



# SocioScope - Final Presentation

Ofir Fichman, Pavel Fadeev, Israel Peled



# Project Motivation & Definition

- **The Problem:** Existing ABSA models are optimized for product reviews and fail to interpret abstract socio-economic discourse.
- **The Solution:** A high-fidelity synthetic pipeline (Llama 3.2) utilizing Negative Sampling to master noise filtration.
- **The Goal:** To build a specialized model capable of disentangling complex, multi-faceted public sentiment.



# Project Achievements & Novelty

- **GenAI Data Pipeline:** Developed a "Ground-Truth First" pipeline using Llama 3.2 to generate 10,000+ tweets with 100% label accuracy, simulating realistic discourse via 7 personas and 6 writing styles.
- **Domain Adaptation:** Applied Aspect-Based Sentiment Analysis (ABSA) to complex socio-economic domains like Taxation and Security, moving beyond standard product review datasets.
- **Robust Architecture:** Implemented a multi-aspect, 3-class system that treats "Not Mentioned" as a learned feature, enabling precise noise filtering and independent evaluation of topics within a single tweet.



# Review Methodology

- Created 10,000 samples via Llama 3.2, applying rigorous EDA to prune logical inconsistencies between text and labels.
- Transitioned to a Sentence-Pair format ([Text + Aspect]), utilizing DeBERTa-v3's attention mechanism to specifically target relevant aspects.
- Countered synthetic overfitting by moving to Step-Based Evaluation (every 500 steps) with aggressive Early Stopping.
- Moved beyond Accuracy to Macro F1-Score and ROC AUC to ensure performance on minority classes.



# Example of output

Tweet: 'The taxes are killing us, but at least the streets are safe at night.'

Aspect: Personal Security -> Positive (0.78)

Aspect: Taxation -> Negative (0.72)

Tweet: 'They turned that ugly parking lot into a public park with ramps and access for everyone. Feels like the city finally cares.'

Aspect: Environment -> Positive (0.95)

Tweet: 'I don't feel safe walking home at night, and the local police station is always 'investigating' but never doing anything.'

Aspect: Personal Security -> Negative (0.93)

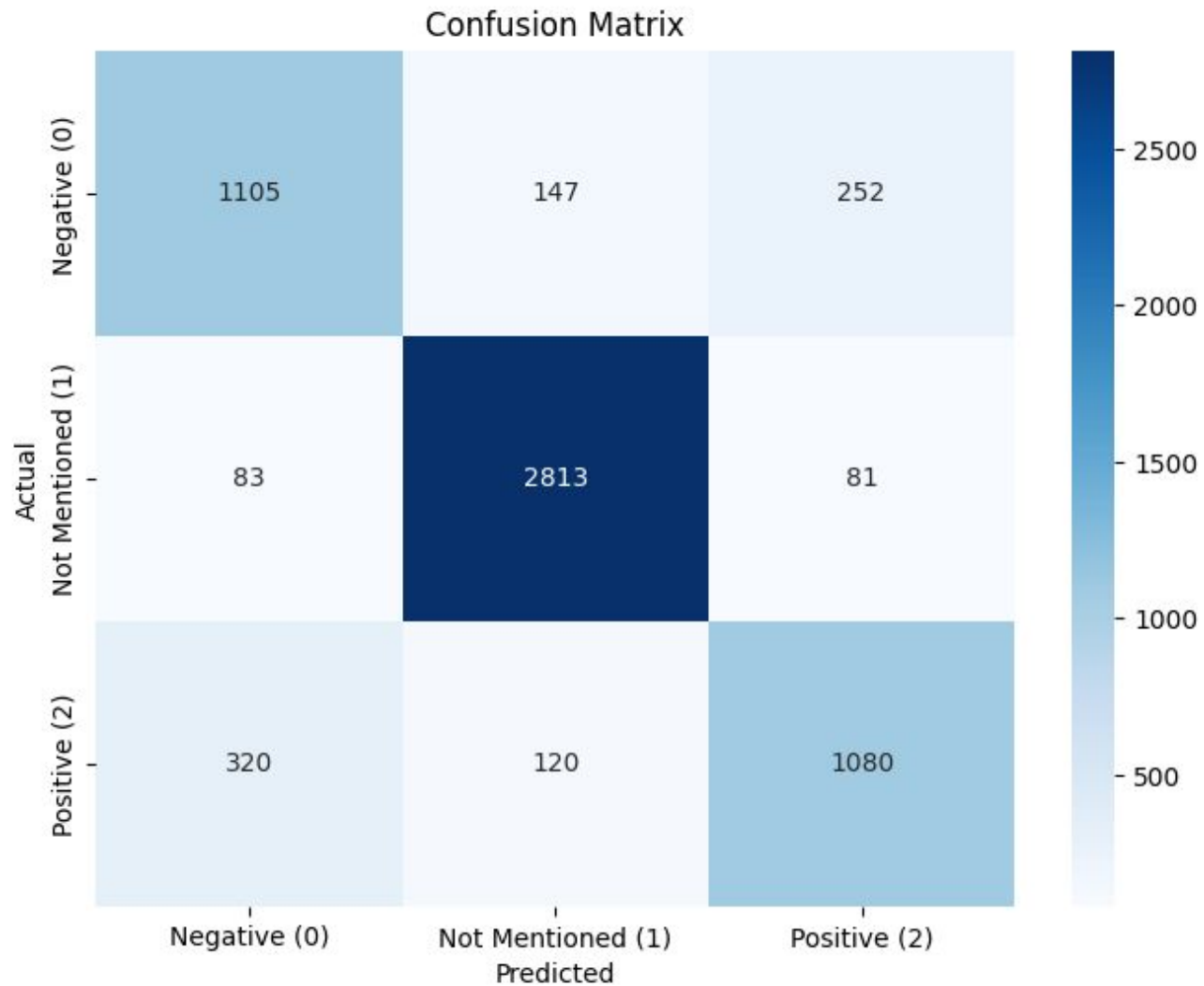


# Results: DeBERTa Model

--- Detailed Classification Report ---

	precision	recall	f1-score	support
Negative (0)	0.73	0.73	0.73	1504
Not Mentioned (1)	0.91	0.94	0.93	2977
Positive (2)	0.76	0.71	0.74	1520
accuracy			0.83	6001
macro avg	0.80	0.80	0.80	6001
weighted avg	0.83	0.83	0.83	6001

# Results: DeBERTa Model





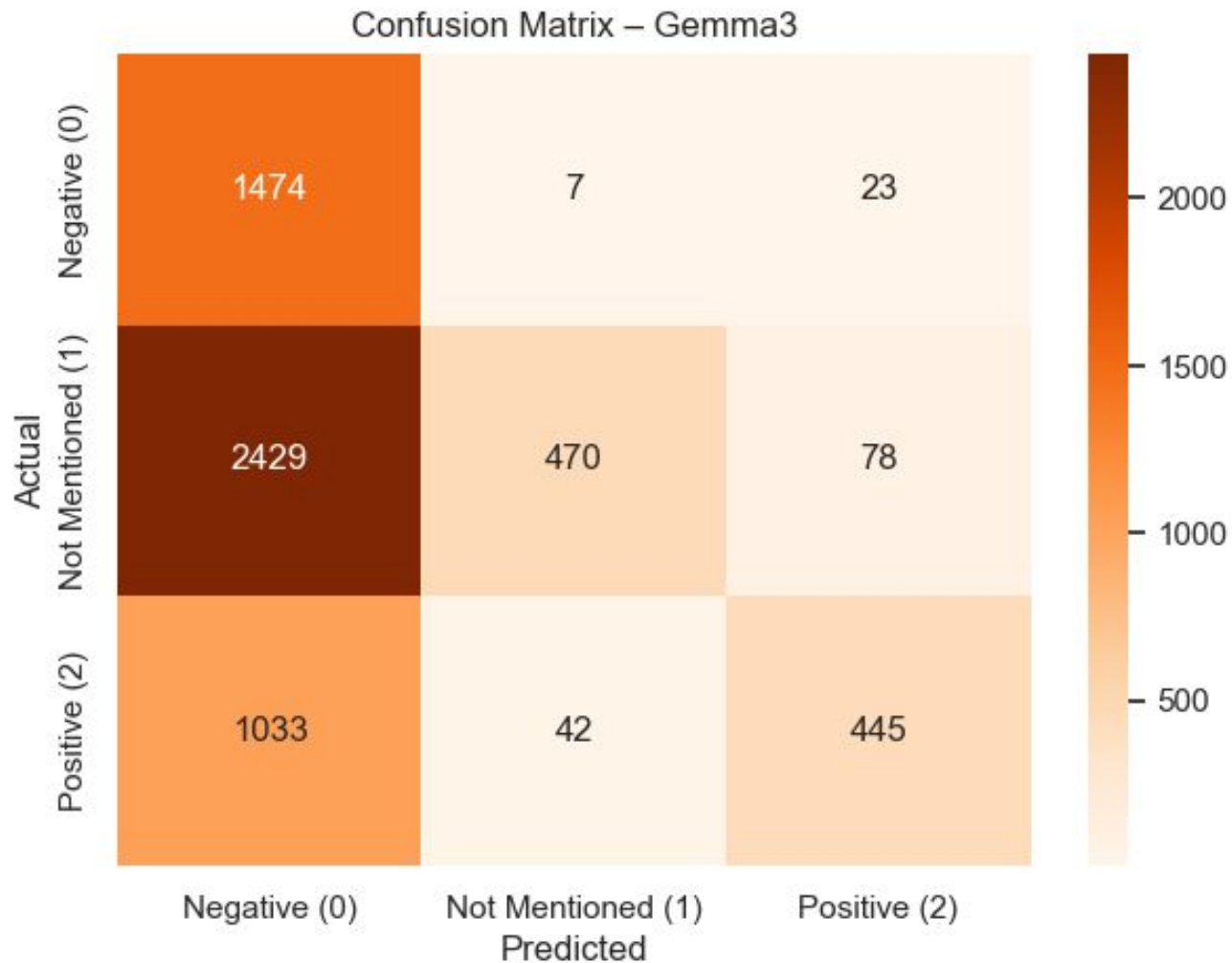
# Results: Gemma3

=== Gemma3: classification report ===

	precision	recall	f1-score	support
Negative (0)	0.30	0.98	0.46	1504
Not Mentioned (1)	0.91	0.16	0.27	2977
Positive (2)	0.82	0.29	0.43	1520
accuracy			0.40	6001
macro avg	0.67	0.48	0.39	6001
weighted avg	0.73	0.40	0.36	6001

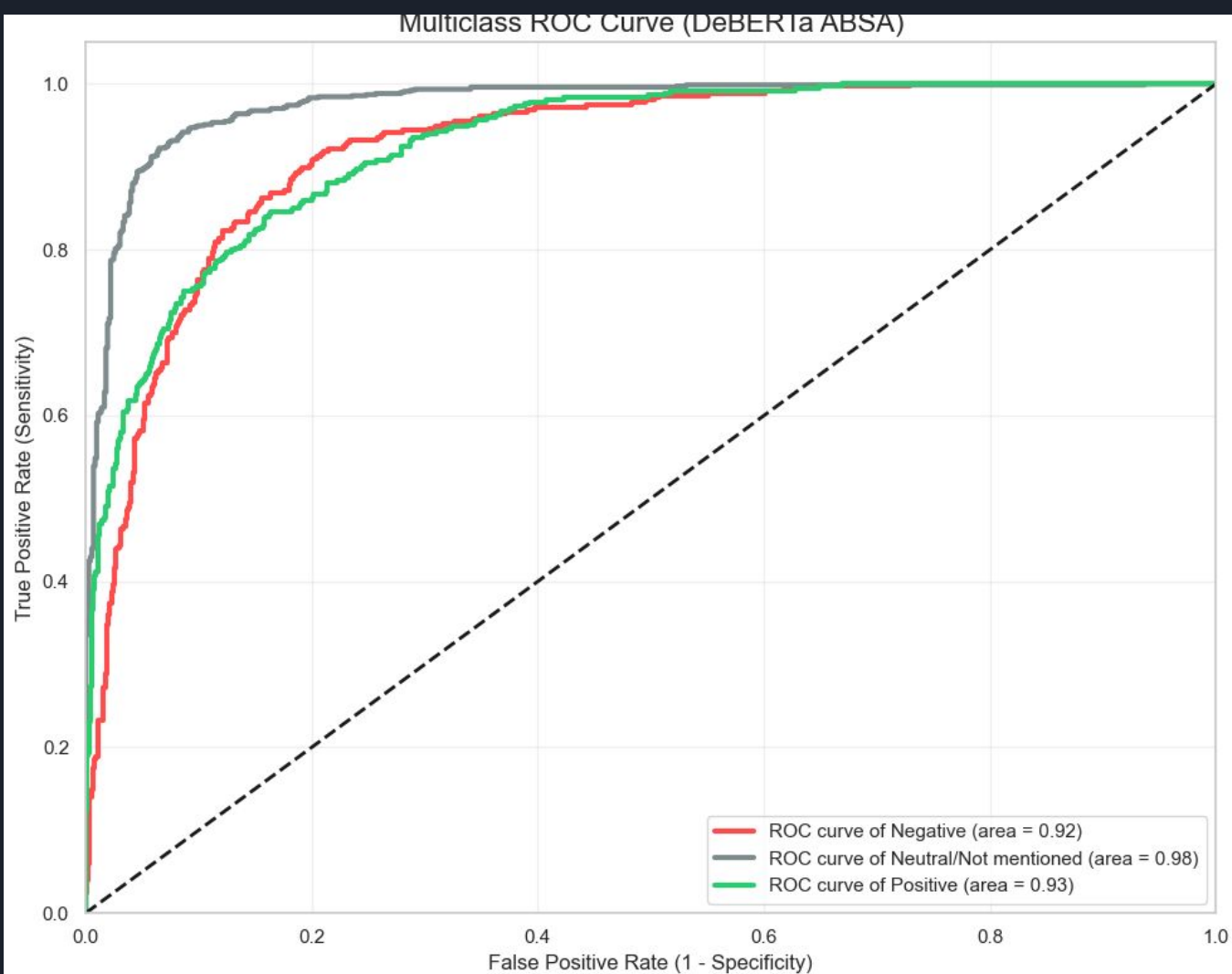


# Results: Gemma3

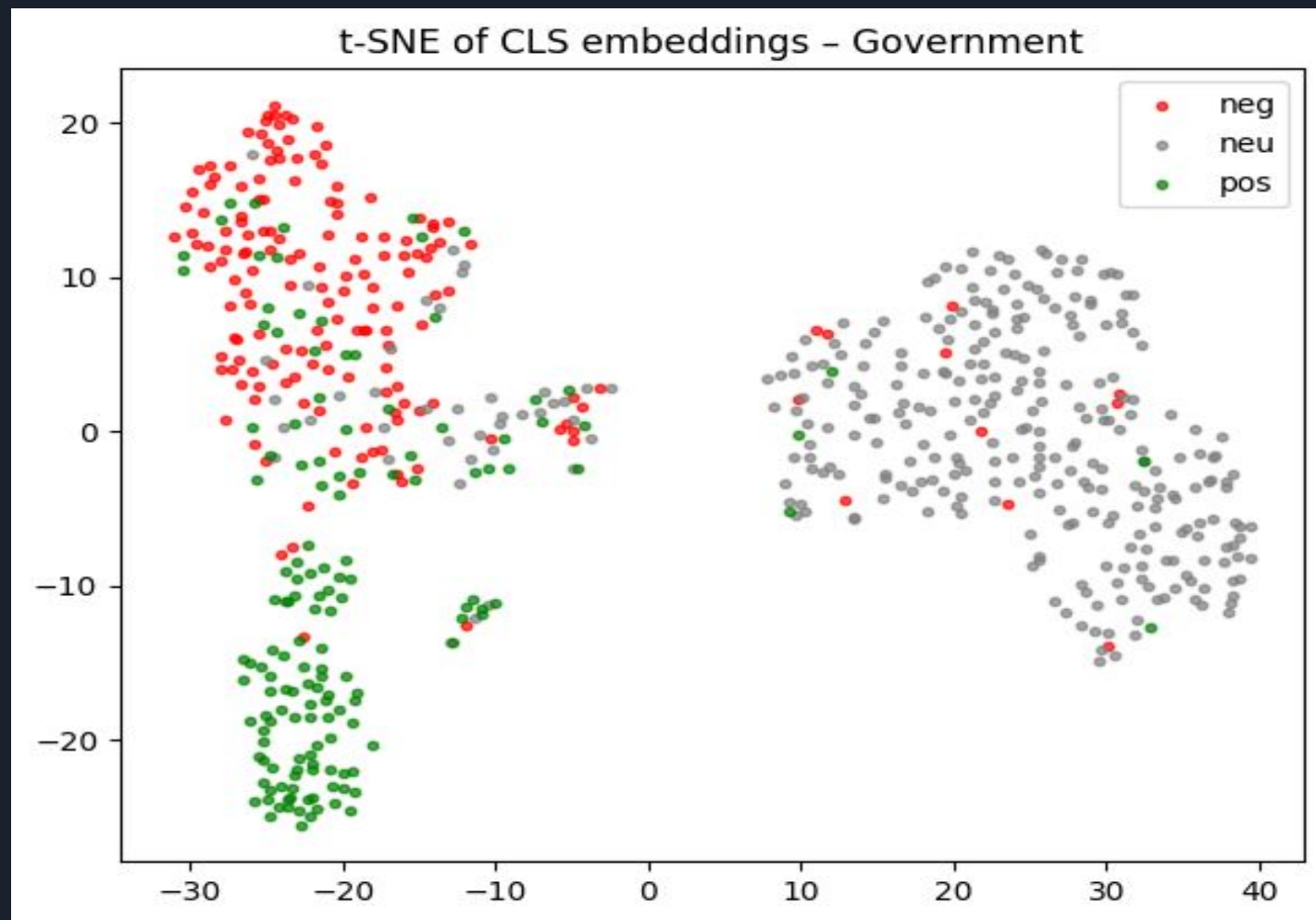




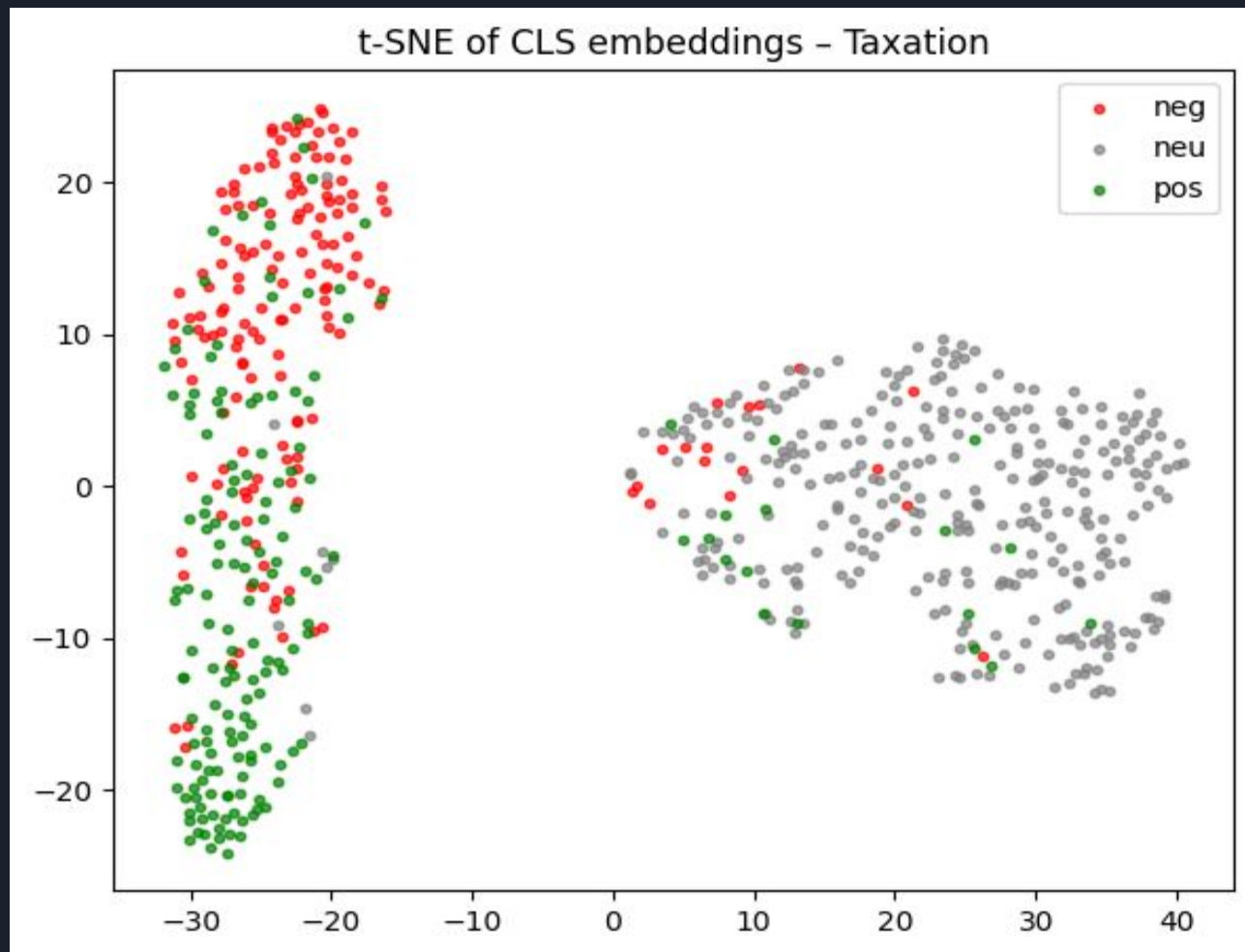
# ROC Curve



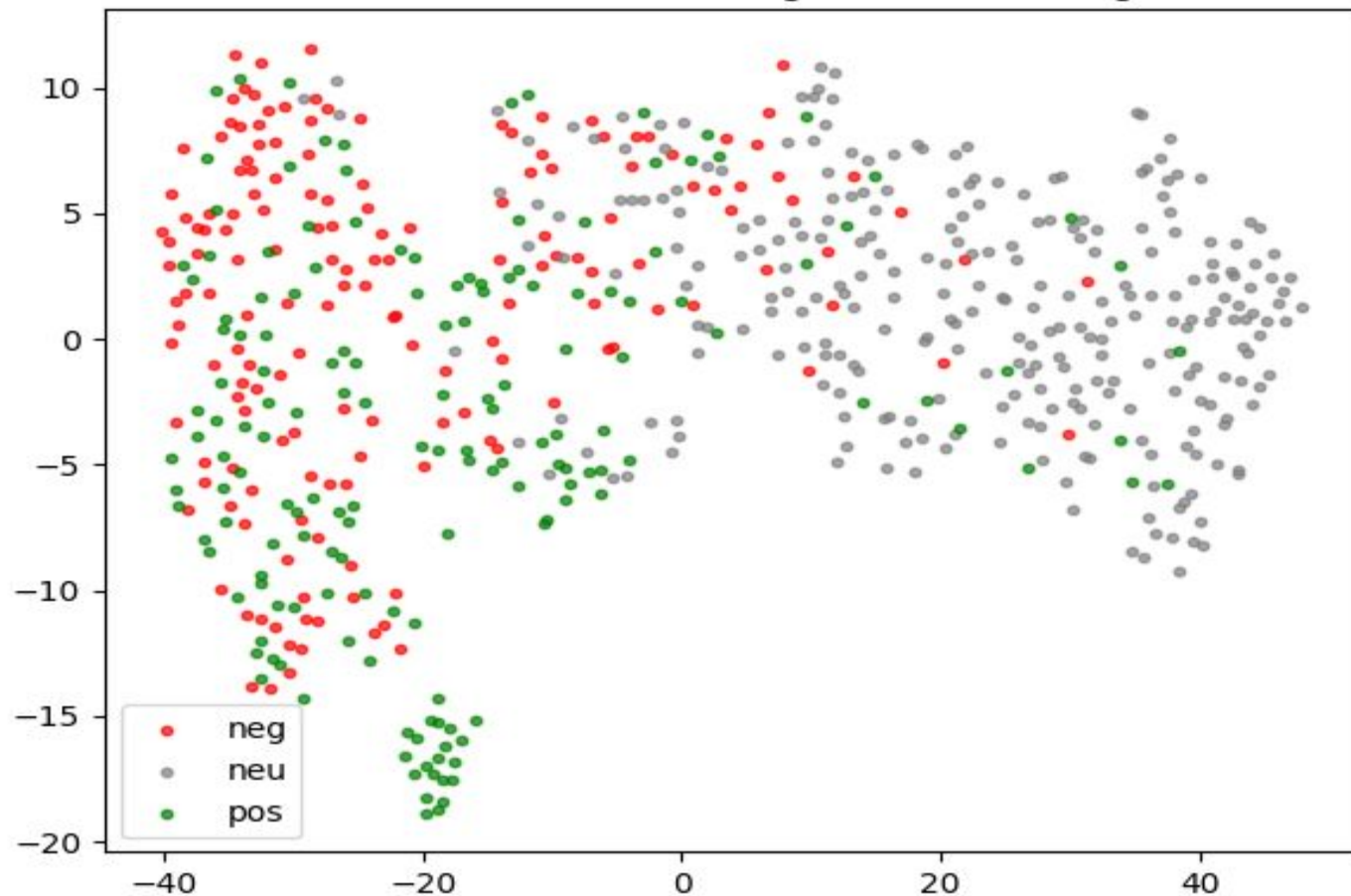
Results:



Results:



t-SNE of CLS embeddings - Cost of Living





# Conclusions & Future Work

- Successfully delivered a robust Socio-Economic ABSA model achieving  $>0.92$  AUC for sentiment and 0.98 AUC for noise filtering.
- Overcame synthetic data overfitting and the sparse-data challenge by implementing aggressive regularization (Weight Decay 0.1) and treating "Not Mentioned" as an active learned class via Negative Sampling.
- Plans to expand the pipeline to multi-lingual support (Hebrew) and deploy the model on live Twitter streams to validate performance against real-world human slang and evolving trends.