



# Amazon Review Analysis

Presented by Varun Jani | CS5804 Mini-Project

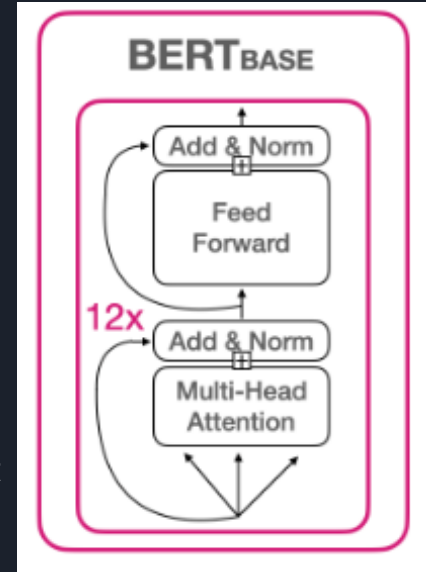


# Problem Statement & Motivation

- Leverage open-source Amazon reviews dataset (US only) to analyze sentiment of people sharing their reviews on Amazon
- Improve customer interaction with sellers/businesses on Amazon
  - Provides insight on product
  - Helps improve seller sales
  - Gather shortcomings for a product
- This project focuses on electronic products on Amazon but versatile as it can be applied to any product Amazon has to offer
- Motivation: Showcase reviewers' feelings and provide insight on the reviews (topic modeling, word clouds and sentiment analysis) as well as score reviews
  - Improves marketing and advertising as it allows Amazon to promote those products more and ideally increase sales

# BERT Fine-tuning on Reviews

- **BERT:** Bidirectional Encoder Representations from Transformers
  - Trained on large amount of data (~3.3 billion words)
  - BERT Base (~110 million parameters)
  - SOTA model for sentiment analysis
- BERT Transformer uses bidirectional self-attention
  - allows the model to look in both directions, forward and backward and extract information the model could have missed on
- Model hosted on Hugging Face and developed by LiYuan: **amazon-review-sentiment-analysis**
  - Predicts review rating star and can also be used for sentiment scoring
- Model is a fine-tuned version of bert-base-multilingual-uncased-sentiment



# Sentiment Analysis & Topic Modeling



- Sentiment scoring: Provides a score to the review
  - This can help understand user sentiment, allowing for better engagement between customer and seller/business
- Word Clouds: Points out most used keywords
  - Can help seller/business filter review based on most sought out keywords
- Topic Modeling: For each topic, we can find the most used keywords
  - This can be expanded for each product - this also makes computation easier
  - Can also be expanded for each type of product
  - Uses LDA (Latent Dirichlet allocation)



# Results & Live Demo

# Lessons Learned & Review Bot Use-Case

- BERT fine-tuned model (amazon-review-sentiment-analysis) has an accuracy of 0.8
  - Training from scratch or fine-tuning a better model could improve accuracy
  - Train with additional data (needs more computer power)
  - Use existing SOTA models (LLMs) - also need more compute power
- The dataset is also useful in training language models to write reviews
- With more time and compute power, we can train large language models to write reviews
  - Although this would promote spam bots, the idea is to create bots that can identify bot reviews





# QUESTIONS?

Interesting Reads: <https://www.securitymagazine.com/articles/99339-47-of-all-internet-traffic-came-from-bots-in-2022#:~:text=A%20new%20report%20reveals%20that,lowest%20level%20in%20eight%20years>  
<https://www.cnbc.com/2023/04/25/amazon-reviews-are-being-written-by-ai-chatbots.html>  
<https://www.theguardian.com/money/2023/jul/15/fake-reviews-ai-artificial-intelligence-hotels-restaurants-products>