Sentiment Analysis using LLM Prompt Tuning

Model Used: Mistral (via Ollama)

Dataset: Custom CSV with text and sentiment columns

Goal: Classify text as positive, neutral, or negative

Manual Prompt Tuning

Method:
The structured prompt was crafted with:

- Explicit instructions
- Three labeled examples (few-shot learning)
- Defined output format

Process:

- The prompt was reused for each input text.
- The model's response was parsed and compared to ground truth.

Observation:

- High interpretability
- Consistent output structure
- Limited flexibility for varied inputs

Automated Prompt Tuning (Monte Carlo Search)

Method:
We applied a **Monte Carlo strategy** to explore prompt variations:

- Created a pool of 5 semantically equivalent templates
- Randomly sampled 3 templates per input
- Queried the model and recorded outputs
- Selected the majority prediction per input

Prompt Variation Examples:

- Classify sentiment as 'positive', 'neutral', or 'negative'.
- What is the sentiment? (One word only)
- Determine tone of: "{text}"

Process:

- Applied random template sampling per row
- Allowed the model flexibility in interpreting prompts
- Returned the most frequent response across the sampled variants

Observation:

- Captures variability in LLM responses
- More robust to prompt phrasing changes
- Slightly higher tolerance to ambiguous inputs

Conclusion

- **Manual tuning** yields reliable and controlled results, especially for known input structures.

- **Monte Carlo prompt search** offers a simple yet effective automated tuning approach, discovering phrasing variations that improve accuracy or model alignment.