Decoding Psychedelic Experiences: An Analysis of Trip Reports with LLMs

Yang Ong
STATS418, Spring 2025

Background and Research Questions

Data Source: Erowid Experience Vault
- Erowid is a non-profit educational organization that provides information about psychoactive plants and chemicals.
- The Experience Vault is a repository of "trip reports", where people write about their experiences with psychoactive substances.

Primary Question: What factors are most correlated with positive trip outcomes?

Secondary Question: How do results differ between small and large parameter models when analyzing large amounts of unstructured text data?
- **Small model:**   DeepSeek R1-8b   (8 billion parameters)
- **Large model**:   GPT4.1   (1.76 trillion parameters)

Methodology

Step 1: Scrape data from Erowid into a CSV file (Selenium, Beautiful Soup)

Step 2: Clean data (Remove short, mixed-substance, and messy reports)

Step 3: Summarize report with a LLM, identifying 6 key categories (Ollama, LangChain, OpenAI APIs)
1. **Experience Level:** How experienced was the author with psychedelics?
2. **Control/Environmental Safety:** How much control or environmental safety did the author have during the experience?
3. **Mindset:** How much guiding mental, spiritual, or philosophical framework did the author have going into the experience?
4. **Intention:** How clear and meaningful was the author's purpose for taking the substance?
5. **Integration:** Did the author take steps to reflect on or apply the experience afterward?
6. **Outcome:** How positive or negative was the experience overall?

Step 4: Classify summarized reports with a LLM, rating each category between 1-5, Integration as a 1 or 0 for yes/no.
- 1 being low/poor to 5 being high/positive
- This is a necessary step as summarization + categorization in one prompt gives very inconsistent results

Step 5: Identify patterns between the category and outcome scores

Expected Insights and Limitations

Insights:

- Key predictors of positive experiences
- LLM scoring reliability
- Differences in output between models

Limitations:

- Subjective, unverifiable data
- LLM hallucinations
- Some important details might be lost during the summarization step
- Limited ground truth labels in LLM prompts due to context size restrictions