

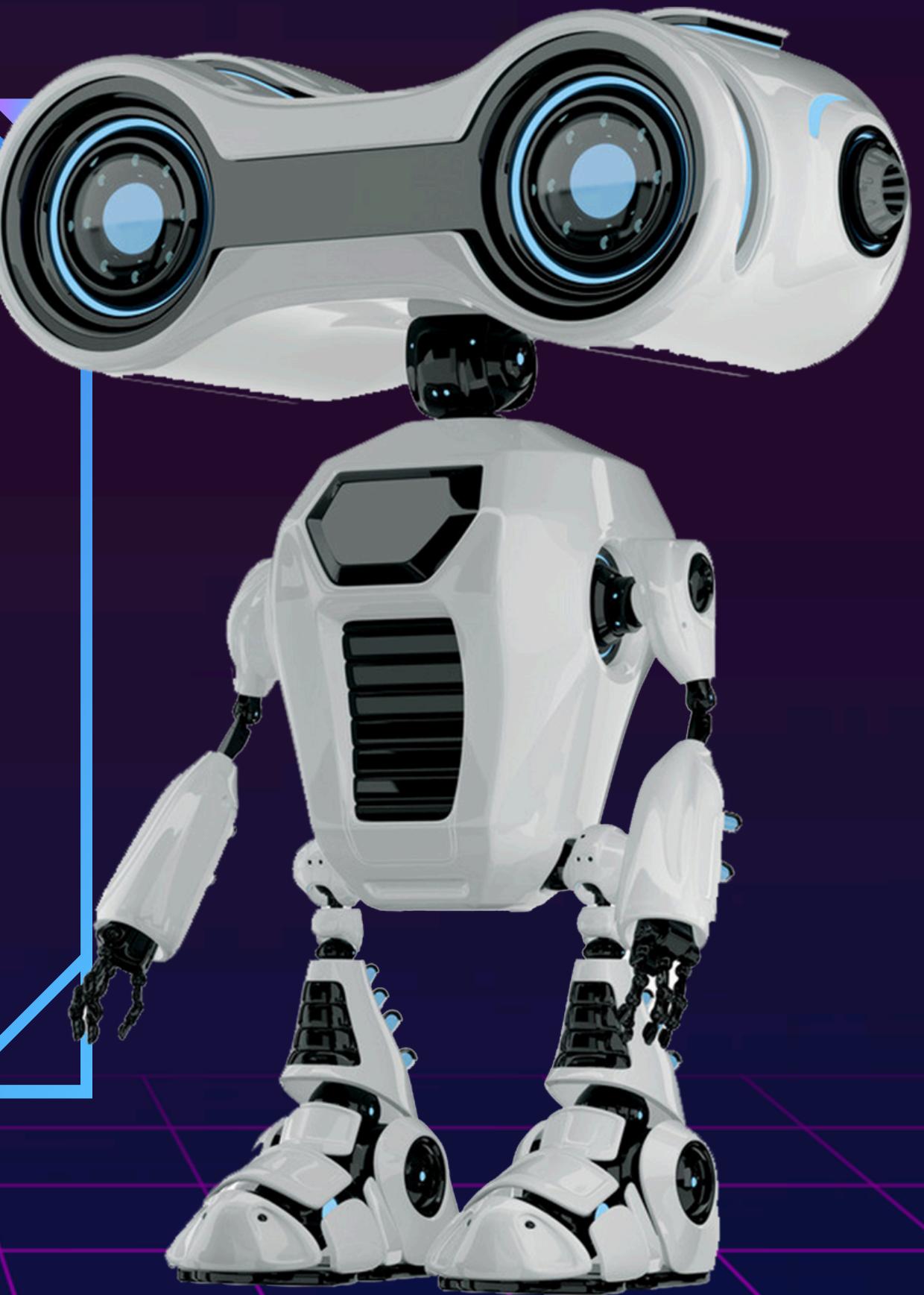
PAUL MÜLLER

ROBO-REVIEWS

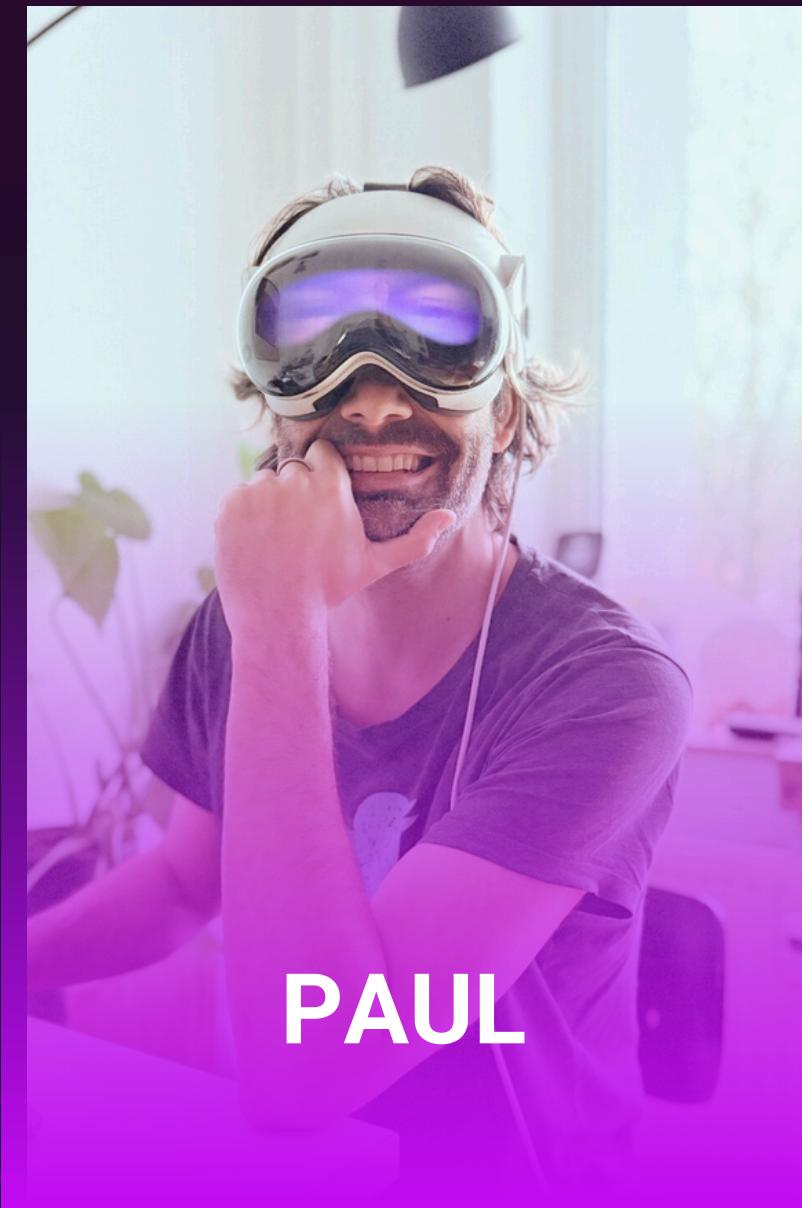
THE BATTLE OF THE DATASET

THE PLOT AND GOAL

- Development of a product review website powered by an NLP model(s)
- The model(s) aggregates customer feedback from different sources
- Main tasks:
 - a. Classify customer reviews into positive, negative, or neutral categories to help the company improve its products and services.
 - b. Cluster product categories into just 4-6 of them
 - c. Use Generative AI to summarize reviews into an article which recommends the top products for each category

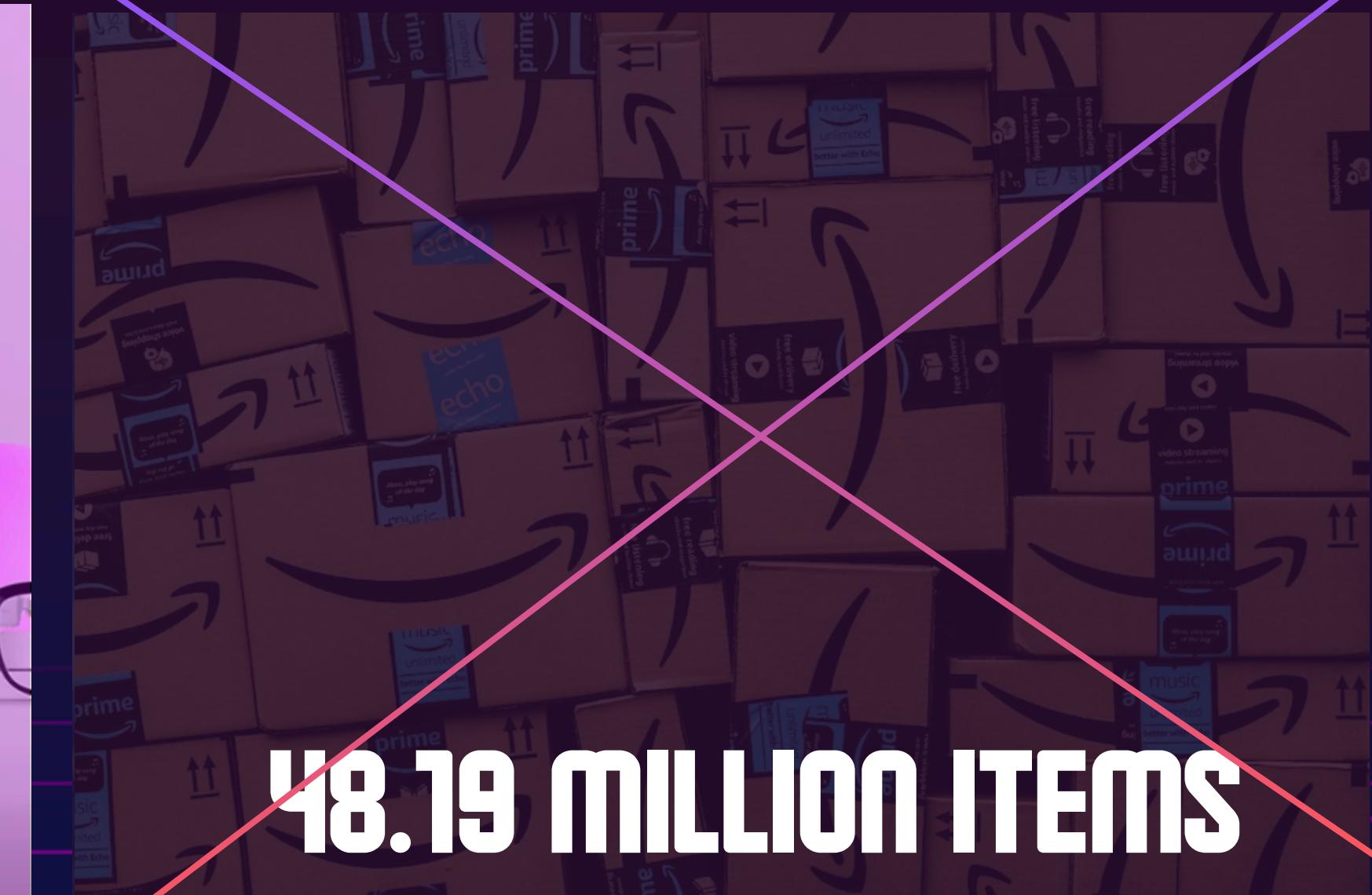


MEET THE WARRIORS



OPPONENT - THE DATASET

We have chosen the small but nasty 34,000 consumer reviews for Amazon products like the Kindle, Fire TV Stick, and more. The dataset includes basic product information, rating, review text, and more for each product.

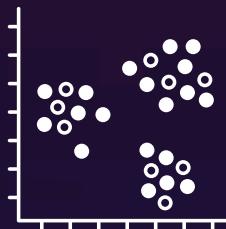


#SUMMARY OF THE STORY



- Quick Achievement of good results with **Vader model** 65% with 1-5 Scoring
- and **85% accuracy of Positive / Neutral / Negative Segmentation**, with the Vader model.

**85 % ACCURACY VS
USER RANKING**



K-MEAN IS KEY.

- Shitty data showed it faces. The dataset is very unclean and probably very realistic.
- a lot of approaches until Thursday to achieve good result.
- post-processing solved the issues - lowers adaptability for other datasets.

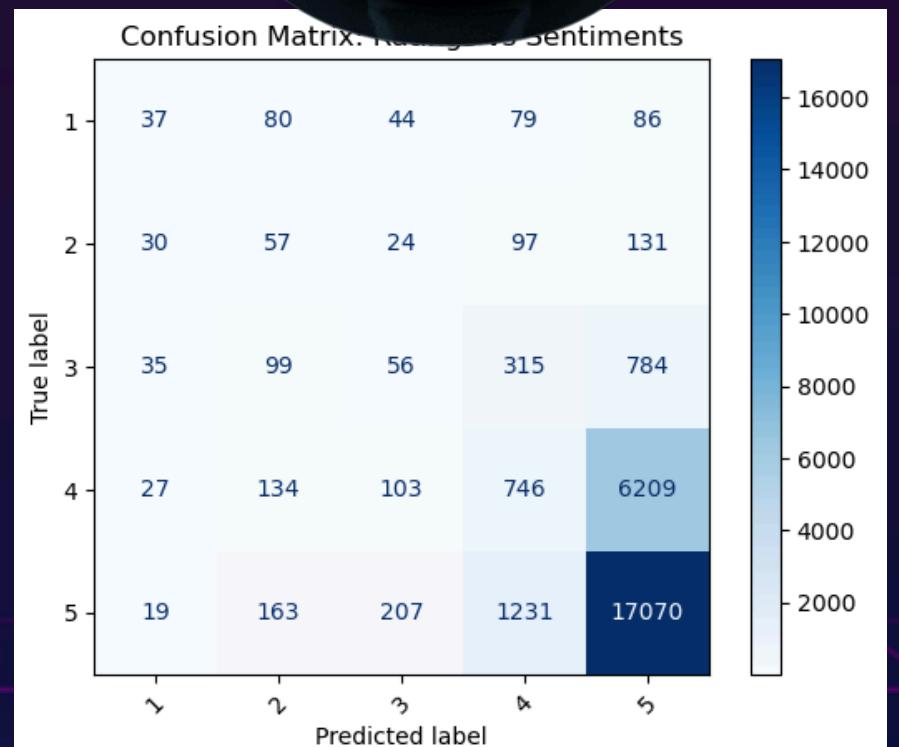


- **Header (LLM)**
- **Product teaser (LLM)**
- **Ranking by sentiment TOP 1-3 products of the category (Vader & K-Mean)**
- **Feature Pro and Cons TOP 1 product (LLM)**
- **Example user quote for emotional attention**
- **Summary of Highs and Lows of product by customer reviews (LLM)**
- **Wrapup and Buy Suggestion (LLM)**

OUTPUT

CHAPTER 1

- simple data cleaning,
 - source reviews.text and reviews.title
 - reduce double, NaN entries, non usefull content (e.g. date)
- Pre-processing
 - TfidfVectorizer, stopworrds
- model decision
 - provided ranking not accurate --> unsupervised learning
 - research for specialized models
 - Vader vs. DistilBert - dark side wins, immediately good results
 - 65 % of accuracy vs. given ranking & Pos / Neu / Neg Segmentation 85% matching



DARK SIDE WINS

CHAPTER 2

- difficult data cleaning,
 - source product_name, categories
 - renaming of products (work-in-process while clustering and LLM setup)
 - deleting non named products not usefull for main task --> Result: 34k -> 27k datat
- Pre-processing
 - visual inspection for clustering,
 - dropping approch for model use for category name creation
- Post-processing
 - inflationary words in name and categories (e.g. tablets, amazon, fire, TV, Echo)
 - missmatch customer reviews and product names (e.g. TV vs Echo)
- model decision
 - no labels --> unsupervised learning
 - K-Mean vs. Distilbert - **alien wins**, immediately good results
 - 2/3 time and resources for okish clustering



THE ALIEN WINS

CHAPTER 3

- data cleaning,
 - renaming of products (work-in-process while clustering and LLM setup)
- concept phase
 - concept for deployment
 - concept of marketing and visual approach (customer centric)
 - development example result for each category as prompt
- model decision
 - focus saving time and resources -- no LLama or GPT
 - T5 model vs. Cohere - **Marketing wins** (specialised model chosen (for specific task))
 - 2/3 time and resources for okish clustering again
 - few-prompting and example file

MARKETING WINS



TOP 10 TABLETS

*CONTENT
LLM CREATED



UNLEASH LEARNING AND FUN WITH THE FIRE KIDS EDITION TABLET - ADVENTURE AWAITS!*

Step into the world of fun and learning with the Fire Kids Edition Tablet! This 7-inch display powerhouse packs a punch with Wi-Fi connectivity and ample 16 GB storage. Designed with kids in mind, it comes in a vibrant green kid-proof case, ensuring durability and peace of mind for parents. Perfect for little hands, it's lightweight and easy to hold, making streaming and learning an immersive experience.*

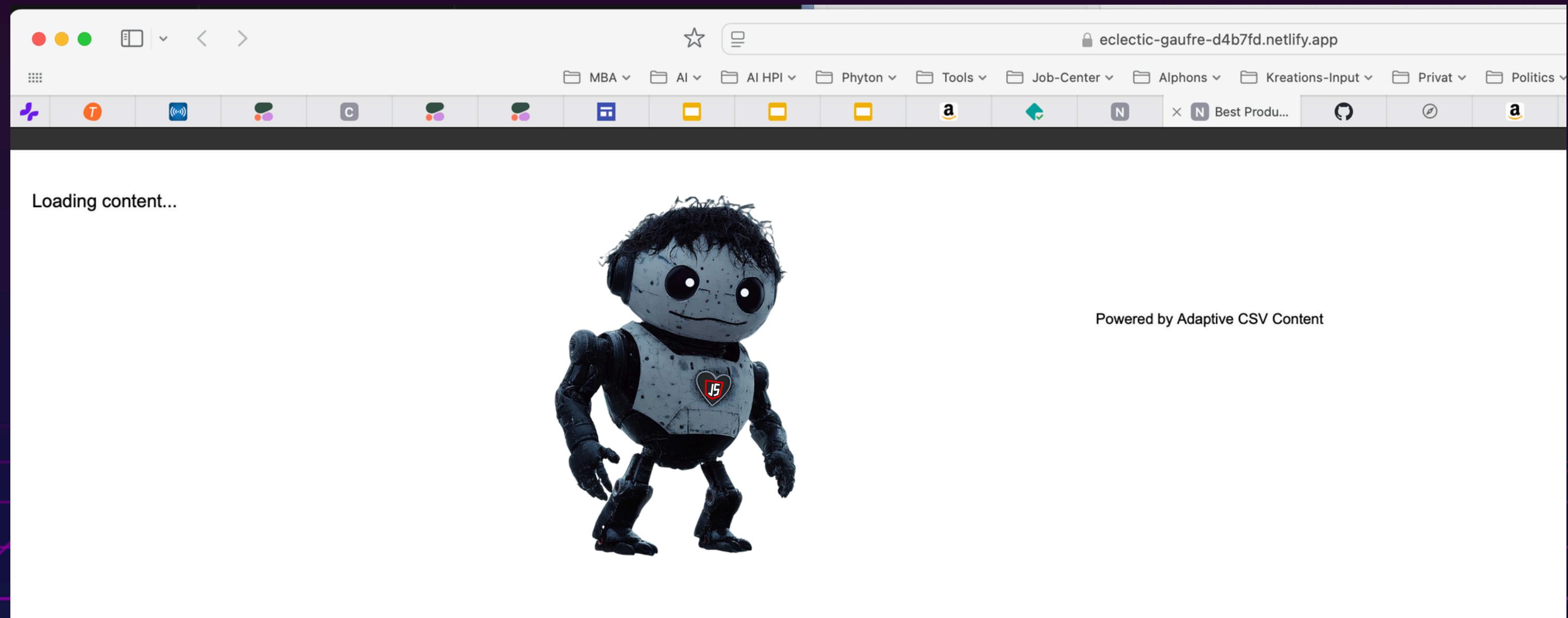
CUSTOMERS VOICE:

"GREAT PRODUCT THE TABLET IS VERY LIGHT AND STREAMS WELL"

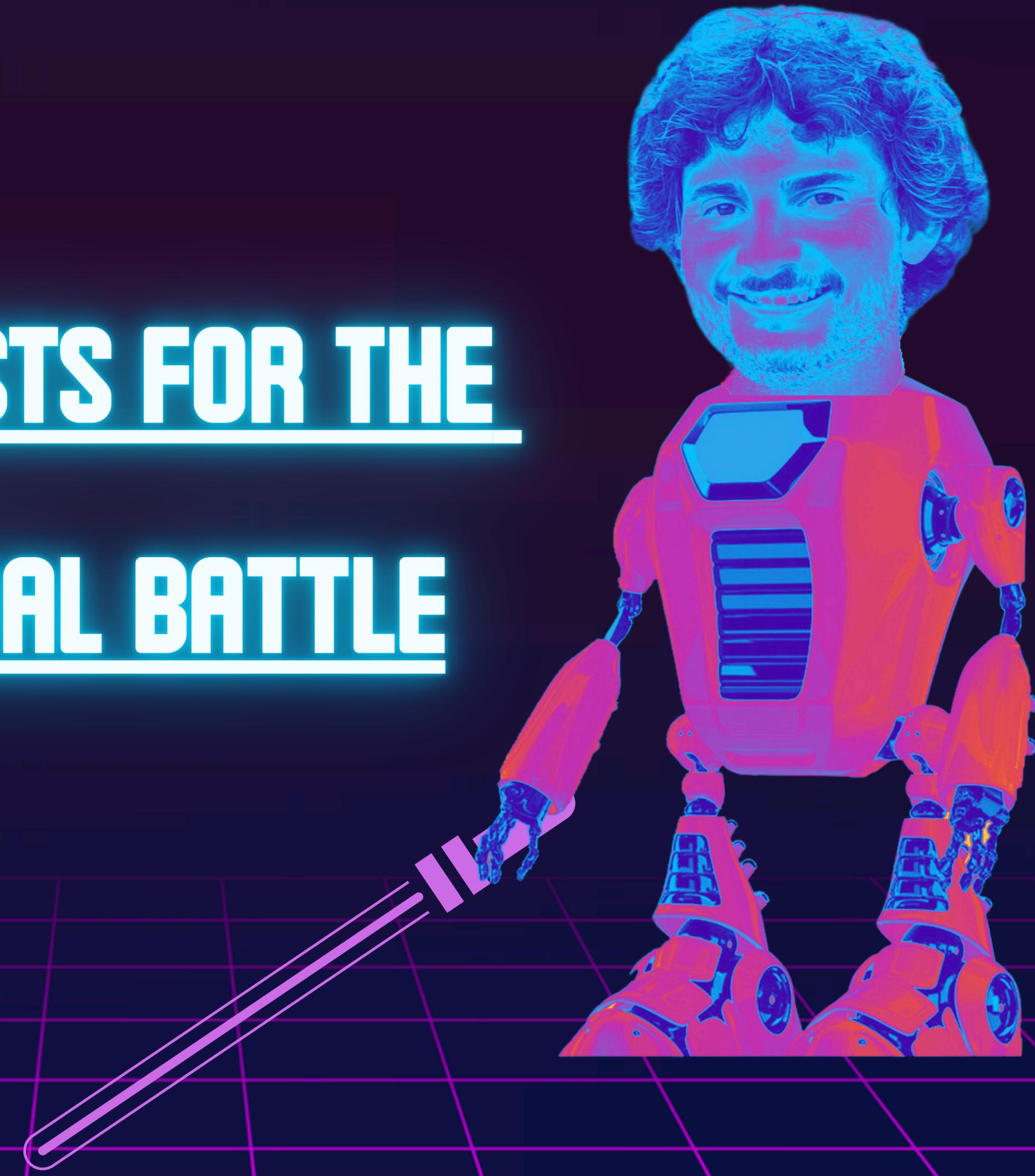
TOP 3 PRODUCTS	CUSTOMER RANKING
Fire Kids Edition Tablet, 7 Display, Wi-Fi, 16 GB, Green Kid-Proof Case	4.81
All-New Fire HD 8 Tablet, 8 HD Display, Wi-Fi, 16 GB - Includes Special Offers, Magenta	4.81
Fire Tablet, 7 Display, Wi-Fi, 8 GB - Includes Special Offers, Magenta	4.78

THIS TABLET OFFERS AN IMPRESSIVE BLEND OF PERFORMANCE AND VALUE, CATERING TO VARIOUS USER NEEDS. WHILE PRAISED FOR ITS LIGHTWEIGHT DESIGN, STREAMING CAPABILITIES, AND AMPLE MEMORY, SOME USERS HAVE ENCOUNTERED ISSUES WITH TOUCH SENSITIVITY AND AN UNUSUAL CHARGING BEHAVIOR. IT'S A GREAT LOW-COST OPTION, BUT SOME QUALITY CONTROL CONCERN HAVE BEEN NOTED.*

END BOSS - DEPLOYMENT

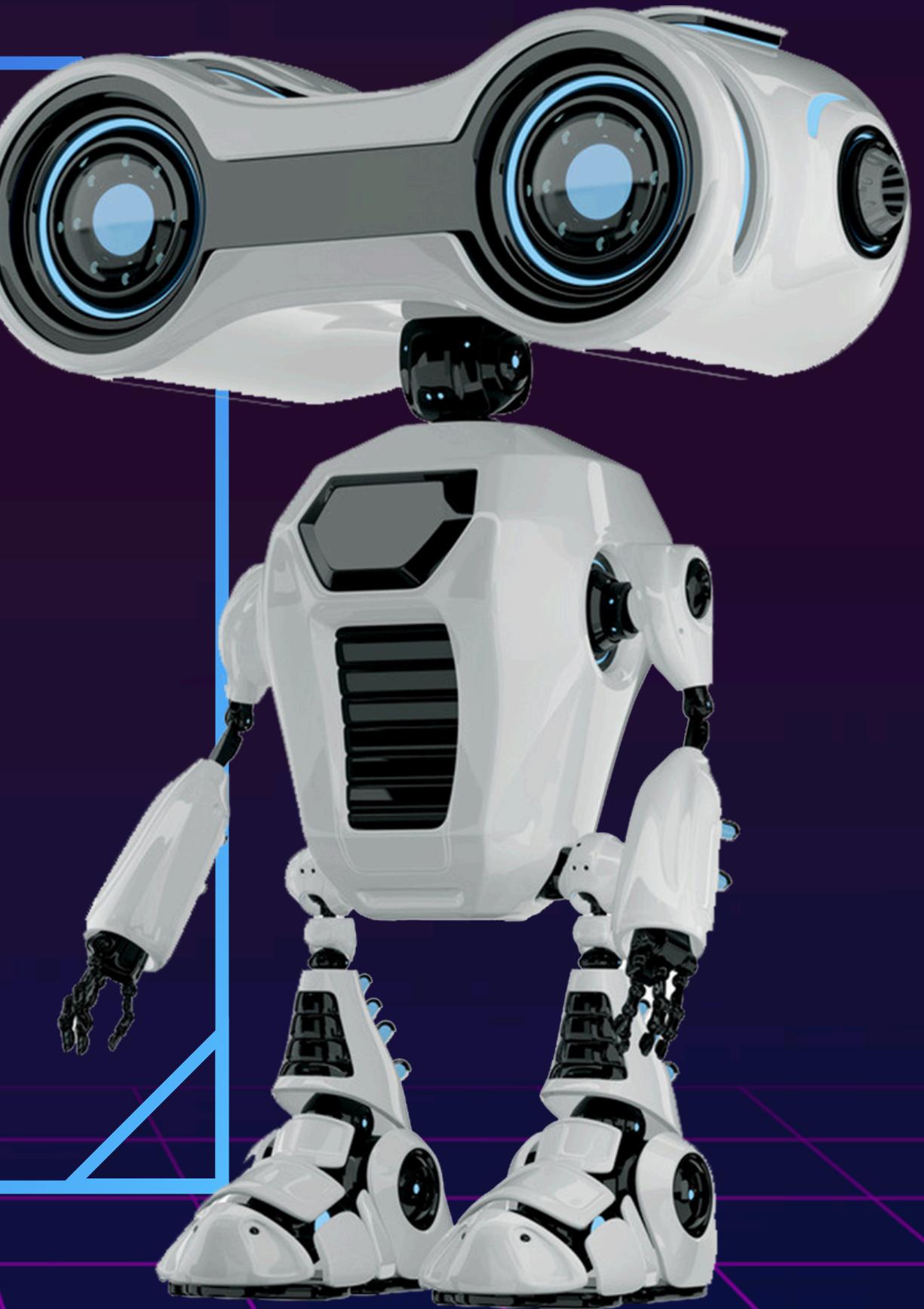


BEASTS FOR THE
FINAL BATTLE



TAKE AWAYS

- DATASET IS KING
- MODELLING IS OKISH DEPENDING THE DATASET
- WITHOUT CHAT GPT I WOULD BE NOTHING
- FRUSTRATION IS HIGH IF I FOCUS TO MUCH ON THE OUTCOME
- JAVA AND HTML IS NOT PYTHON



“MEDAL APPLICATION”

TURNING FRUSTRATION INTO CREATIVITY

CTA-DRIVEN AFFILIATE APPROACH OVER A REVIEW PAGE

THE PROFESSOR BECAME PART OF HERO STORY





THANK YOU

