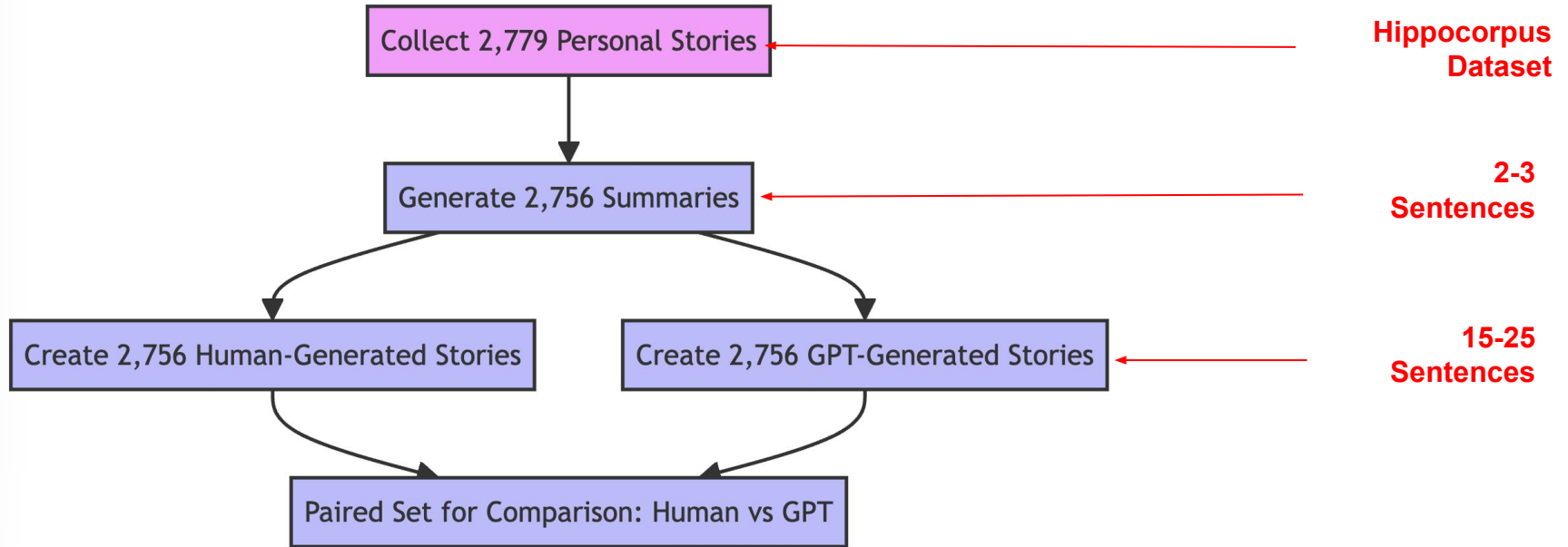


Deciphering the Narrative: Distinguishing Human and GPT Generated Fictitious Stories in the Hippocorpus Dataset

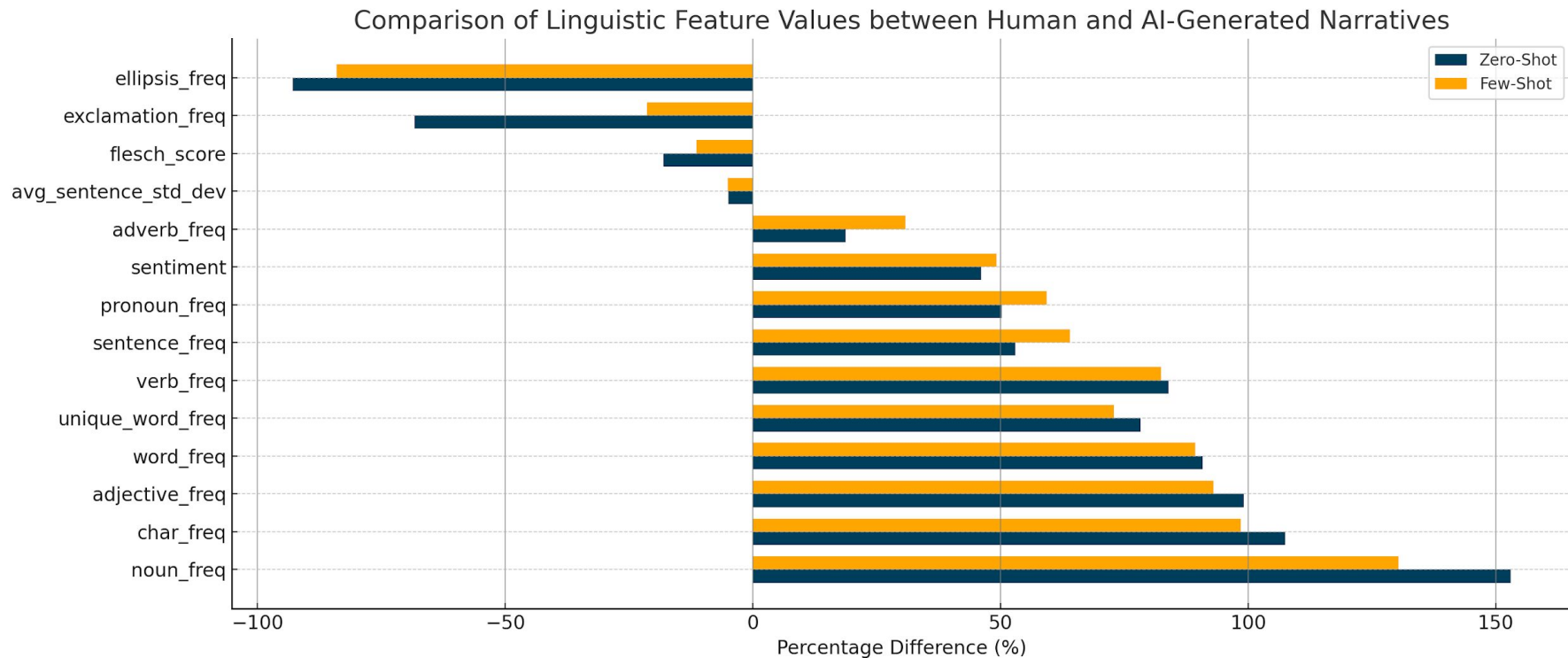
CS598 Fall 2023 Project 5 - Tej Shah

How do human generated and GPT generated stories differ in various aspects?

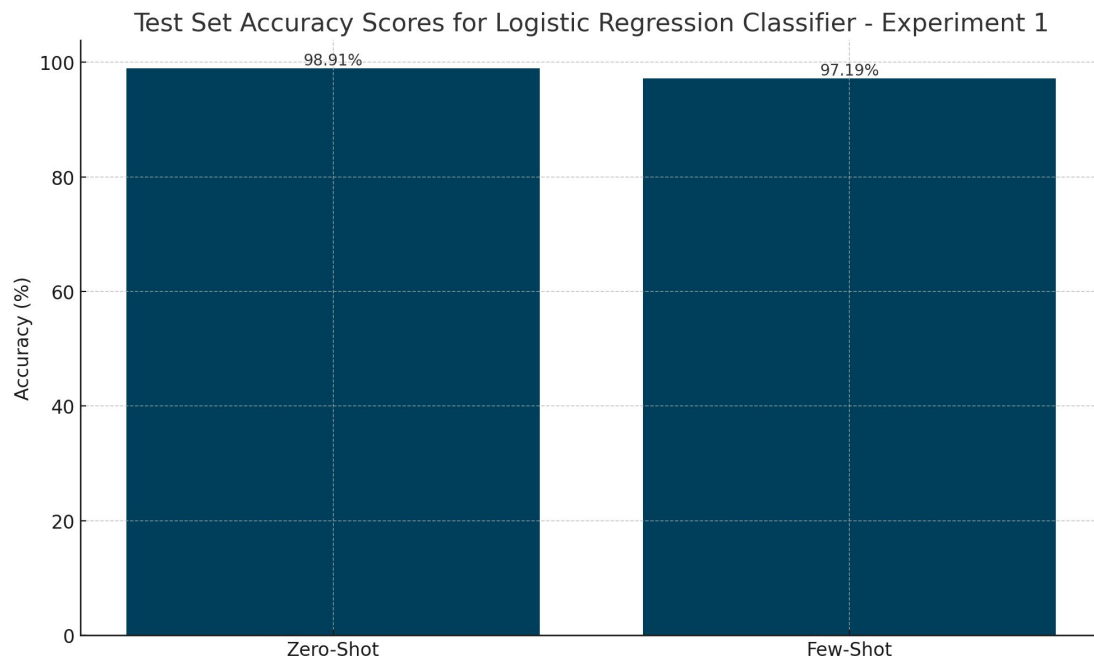
Data Collection



Do linguistic features differentiate
human and GPT-generated stories?

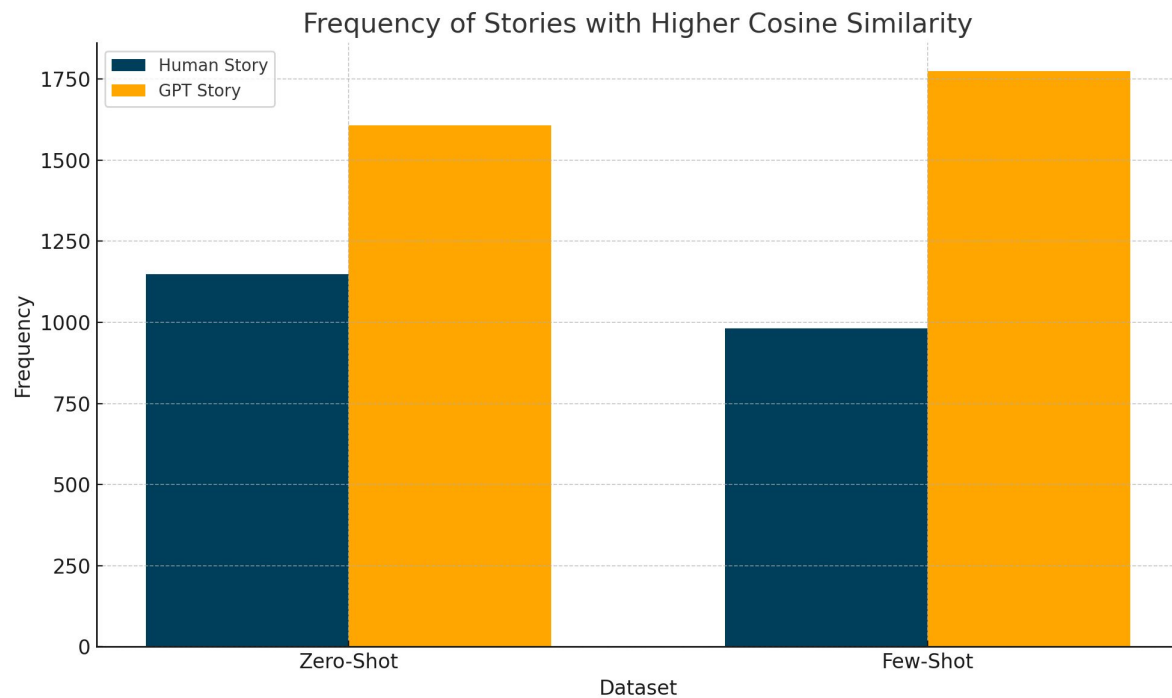


GPT Narratives Showcase Unique Linguistic Traits Compared to Human Stories

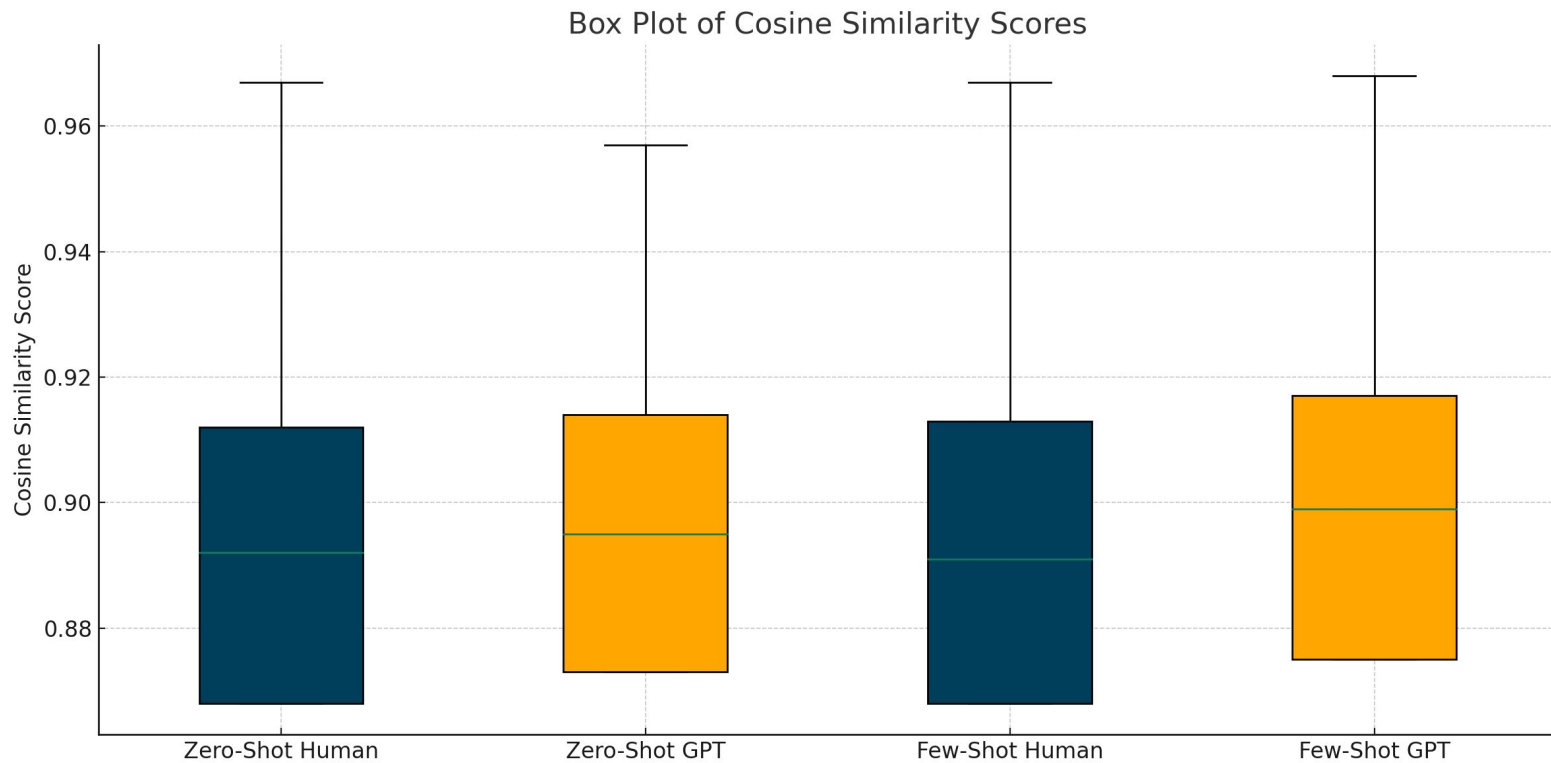


Linguistic Features Predict Story Origin with High Accuracy

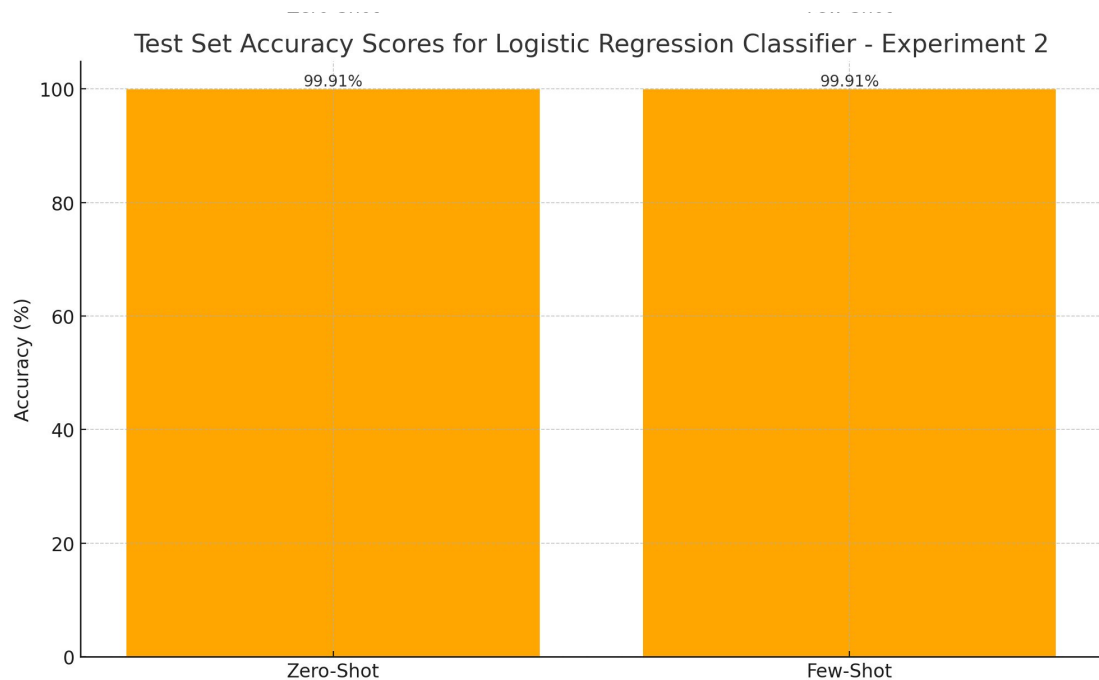
Which story type – human or GPT generated – more accurately reflects the intent of their original summaries?



GPT Stories Triumph in Matching Intended Narrative Summaries

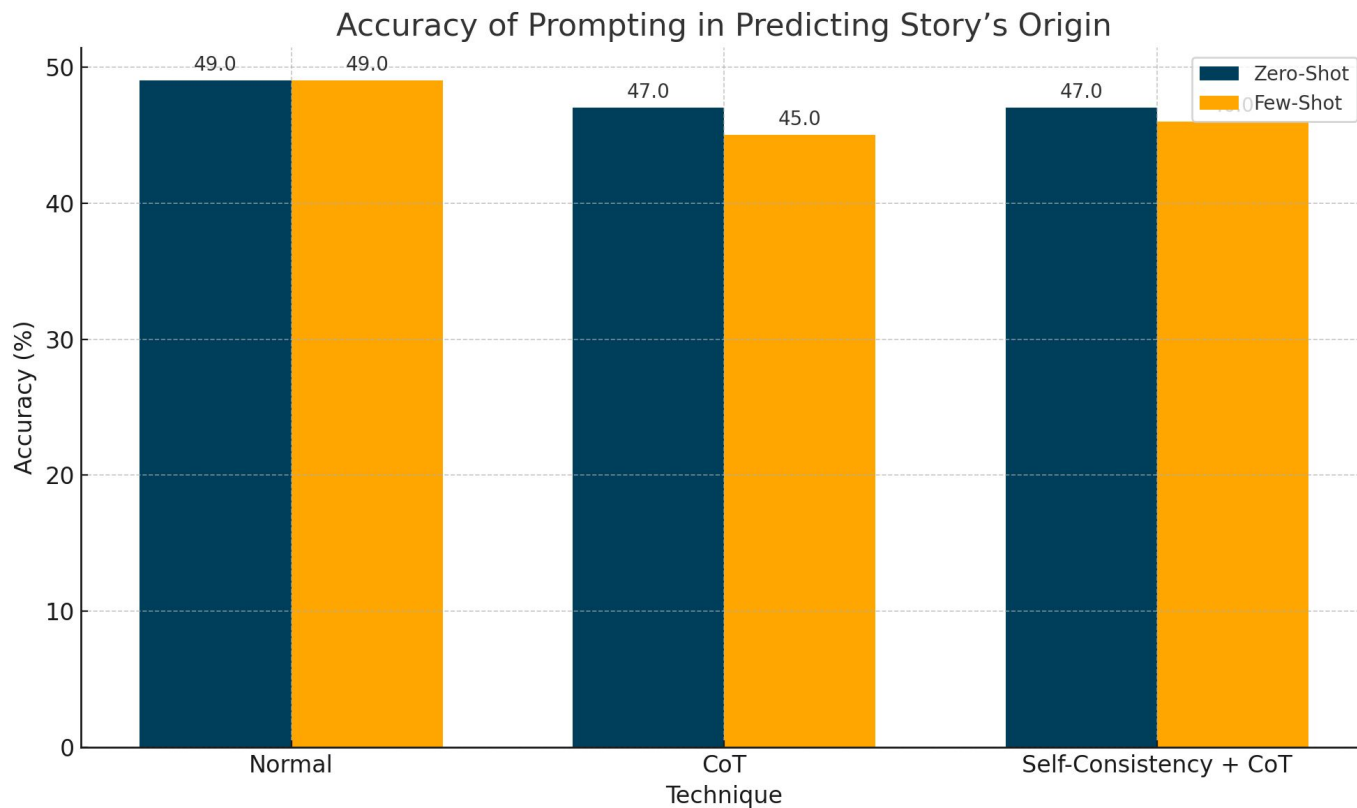


GPT-Generated Stories have Less Variability and Higher Median Similarity



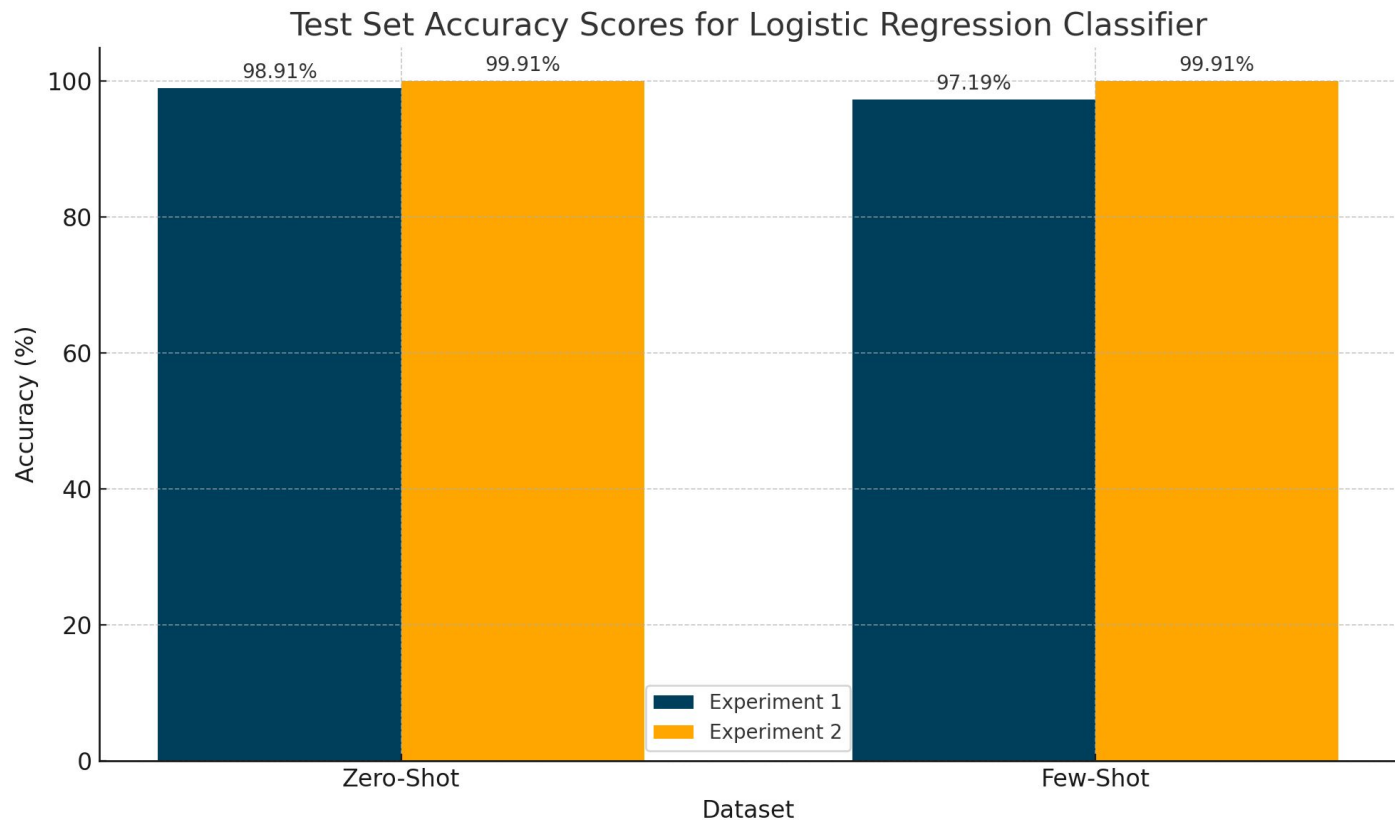
Embedding Analysis: A Clear Marker of Human and AI-Generated Stories

Do different in-context learning prompting techniques affect the ability for **gpt-3.5-turbo** to identify a story's origin?



Prompting Techniques Fall Short of Random Baseline in Origin Prediction

Does incorporating human examples in few-shot generation alter experimental outcomes?



Consistency Across Few-Shot Learning: Human Examples Don't Shift Results

Results

1. GPT and human stories show unique linguistic distributions
2. GPT narratives align closer to human intent, with less variability
3. GPT captures intent better, even in few-shot scenarios with human stories
4. Prompt engineering techniques cannot classify story provenance alone
5. Provenance identification still effective in few-shot learning dataset

Conclusion: Human-Machine Comparison

1. **Distinct Storytelling Approaches:** GPT excels in aligning with prompts and having consistency, whereas human stories are variable and rich in creativity.
2. **Learning and Adaptability:** Few-shot learning boosts GPT's prompt alignment, heightening the divide between human and GPT narrative styles.
3. **Storytelling:** Human narratives are enriched with personal experiences and socio-cultural contexts, offering depth and creativity; GPT, driven by data patterns, produces structurally uniform stories with limited emotional depth.

Contact Me

Email: tej.shah@rutgers.edu

Github: <https://github.com/tejpshah>