Project 3: Modeling Neurodiverse Experiences

By Joey Notaro

Problem Statement

- A disability advocacy organization is aiming to provide more targeted social support and referral services for new clients who suspect that they are living with neurodiversity.
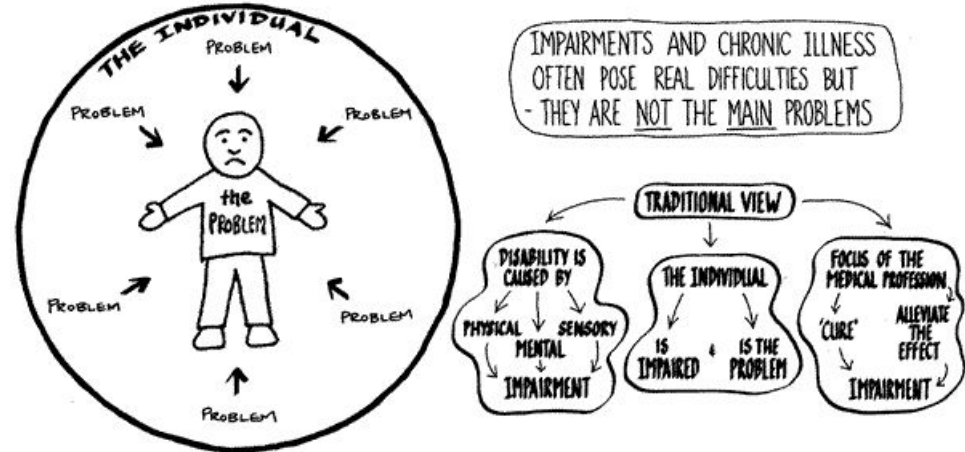
Problem Statement

*For **novice individuals** seeking social support for their neurodiversity, how can we **use the written experiences of others living with ADHD and autism** to build a model which most accurately classifies a novice's written, personal testimony as being one of more likely living with ADHD or with autism?*

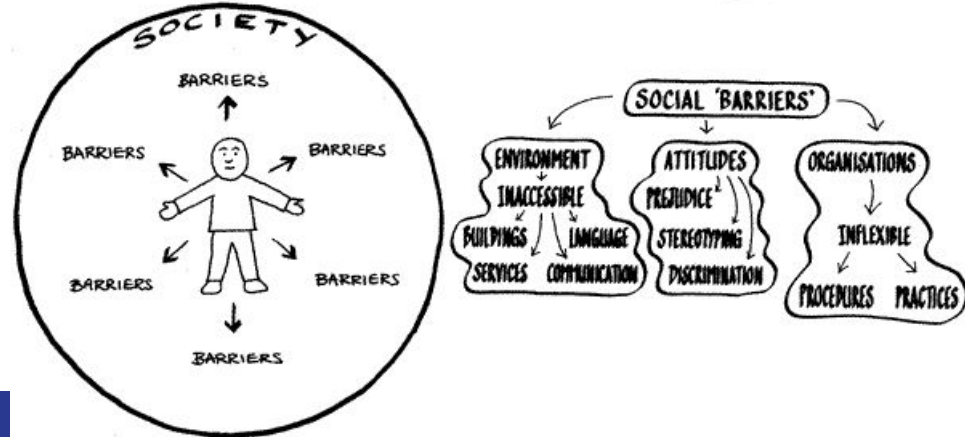- Goal: Classify experiential writing with 90% accuracy

Mission Statement

- The organization believes in a social model of disability.

- Diagnosis and evaluation matter but not the key to liberate people with disabilities.

THE MEDICAL MODEL OF DISABILITY

THE INDIVIDUAL

PROBLEM

PROBLEM → ← PROBLEM

the PROBLEM

PROBLEM → ← PROBLEM

PROBLEM

IMPAIRMENTS AND CHRONIC ILLNESS OFTEN POSE REAL DIFFICULTIES BUT - THEY ARE NOT THE MAIN PROBLEMS

TRADITIONAL VIEW

DISABILITY IS CAUSED BY
PHYSICAL → SENSORY
MENTAL
IMPAIRMENT

THE INDIVIDUAL
IS IMPAIRED & IS THE PROBLEM

FOCUS OF THE MEDICAL PROFESSION
'CURE' → ALLEVIATE THE EFFECT
IMPAIRMENT

THE SOCIAL MODEL OF DISABILITY

SOCIETY

BARRIERS

BARRIERS → ← BARRIERS

BARRIERS ← → BARRIERS

BARRIERS

SOCIAL 'BARRIERS'

ENVIRONMENT
INACCESSIBLE
BUILDINGS → LANGUAGE
SERVICES → COMMUNICATION

ATTITUDES
PREJUDICE
STEREOTYPING
DISCRIMINATION

ORGANISATIONS
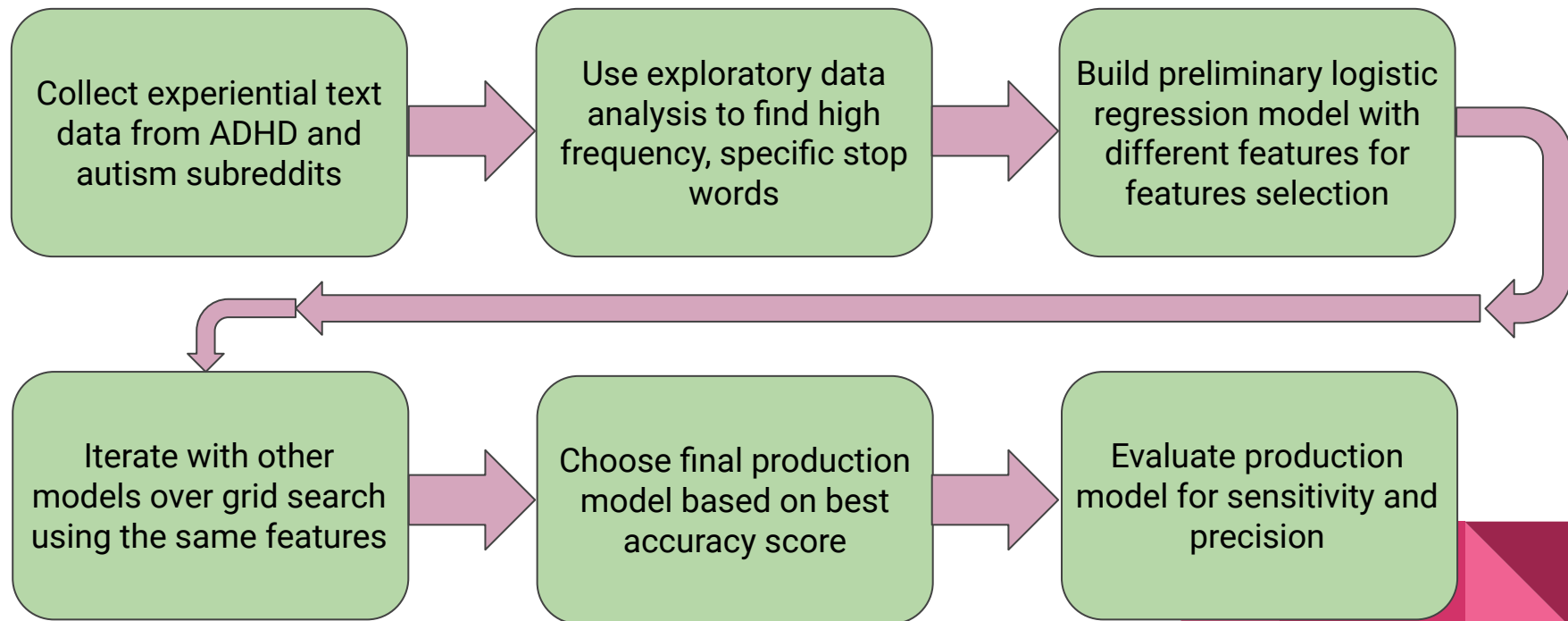INFLEXIBLE
PROCEDURES → PRACTICES

Why This Model?

- More overlap between ADHD and autism than previously believed.

- Many people do not have access to formal evaluation or diagnosis.

Misdiagnosis Monday:
ADHD vs. Autism

ADHD

Autism

Overlap

Craves novelty

Craves familiarity

Stimming

Sensory differences

Interoception issues

Impulse control difficulties

Emotional regulation difficulties

Difficulty regulating attention & focus

Finds comfort in repetitive behaviors and routines

Higher rates of substance abuse, suicidality, depression, eating disorders, anxiety, bipolar & OCD

ND communication patterns (info-dumping, connecting over shared interests, story swapping)

Difficulty with social cues due to focus & attention issues

Webbed/interconnected thought patterns

Difficulty reading allistic social cues intuitively

Executive functioning difficulties

Increased risk of victimization

Special interests & passions

Hyperactivity & impulsivity

Task-switching difficulties

Strict adherence to routines

Neurodivergent masking

Eye contact differences

Inhibition difficulties

Social differences

High need for verbal context

Passionate focus

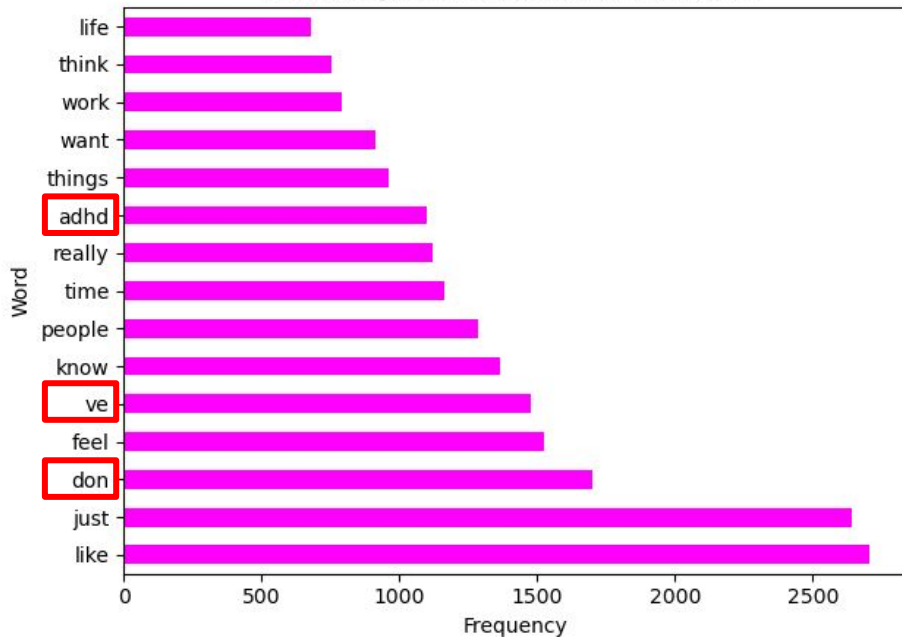**DR. NEFF
MISDIAGNOSIS MONDAY SERIES**

Disclaimer: This diagram represents common patterns or trends associated with autism and ADHD. However, these are generalizations & will not resonate with everyone's individual experiences. There is a lot of diversity among Autistic and ADHD people & each person's journey is different. For more information and full write up please see www.neurodivergentinsights.com
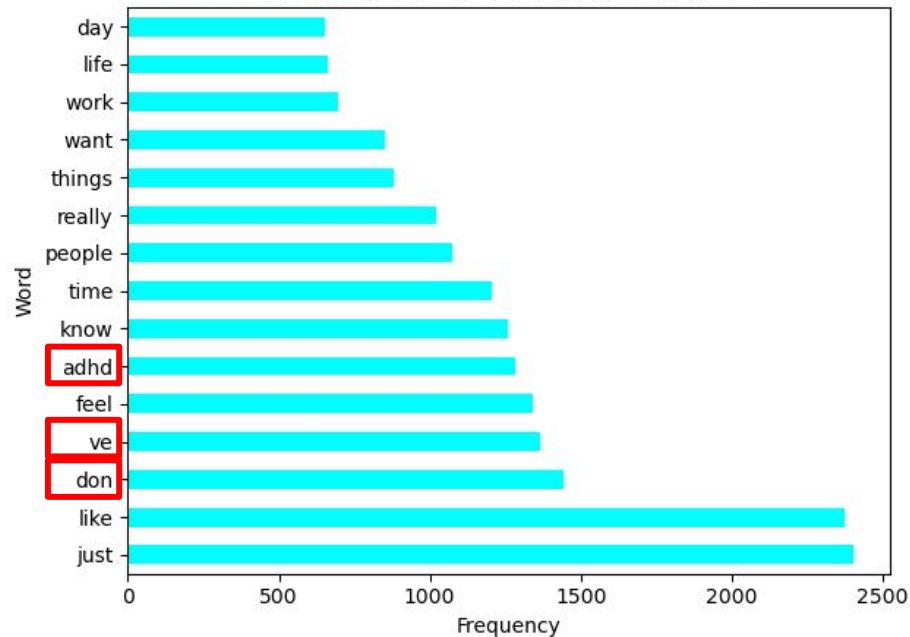
Methodology

Collect experiential text data from ADHD and autism subreddits

→

Use exploratory data analysis to find high frequency, specific stop words

→

Build preliminary logistic regression model with different features for features selection

Iterate with other models over grid search using the same features

→

Choose final production model based on best accuracy score

→

Evaluate production model for sensitivity and precision

High Frequency Words

Most Frequent Words in ADHD Subreddit

Most Frequent Words in Autism Subreddit

High Frequency Bigrams

Most Frequent Bigram Phrases in ADHD Subreddit
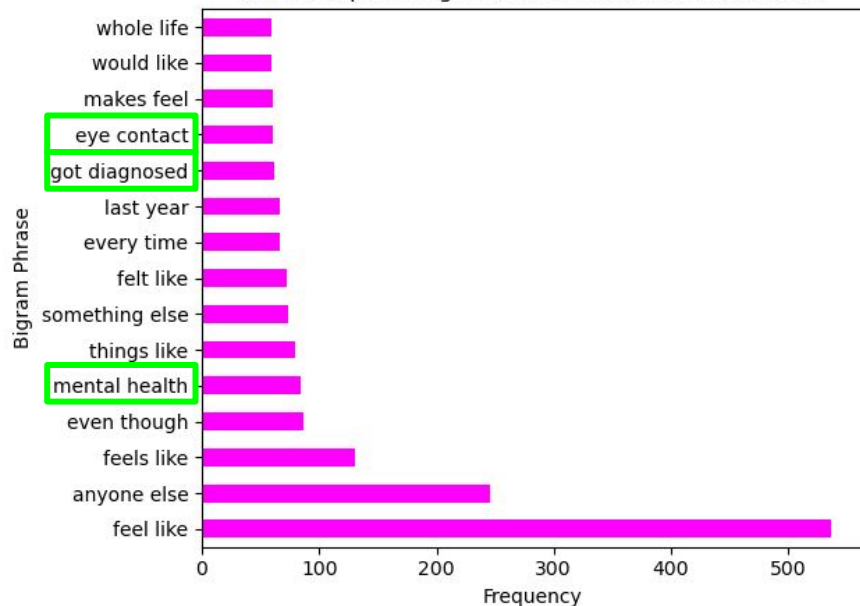
Most Frequent Bigram Phrases in Autism Subreddit
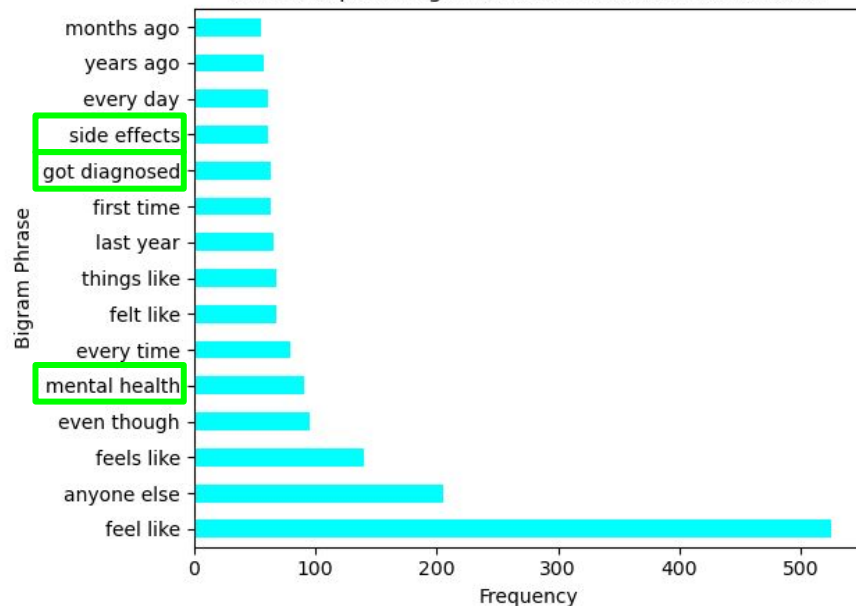
Stop Words

- Much more overlap in themes than expected and overlap increased with subsequent data collection.

- Sought stop words consistently unique to one subreddit

- 'english', 'don', 've', 'adhd', 'autism', 'autistic', 'meds', 'medication'

Preliminary Model

- Null Baseline Accuracy - 50%
- Logistic Regression Model with Four Different Features
    - Countvectorizer of **original** post text
    - Countvectorizer of **lemmatized** post text
    - Countvectorizer of **bigrams** post text
    - Countvectorizer of **lemmatized bigrams** post text

Feature Selection

- Null Baseline Accuracy - 50%
- Logistic Regression Model with Four Different Features
  - Countvectorizer of **original** post text - 0.82
  - Countvectorizer of **lemmatized** post text - 0.82
  - Countvectorizer of **bigrams** post text - 0.71
  - Countvectorizer of **lemmatized bigrams** post text - 0.71

Model Iteration

Model	Accuracy Test Score
K-Nearest Neighbors	~0.55
Decision Tree	~0.71
Bagging Classifier	~0.82
Random Forest Classifier	~0.81
AdaBoost Classifier	~0.78

Production Model

- Features: Lemmatized Text with English and Modified Stop Words
- Max Document Frequency: 100%
- Maximum Features: 5,000
- Minimum Document Frequency: 2
- Logistic Regression C-value: 0.1
- Logistic Regression Penalty: L2 "Ridge"

Production Model	
Cross-Validation Score	~0.82
Training Score	~0.96
Test Score	~0.82

Evaluation of Production Model

- Sensitivity/Recall: ~0.85

- Precision: ~ 0.80

- More sensitive than precise. More likely to predict autism erroneously.
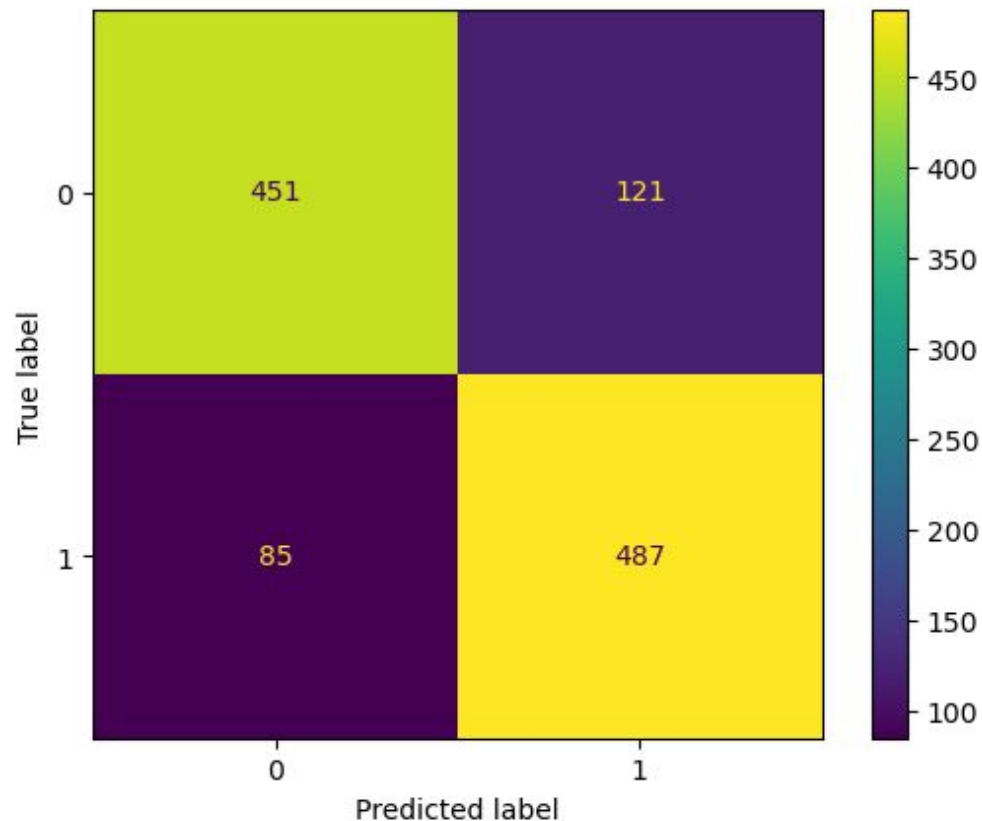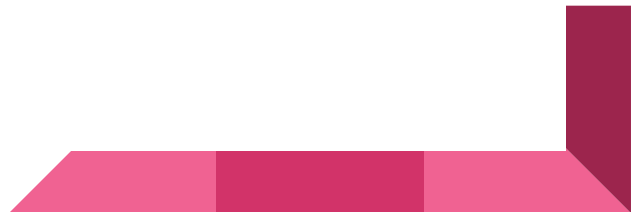
Finding #1

- The final logistic regression production model after multiple model iterations only achieved about 82% accuracy, which failed to meet our target of a model which could predict with 90% accuracy.
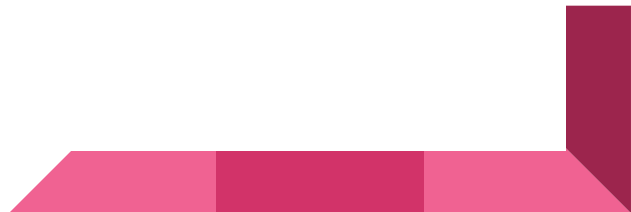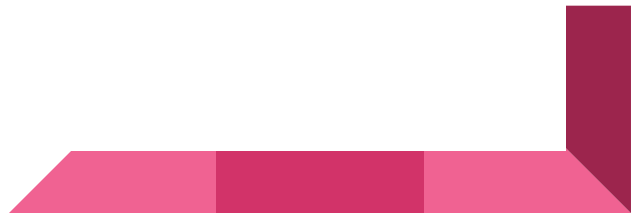
Organization Recommendation #1

- If accuracy is important to the organization, invest in collecting more data from different sources.

- Data may be skewed by specific demographics who use Reddit. Collect from APIs on other social media websites.

Finding #2

- The final logistic regression production model is more sensitive than it is precise. In this case, false positives were more common than false negatives, so our production model is more likely to predict a post is falsely from the autism subreddit.

Organization Recommendation #2

- Use the model as a preliminary referral tool to direct new clients to peer support groups and services.

- Invest in an added layer of an expert professional being present to facilitate group meetings and identify potential new clients who may be misclassified.

- Focus on more misclassifications in the autism support groups.

Finding #3

- In our exploratory data analysis, corroborated by outside research, we see extremely high overlap in the bag of words used most frequently to describe the lived experiences of people with ADHD and autism, both in common social language, but also in themes of specific topics which appear similar across the neurodiverse spectrum of experiences.

Organization Recommendation #3

- Begin offering a mixed neurodiversity support group for new clients.

- Good first step in exploring disability identity and may not yet have the specific needs of belonging to one group or the other.