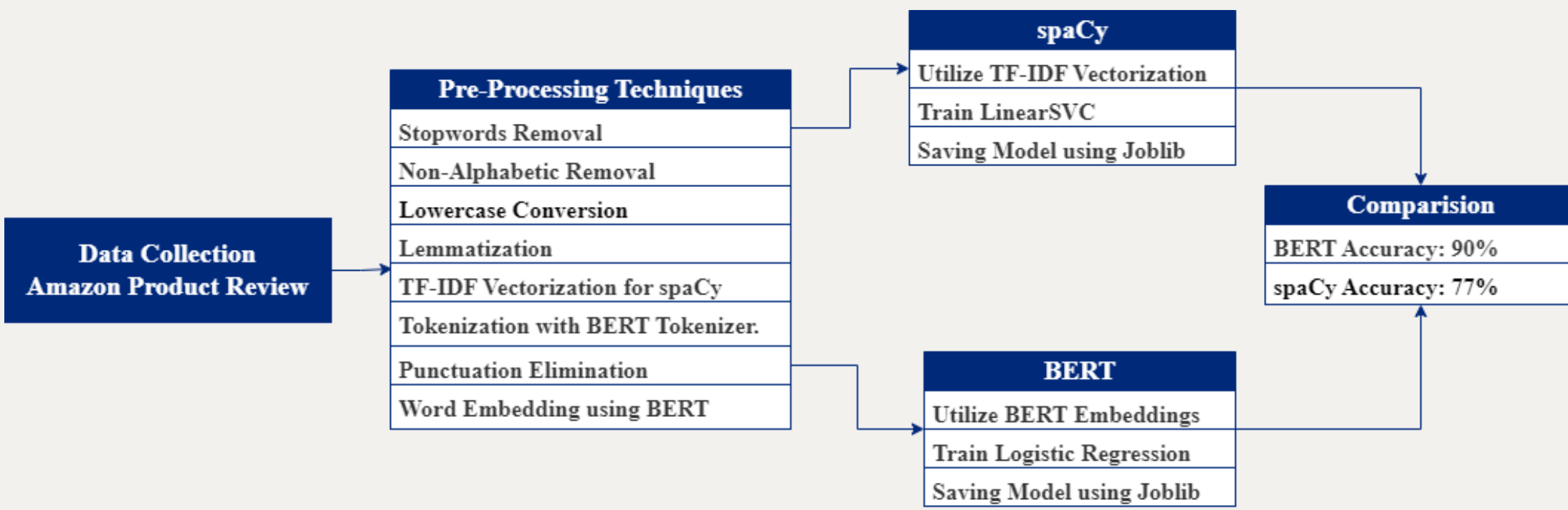
## MOTIVATION

Leveraging insights from the research[1], sentiment analysis emerges as a potent tool for extracting valuable information from customer reviews.. Amazon product reviews are a rich source of data that can be used to understand customer sentiment towards products and services. By analyzing Amazon product reviews, businesses can gain insights into customer satisfaction[2], identify product issues, and improve their overall customer experience.

## DATA PRE-PROCESSING

1. Data Loading and Cleaning:
- The dataset used in this project was sourced from Kaggle's open-source platform. Review text was cleaned to remove irrelevant characters, punctuation, and stop words.
2. Spacy Pre-processing and clean-up:
- Defined a function for lemmatization and removed stopwords and punctuation to pre-process the data.
  - TF-IDF Vectorization:* Used scikit-learn's TF-IDF Vectorizer to convert cleaned text into numerical vectors.
3. BERT Pre-processing and cleanup:
- The review text was decoded from bytes to UTF-8 format to ensure consistent text representation. The review text was cleaned by removing non-alphabetic characters, converting text to lowercase, and converting to a standard encoding format. The cleaned review text was broken down into individual tokens using the BERT tokenizer. BERT embeddings were generated for each token sequence representing the reviews.

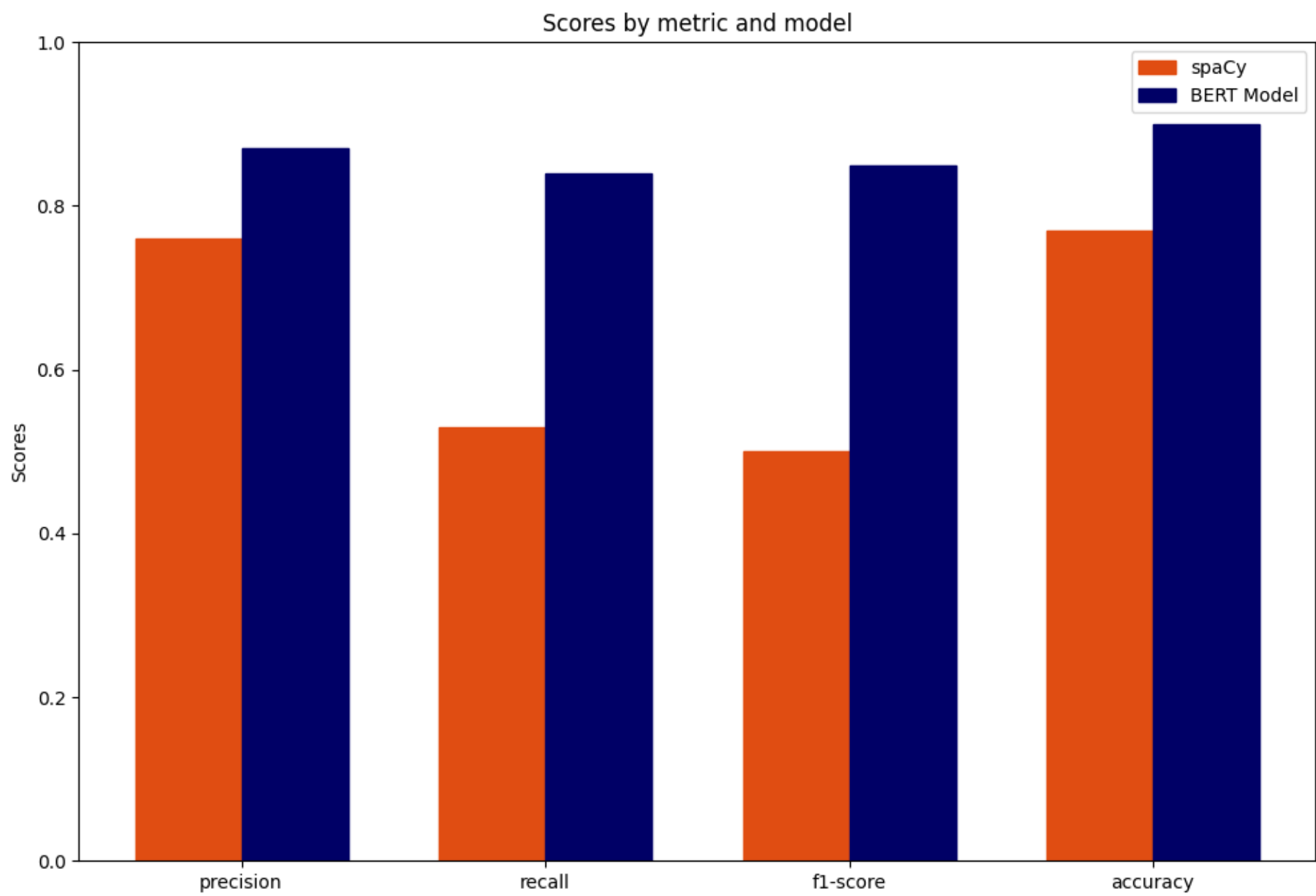
## MODELLING



This process involved two approaches: leveraging traditional NLP methods with spaCy and employing advanced NLP techniques with BERT.

1. SpaCy Model:
- The spaCy model employed a feature-based approach to sentiment analysis with referened to the research[3]. Loaded spaCy's English language model ('en\_core\_web\_sm'). It extracted meaningful features from the cleaned review text using TF-IDF vectorization. These features were then used to train a LinearSVC classifier, which assigned sentiment labels (positive or negative) to the reviews.
2. BERT Model:
- To better grasp subtle nuances in sentiment, the BERT model, trained through sophisticated natural language processing techniques as detailed in the research[4] was employed. It converted review text into BERT embeddings, numerical representations that captured the meaning and context of words. These embeddings were then used to train a logistic regression classifier, which predicted the sentiment of new reviews based on their embedding[5].

## DATA ANALYSIS



- The spaCy model achieved an accuracy of 0.77, while the BERT model achieved an accuracy of 0.90. This substantial difference highlights the superior ability of BERT to capture sentiment nuances compared to traditional NLP methods. BERT's ability to consider contextual relationships between words and the overall sentiment of the review allows it to make more accurate sentiment predictions.
- The bar graph demonstrates BERT's superior performance with precision, recall, f1-score, and accuracy scores of 0.87, 0.84, 0.85, and 0.90, respectively, which are significantly higher than spaCy's 0.76, 0.53, 0.50, and 0.77.

	Review	Predicted_Bert	Predicted_SpaCy	Actual
	this bundle was a great value for the money i couldn't believe everything that came included in this package for the price	Positive	Positive	Positive
	this was a cool idea but it didn't understand my commands wouldn't even turn on the nd day of trying to get it to work properly i sent it back amazon was very nice no refund problem	Negative	Positive	Negative
	unique excellent sounds great preformance just dont get robbed going down the street cause everybody gonna want it pretty neat and awesome	Positive	Positive	Positive
	we have a samsung galaxy the phone doesn't sit well on the docking station and isn't able to recharge we also had to remove the otter box to try this all out we had some issues in the beginning with the fidelio app but it is now working the actual sound is great we are just disappointed about the charging issue	Negative	Positive	Positive
	it can be difficult to get your phone in and out but once it is there it's there the arm wallet is comfortable and fits well	Positive	Positive	Positive
	i pre ordered this item when i pre ordered my kindle fire wish i could have seen all the other reviews because like alot of others i do not like the blurry pictures you get with this screen protector watching video or playing games the picture quality is grainy i took mine off and threw it away not what i expected	Negative	Positive	Negative
	way thin and cheaply made searched far and wide quality vhs clamshell cases must be a thing of the past	Negative	Negative	Negative
	it works i am happy with it thank you	Positive	Positive	Positive
	the radio doesn't come in at all the instructions that came with it suck we can't figure out how to use anything except the radio which doesn't come in clear at all there are other radios in our home that work fine so it's not the signal br i wish i could return it waste of money	Negative	Positive	Negative
	i did not like it the leather came off	Negative	Positive	Negative

BERT's predictions match the actual sentiments for 9 out of 10 reviews, and SpaCy's for 6 out of 10, in a random selection of reviews from the test dataset.

## SENTIMENT ANALYSIS USE CASES

- Lexalytics: Provides text analytics and sentiment analysis services for analyzing customer feedback, market research, and social media monitoring.

Salesforce Service Cloud: Part of the Salesforce suite, this customer service application uses sentiment analysis to glean insights from customer interactions and feedback.
- HubSpot: Offers CRM software that includes sentiment analysis features, especially useful in analyzing customer feedback and improving customer service.

Talkwalker: Talkwalker is another social media monitoring tool that uses sentiment analysis to track brand mentions. It offers a wider range of features than Brand24, including competitive analysis and influencer marketing.

## CONCLUSION

Demonstrating the significance of sentiment analysis in the current customer-centric business world, this project utilized it to successfully glean valuable insights from Amazon product reviews. This project successfully employed sentiment analysis to extract valuable insights from Amazon product reviews. The BERT model outperformed the spaCy model, highlighting the effectiveness of advanced NLP techniques in sentiment analysis tasks. By leveraging sentiment analysis, businesses can gain deeper insights into customer feedback, leading to improved product offerings and enhanced customer satisfaction. As NLP technologies continue to evolve, sentiment analysis will undoubtedly play an even more crucial role in helping businesses connect with their customers and thrive in the competitive marketplace.

## REFERENCES

- Nur Azizah Vidya, Mohamad Ivan Fanany, Indra Budi, Twitter Sentiment to Analyze Net Brand Reputation of Mobile Phone Providers,Procedia Computer Science,Volume 72,2015,Pages 519-526,ISSN 1877-0509
- Capuano, N., Greco, L., Ritrovato, P. et al. Sentiment analysis for customer relationship management: an incremental learning approach. Appl Intell 51, 3339–3352 (2021).
- S. JUGRAN, A. KUMAR, B. S. TYAGI and V. ANAND, "Extractive Automatic Text Summarization using SpaCy in Python & NLP," 2021 International Conference on Advance Computing and Innovative Technologies in Engineering (ICACITE), Greater Noida, India, 2021, pp. 582-585, doi: 10.1109/ICACITE51222.2021.9404712.
- Alaparthi, S., Mishra, M. BERT: a sentiment analysis odyssey. J Market Anal 9, 118–126 (2021).
- M.P. Geetha, D. Karthika Renuka,Improving the performance of aspect based sentiment analysis using fine-tuned Bert Base Uncased model, International Journal of Intelligent Networks, Volume 2,2021,Pages 64-69,ISSN 2666-6030
- Pan R, García-Díaz JA, García-Sánchez F, Valencia-García R. Evaluation of transformer models for financial targeted sentiment analysis in Spanish. PeerJ Comput Sci. 2023 May 9;9:e1377. doi: 10.7717/peerj-cs.1377. PMID: 37346571; PMCID: PMC10280559.