Psychedelic Trip Report Generator

A Flatiron Phase 5 Project by Jordan Loewen-Colón

Today’s Agenda

BUSINESS PROBLEM

GOALS

DATA

METHODS

RESULTS

The Problem

The (fictional) MindSpectrum Research Institute, is deeply engaged in groundbreaking work involving the clinical trials of **psychedelic-assisted therapies**. This research aims to contribute to **discovering new medicines** by understanding how different drugs can affect the mind. However, ***a major challenge in this area is the scarcity of firsthand accounts or personal descriptions of these experiences.***

To address this gap, our objective is to leverage existing personal narratives and use artificial intelligence (AI) to **create realistic 'trip reports' categorized by different types of drugs.** These AI-generated descriptions will then be used to enhance future research in the exciting and evolving field of psychedelics, potentially leading to innovative therapeutic applications.

Primary Goals

Develop a text generator model

Aim Produce human-like trip reports

Focus on minimizing false positives

The Data

- The dataset has 70k+ entries
- The ‘reports’ vary from 3 to 32k words.
- Other data points include:
  - Drug (11728 unique entries)
  - Dosage
  - Delivery
  - Weight
  - Year

Methods

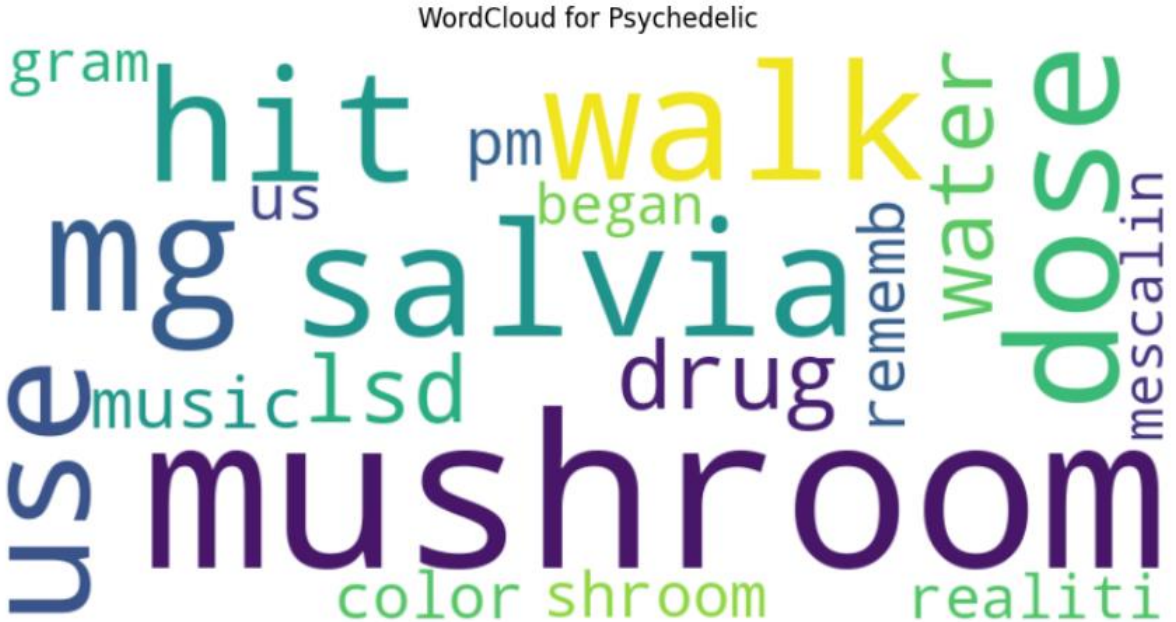
Vectorization and Unsupervised Learning To Discover Patterns

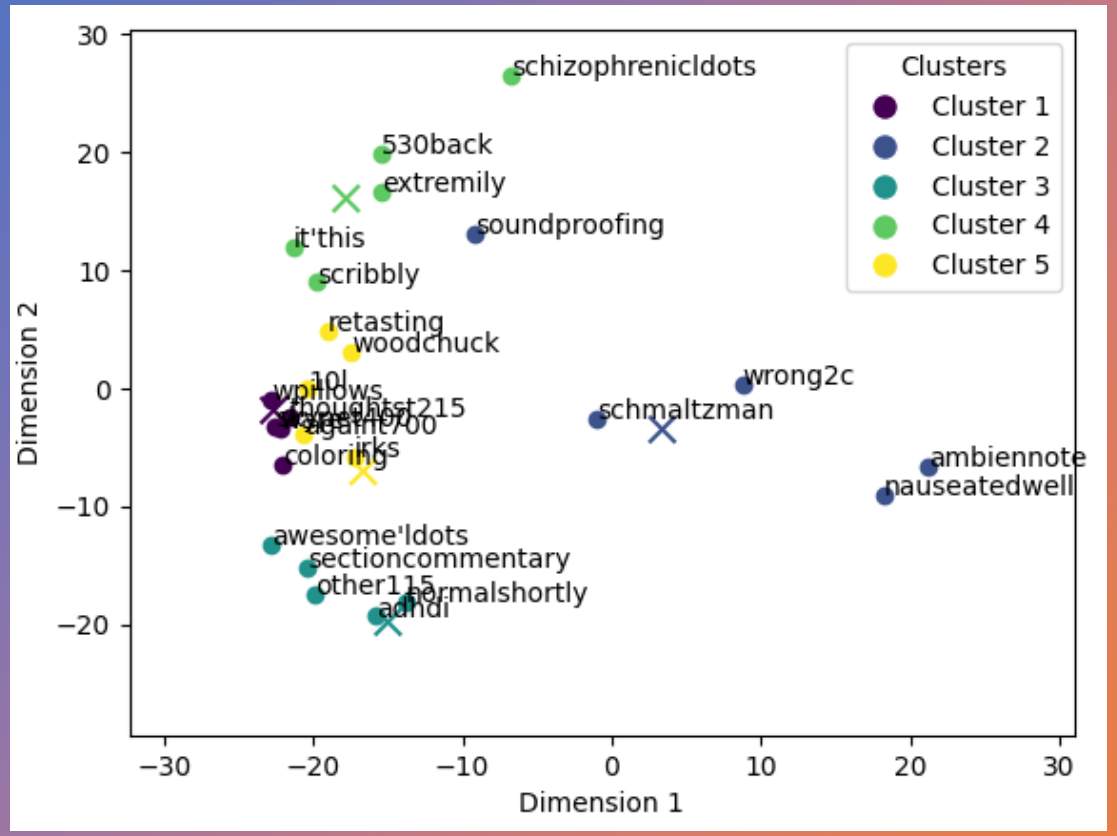
Multiclassification for drug prediction

Large Language Model to help verify text output

GPT-2 for Text Generation

Word Clouds from Unsupervised Learning Model (LDA)

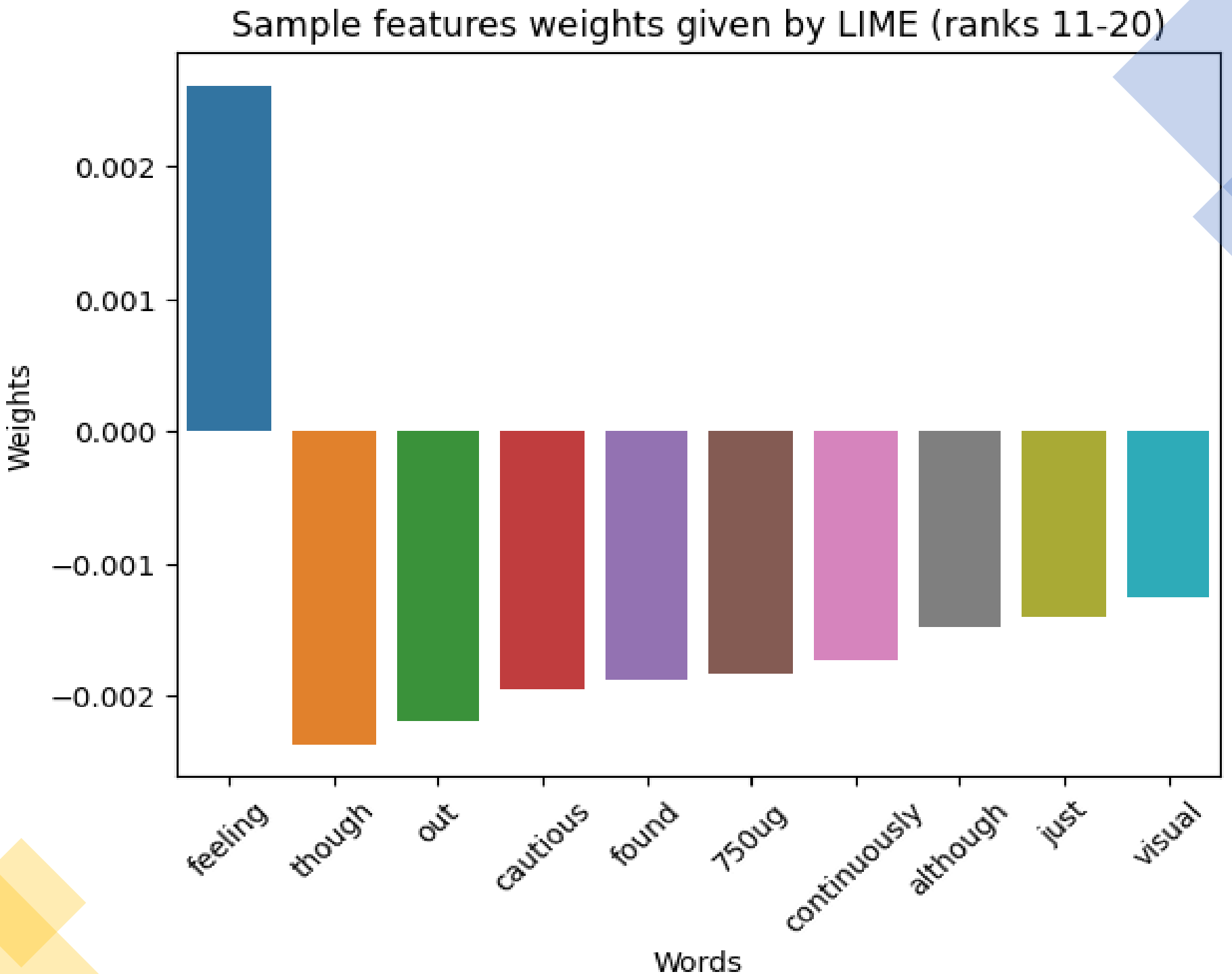




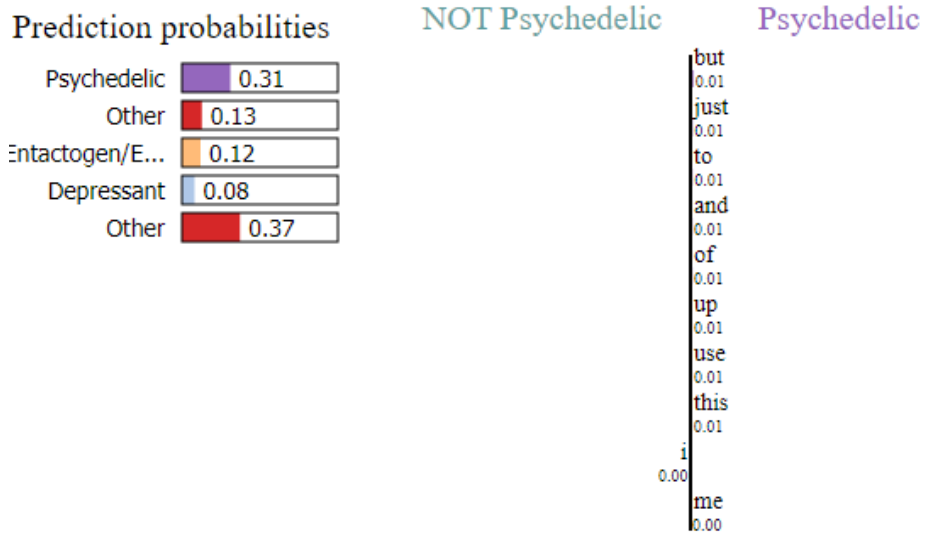
- 5 Clusters
- Words do not seem to share any clear connection or pattern
- Many of the words look like they are combinations of multiple words and are thus contributing “noise” to our data processing!

The weight of importance for how a word would contribute to a positive prediction of “Psychedelic”

The words “feeling” and “though” are the most weighty, but inversely.



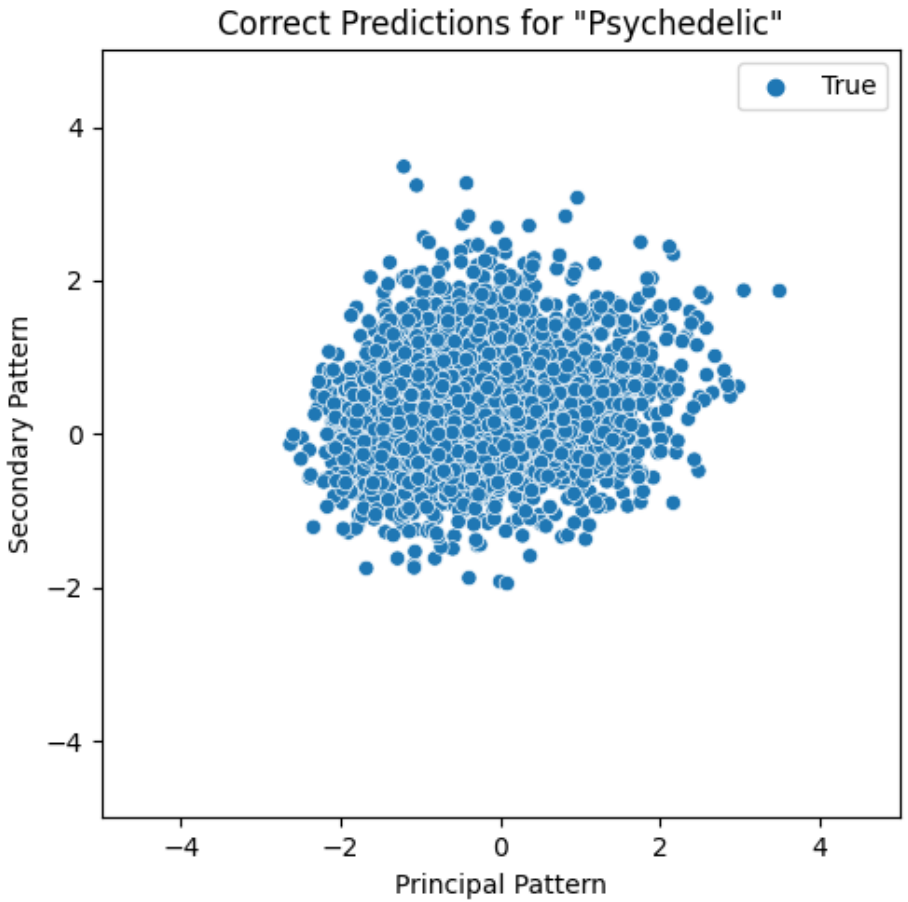
LIME Explainer Report

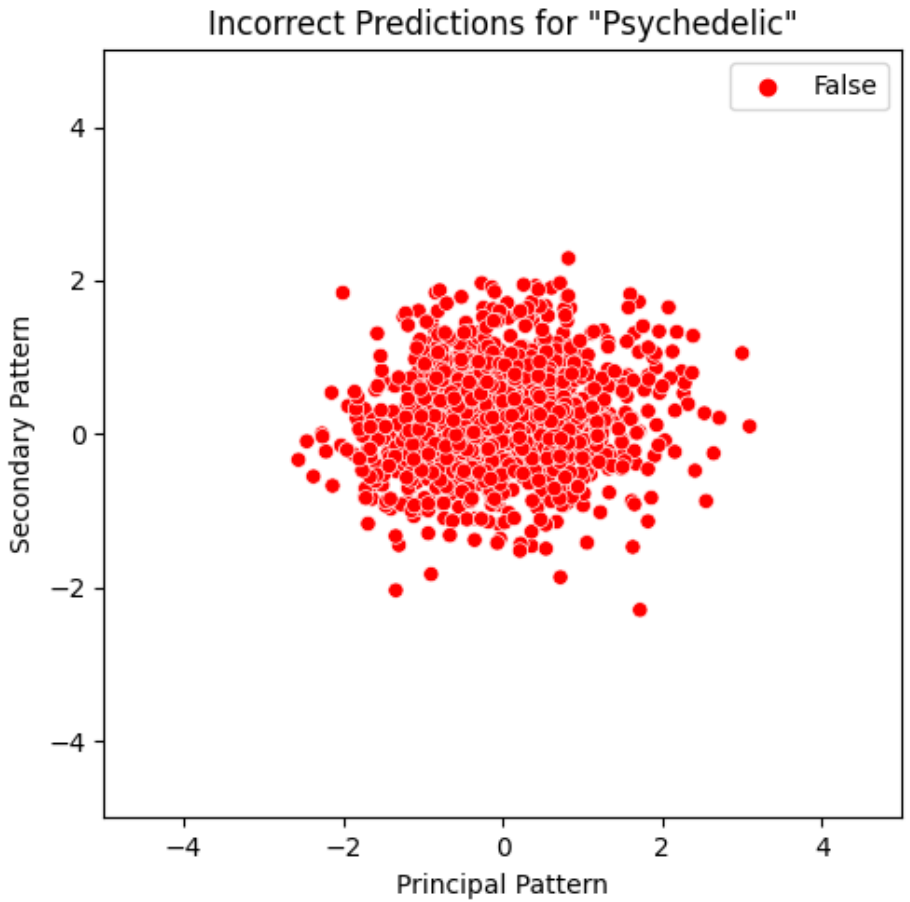


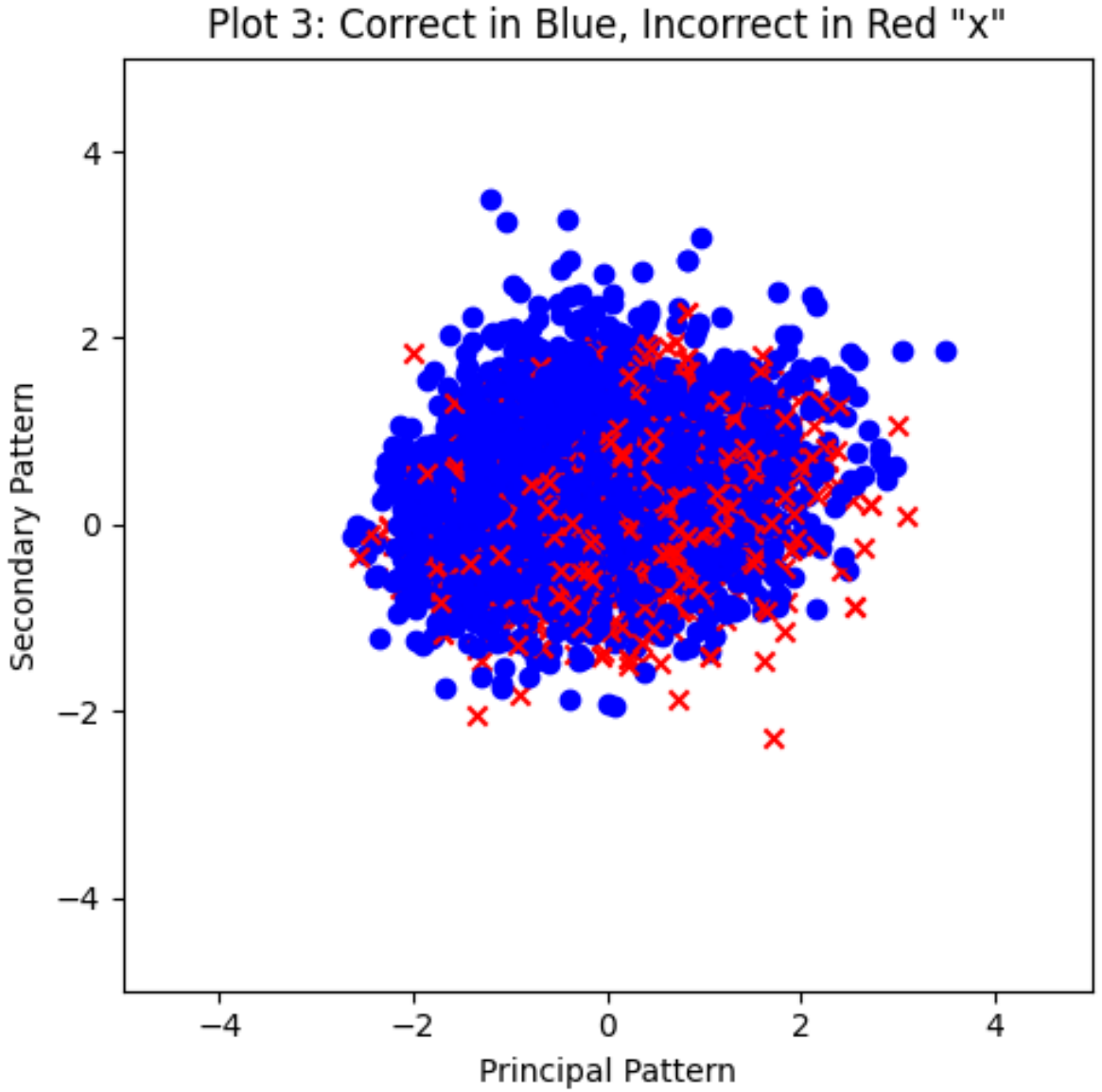
Text with highlighted words

first i would like to say that i am somewhat seasoned explorer " of the mind i have use many substance for purpose of recreation or exploration i 'm not boasting or anything i just want to be thorough i have been known to be able to keep my cool in le than desirable phsychedelic condition i have watched many friend cause long term damage to their psyche by misusing psychedelics and i try to remember this and repect the chemical i use i just want to imply that i can not compare the salvia experience to anything else i wa truly terrified like never before in my life the day of my last experience my girlfriend and i were scheduled to leave for week vacation the following day so there wa greatly relaxed atmosphere i wa in great mood maybe slightly apprehensive i had only smoked marijuana and am regular smoker i wa not stoned at the time i had experienced mild effect from salvia 10x but i did n't hold it in before but nothing like the one i wa about to i think this led to over eagerness and lack of respect i just did n't think it wa this

Underlying Patterns






Plot 3: Correct in Blue, Incorrect in Red "x"
Secondary Pattern
Principal Pattern
4
2
0
−2
−4
−4
−2
0
2
4

Conclusions

Significance of the word "Feeling” (and stop words)

Improvement with Larger Data Sample for Psychedelics
(more data leads to higher precision)

Underlying Patterns in Predictions
(More time could allow for more understanding)

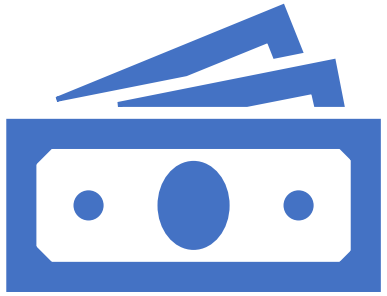
Recommendations

Experiment with removing stop words from reports to see if the model is able to focus more clearly on key words in reports.

Incorporate BigBird feature embeddings into GPT-2 model and use the trained RFC model to validate GPT-2 outputs.

Concerns & Interest

Computational efficiency limited pattern recognition



Findings indicate there is value in further study!

Thank You

Jordan Loewen-Colón

jbloewen@syr.edu