ROBO-REVIEWS

Diego Alonso

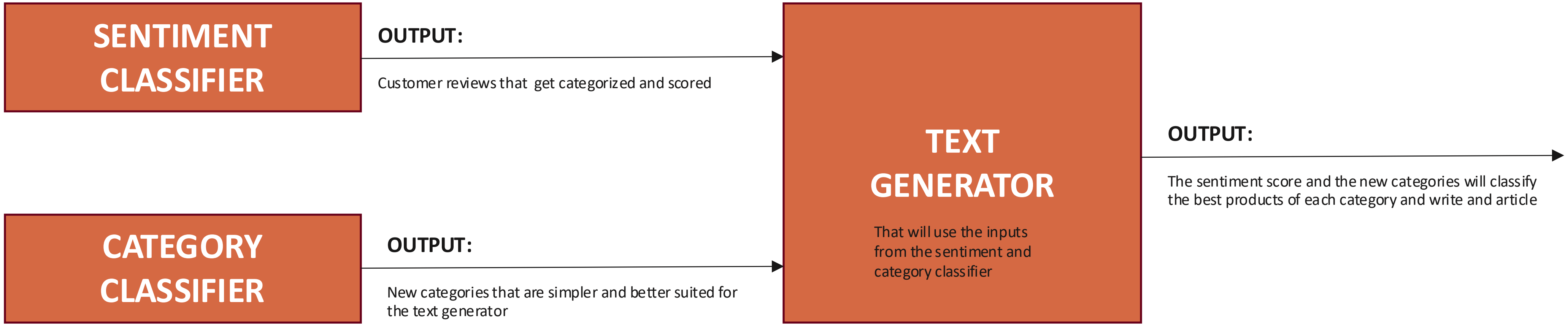
1.- PROJECT OVERVIEW

- Used NLP techniques to **classify consumer reviews** by sentiment (positive, neutral, negative) and **grouped products into broader categories**. I also tried to identify the top 3 best and worst products per category based on sentiment and ratings..

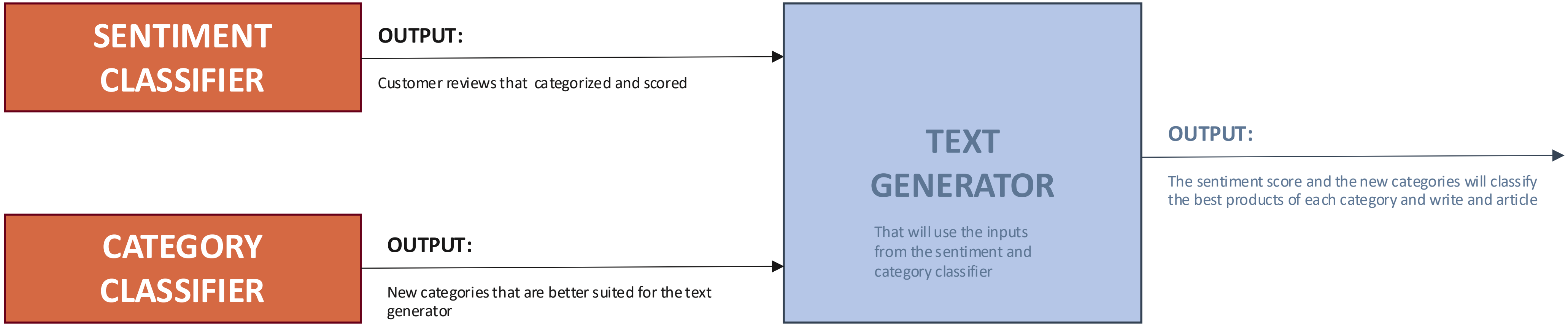
- FINAL MODELS ➔ Electra_discriminator / all-MiniLM-L6-v2 / And…

- **Why this is awesome?** It **automates insights** from large datasets, helps with **decision-making**, and provides quick, user-friendly product summaries for **smarter shopping**. If it worked properly…

2.- STORY


SENTIMENT CLASSIFIER
OUTPUT:
Customer reviews that get categorized and scored
CATEGORY CLASSIFIER
OUTPUT:
New categories that are simpler and better suited for the text generator
TEXT GENERATOR
That will use the inputs from the sentiment and category classifier
OUTPUT:
The sentiment score and the new categories will classify the best products of each category and write and article

2.- STORY



SENTIMENT CLASSIFIER
OUTPUT:
Customer reviews that categorized and scored
CATEGORY CLASSIFIER
OUTPUT:
New categories that are better suited for the text generator
TEXT GENERATOR
That will use the inputs from the sentiment and category classifier
OUTPUT:
The sentiment score and the new categories will classify the best products of each category and write and article

2.- STORY

REALITY

SENTIMENT CLASSIFIER

OUTPUT:

Customer reviews that categorized and scored

CATEGORY CLASSIFIER

OUTPUT:

New categories that are better suited for the text generator

The amazon echo show alexa-enabled bluetooth speaker with 7" screen has received some criticism from customers. Here's why they don't recommend it: not ery ancy ut not ow If ou ave am Am Am Am Am Am Am...

3.- STORY

Review: This is the worst product ever. Do not buy it. Predicted Sentiment: **negative**

Review: Not a bad product but not great either Predicted Sentiment: **positive**

Review: Im in love with this tablet. Predicted Sentiment: **positive**

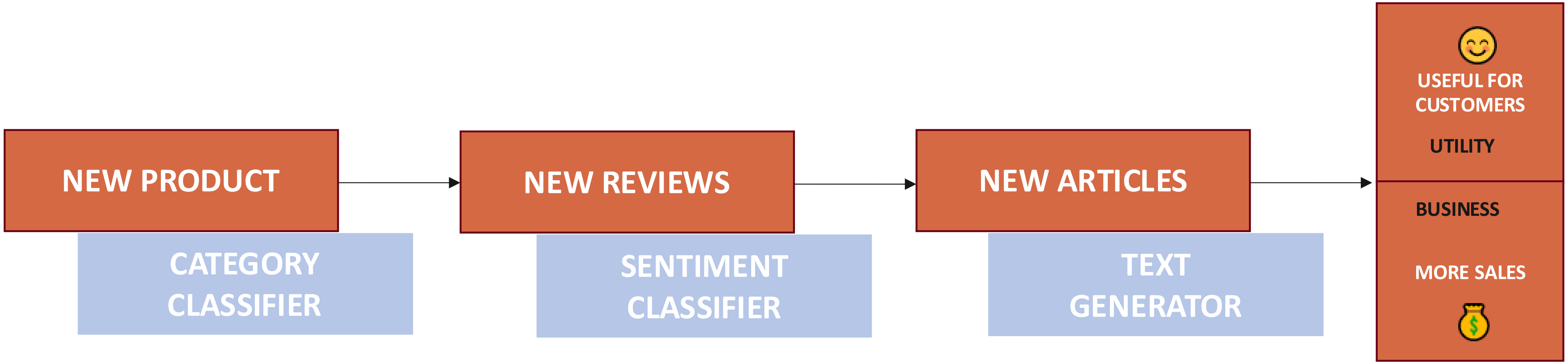
Review: Just an ok product, not good but not bad. Predicted Sentiment: **neutral**

Review: I think this is not a good product at all. Predicted Sentiment: **negative**

Review: I´m going to ask for a refund, I´m not satisfied. Predicted Sentiment: **negative**

4.- INTRODUCTION

- If it worked, it would have ben useful to get a quick overview of a best product of a category without having to read tons of reviews.
- This solution is implemented through NLP methodology, which combines **sentiment analysis**, **category grouping**, and **text generation** using transformer models to classify reviews and generate product summaries.



5.- PIPELINE - DATA

DATASET

Sparse categories

- Missing values
- Useless columns
- Sparse categories
- Unbalanced

SENTIMENT CLASSIFIER

- Lowercasing
- Punctuation
- Oversampling
- Naan
- Stemming
- Lemmatizing
- **TOKENIZATION**

CATEGORY CLASSIFIER

- Lowercasing
- Punctuation
- Oversampling
- Naan
- Stemming
- Lemmatizing
- **TOKENIZATION**

TEXT GENERATOR

- new_categories
- sentiment_score

Implemented
Tried but worked worse
Tried but didn´t work

5.- PIPELINE - TECHNIQUES

DATASET

SENTIMENT CLASSIFIER

- 3 LLMs
- Fine tuning

CATEGORY CLASSIFIER

- Kmeans
- LDA
- LoReg
- 2 LLMs
- Fine tuning
- Manual labeling

TEXT GENERATOR

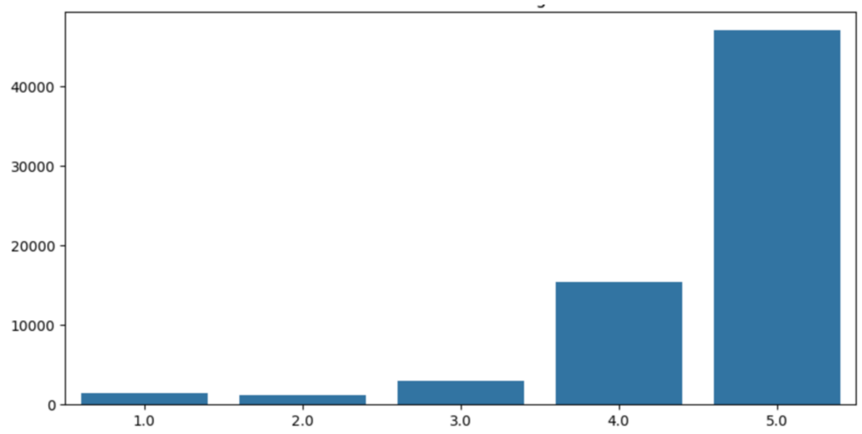
- new_categories
- sentiment_score

6.- SENTIMENT ANALSYSIS MODEL

- Preprocessing
  - Selectively remove Naan, lowercase, tried oversampling for negative feeling but was impossible.

- LLMs:
  - **BERT** ➔ High Accuracy, but really heavy on computing.
  - **DistilBert** ➔ A bit faster than BERT but failed in accuracy but much faster.
  - **Electra_Discriminator** ➔ Good Accuracy (almost like BERT) but much faster.

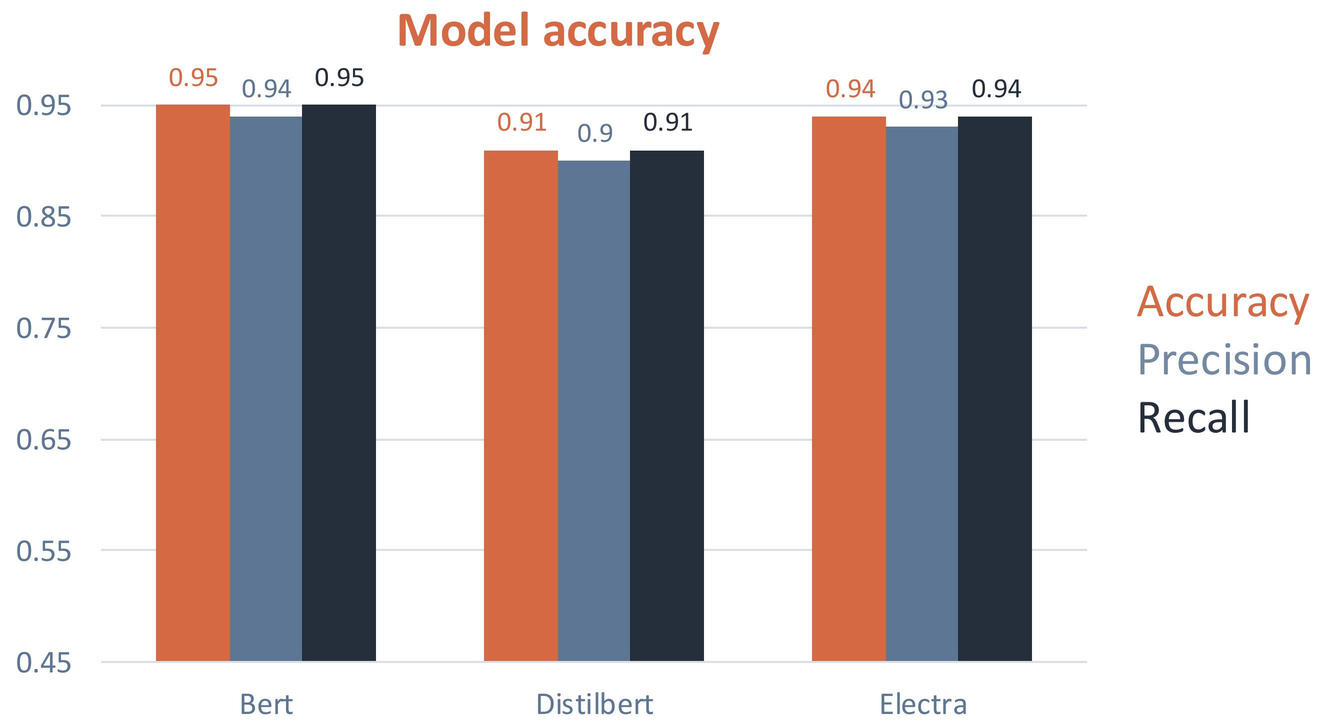
- Implementation:
  - Trained over review_text for fine tunning and used ratings as a validations dataset. And fine tuned the over training. 2 Epochs, drop & learning rate



6.- SENTIMENT ANALSYSIS MODEL

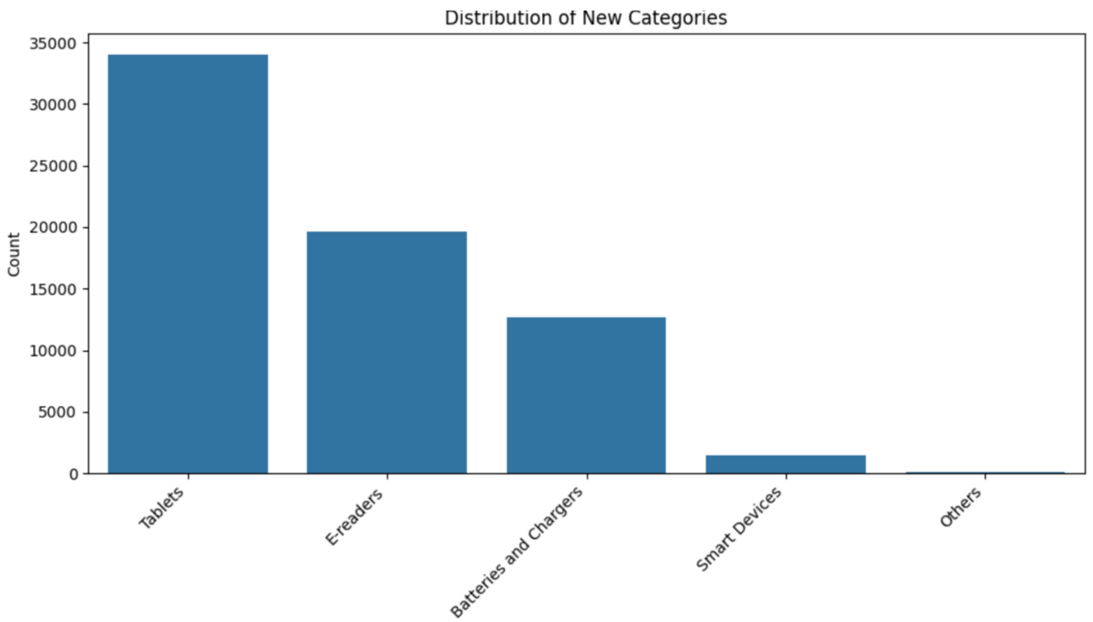
EVALUATION

- Evaluation:
  - Inference with examples out of the dataset and random samples of the dataset. Classic evaluation metrics F1, accuracy etc.



7.- CATEGORY CLASSIFICATION MODEL

- Preprocessing
  - Tokenize data, configurated max length and batch size.
  - Tried lemmatization and stemming but didn´t work.
  - Everything started to work after I used custom ‘manual’ mapping on the categories column.

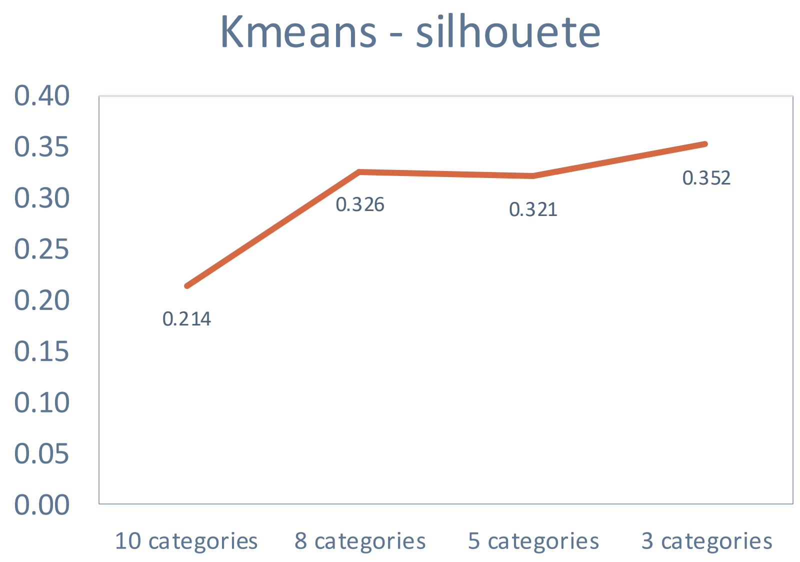


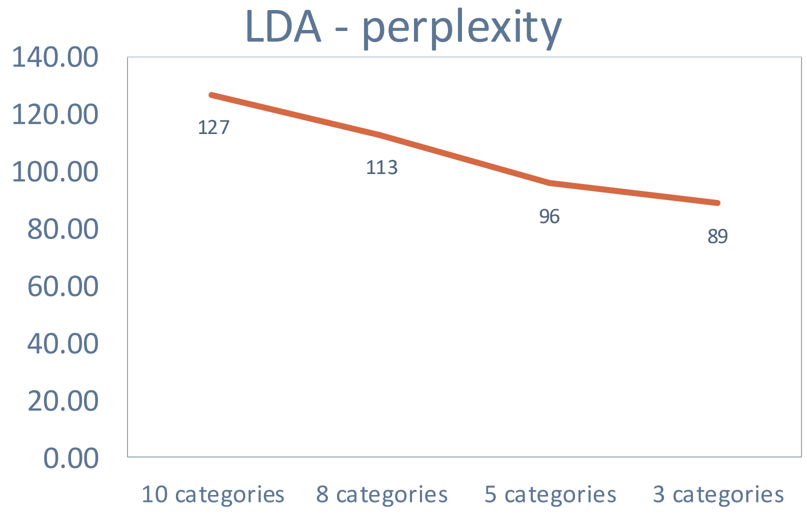
- LLMs:
  - **Electra_Discriminator** ➔ Didn´t performed as well as with sentiment analysis specially because I used it with the clustering technique that didn´t worked very well.
  - **all-MiniLM-L6-v2** ➔ More or the same in terms of speed and Accuracy.

- Implementation and fine tunning:
  - Same approach that with sentiment but using custom mapping instead of ratings.

7.- CATEGORY CLASSIFICATION MODEL

EVALUATION





- Evaluation:
  - Inference with random samples of the dataset of the name column. **Classic** evaluation metrics F1, accuracy etc.

7.- CATEGORY CLASSIFICATION MODEL

EVALUATION

Product Name: Fire Tablet, 7 Display, Wi-Fi, 8 GB - Includes Special Offers, Magenta
Predicted category: Tablets 👍

Product Name: AmazonBasics AA Performance Alkaline Batteries (48 Count) - Packaging May Vary
Predicted category: Batteries and Chargers 👍

Product Name: Amazon Fire Tv,,, Amazon Fire Tv,,, Predicted category: Home Electronics 👍

Product Name: AmazonBasics AAA Performance Alkaline Batteries (36 Count)
Predicted category: Batteries and Chargers 👍

Product Name: Amazon Kindle Paperwhite - eBook reader - 4 GB - 6 monochrome Paperwhite - touchscreen - Wi-Fi - black,,,
Predicted category: E-readers 👍

8.- TEXT GENERATOR

- Preprocessing
  - Created a new dateset based on the “new_categories” and ”sentiment”, in an ideal a would have implemented nother column with sentiment score but had Colab problems.

- Implementation:
  - **TE-RRI-BLE.** Left it to focus on the two previous models.

8.- TEXT GENERATOR

- GPT 2 Says:

…………..

- T5 Small Says:

"not ery ancy ut not ow If ou ave am Am Am Am Am Am Am…"

9.- TAKEAWAYS

- Need better **planning** from my side.
- Implement **sooner manual inference** to check the results as classical metrics could “lie”.
- Don´t try many things in big projects **grab one and make it work.**
- Define **clearer goals**.
- Use **cloud computing resources** wisely (spend 34€ this week in Colab).
- **Save** frequently.

Thank you