UC San Diego

Myers Briggs Type Indicator (MBTI) Analysis and Classification

ECE 143 Winter 2021 Group 14

Fengqi Yang, Luis Herrera, Blue Luu

Myers Briggs Type Indicator (MBTI)

Personality Category:

1. Extraversion (E) or Introversion (I)
2. Sensing (S) or Intuition (N)
3. Thinking (T) or Feeling (F)
4. Judging (J) or Perceiving (P)

About Dataset

MBTI Type	Forum posts
ENTJ	You're fired.|||That's another silly misconception…
ENFJ	https://www.youtube.com/watch?v=AwgF14ySLpw I do stuff like dis.|||They...
INTP	I have many social screw ups. None of them funny though, ...
INTJ	You clearly don't want to move there. Don't.|||Want it or don't do it...

8675 people's 50 forum posts from personalitycafe.com

provided by Kaggle.com

Objectives

- Understand distribution of different MBTIs.
- Analyze the correlations between MBTI and forum posts
  - Detecting patterns of MBTIs and writing styles
  - Classifying MBTIs using forum posts.

Data Preprocess

Raw Posts:

"Yes! you are quite right. But many times this help has to come from outside the relationship." |||

Noisy Entity Removal:

Stop-Words, Punctuations, URLs, special character etc.

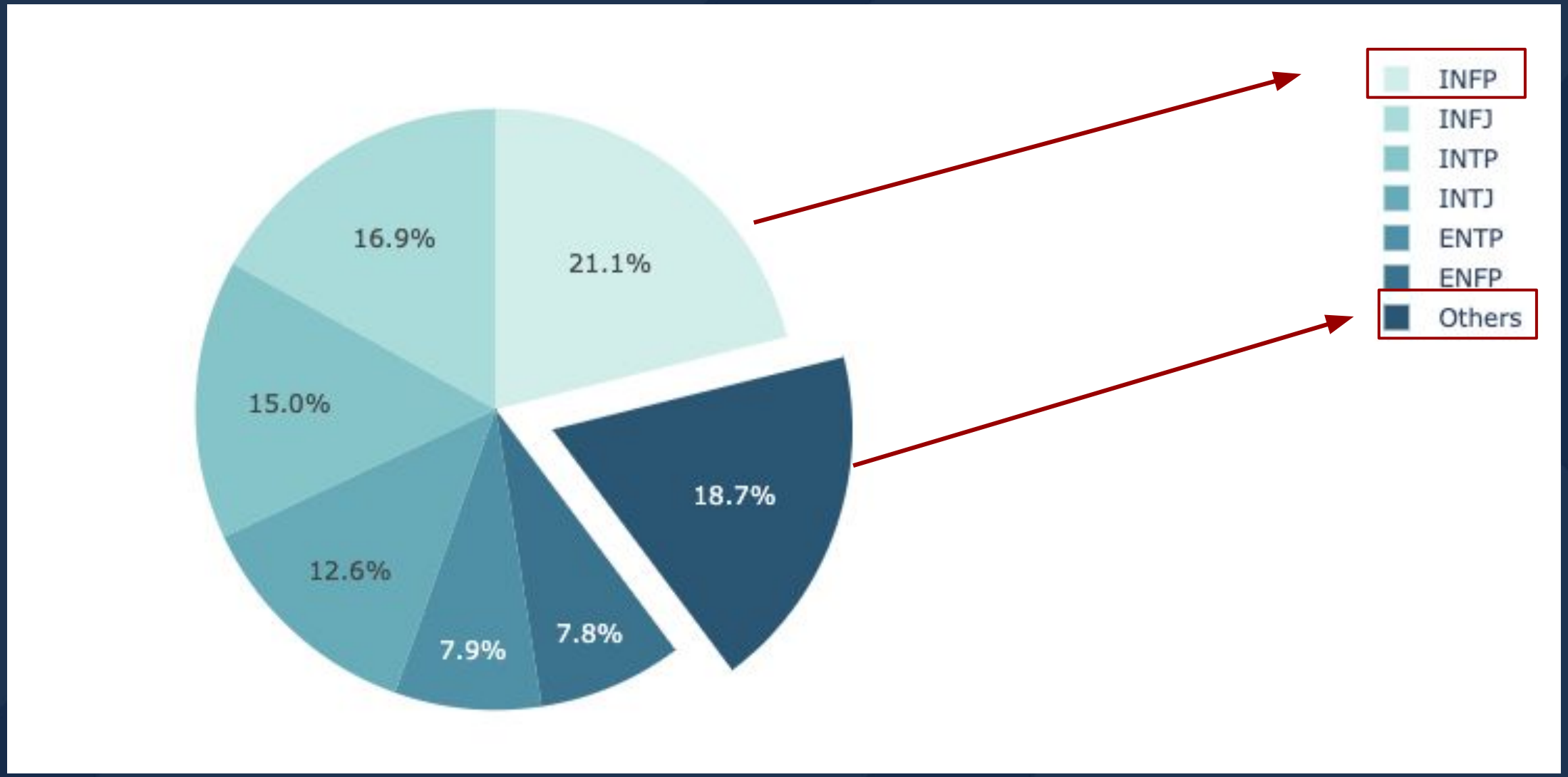
Text Normalization:

Tokenization,Lowercase, Sentence Segmentation etc.

Processed Posts:

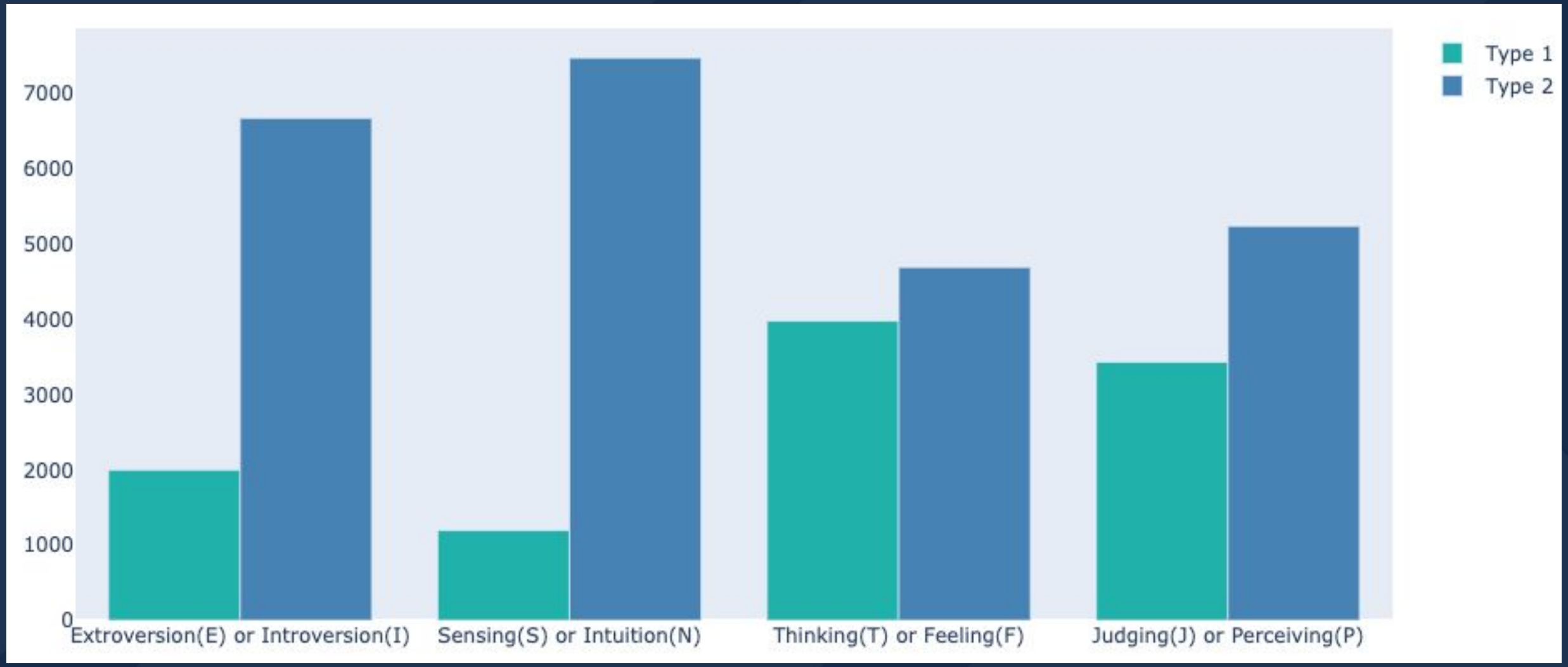
['yes', 'quite', 'right', 'many' 'times', 'help', 'come' 'outside','relationship']

MBTI Distribution



Extraversion (E) or Introversion (I) Thinking (T) or Feeling (F) Judging (J) or Perceiving (P) Sensing (S) or Intuition (N)

MBTI Category Distribution

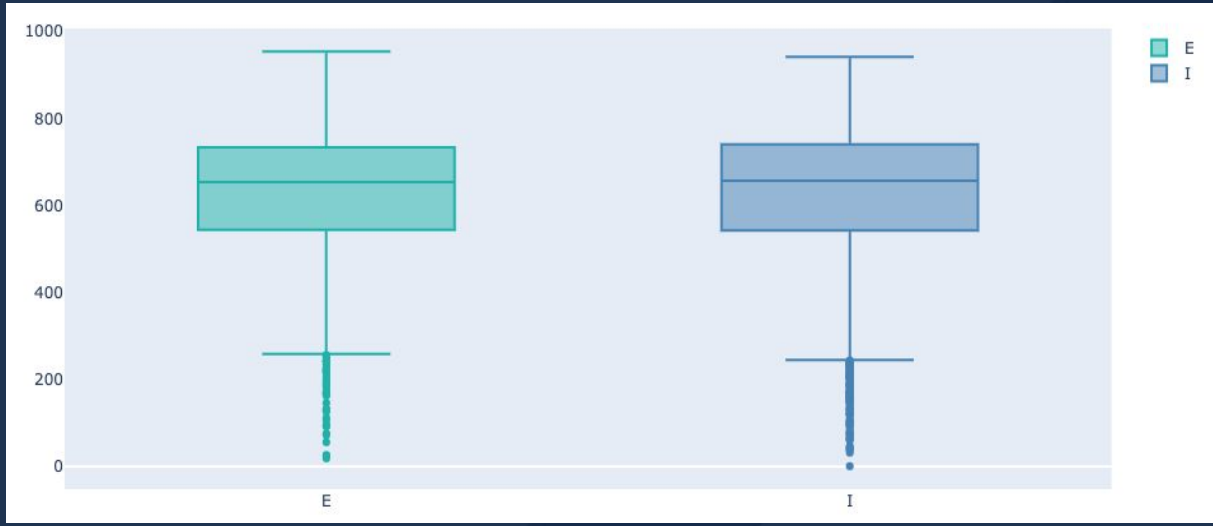


UC San Diego

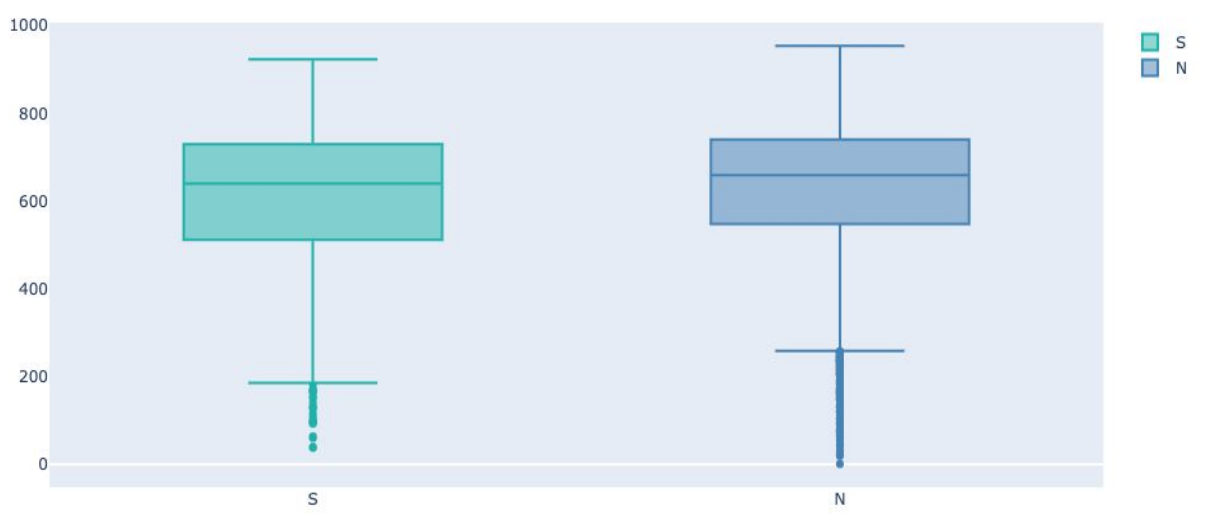
Part 2: Detecting Patterns

Length of Posts

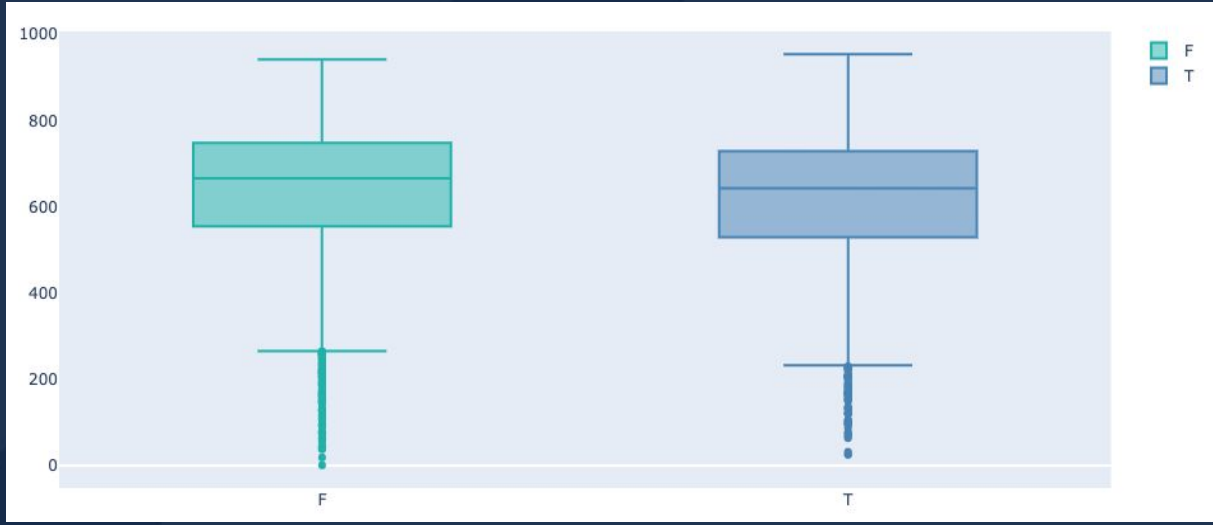
Extraversion (E) or Introversion (I)



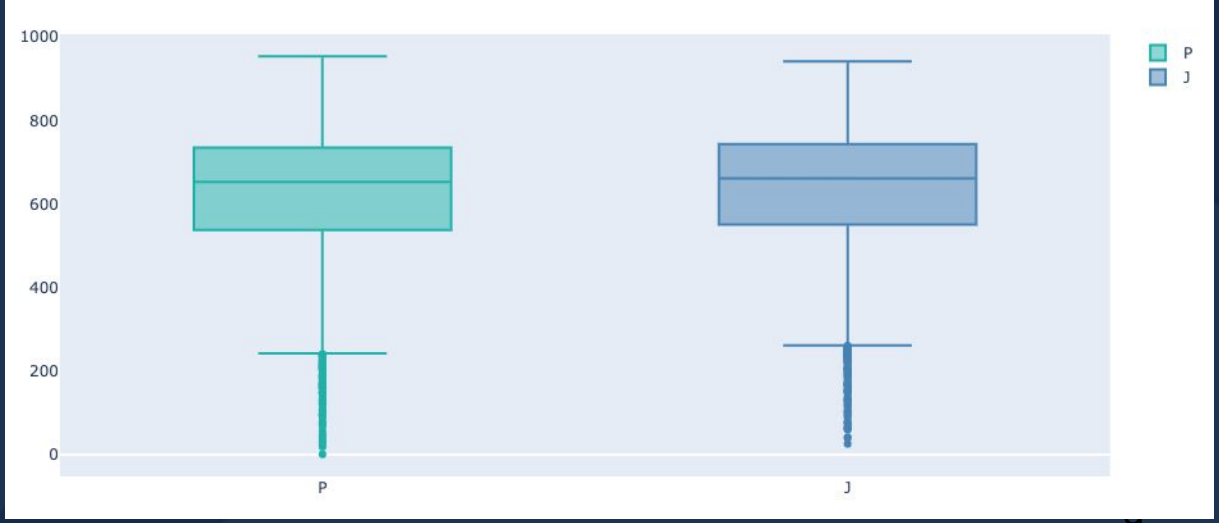
Sensing (S) or Intuition (N)



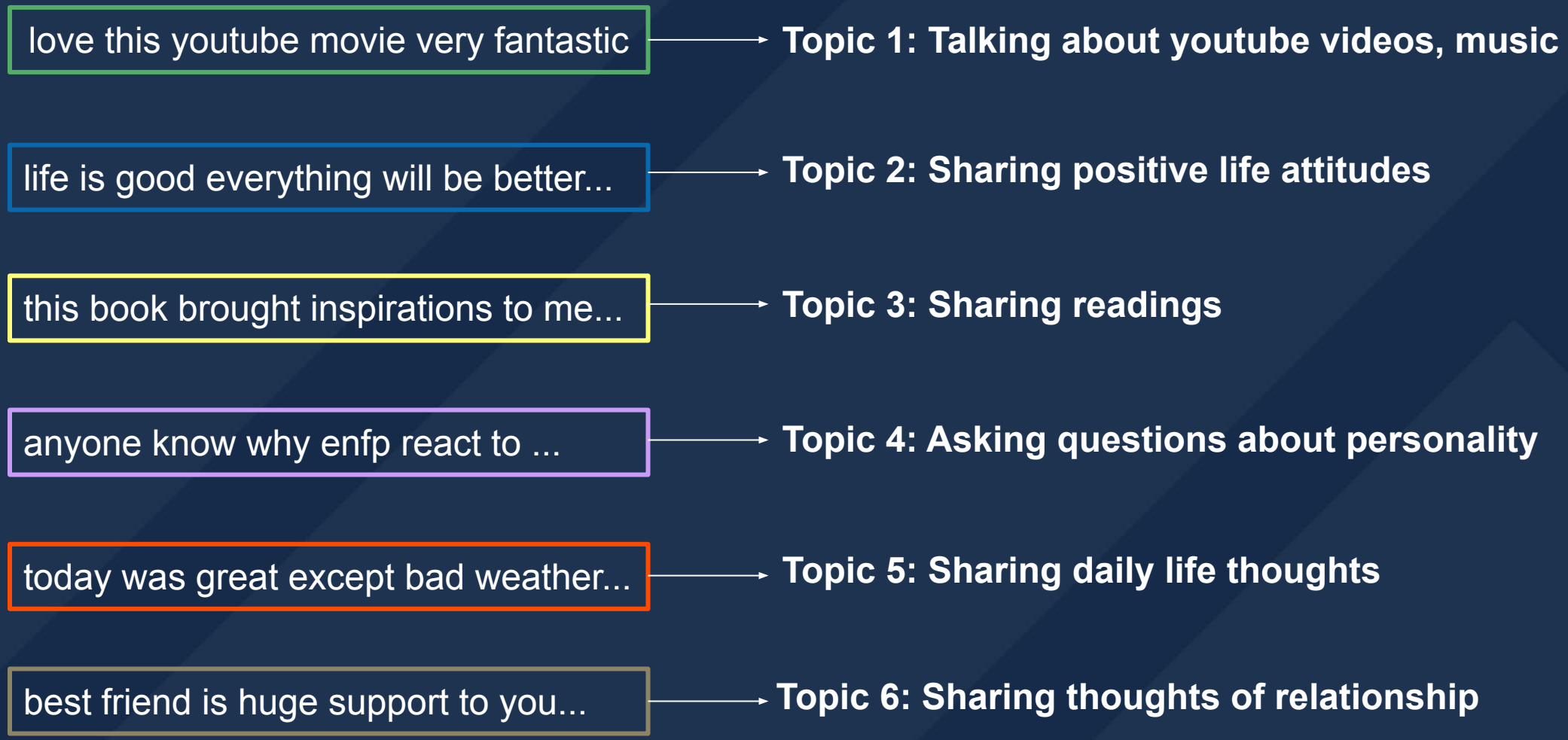
Thinking (T) or Feeling (F)



Judging (J) or Perceiving (P)

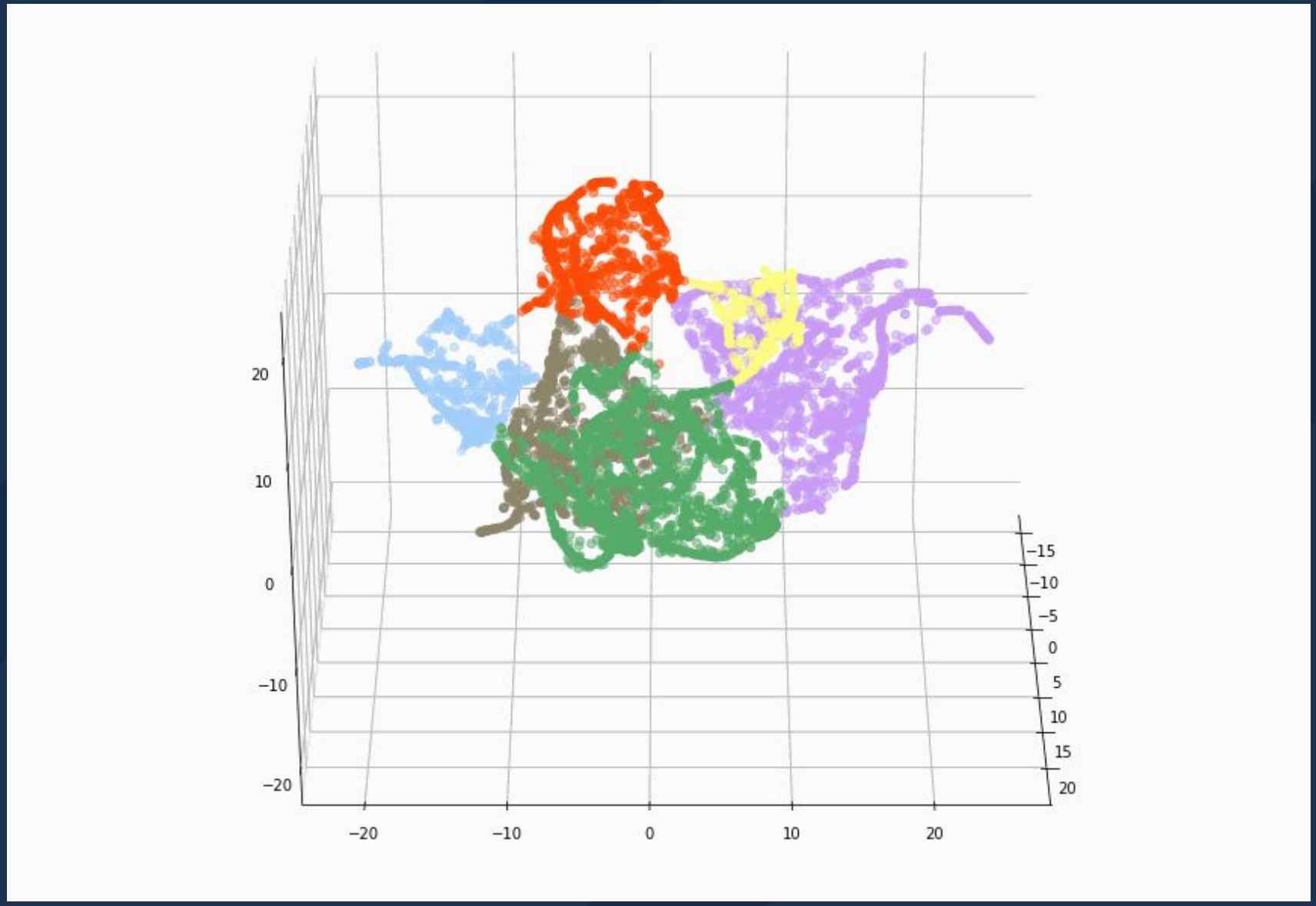


LDA Topic Modeling



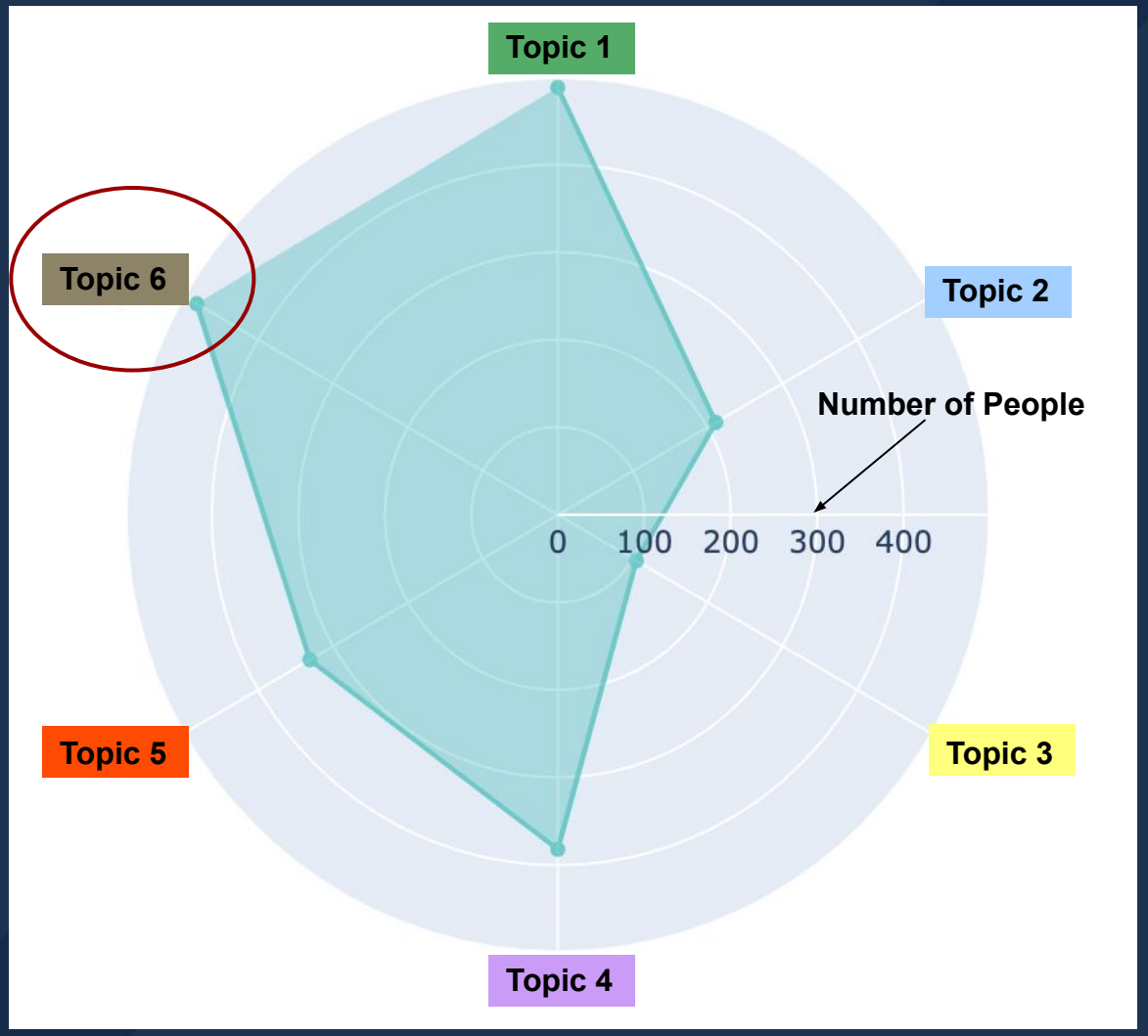
Topics Modelling Visualization

Topic 1	youtube
Topic 2	positive life
Topic 3	reading
Topic 4	questions
Topic 5	daily-life
Topic 6	relationship

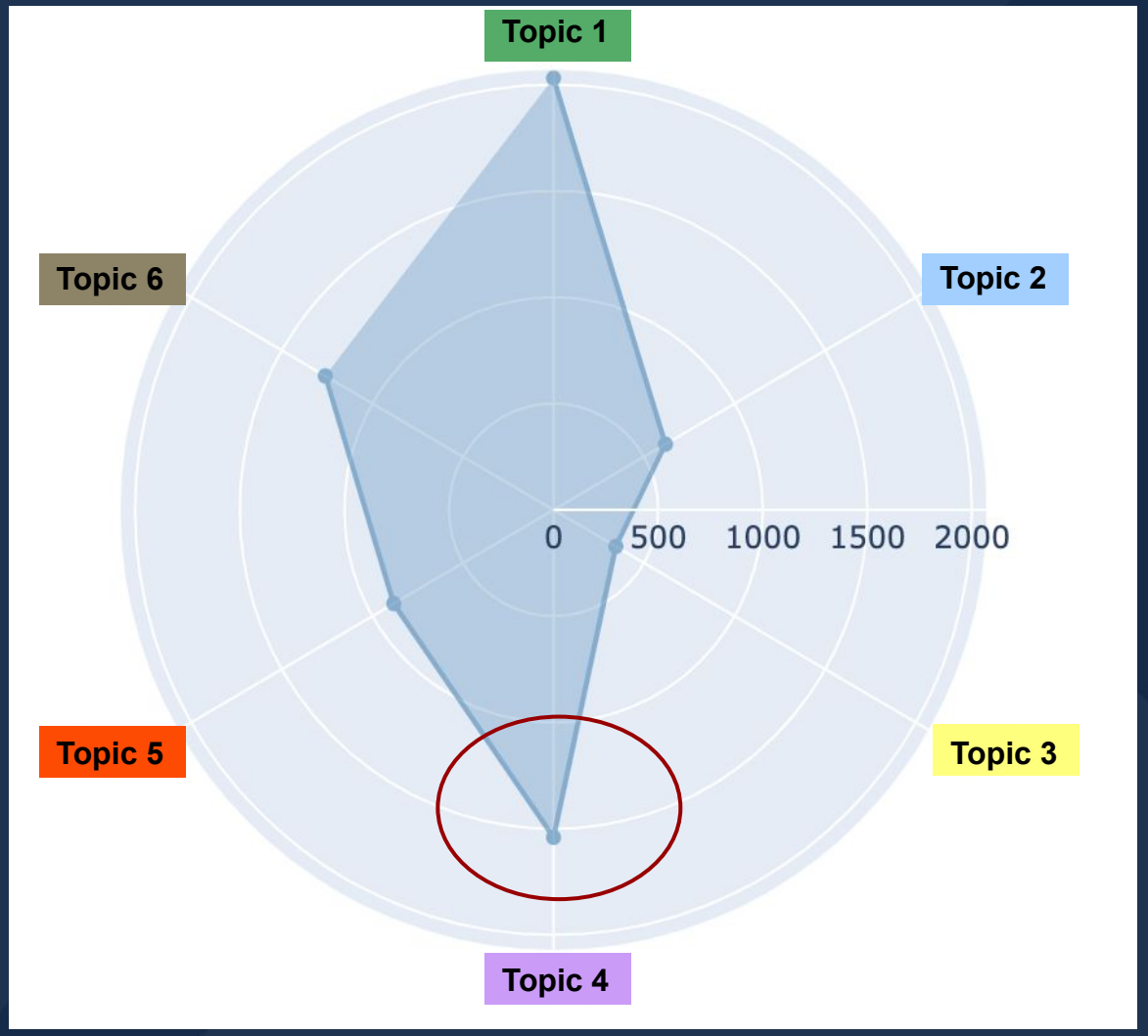


Most Frequently Discussed Topics

Extraversion (E)



Introversion (I)



UC San Diego

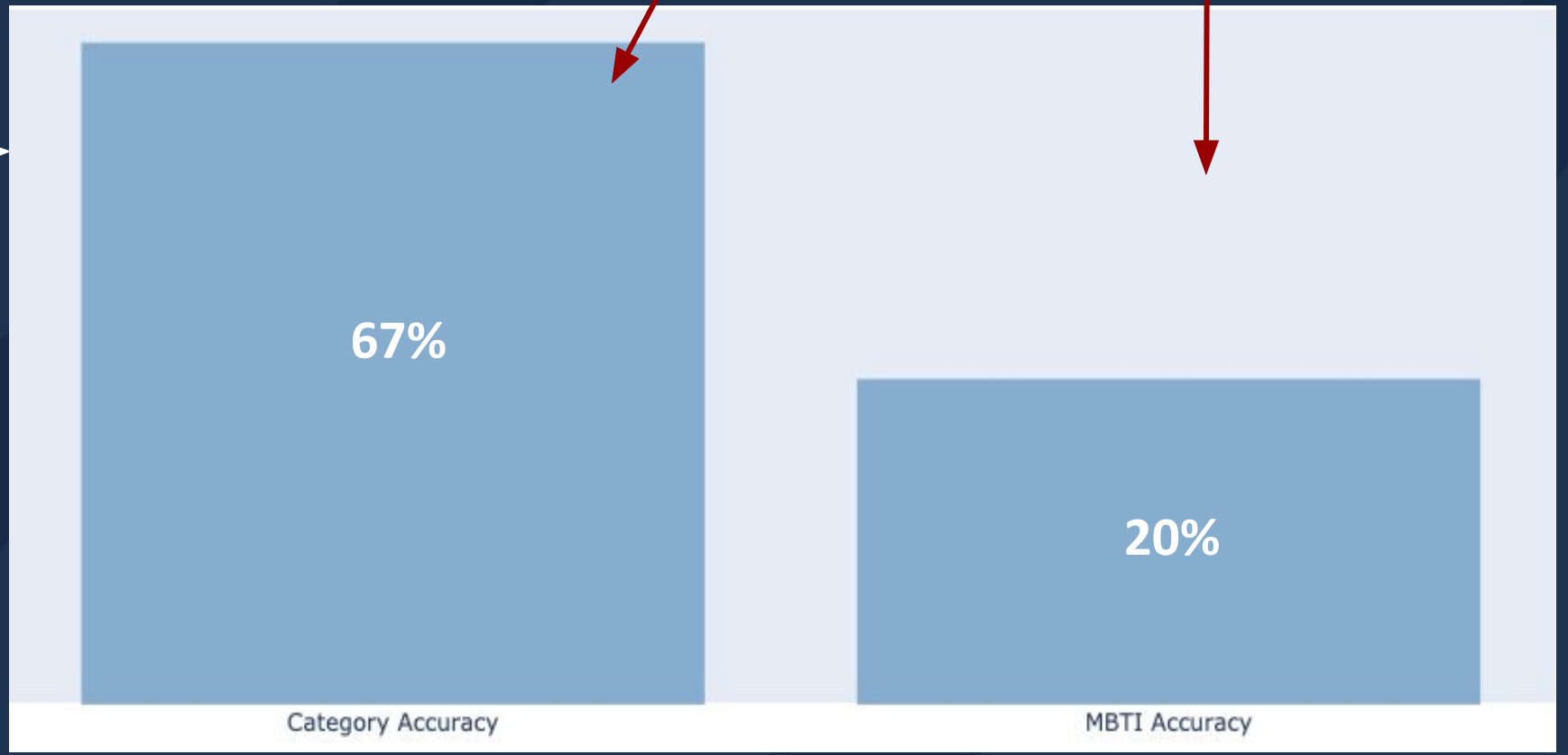
Part 3: Classification

RNN Model Classification Accuracy

Data Label		Two Accuracy Metrics	
True Label	Prediction	Category Accuracy	MBTI Accuracy
“ETPS”	“ETPN”	75%	0%

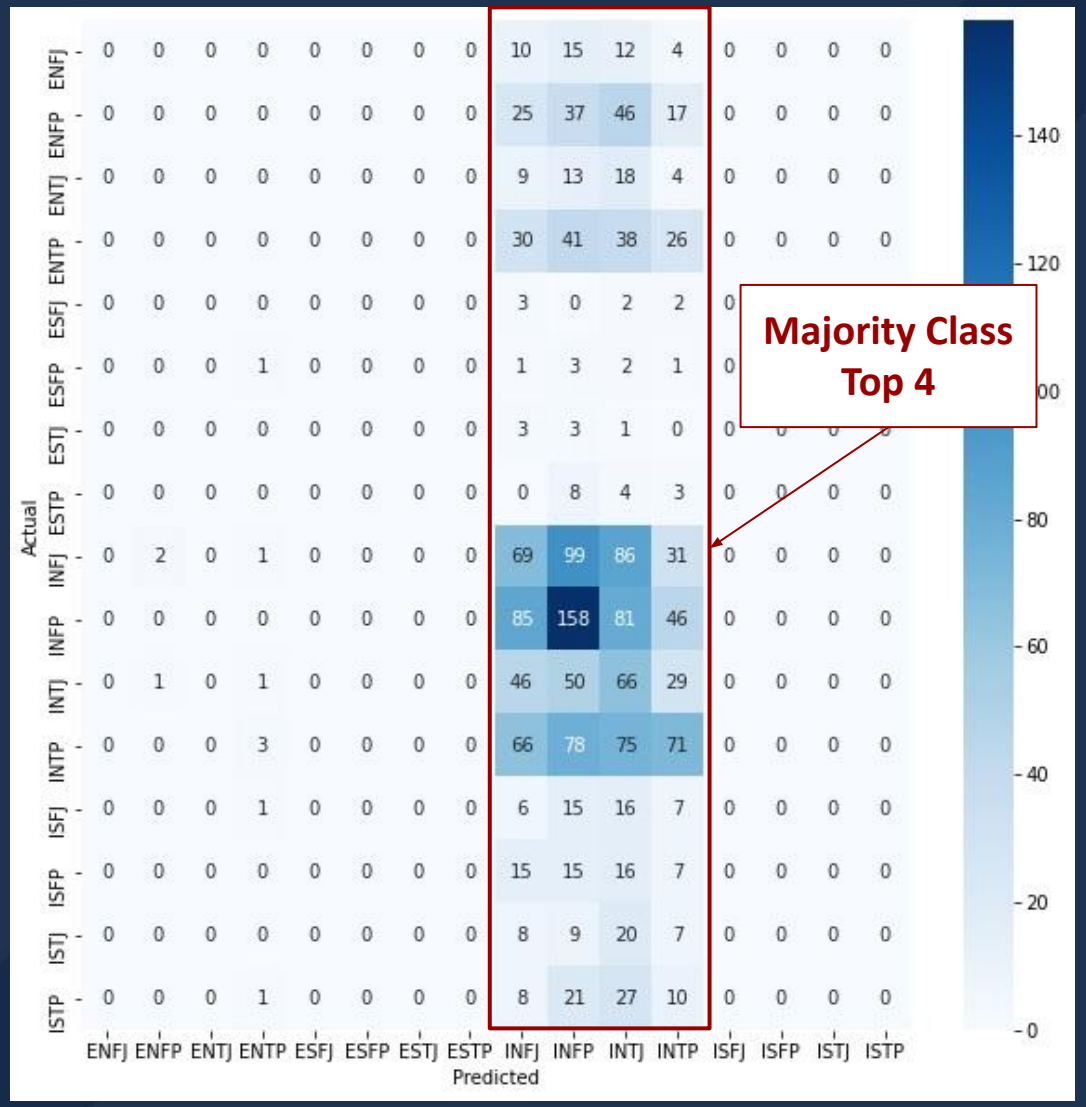
Example →

Test Data →

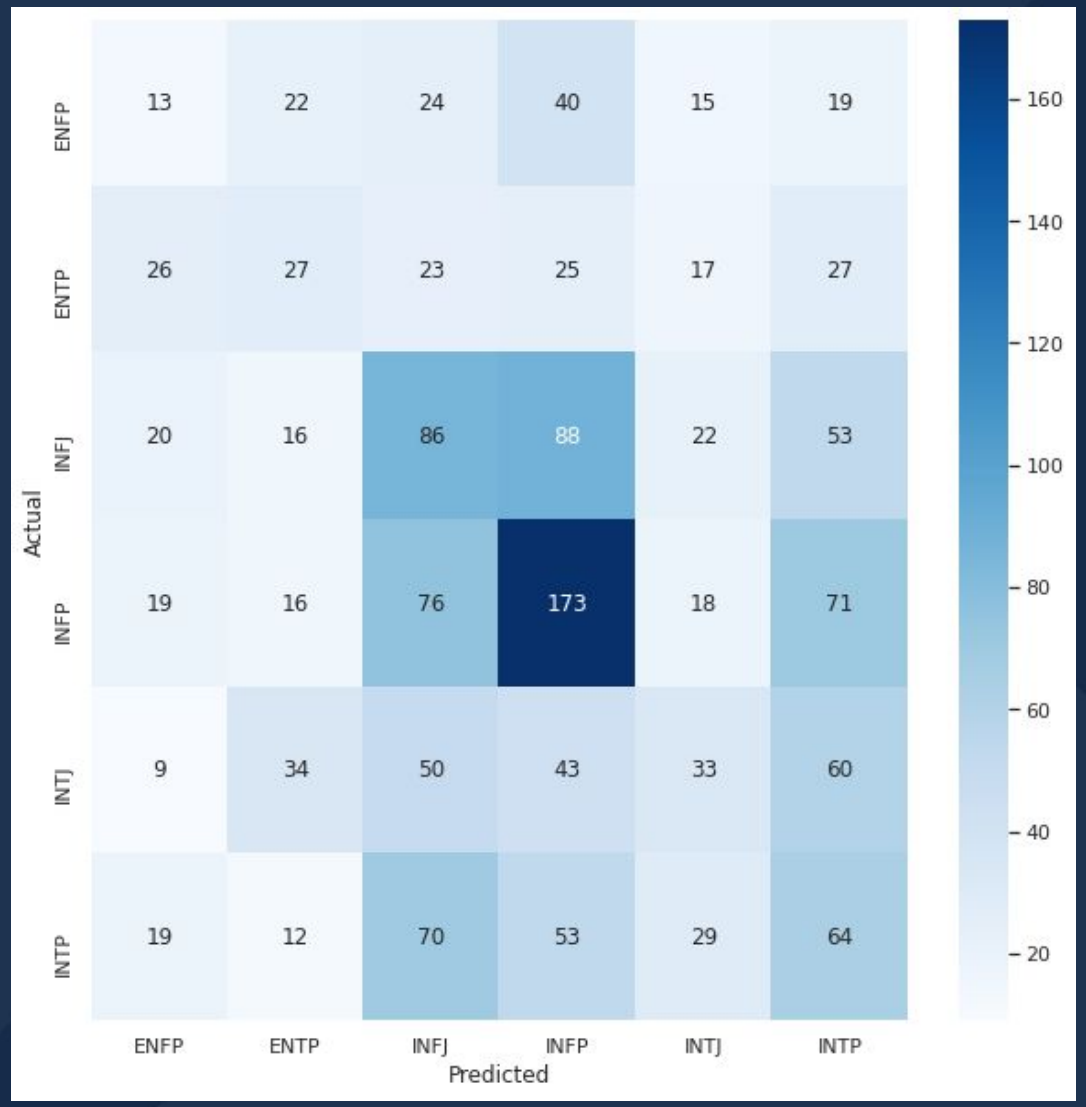


Actual vs Predicted

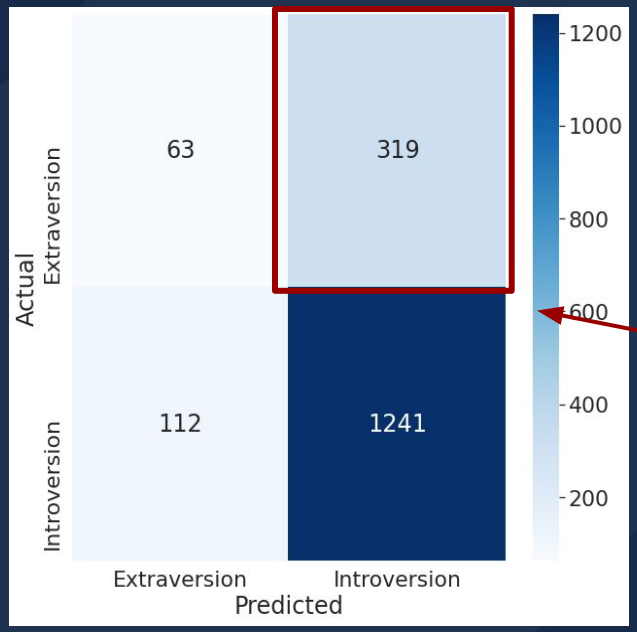
16 Classes: Accuracy 67%/20%



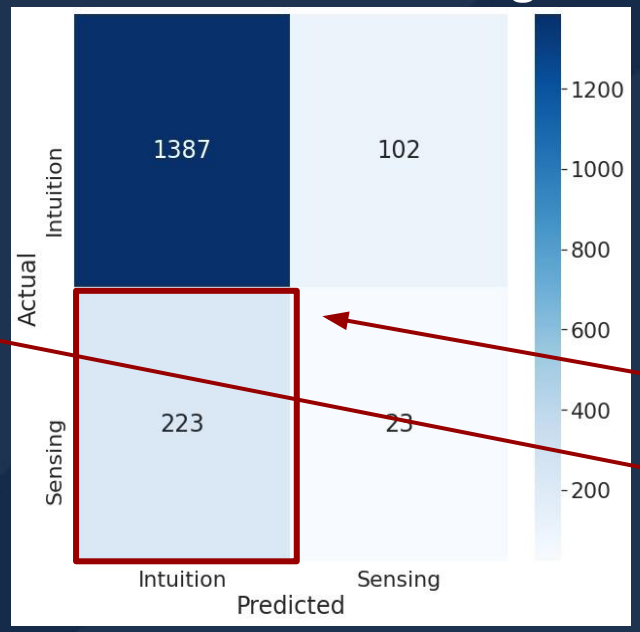
Top 6 Classes: Accuracy 73%/28%



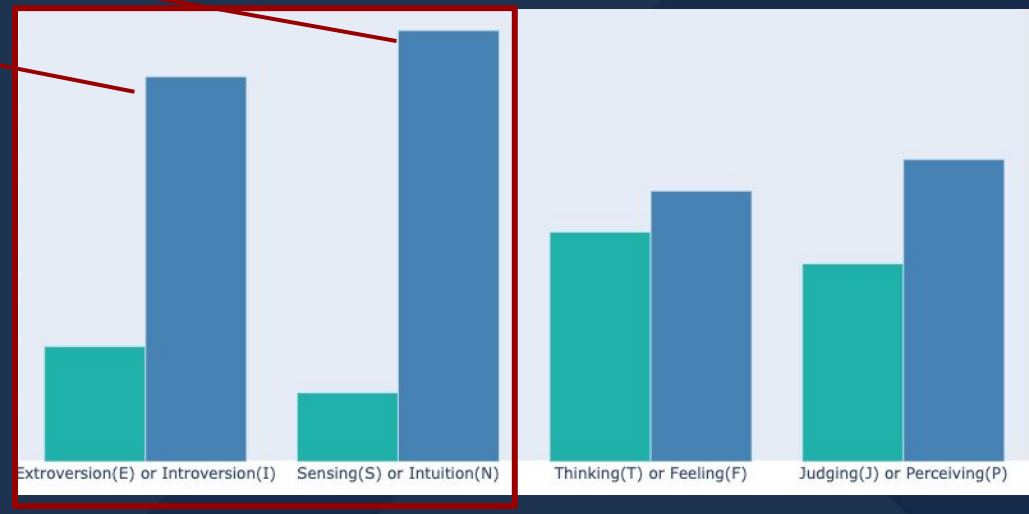
Extraversion vs Introversion



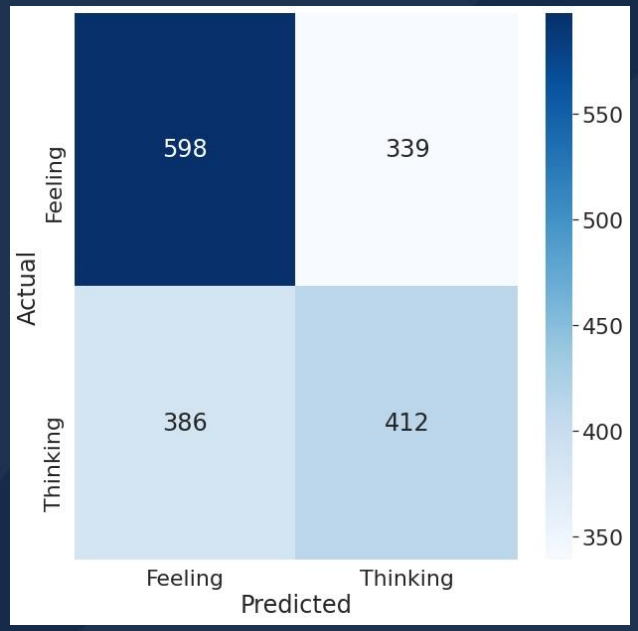
Intuition vs Sensing



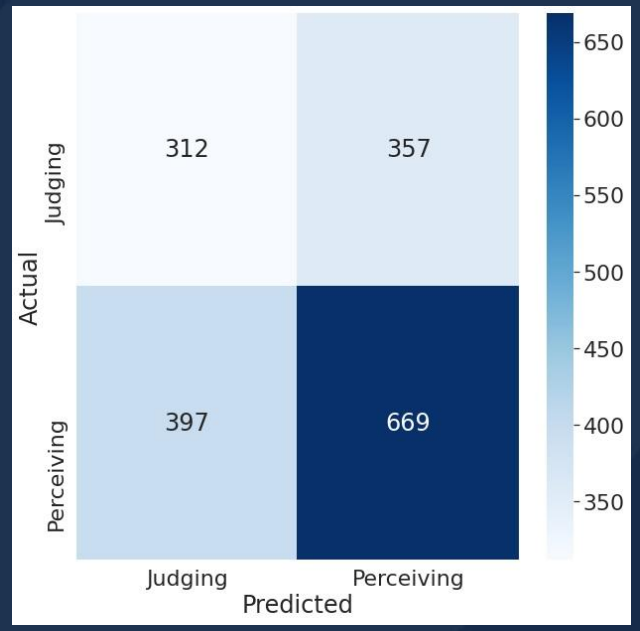
Category Distribution from Slide 7



Feeling vs Thinking



Judging vs Perceiving



Naive Bayes Model Classification

Limitations and Future Works

- Limitations:
  - Unbalanced distribution
  - Small sample size
- Future Works
  - Collect more data
  - Analyze user online behaviours

UC San Diego

Thank You

ECE 143 Winter 2021 Group 14

Fengqi Yang, Luis Herrera, Blue Luu