

The background is a dark, atmospheric landscape at night. A large, full moon hangs in the sky, casting a soft glow. To the right, a large, gnarled tree with bare branches stands prominently. The ground is rocky and uneven, with some sparse vegetation. The overall mood is mysterious and eerie.

The Incredible Mystery of the Great Theft of Time

Marina Castillo Ariza

The Case

Two models:

- 1 Classifier
- 1 Clustering

They work! But very few iterations,
no summarizer, no deployment...

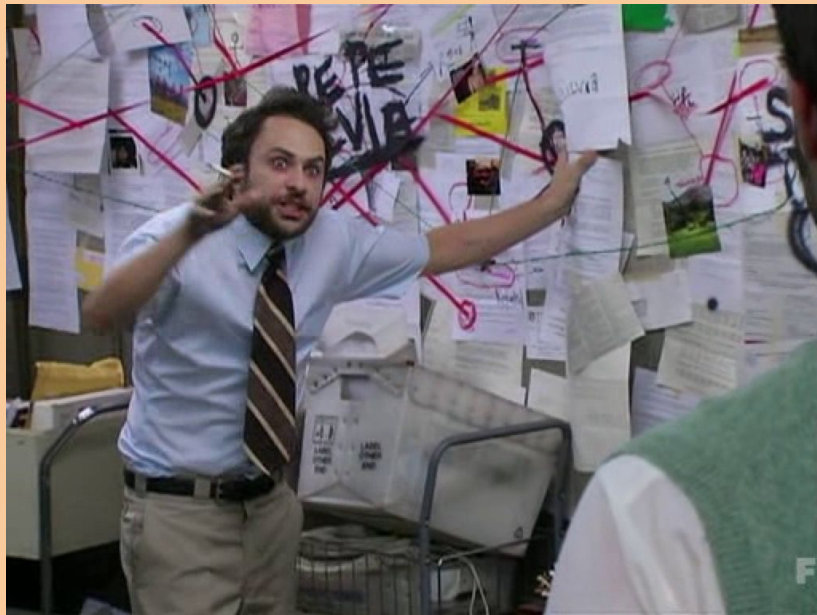
What happened during this week?!



It all started one Monday... Last Monday 😊💧

- Planning for the week
- Choosing models to try
- Even feeling optimistic!

Let's trace back...



Initial clues

1. Data Collection & Preparation:

- 34,000 Amazon reviews: **Unbalanced**

2. Preprocessing:

- Text cleaning:
 - a. Missing values
 - b. For clustering, only 'name' column
- New features: Mapped star ratings for **sentiment labels**.

3. Embedding Generation:

- Utilized DistilBERT:

text data >>> numerical vectors.

4. Modeling:

- **Clustering:** K-Means to categorize products.
- **Classification:** Trained DistilBERT-based classifier for sentiment analysis.

5. Integration:

- Merged both clustering and classifying at the end

Let's look into more detail...



Model 1, or should I say 2? Anyway, **Clustering**

- Embedding Generation with **DistilBERT**
- Clustering with **K-Means**

Initial Clusters

5 clusters + Silhouette Score of **0.5883**.

1st Iteration

6 clusters + Silhouette Score of **0.7974**

!!! Visual Inspection: **weird category**

“Tablets & Bluetooth devices” mix

2nd Iteration

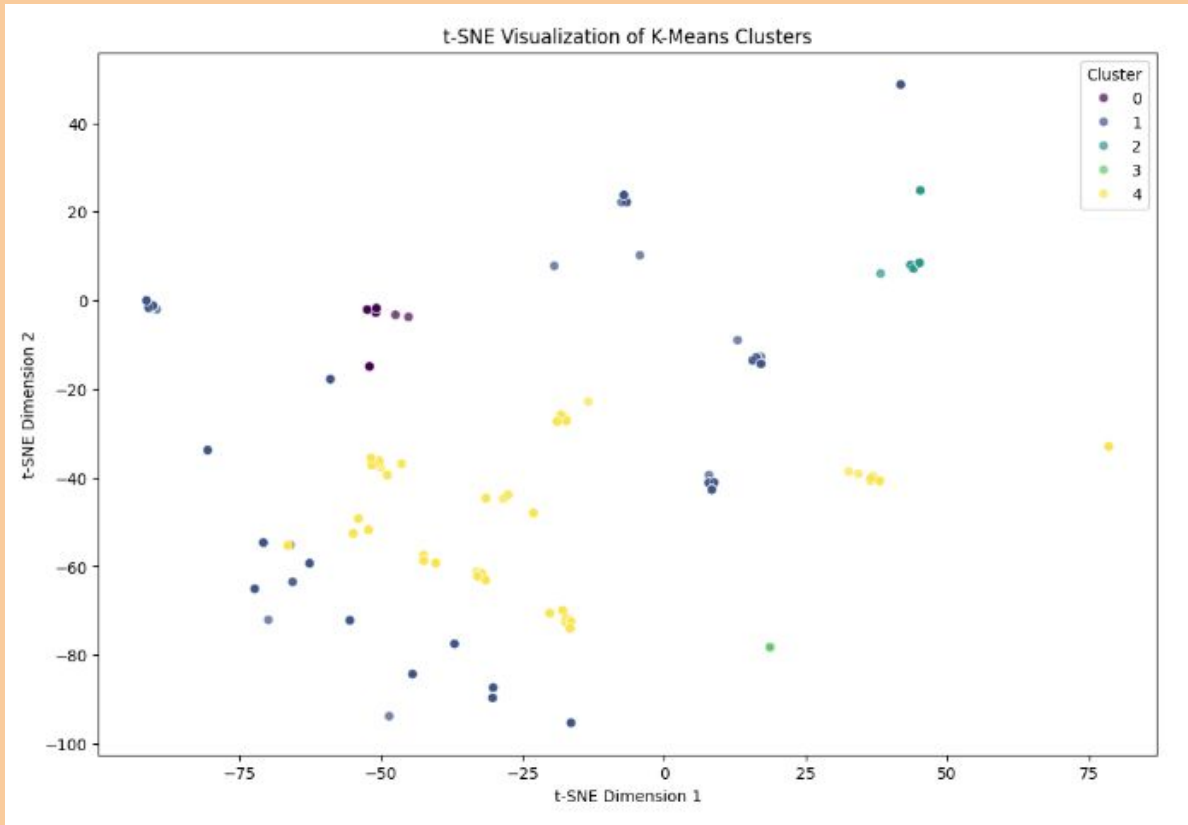
5 clusters + Silhouette Score improved to **0.7320**

Solutions

Employ PCA before t-SNE:

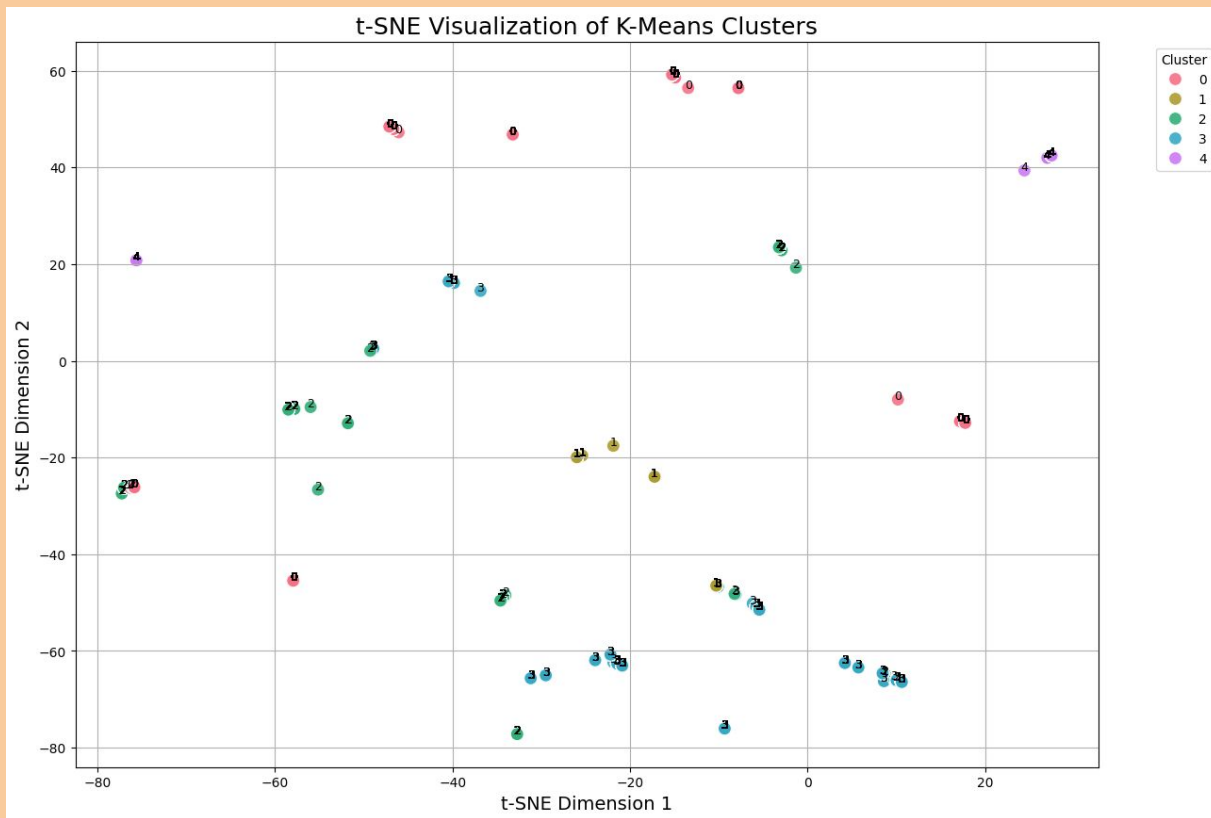
- **Simplifies Data:**
Makes the data less complex.
- **Improves t-SNE Performance:**
Helps t-SNE create clearer and more distinct groups.

Clustering Model: Evaluation



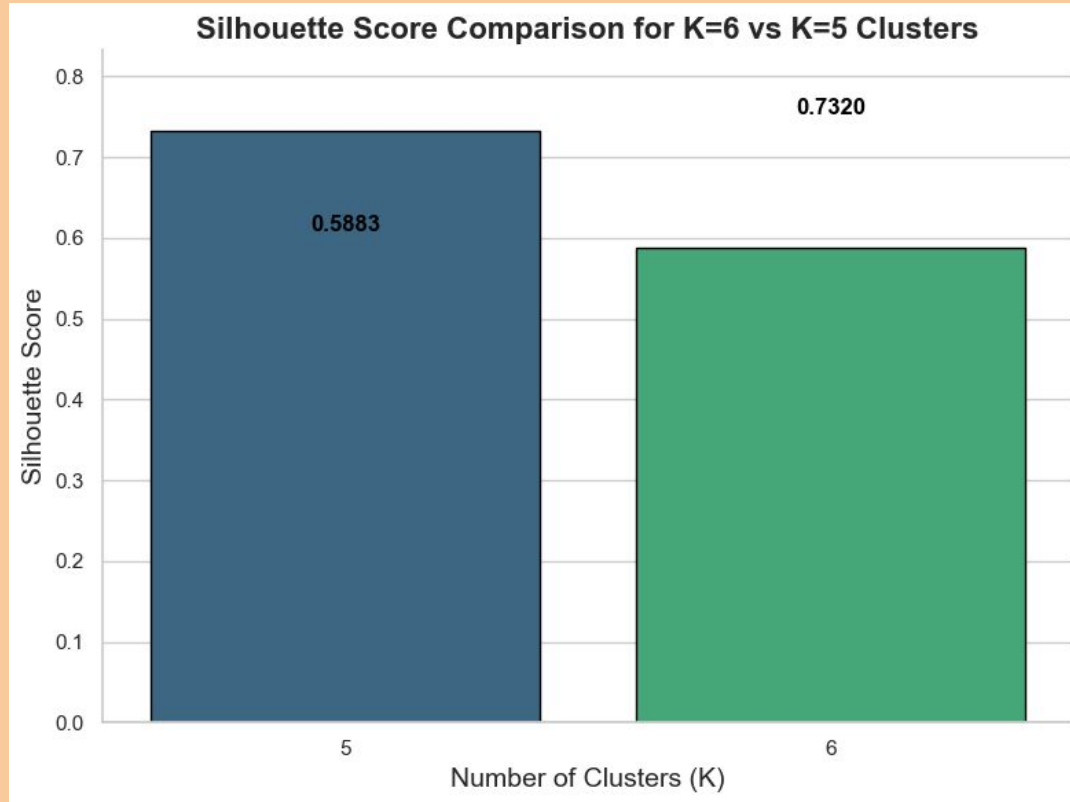
Silhouette Score
of **0.5883**

Clustering Model: Evaluation



Silhouette Score of
0.7320

Clustering Model: Evaluation



Clustering Model: Evaluation



Tablets & E-Readers:

- "Fire Tablet, 7 Display, Wi-Fi, 8 GB - Includes Special Offers, Magenta"
- "Kindle Voyage E-reader, 6 High-Resolution Display (300 ppi) with Adaptive Built-in Light"
- "Amazon Kindle Paperwhite - eBook reader - 4 GB - 6 monochrome Paperwhite"
- "All-New Fire HD 8 Tablet, 8 HD Display, Wi-Fi, 32 GB - Includes Special Offers, Magenta"



Smart Speakers:

- "Echo (White)"
- "Echo (White)"
- "Amazon Echo Dot (4th Gen) - Smart Speaker with Alexa"
- "Echo Show 8 - Smart Display with Alexa"

Clustering Model: Evaluation

Kids Edition Tablets:

- **"Amazon Kindle Paperwhite - eBook reader - 4 GB - 6 monochrome Paperwhite"**
- "Fire Kids Edition Tablet, 7 Display, Wi-Fi, 16 GB - Green Kid-Proof Case"
- "Fire Kids Edition Tablet, 7 Display, Wi-Fi, 32 GB - Blue Kid-Proof Case"
- "Fire 7 Kids Tablet - 16 GB - Includes 1 Year of Amazon Kids+"

Entertainment Devices:

- "Amazon Fire TV Stick - Streaming Device"
- "Amazon Fire TV Cube - Hands-Free Streaming Device"
- "All-New Fire HD 10 Tablet, 10.1" - Wi-Fi, 32 GB"
- "Fire TV Stick 4K - Streaming Device with Alexa"

Bluetooth & Charging Accessories:

- **"All-New Fire HD 8 Tablet, 8 HD Display, Wi-Fi, 16 GB - Includes Special Offers, Magenta"**
- "Amazon Fire HD 10 Tablet - 10.1", 64 GB"
- "Amazon 5W USB Official OEM Charger for Fire Tablets"
- "Amazon - Amazon Tap Portable Bluetooth and Wi-Fi Speaker - Black"

Classifying Model

Initial Attempt:

- Started with BERT, took ages.
- Pretrained RoBERTa unsupervised + K-Means
- **Challenges:**
 - chose **wrong** model: only positive and negative AND **no ground truths**

Change to Supervised Learning:

- Mapped star ratings to sentiments:
 - Ratings ≥ 4 : Positive
 - Rating = 3: Neutral
 - Ratings < 3 : Negative
- Added a sentiment column in the dataset.

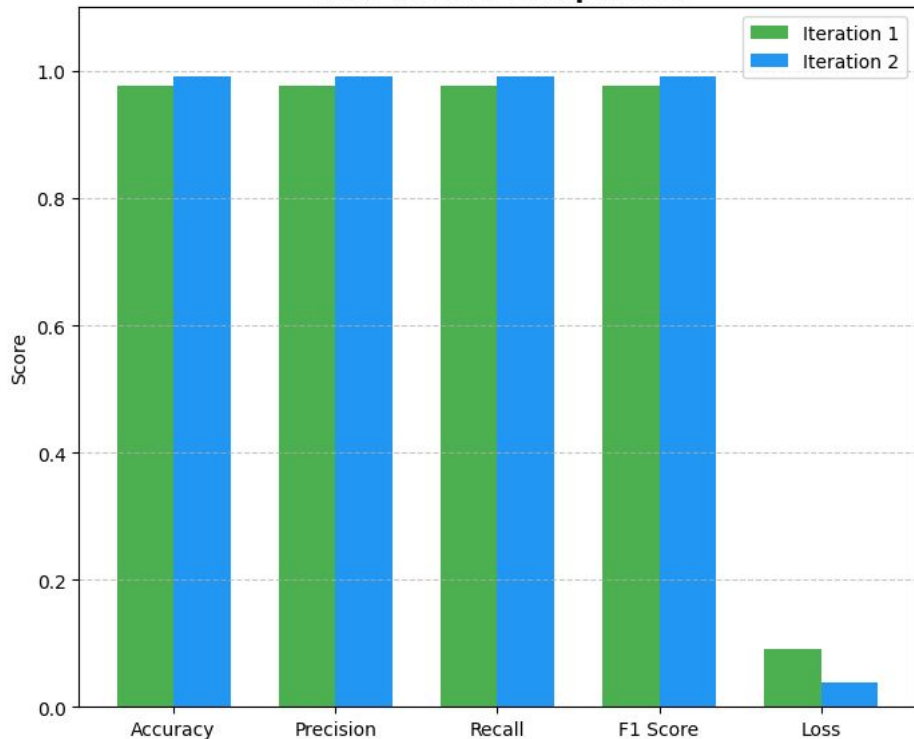
Classifying Model: Evaluation

- **Higher Train Accuracy** in the 2nd iteration: overfitting?
- **Test Accuracy** consistent: good generalization!
- **Improved Test F1 Score** in the second iteration
- **First iteration** best balance between Train and Test.

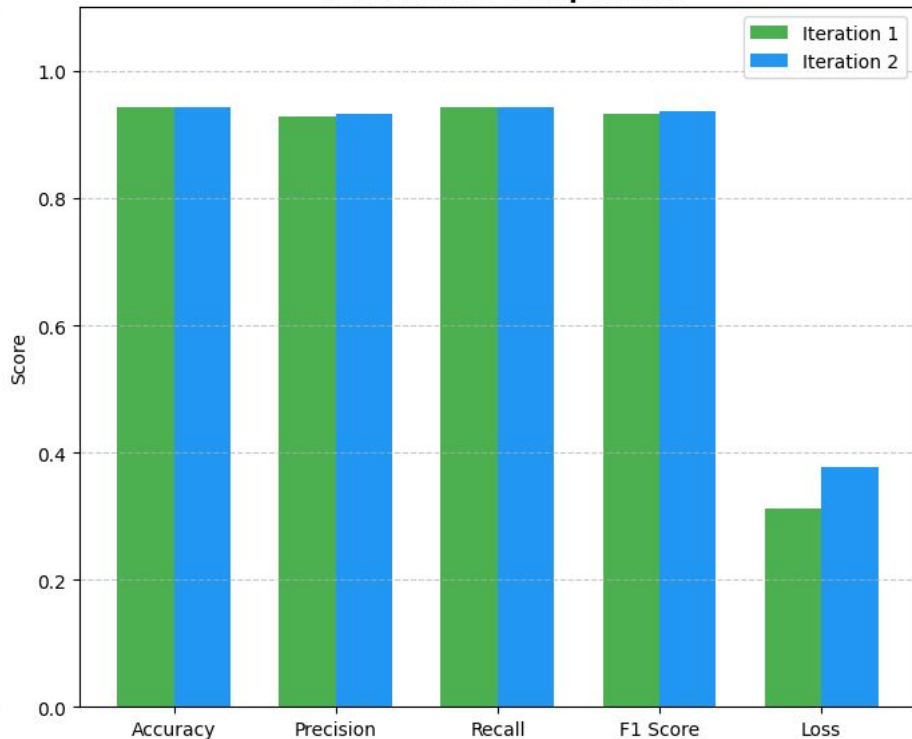
Metric	lr=5e-5	lr=2e-5	Change
Train Accuracy	96.89%	99.11%	↑ +2.22%
Train Precision	96.55%	96.55%	→ No Change
Train Recall	96.89%	96.89%	→ No Change
Train F1 Score	96.49%	96.49%	→ No Change
Train Loss	0.1044	0.1044	→ No Change
Test Accuracy	94.26%	94.25%	→ -0.01%
Test Precision	92.79%	92.79%	→ No Change
Test Recall	94.26%	94.26%	→ No Change
Test F1 Score	93.21%	93.67%	↑ +0.46%
Test Loss	0.1983	0.1983	→ No Change

Classifying Model: Evaluation

Train Metrics Comparison



Test Metrics Comparison



Classifying Model: Evaluation

Positive Reviews: 😄

1. ***"so far so good,,,, wife very happy. misses her old one...."***
2. *"Its awesome. I love it. The price is worth it to buy it."*
3. *"Easy to use and install Love it I get to stream almost everything"*
4. *"This is an awesome kindle easy to transport and love the clear screen."*

Neutral Reviews: 😐

1. *"not as smart as the commercials show. cant answer a lot of questions"*
2. *"Easy to use.Easy to connect with cellphone and tablet to play YouTube. The product said it will upscale to 4k quality but the picture show on the TV is the same as regular product without 4K."*
3. ***"It is a good tablet for beginner kids and great for the price"***
4. *"Works well for the purpose it was used for not fancy but good quality tablet"*

Classifying Model: Evaluation

Negative Reviews: 🤡

1. *"This is a good tablet for children and adults too. However there is an issue with the port and charger cable. We did not even have this tablet 3 weeks before it would barely charge. Took it back to best buy and they said they could not fix it, and imagine the port is not covered under the manufacturer warranty. So now I have a tablet that takes a couple days to charge. Its really frustrating."*
2. *"I'm use to android so this was hard for me but I returned it and got something andriod"*
3. *"You need goooooood wifi for this. If the wifi isn't good like it's in and out this will not work that good. You also need the app to set up alexa. I would only recommend this if you have great wifi"*
4. *"just okay! Product was for my girlfriend. Doesn't appear she uses it much except to steam music. Most of the questions I ask it I get nothing. I like Siri better."*

Perpetrators

Nonstop

Troubleshooting!



Perpetrators

Trusting Chat
GPT too much!



Perpetrators

Time itself!!!



Perpetrators




Ultimately...



Learnings

- Data Label is makes a difference!
- Watch out for Overfitting
- Iterative Approach
- Clustering Challenges
- Model Selection **takes time!**
- Preprocessing has a **high impact**

Learnings

- “Use your time wisely!” 
- Study, refresh, research... 
- Share and ask for help! 

<https://github.com/MarinaJuni/1Mystery2Models.git> 

Thank you for your
attention and your help!

