SEMANTIC TITLE GENERATION FOR SONG LYRICS WITH AI MODELS AND PROMPT ENGINEERING

Problem Description

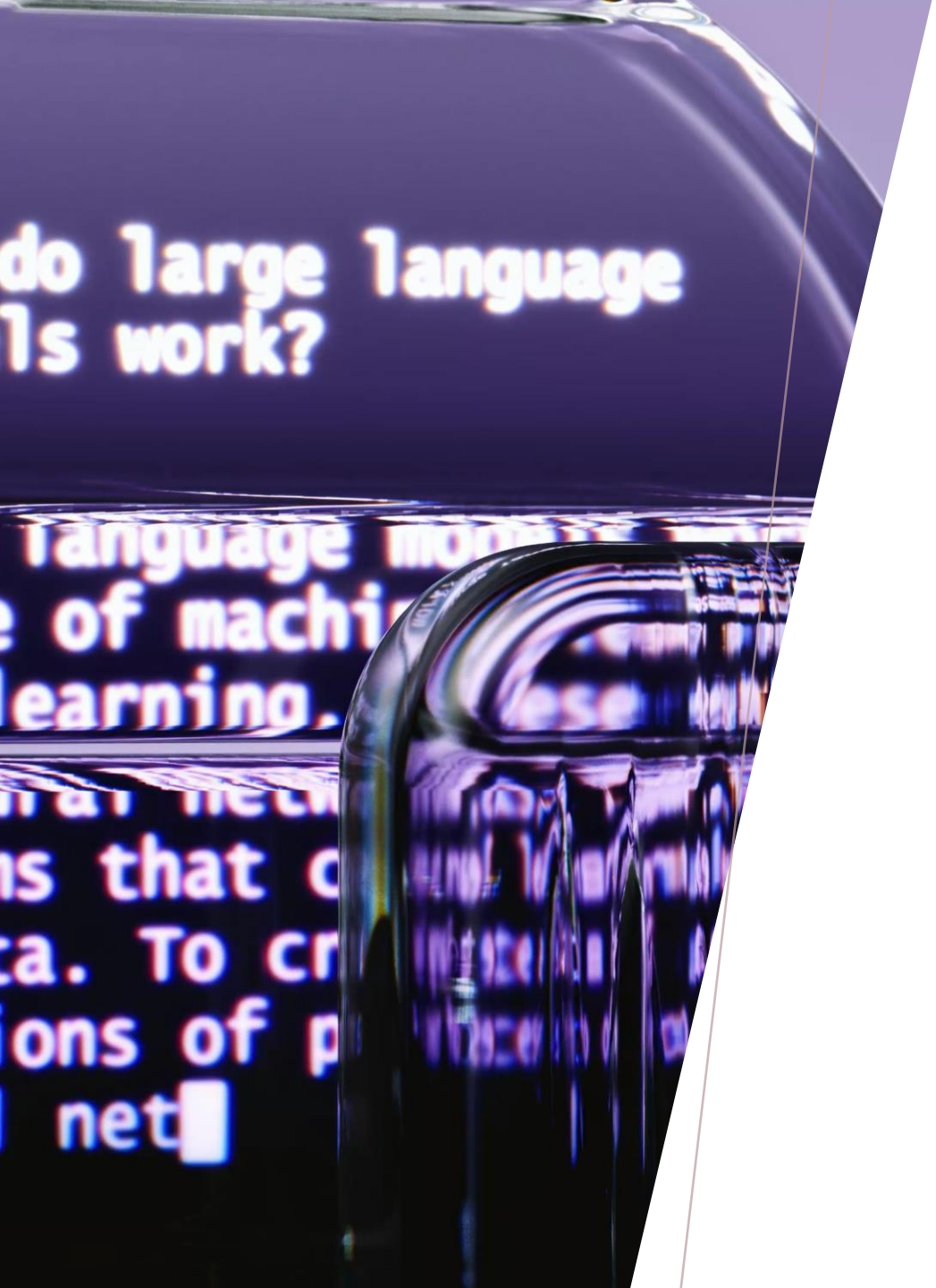
Dataset and Model

Methodology

Results and Evaluation

Conclusion

Future Work



PROBLEM DESCRIPTION

- **Problem:**
  Every song has a message, a meaning, a feeling, a story to be told. But **how do we condense that story into just a few words?**

- **Why is it important?**
  Titles provide the first impression and transmit a certain motion, hence making a good title crucial for the listeners experience (and therefore marketing)

- **Challenges**
  Songs are often abstract and include figures of speech (e.g. metaphors, similes, etc.) which are difficult to capture by AI

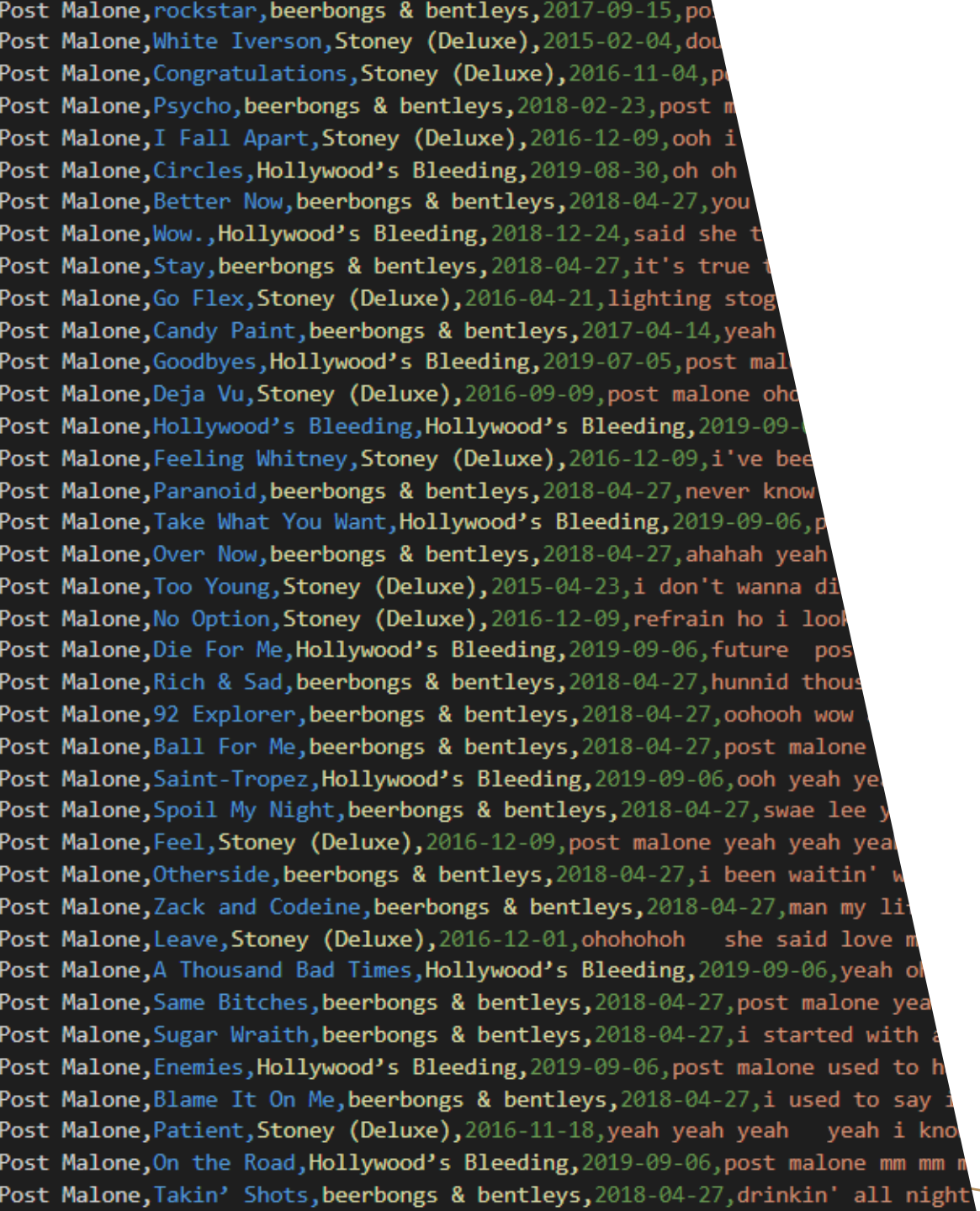
  Balancing creativity and semantic relevance

THE GOAL

1. Predict song titles based on lyrics using AI and different prompts
2. Evaluate the performances of both in terms of semantic similarity
3. Compare and contrast the emotional resonance of the (predicted) title

RESEARCH QUESTION

What is the impact of prompt engineering on the semantic accuracy, emotional resonance, and strength of AI-generated song titles?



DATASET & MODEL

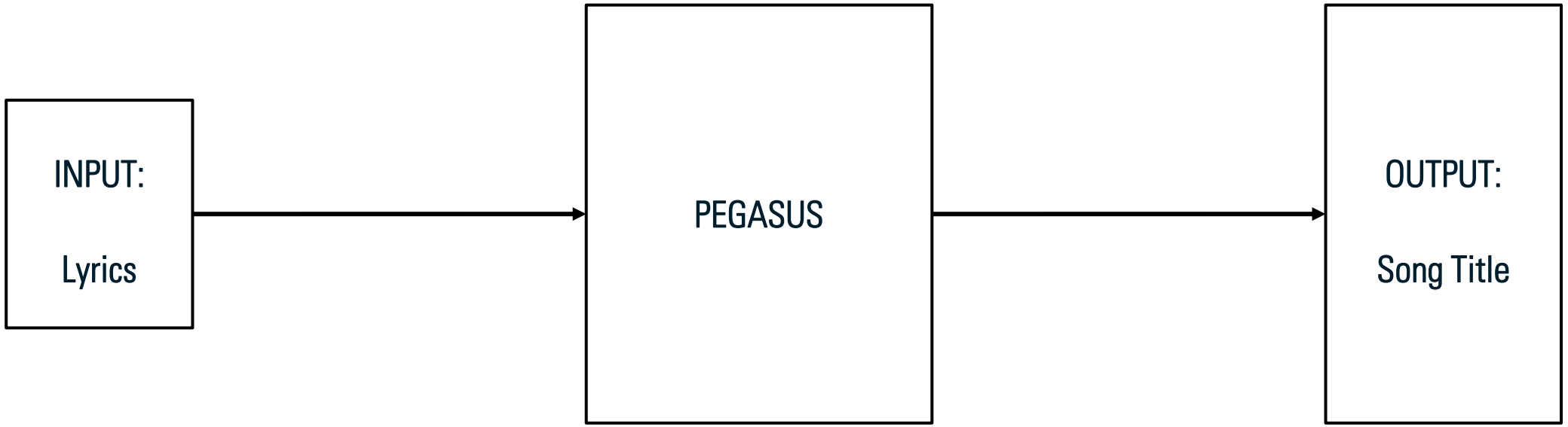
DATA

- 21 ARTISTS
- 12,054 TOTAL SONGS
- ARTIST, TITLE, ALBUM, DATE, LYRIC, YEAR
- SIZE: 22MB

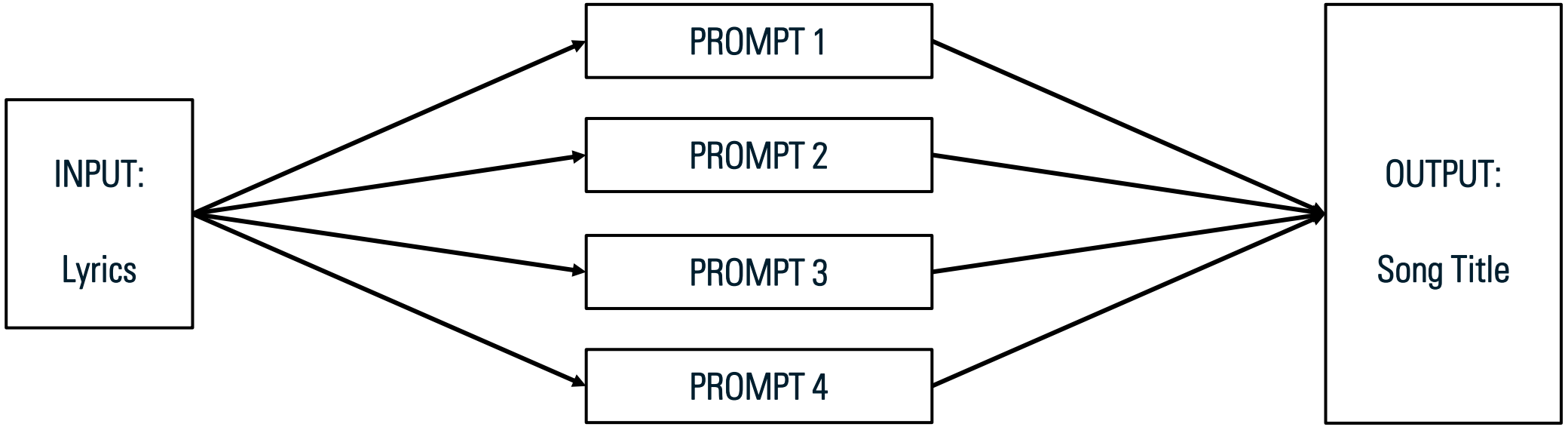
MODEL: PEGASUS-7B-SLERP.Q4_K_S.GGUF

- PARAMETERS: 7 BILLION
- TRANSFORMER - BASED LM
- TRAINED ON WEB PAGES, BOOKS, ARTICLES, ETC.
- SIZE: 3.9GB
- SLERP (SPHERICAL LINEAR INTERPOLATION)

METHODOLOGY

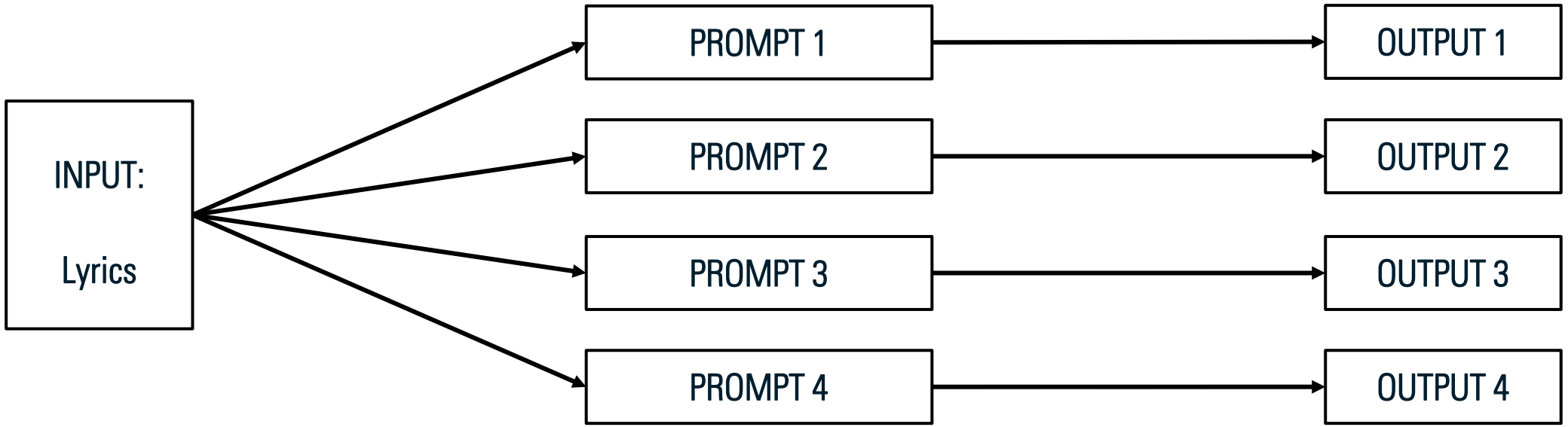


METHODOLOGY



	Prompt	Idea
Prompt 1	"Analyze the following song lyrics and generate a concise title using only words or phrases that appear directly in the lyrics: {lyrics}"	Very detailed and instructive query
Prompt 2	"Find a title: {lyrics}"	Very vague
Prompt 3	"What would you name the following songs based on its lyrics: {lyrics}"	Instructive query giving it some freedom
Prompt 4	"Summarize the following lyrics in a few key words: {lyrics}"	Titles are often sort of summaries

METHODOLOGY



METHODOLOGY

	Predicted Title	Semantic Similarity Score	Actual Title
OUTPUT 1			
OUTPUT 2			
OUTPUT 3			
OUTPUT 4			

METHODOLOGY: SEMANTIC SIMILARITY SCORE

	Semantic Similarity Score
Why?	Allow for a measure of performance
Step 1	Generate embeddings for the original title and predicted title using the SentenceTransformer model
Step 2	Use cosine similarity to compute the similarity score: $$\text{cosine } similarity = \frac{A \cdot B}{\|A\| \cdot \|B\|}$$

Title	Predicted Title	Similarity Score
Rockstar	Party Animal Music	0.243345

- Entirely different words
- Not too different in terms of meaning

METHODOLOGY: EMOTIONAL RESONANCE

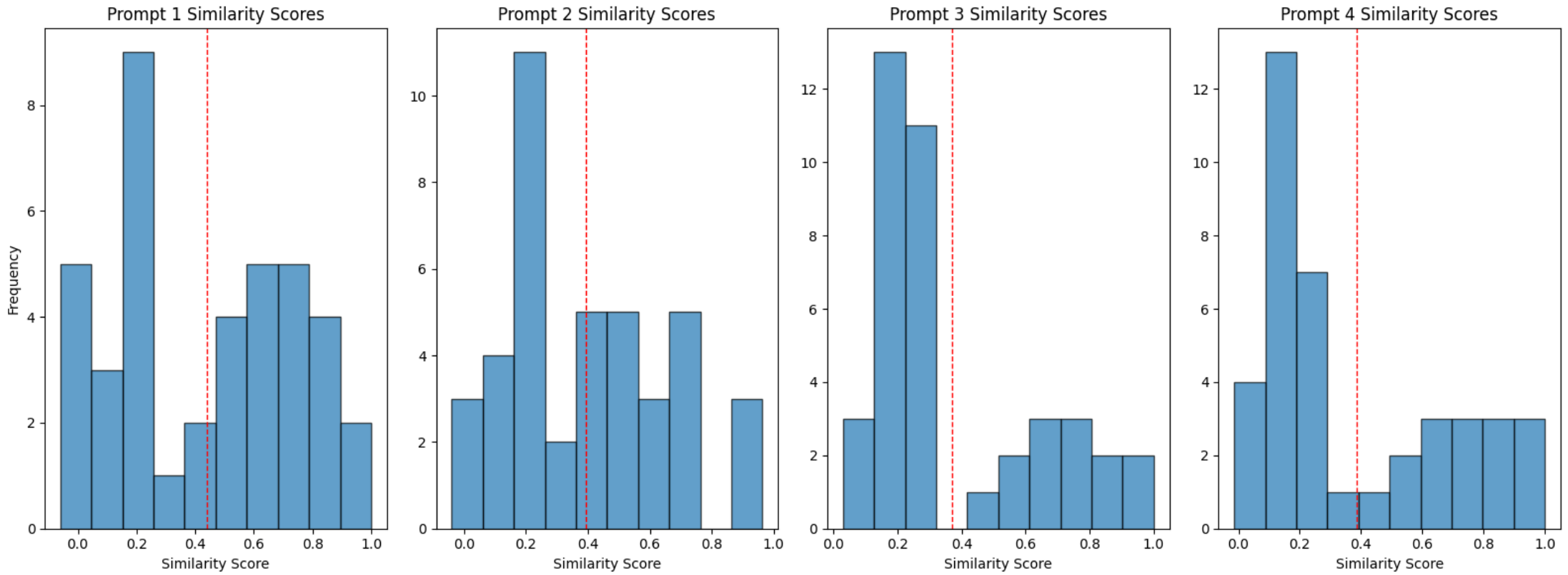
	Emotional Resonance Score
Why?	To quantify the sentiment expressed by the title, allowing for an additional measure of performance.
Step 1	Perform sentiment analysis on the title using nlptown/bert-base-multilingual-uncased-sentiment
Step 2	Transform sentiment into continuous values in $[-1, 1]$

$$sentiment\ score = \frac{output\ - 3}{2} \cdot confidence$$

e.g.: $sentiment\ score = \frac{4-3}{2} \cdot 0.24 = 0.12$

Artist	Title	Stars (score)	Emotional Resonance Score
Ed Sheeran	Homeless	1 (0.65)	-0.65
Coldplay	Glass of Water	4 (0.24)	0.12
BTS	First Love	5 (0.59)	0.59

RESULTS: SIMILARITY SCORE

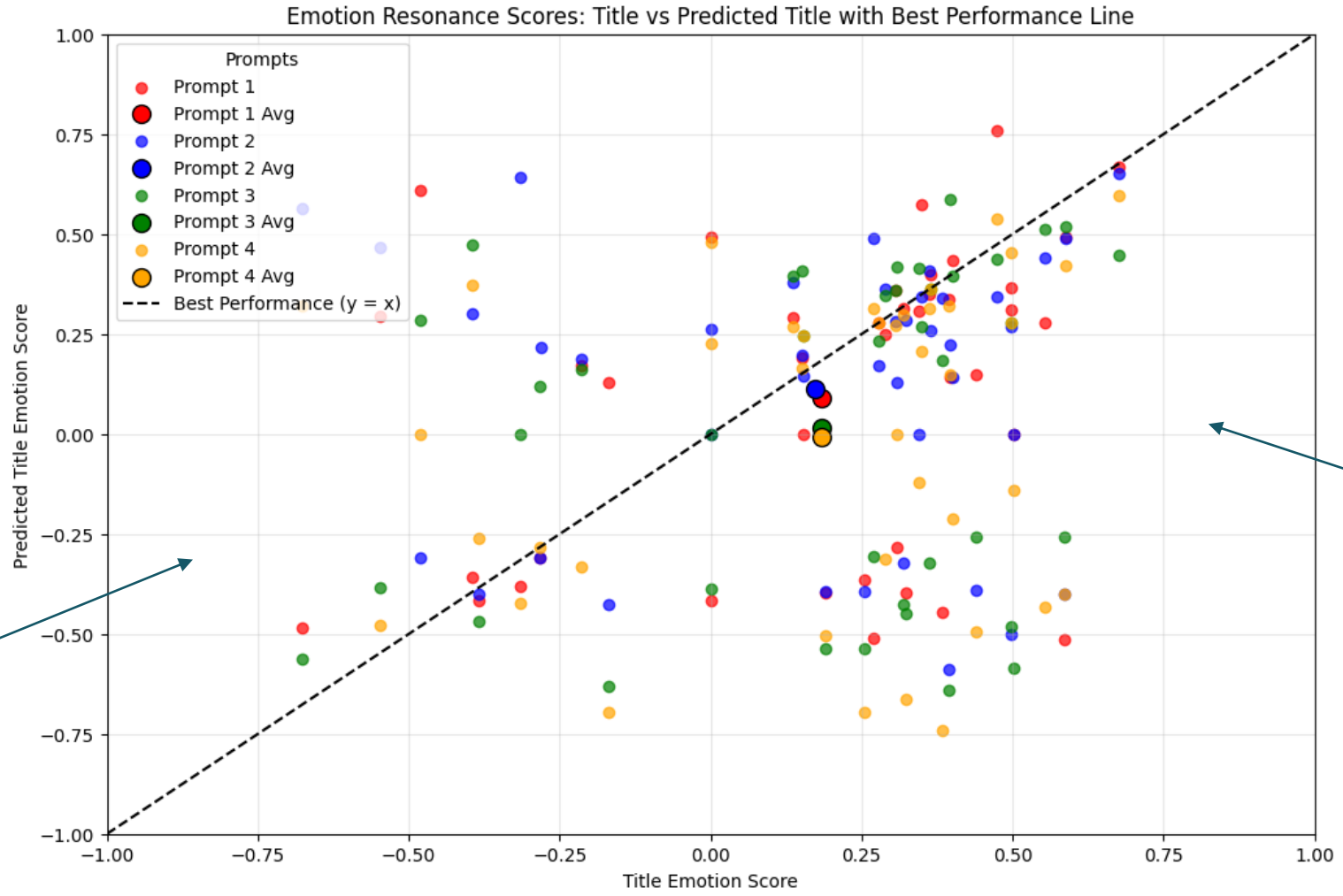


Avg similarity:	0.48	0.39	0.37	0.38

RESULTS: EXAMPLES

Artist	Title	Predicted Title	Semantic Similarity
Ariana Grande	my heart belongs to daddy (live)	Heart Belongs to Daddy	0.862453

RESULTS: EMOTIONAL RESONANCE



True Title Sentiment < Predicted Title Sentiment

Predicted Title Sentiment < True Title Sentiment

RESULTS: EXAMPLES

Artist	Title	Predicted Title	Title Emotion Score	Predicted Title Emotion Score
Taylor Swift	Closest To A Cowboy	Cowboy Caviar	0.394301	0.320776

RESULTS

	Title Average Absolute Emotional Resonance Score	Predicted Title Average Absolute Emotional Resonance Score
Prompt 1	**0.33**	0.32
Prompt 2	**0.35**	0.33
Prompt 3	0.36	**0.38**
Prompt 4	**0.36**	0.35

CONCLUSION

- Small (7b parameter) models can achieve reasonable degree of semantic similarity when predicting titles
- Phrasing prompts can have significant impact on outcome: Queries that are more descriptive and clear, without much room of freedom, generally produce better outcomes.
- Humans remain better at expressing emotion: AI model struggles to express emotions even when explicitly told to do so.

FUTURE WORK

- Apply same approach to larger models (e.g.GPT4o)
- Explore adding audio features (intonation, tempo, instrumentation)
- Fine-tune model: Training on song data, integrate emotional resonance score
- Expand application to movies, books, research papers, advertisements, etc. based on perceived emotional resonance

THANK YOU