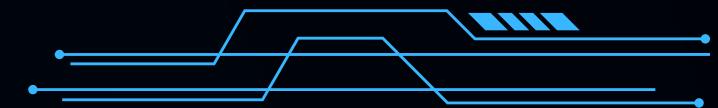


CADENCE 2B

UNLOCKING INSIGHTS FROM MILLIONS OF PRODUCT REVIEWS



OVERVIEW



- As Principal Product Designer, you face endless decisions to build the best product...
- Imagine wading through thousands of product reviews.
 - The star rating looks fine but do people secretly dislike the battery life?
 - Are hidden frustrations going undetected?
- Our project uses machine learning to look beyond ratings and extract what customers really think about key product features.
- Automated, scalable analysis which means no more manual review!

WHO ARE WE



Aaryaa Moharir



Jianhua Deng



Hannah Rauch



Darlyn Gomez

OUR AI STUDIO COACH & ADVISOR



Matt Brems



Farhan Rasheed

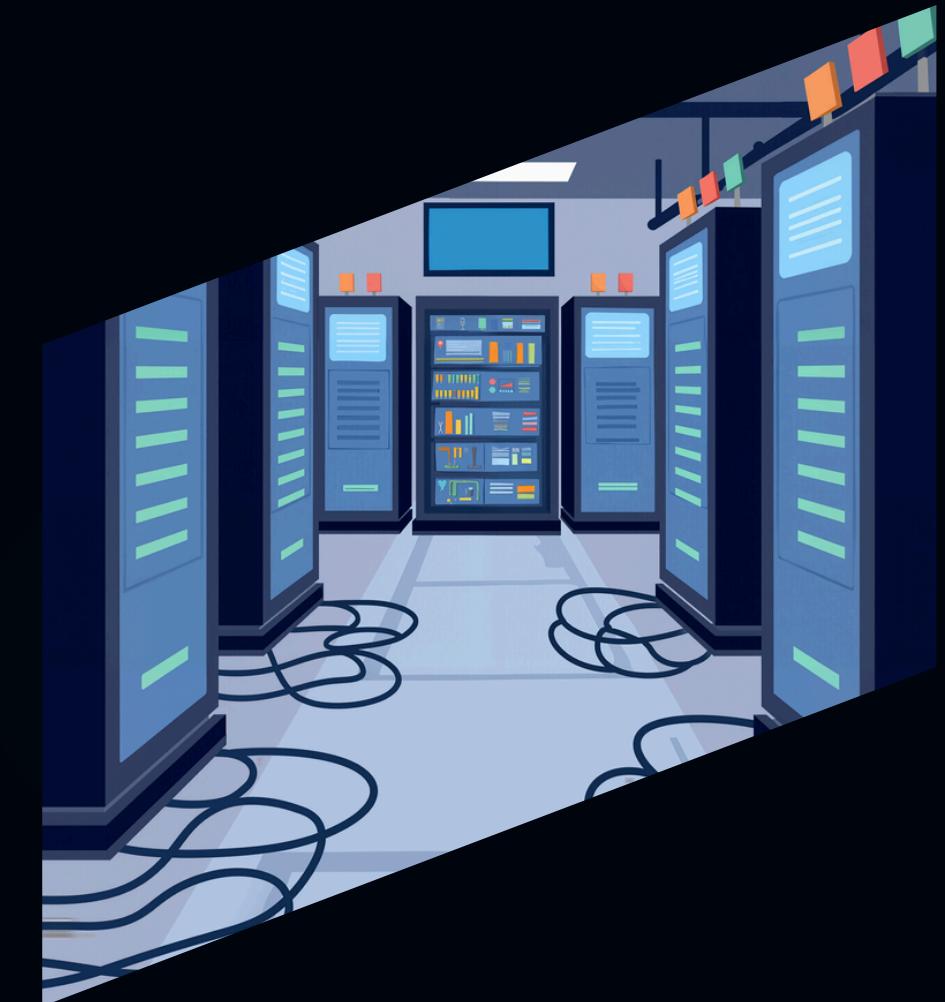
VISION

PROBLEM

- Companies have lots of reviews, but little clarity on specific product strengths/weaknesses.

SOLUTION

- AI system for extracting feature-level insights from Amazon reviews.
- Initial models separated sentiment and feature extraction, but the ABBA model combines both.
- Had to solve for long runtimes—used random sampling to accelerate testing.
- Early results show strengths and pain points for battery life, usability, and more.



HOW

Using Data - Preprocessing and a Amazon Reviews Dataset

MISSION

DATASET

- Amazon reviews, 2023
- 43 million entries (text, ratings, timestamps), shrunk to the Electronic categories with 1 million entries, stored as Parquet

CLEANING

- Removed HTML, duplicates, short/garbage reviews
- Got rid of filler words, conjunctions
- Major class imbalance since we had overwhelmingly positive reviews

EDA

- Distribution patterns by product type
- What features are being extracted
- How does class imbalance change after sampling

MODLING & EVALUATION

Data
Preprocessing

Explored Feature
Extraction +
BERT

Pivoted to ABSA
Model

Built a Streamlit
Dashboard



MODELS

BERT + FEATURE EXTRACTION

1. The Bert model is a Foundational Transformer model from Google
2. It is capable of text classification, sentiment classification and much more
3. It is very capable of being adapted and fine-tuned to do different text related tasks

ASPECT-BASED SENTIMENT ANALYSIS (ABSA)

1. The model identifies aspects (features) in the data and determines whether the sentiment is positive, negative, neutral
2. Parse review with spaCy
3. For each aspect, build an input pair (review + aspect) and tokenize with AutoTokenizer
4. Run tokens through AutoModelForSequenceClassification for sentiment label
5. Load pretrained checkpoint from Hugging Face and evaluate

RESULTS

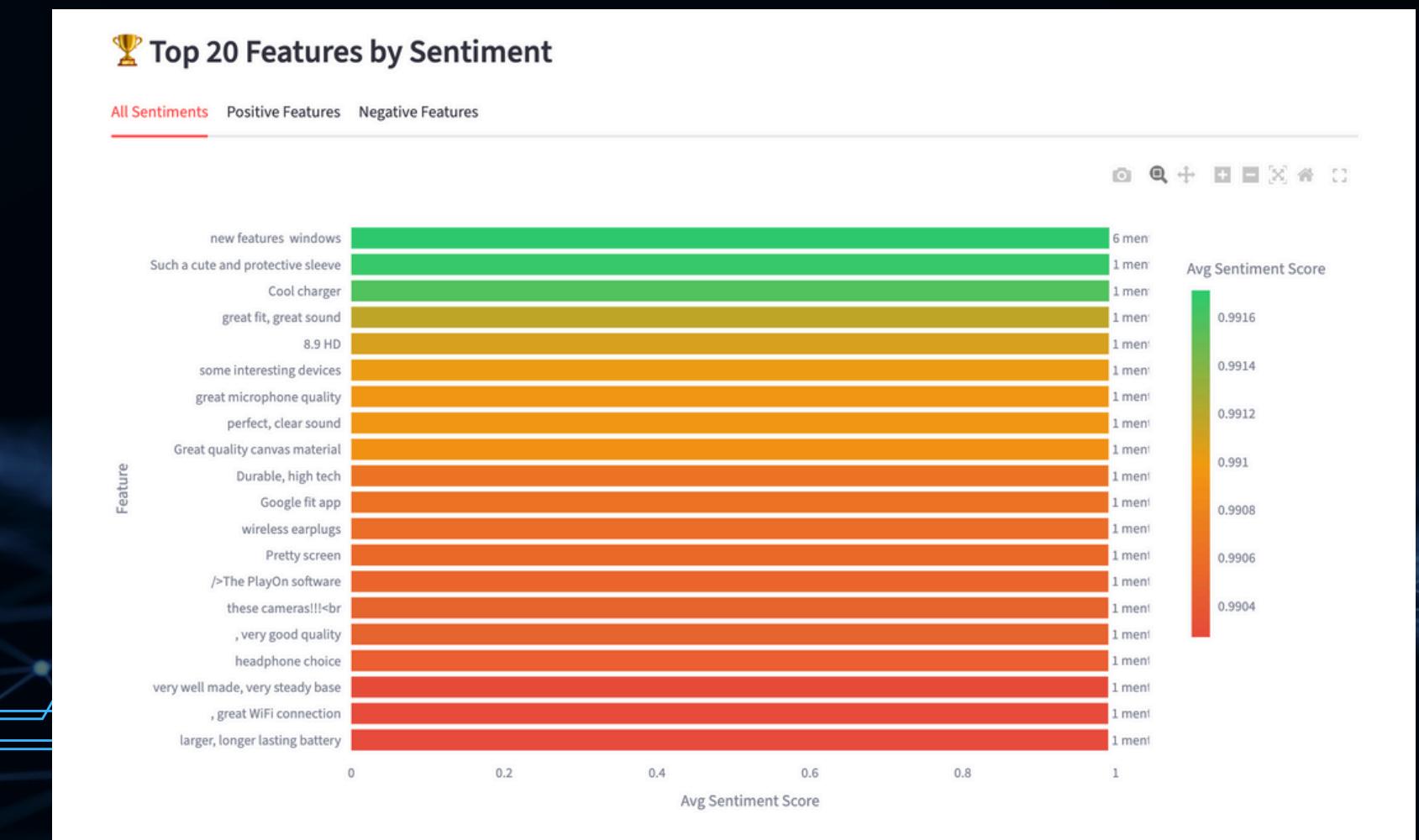
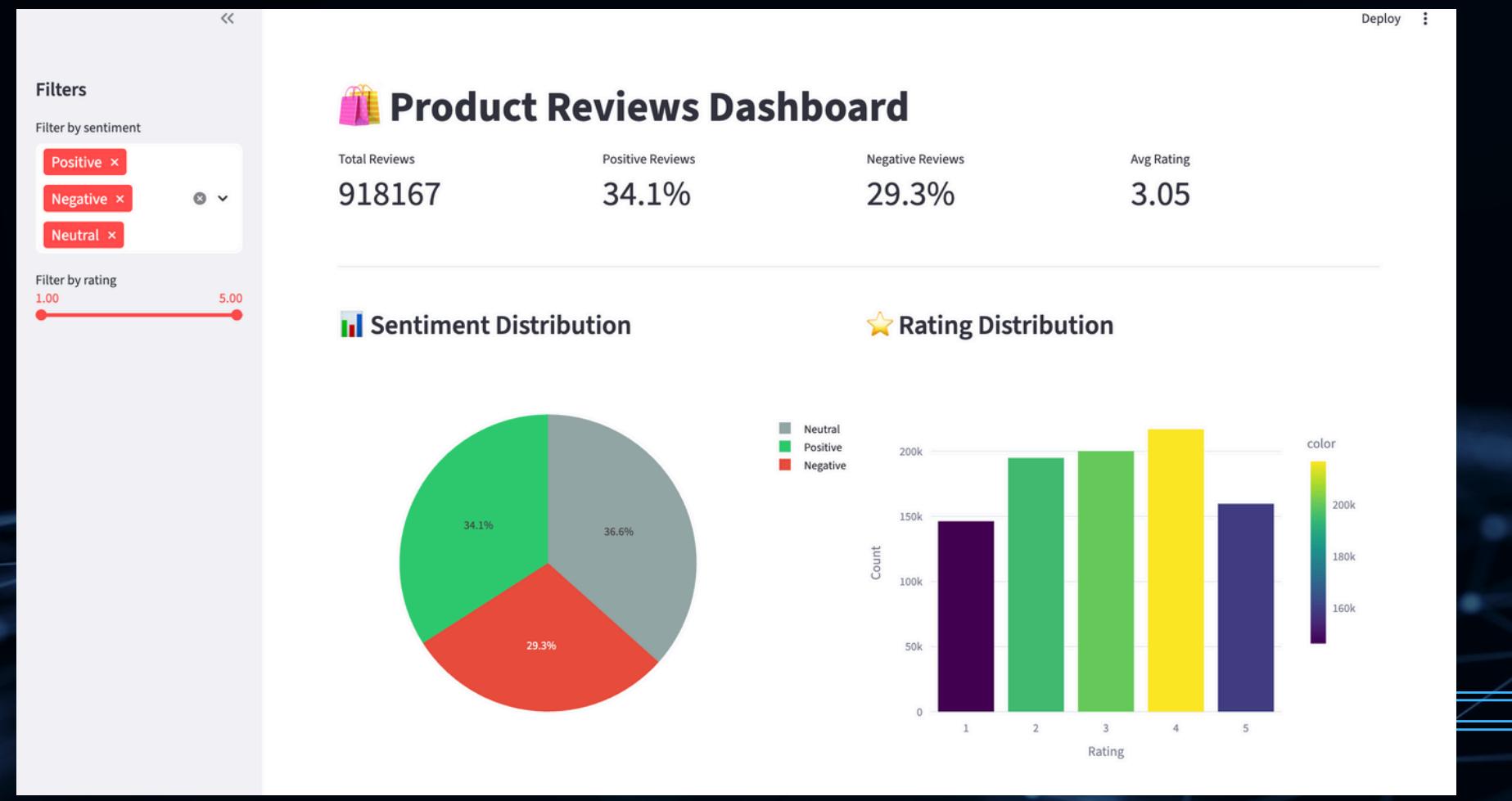
BERT + FEATURE EXTRACTION

1. Good Accuracy on the fine-tuned Bert model, with an 88% accuracy on sentiment analysis
2. The resulting dataset consists of many garbage words, such as: Words, Weeks, Works
3. Basically, by only using the Spacy Library for feature extraction, we weren't able to extract the features correctly

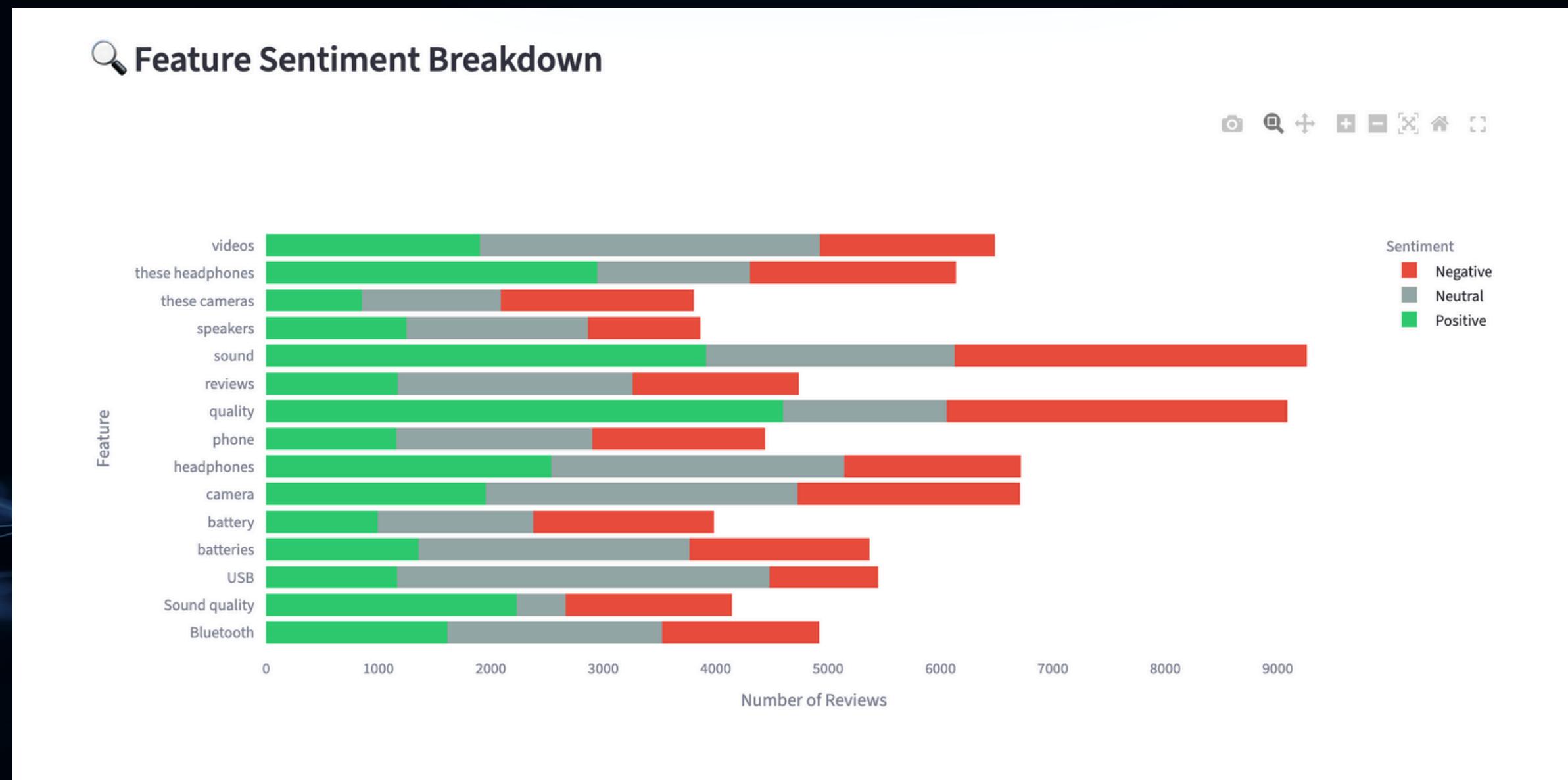
ASPECT-BASED SENTIMENT ANALYSIS (ABSA)

1. Identifies specific product features such as battery, screen, and camera, assigning sentiment to each to solve the mixed-review problem where users praise one feature but dislike another in the same review.
2. Uses a pre-trained DeBERTa model from Hugging Face with 88% accuracy, leveraging transfer learning to cut processing time from days to hours.
3. Delivers clear, useful insights like “camera sentiment 4.8/5, battery 2.1/5” so product teams instantly know what to keep improving and what to fix.

RESULTS



RESULTS





Hugging Face



Google Drive

OUR TOOLS

FINAL THOUGHTS

We grew so much as a team, not only technically but also professionally.



CADENCE 2B

THANK YOU

