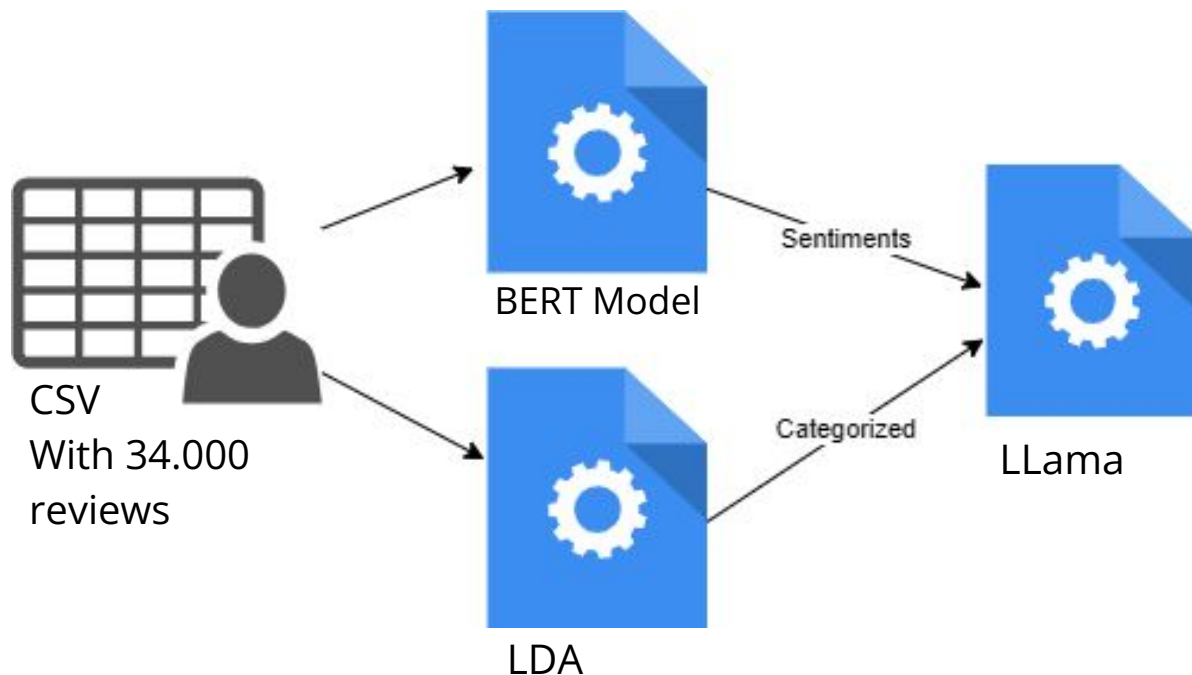# Robo Reviews

Fabian

Tired of writing reviews to push your Amazon score?

- Generate automatically a review for a product
- Define if you're satisfied or not, to generate a review which fits!
- Get free products from Amazon to write reviews for.

What is your story?

CSV
With 34.000 reviews

BERT Model

Sentiments

LDA

Categorized

LLama

➕ Positive

I am happy with the purchase of Google chromcast and it's working fine.
What are you waiting to do? Get one today!

➖ Negative

I bought it to use with my TV and the remote control, but after 2 months of using them together they stopped working at all!

Demo

REVIEW GENERATOR

This is Fabi, the best review generator in the world.

Read my best generated reviews of my favourite categories and generate you own reviews below.

Amazon Fire Tablet ⌄

FIRE TABLET, 7 DISPLAY, WI-FI, 8 GB - INCLUDES SPECIAL OFFERS, MAGENTA

> *I am satisfied with the performance of my Kindle Fire tablet and ...*
>
> *I am satisfied with the performance of my Kindle Fire tablet and it has exceeded all expectations in terms on functionality as*

- Read my generated reviews
- See which products I recommend
- Generate your own review
  (Currently disabled cause alpha state)

https://e045-88-72-105-108.ngrok-free.app

GENERATE YOUR REVIEW

Currently under construction

Enter product name   Positive   ⌄

Your generated review

GENERATE

Project Overview & Data Preprocessing

Dataset:
34.000 consumer reviews for Amazon from Kaggle

Project Goal:
Analyze Amazon product reviews to understand customer sentiment and categorize products based on their descriptions. Create Reviews for Products.

Preprocessing
Text Cleaning: Removed special characters, numbers and stop words from product names and reviews.
Text Transformation: Combined review text with recommendation and rating information.
Category Cleaning: Cleaned and deduplicated product categories, removing irrelevant terms

Methods

Sentiment Analysis:

Utilized pre-trained BERT and Roberta models to classify review sentiment into positive, negative, and neutral categories.

Compared model performance and selected the best performing model for sentiment prediction (BERT).

Product Categorization:

Applied Latent Dirichlet Allocation (LDA) to identify underlying topics within product names.

Assigned dominant topics to products as predicted categories.
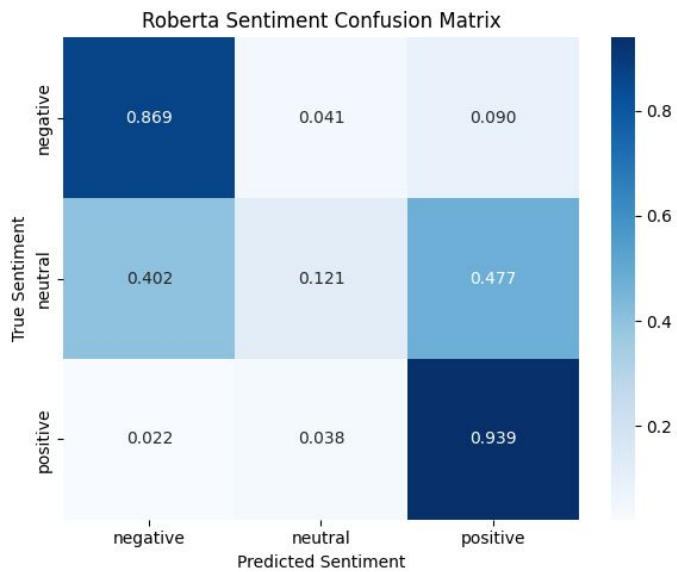
Review generation:

Applied LoRA and a PEFT technique, to efficiently fine-tune a pre-trained Llama 3.2-1B Model
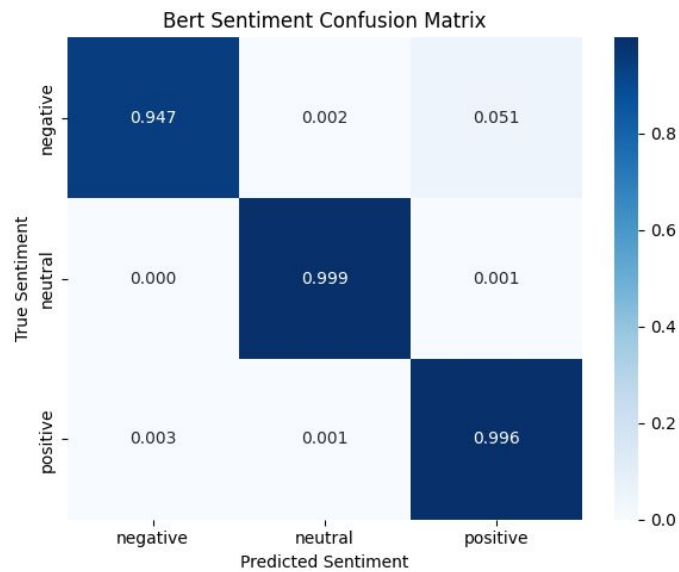Used a Bart Reviews Model for Summarization

BERT vs. Roberta

Classify customer reviews into positive, negative, or neutral
Tested against predictions generated by the ratings
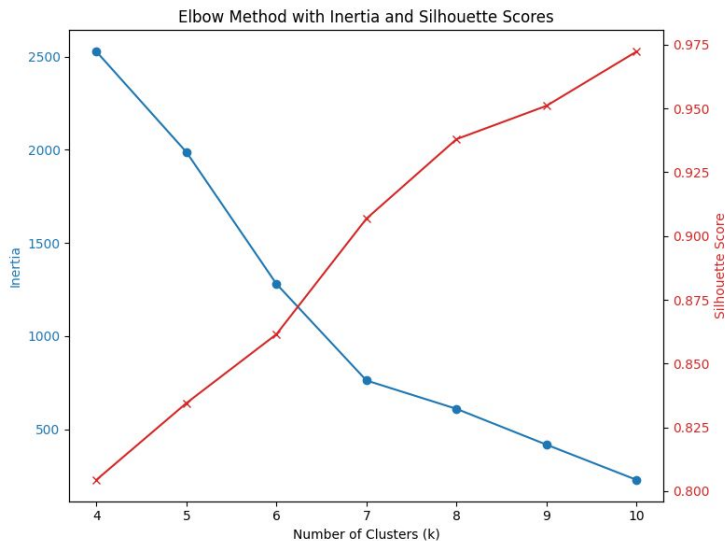

Accuracy: **0.8996**

Accuracy: **0.9951**

KMeans vs. LDA

KMeans

Clustered by categories
n_clusters between 4-10

LDA

Clustered by products with BoW
n_clusters between 6-10

Elbow Method with Inertia and Silhouette Scores

LDA - Evaluation

Generated clusters:

Amazon Echo White

Amazon Fire Kids Edition

Amazon Fire Tablet

Amazon Fire TV

Amazon Kindle Paperwhite

Amazon Portable Speaker

Charger & Accessories

- Removed more and more words
- Tried 10-5000 epochs
- Tried some Hyperparameter
- Cluster naming by visualisation

Perplexity: **-2.3055**

Llama

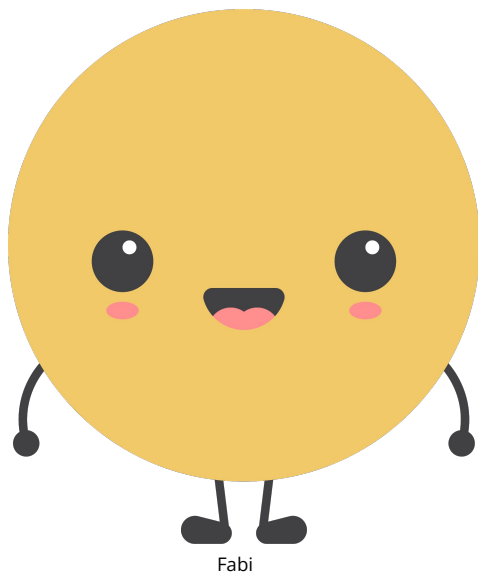After problems with Mistral and GPT2, Llama-3.2-1B was running

Mistral
GPT2

Llama

- LoRA for fine-tuning
- Peft to configure the base model with the LoRA config
- Few-Shot Learning with promts

```
"Write a detailed Amazon review for
'{product_name}'.
My meaning about this product is
{sentiment}"
```

Takeaway

- Transformers safe your day
- Transformers are easy to implement
- Transformers are great!

- LDA is interesting, will use when there is time

- It's not hard to use Flask

- Need to learn more about deploy on Huggingface

Fabi

Questions & Complaints