



Machine Learning for Determining Myers-Briggs Type

CHRIS O'BRIEN, JAY PIKE, CARRIE
SANFORD

Myers-Briggs History

- ▶ Carl Jung – differences in behavior are not random, but due to the different ways people use their perception and judgement
- ▶ Katharine Cook Briggs and Isabel Briggs Myers inspired by those with different ways of seeing the world [8]
- ▶ Briggs & Myers wanted to make Jung's work accessible
- ▶ Believed if people understood one another better, less conflict
- ▶ First published in 1962

Myers-Briggs Type Indicator (MBTI)

- ▶ Purpose – “help people understand personality differences in the general population” to aid in interactions
- ▶ Assigns a personality type to an individual
- ▶ Determined by situational or relational questions

MBTI cont.

- ▶ 16 different types based on following four dimensions:
 - ▶ Introversion (I) vs. Extroversion (E)
 - ▶ Intuition (N) vs. Sensing (S)
 - ▶ Thinking (T) vs. Feeling (F)
 - ▶ Judging (J) vs. Perceiving (P)
- ▶ Example of an ISTP individual
 - ▶ “Tolerant and flexible, quiet observers until a problem appears, then act quickly to find workable solutions”

MBTI cont.

- ▶ Four MBTI forms:
 - ▶ Step I Form M (computer or self-scored): 93 questions
 - ▶ Step II Form Q: 144 questions
 - ▶ Step III: 222 questions
- ▶ **Current methods for determining MBTI are time consuming, a successful machine learning model would significantly reduce user time.**

Relevant Work

| Data | Methods | Accuracy |
|------------------------------------|---|----------|
| Reddit posts, comments, flairs [4] | Linguistic Inquiry and Word Count (LIWC) | 82% |
| Tweets [5] | LIWC, EmoSenticNet, ConceptNet | 88% |
| Social media posts [6] | Naïve-Bayes | 26% |
| | Regularized SVM | 33% |
| | Deep learning | 38% |
| Essays [7] | Mairesse, SenticNet, NRC Emotion Lexicon, VAD Lexicon and BERT, Albert, and Roberts | 61% |
| Social media posts [7] | Mairesse, SenticNet, NRC Emotion Lexicon, VAD Lexicon and BERT, Albert, and Roberts | 77% |

Overview



Objective:

predict MBTI type



Hypothesis:

MBTI type can be predicted from an individual's chat forum post history



Data:

8,675 people

MBTI types

Individuals last 50 posts on PersonalityCafe

Includes text, hyperlinks, GIFs

Problem Split

8



Classify across MB **type**
(n=16)



Classify across
MB **dimension**
(n=2)
E/I, N/S,
TF, J/P

Methods

9

Data Visualization

Data Cleaning

Data Splitting

- Training (70%), validation (15%), Testing (15%)

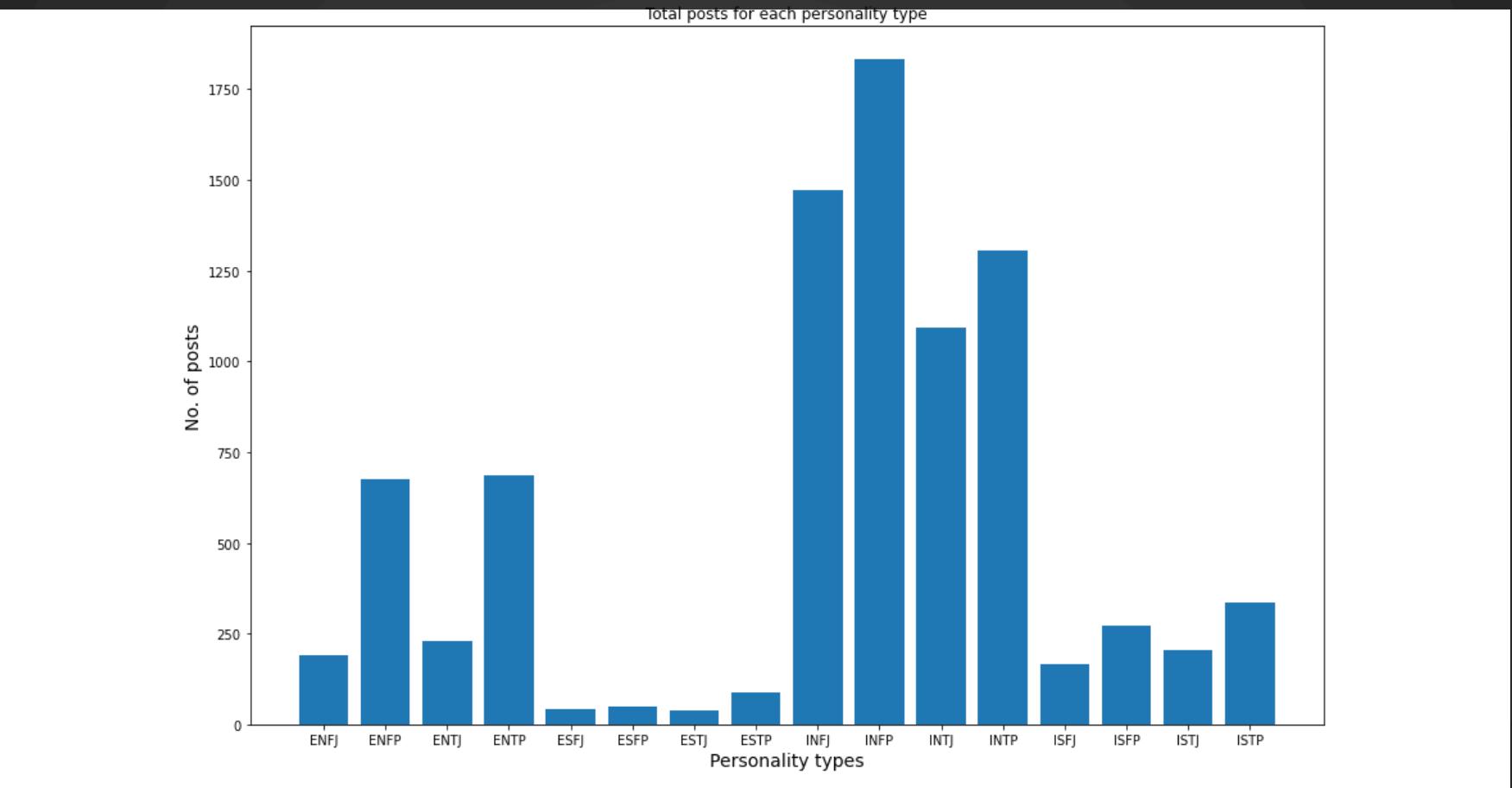
Classification

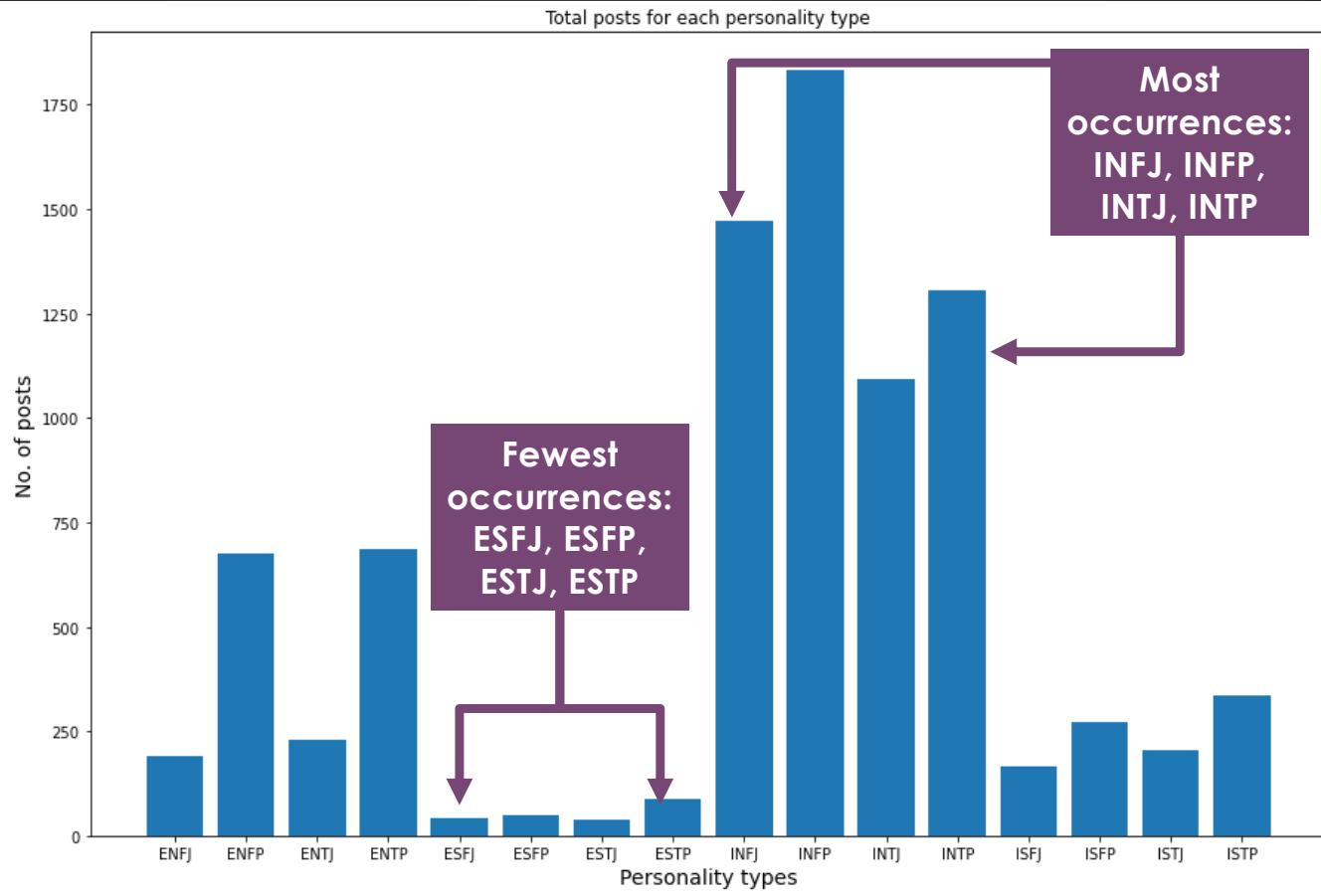
- Logistic Regression, Random Forest, Linear Support Vector Machine, XGBoost

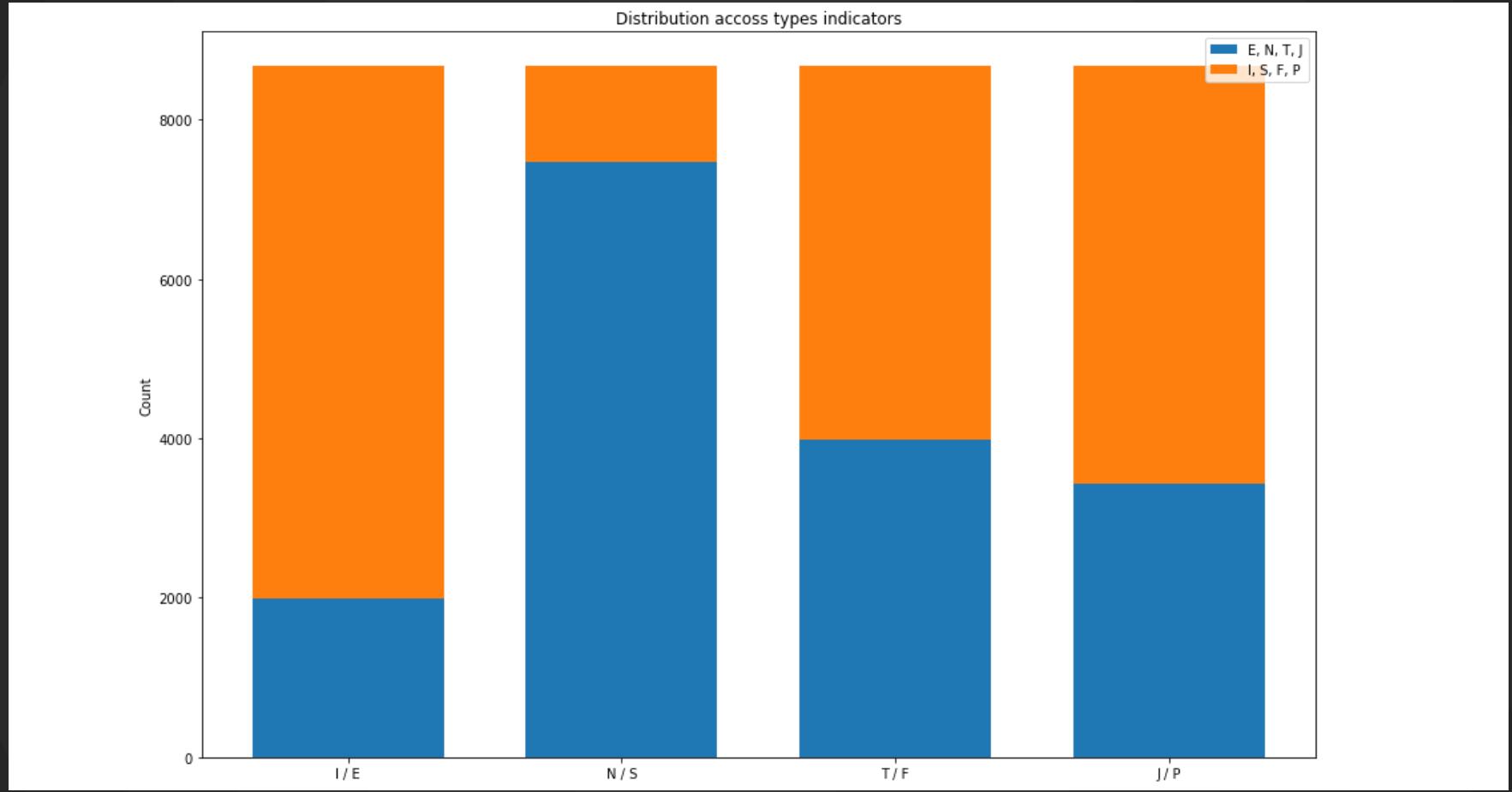
Dataset Description

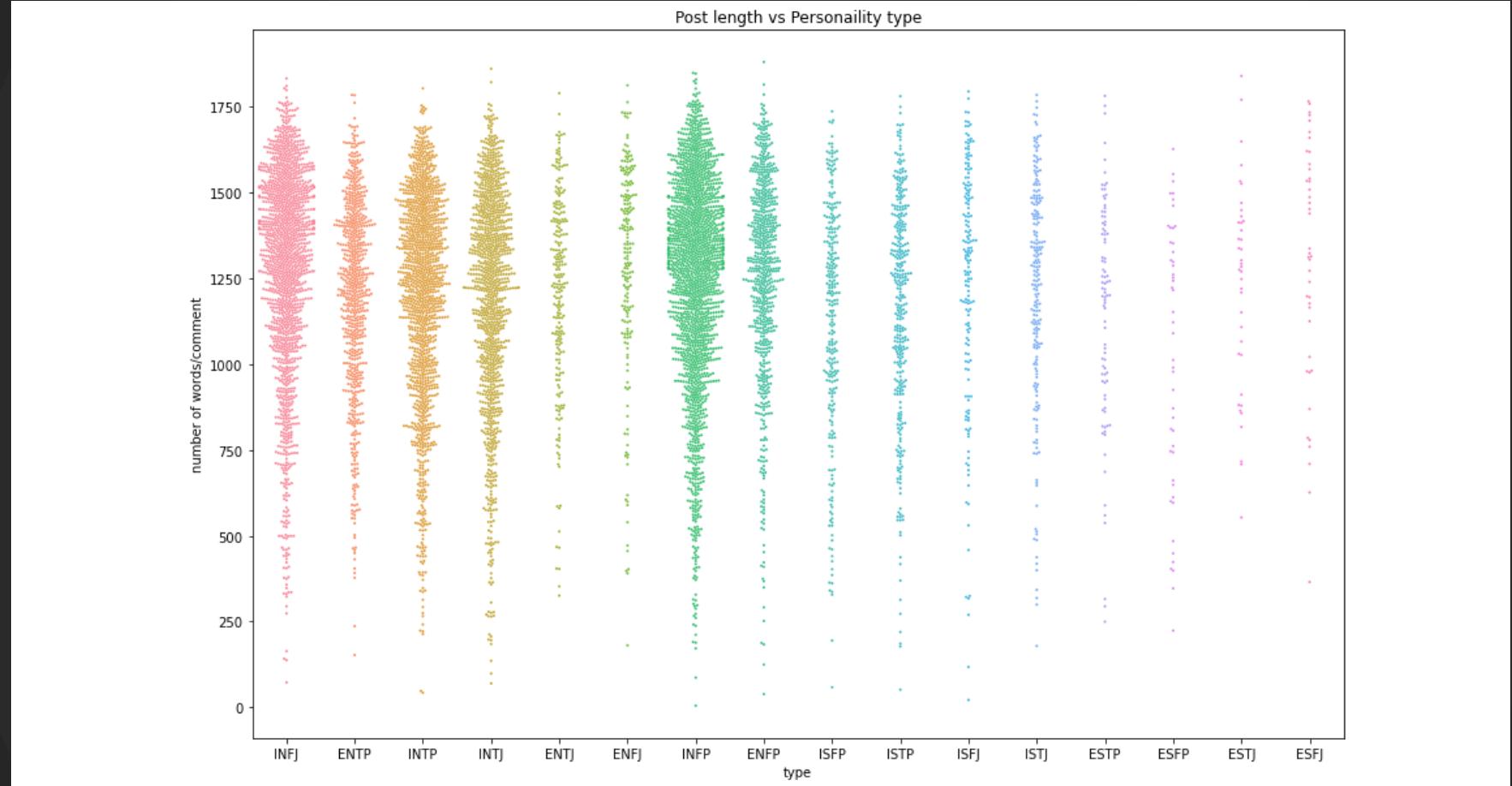
- ▶ Data obtained from Kaggle
- ▶ 8675 rows & 2 columns
 - ▶ 'types' => forum user's MB Type
 - ▶ 'posts' => 50 posts from user
- ▶ Posts separated by | | |
- ▶ ~450,000 posts total

'http://www.youtube.com/watch?v=qsXHcwe3krw|||http://41.media.tumblr.com/tumblr_lfouy03PMA1qa1roo1_500.jpg|||enfp and intj moments https://www.youtube.com/watch?v=iZ7lE1g4XM4 sportscenter not top ten plays https://www.youtube.com/watch?v=uCdfze1etepranks|||What has been the most life-changing experience in your life?|||http://www.youtube.com/watch?v=vXZeYwwRDw8 http://www.youtube.com/watch?v=u8ejam5DP3E On repeat for most of today.|||May the PerC Experience immerse you.|||The last thing my INFJ friend posted on his facebook before committing suicide the next day. Rest in peace~ http://vimeo.com/22842206|||Hello ENFJ. Sorry to hear of your distress. It's only natural for a relationship to not be perfection all the time in every moment of existence. Try to figure the hard times as times of growth, as...|||84389 84390 http://wallpaperpassion.com/upload/23700/friendship-boy-and-girl-wallpaper.jpg http://assets.dornob.com/wp-content/uploads/2010/04/round-home-design.jpg ...|||Welcome and stuff.|||http://playeressence.com/wp-content/uploads/2013/08/RED-red-the-pokemon-master-32560474-450-338.jpg Game, Set, Match.|||Prozac, wellbrutin, at least thirty minutes of moving your legs (and I don't mean moving them while sitting in your same desk chair), weed in moderation (maybe try edibles as a healthier alternative...)|||Basically come up with three items you've determined that each type (or whichever types you want to do) would more than likely use, given each types' cognitive functions and whatnot, when left by...|||All things in moderation. Sims is indeed a video game, and a good one at that. Note: a good one at that is somewhat subjective in that I am not completely promoting the death of any given Sim...|||Dear ENFP: What were your favorite video games growing up and what are your now, current favorite video games? :cool:|||https://www.youtube.com/watch?v=QyPqT8umzrnY|||It appears to be too late. :sad:|||There's someone out there for everyone.|||Wait... I thought confidence was a good thing.|||I just cherish the time of solitude b/c I revel within my inner world more whereas most other time i'd be workin... just enjoy the me time while you can. Don't worry, people will always be around to...|||Yo entp ladies... if you're into a complimentary personality,well, hey.|||... when your main social outlet is xbox live conversations and even then you verbally fatigue quickly.|||http://www.youtube.com/watch?v=gDhy7rdfm14 I real2:50|||http://www.youtube.com/watch?v=msqXffgh7b8|||Banned because this thread requires it of me.|||Get high in backyard, roast and eat marshmallows in backyard while conversing over something intellectkisses.|||http://www.youtube.com/watch?v=Mw7eoU3BMBE|||http://www.youtube.com/watch?v=4V2uYORhQOk|||http://www.youtube.com/watch?v=SlVmgFQQ0T|||Banned for too many b's in that sentence. How could you! Think of the B!|||Banned for watching movies in the corner with the dunces.|||Banned because Health class clearly taught you nothing about peer pressure.|||Banned for a whole host of reasons!|||http://www.youtube.com/watch?v=lRcrv41hgz4|||1) Two baby deer on left and right munching on a beetle in the middle.|||2) Using their own blood, two cavemen diary today's latest happenings on their designated cave diary wall.|||3) I see it as...|||pokemon world an infj society everyone beoptimist|||49142|||http://www.youtube.com/watch?v=zRCEq_JFeFM|||http://discovermagazine.com/2012/jul-aug/20-things-you-didnt-know-about-deserts/desert.jpg|||http://oyster.ignimgs.com/mediawiki/apis.ign.com/pokemon-silver-version/d/dd/Ditto.gif|||http://www.serebii.net/potw-dp/Scizor.jpg|||Not all artists are artists because they draw. It's the idea that counts in forming something of your own... like a signature.|||Welcome to the robot ranks, person who downed my self-esteem cuz I'm not an avid signature artist like herself. :proud:|||Banned for taking all the room under my bed. Ya gotta learn to share with troaches.|||http://www.youtube.com/watch?v=w8lgmn57aQ|||Banned for being too much of a thundering, grumbling kind of storm... yep.|||Ahh... old high school music I haven't heard in ages.|||http://www.youtube.com/watch?v=dcCRUPCdB1w|||I failed a public speaking class a few years ago and I've sort of learned what I could do better were I to be in that position again. A big part of my failure was just overloading myself with too...|||I like this person's mentality. He's a confirmed INTJ by the way. http://www.youtube.com/watch?v=hGKLIGEc6M|||Move to the Denver area and start a new life for myself.'









Discuss Dataset Limitations

| Personality Type | Actual Percentage (%) | Dataset Percentage (%) | Difference |
|------------------|-----------------------|------------------------|------------|
| INTJ | 2.1 | 12.6 | 10.5 |
| INTP | 3.3 | 15 | 11.7 |
| ENTJ | 1.8 | 2.7 | 0.9 |
| ENTP | 3.2 | 7.9 | 4.7 |
| INFJ | 1.5 | 16.9 | 15.4 |
| INFP | 4.4 | 21.1 | 16.7 |
| ENFJ | 2.5 | 2.2 | 0.3 |
| ENFP | 8.1 | 7.8 | 0.3 |
| ISTJ | 11.6 | 2.4 | 9.2 |
| ISFJ | 13.8 | 1.9 | 11.9 |
| ESTJ | 8.7 | 0.4 | 8.3 |
| ESFJ | 12 | 0.5 | 11.5 |
| ISTP | 5.4 | 3.9 | 1.5 |
| ISFP | 8.8 | 3.1 | 5.7 |
| ESTP | 4.3 | 1.0 | 3.3 |
| ESFP | 8.5 | 0.6 | 7.9 |

Data Cleaning

Misleading Use of Personality Types

ENTP - "Your comment screams INTJ, bro."

INTP - "I'm dating a guy that I'm about 95% sure is an INFJ"

Careful Sequence of Preprocessing

"The INFp sounds exactly like me :) Thank you."

"I think the real question is whether INFJs can be NEAT."

Data Cleaning / Feature Extraction

Cleaning

- ▶ Lowercase
- ▶ Remove personality types
- ▶ Remove Links
- ▶ Remove non-words (numbers, etc)
- ▶ Remove punctuation
- ▶ Remove extra spaces

Label encoding

Feature Extraction

Performed TFIDF

N-grams 1, 2

▶ Minimum Word Frequency = 3

Before Cleaning

'http://www.youtube.com/watch?v=qSXHcwe3krw' || http://41.media.tumblr.com/tumblr_lfouy03PMA1qa1roo1_500.jpg' || enfp and intj moments https://www.youtube.com/watch?v=iZ7lE1gXMA4 sportscenter not top ten plays https://www.youtube.com/watch?v=uCdfze1teec pranks' || What has been the most life-changing experience in your life?' || http://www.youtube.com/watch?v=vXZeYwwrDw8 http://www.youtube.com/watch?v=u8ejam5DP3E On repeat for most of today.' || May the PerC Experience immerse you.' || The last thing my INFJ friend posted on his facebook before committing suicide the next day. Rest in peace' http://vimeo.com/22842206' || Hello ENFJ. Sorry to hear of your distress. It's only natural for a relationship to not be perfection all the time in every moment of existence. Try to figure the hard times as times of growth, as... ||| 84389 84390 http://wallpaperpassion.com/upload/23700/friendship-boy-and-girl-wallpaper.jpg http://assets.dornob.com/wp-content/uploads/2010/04/round-home-design.jpg ... ||| Welcome and stuff.' ||| http://playeressence.com/wp-content/uploads/2013/08/RED-red-the-pokemon-master-32560474-450-338.jpg Game. Set. Match.' ||| Prozac, wellbutrin, at least thirty minutes of moving your legs (and I don't mean moving them while sitting in your same desk chair), weed in moderation (maybe try edibles as a healthier alternative... ||| Basically come up with three items you've determined that each type (or whichever types you want to do) would more than likely use, given each types' cognitive functions and whatnot, when left by... ||| All things in moderation. Sims is indeed a video game, and a good one at that. Note: a good one at that is somewhat subjective in that I am not completely promoting the death of any given Sim... ||| Dear ENFP: What were your favorite video games growing up and what are your now, current favorite video games? :cool: ||| https://www.youtube.com/watch?v=QyPqt8umzny' ||| It appears to be too late. ssd: ||| There's someone out there for everyone.' ||| Wait... I thought confidence was a good thing.' ||| I just cherish the time of solitude b/c revel within my inner world more whereas most other time I'd be workin... just enjoy the me time while you can. Don't worry, people will always be around to... ||| Yo ent ladies... if you're into a complimentary personality, well, hey.' ||| when your main social outlet is xbox live conversations and even then you verbally fatigue quickly.' ||| http://www.youtube.com/watch?v=gDhy7rdfm14 I real2:50' ||| http://www.youtube.com/watch?v=msqffgh/b8' ||| Banned because this thread requires it of me.' ||| Get high in backyard, roast and eat marshmallows in backyard while conversing over something intellectus.' ||| http://www.youtube.com/watch?v=Mw7eoU3BMe' ||| http://www.youtube.com/watch?v=4V2uYOrhQOk' ||| http://www.youtube.com/watch?v=SlvMgFQQ0T' ||| Banned for too many b's in that sentence. How could you! Think of the B!' ||| Banned for watching movies in the corner with the dunces.' ||| Banned because Health class clearly taught you nothing about peer pressure.' ||| Banned for a whole host of reasons! ||| http://www.youtube.com/watch?v=Ircv41hg24' ||| Two baby deer on left and right munching on a beetle in the middle.' 2) Using their own blood, two cavemen diary today's latest happenings on their designated cave diary wall.' 3) I see it as... ||| a pokemon world an infj society everyone beoptimist' ||| 49142' ||| http://www.youtube.com/watch?v=ZRCEq_JFeFM' ||| http://discovermagazine.com/2012/jul-aug/20-things-you-didnt-know-about-deserts/desert.jpg' ||| http://oyster.gnring.com/mediawiki/apis.ign.com/pokemon-silver-version/d/dd/Ditto.gif' ||| http://www.serebii.net/potw-dp/Scizor.jpg' ||| Not all artists are artists because they draw, it's the idea that counts in forming something of your own... like a signature.' ||| Welcome to the robot ranks, person who downed my self-esteem cuz I'm not an avid signature artist like herself. :proud:' ||| Banned for taking all the room under my bed. Ya gotta learn to share with troaches.' ||| http://www.youtube.com/watch?v=w8lgimn57aQ' ||| Banned for being too much of a thundering, grumbling kind of storm... yep.' ||| Ahh... old high school music I haven't heard in ages.' http://www.youtube.com/watch?v=dCCRUPCb1w' ||| failed a public speaking class a few years ago and I've sort of learned what I could do better were I to be in that position again. A big part of my failure was just overloading myself with too... ||| I like this person's mentality. He's a confirmed INTJ by the way.' http://www.youtube.com/watch?v=hGKL-GEc6M' ||| Move to the Denver area and start a new life for myself.'

After Cleaning



moments sportscenter plays pranks lifechanging experience life repeat today perc experience immerse thing friend posted facebook committin g suicide day rest peace hello sorry hear distress natural relationship perfection time moment existence try figure hard times times growth welc ome stuff game set match prozac wellbutrin thirty minutes moving legs dont mean moving sitting desk chair weed moderation maybe try edible healthier alternative basically come items youve determined type whichever types want likely use given types cognitive functions whatnot left t things moderation sims video game good note good somewhat subjective completely promoting death given sin dear favorite video games gro wing current favorite video games cool appears late sad theres wait thought confidence good thing cherish time solitude bc revel inner world ti me id workin enjoy time dont worry people yo ladies youre complimentary personalitywell hey main social outlet xbox live conversations verbal ly fatigue quickly dig banned thread requires high backyard roast eat marshmallows backyard conversing intellectual followed massages kisses banned bs sentence thinkbanned watching movies corner dunces banned health class clearly taught peer pressure banned host reasons baby d eer left right munching beetle middle blood cavemen diary todays latest happenings designated cave diary wall pokemon world society optimist artists artists draw idea counts forming like signature welcome robot ranks person downed selfesteem cuz im avid signature artist like proud ba nned taking room bed ya gotta learn share roaches banned thundering grumbling kind storm yep ah old high school music havent heard age fa illed public speaking class years ago ive sort learned better position big failure overloading like persons mentality hes confirmed way denver are a start new life

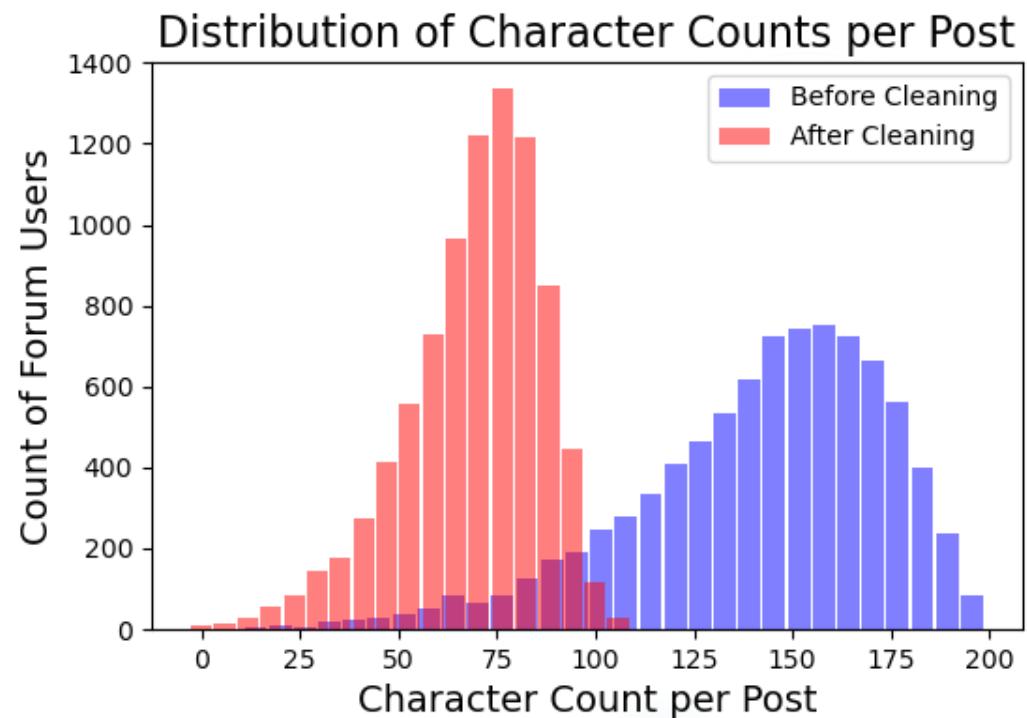
Data Cleaning

Before Cleaning:

- Avg Number of Characters per Post = 145 (~25 words)

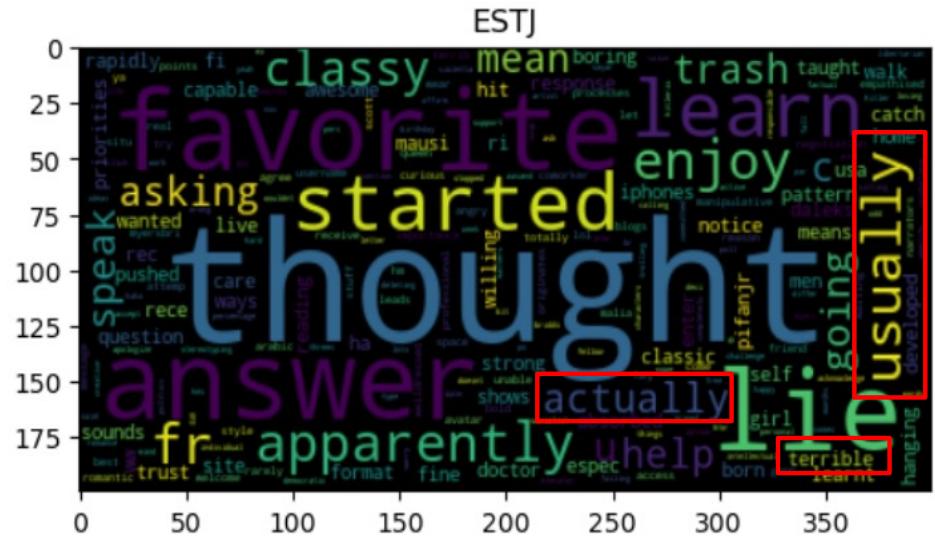
After Cleaning:

- Avg Number of Characters per Post = 72 (~12 words)



Common Words for ESTJ's

- ▶ Typically leaders
- ▶ Analytical, practical
- ▶ Vocal about opinion



ML Models

21

Logistic Regression

Linear SVM

Random Forest

XGBoost

Hyperparameter Tuning

22

sklearns GridsearchCV
using the validation set

20 Total Tuned Models

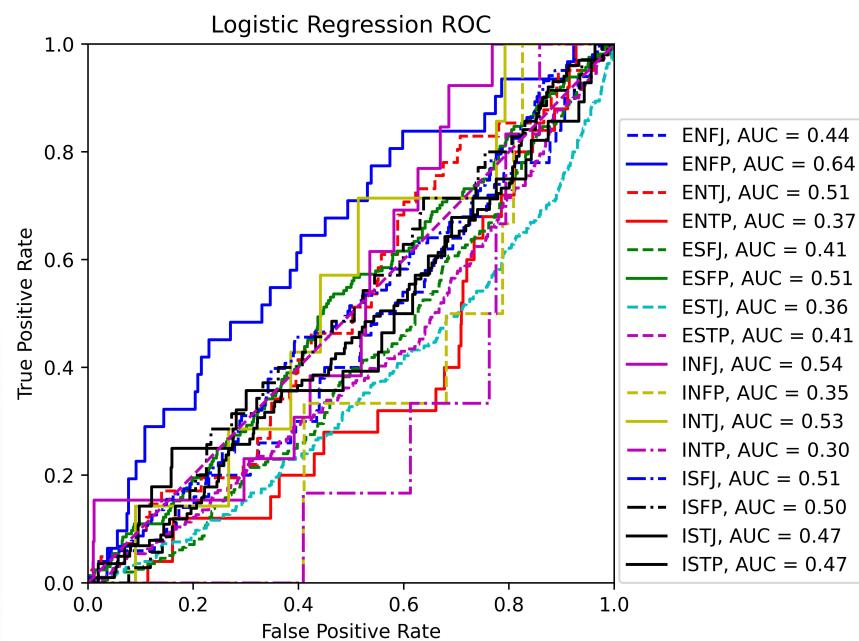
N = 16 (4
models)

N = 2 (16
models)

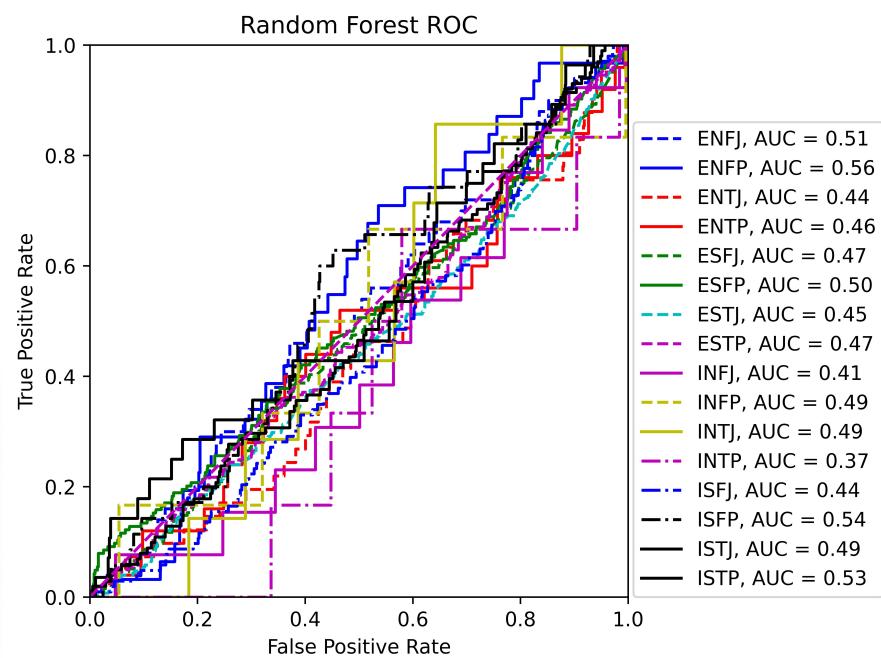
Classification Results (n = 16)

| Algorithm | Accuracy | Train Time (s) | Test Time (s) |
|---------------------|---------------|----------------|---------------|
| Logistic Regression | 0.4155 | 13.86 | 0.0221 |
| Linear SVM | 0.4237 | 19.97 | 0.0237 |
| Random Forest | 0.2919 | 25.39 | 0.4674 |
| XGBoost | 0.3648 | 186.8 | 0.132 |

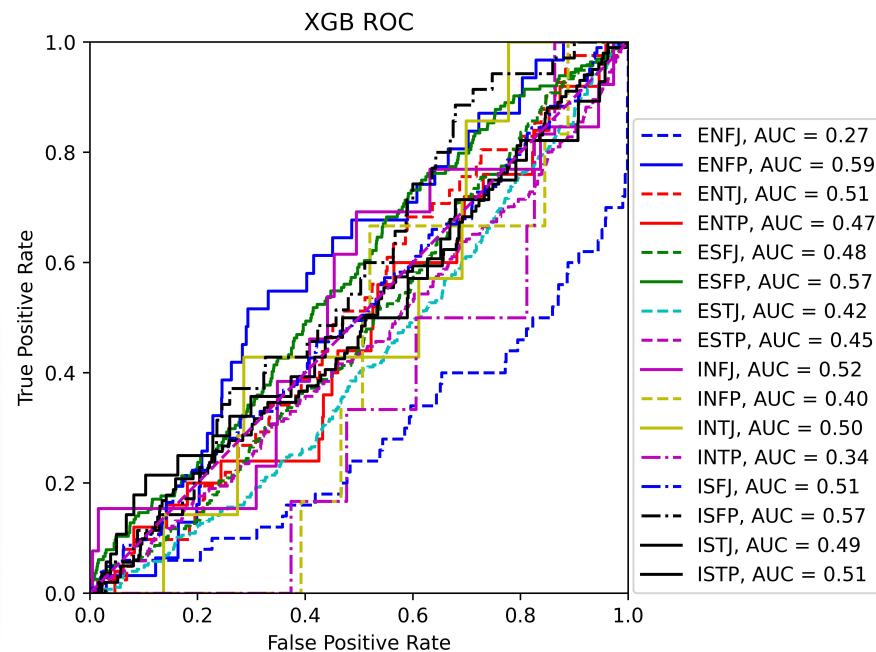
ROC (n = 16)



ROC (n = 16) cont.



ROC (n = 16) cont.



Classification Results (n = 2)

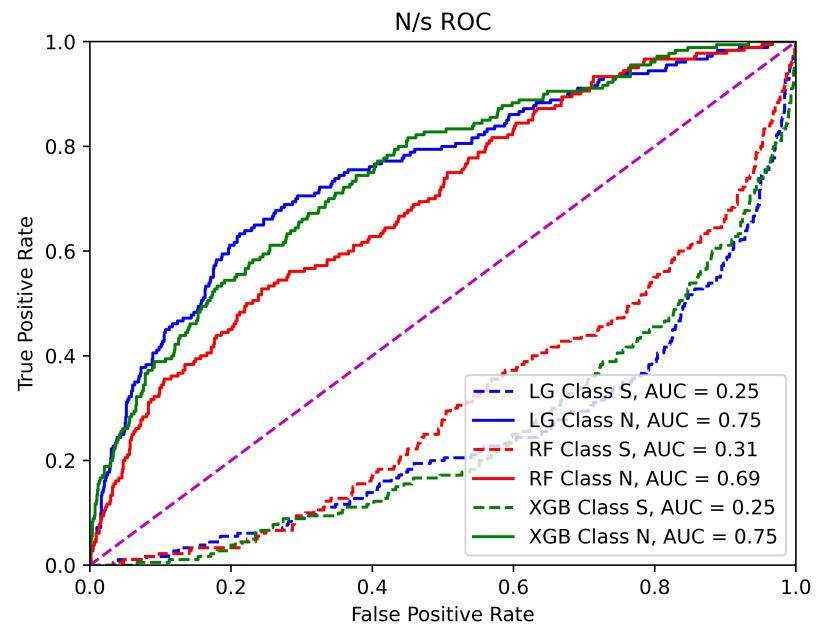
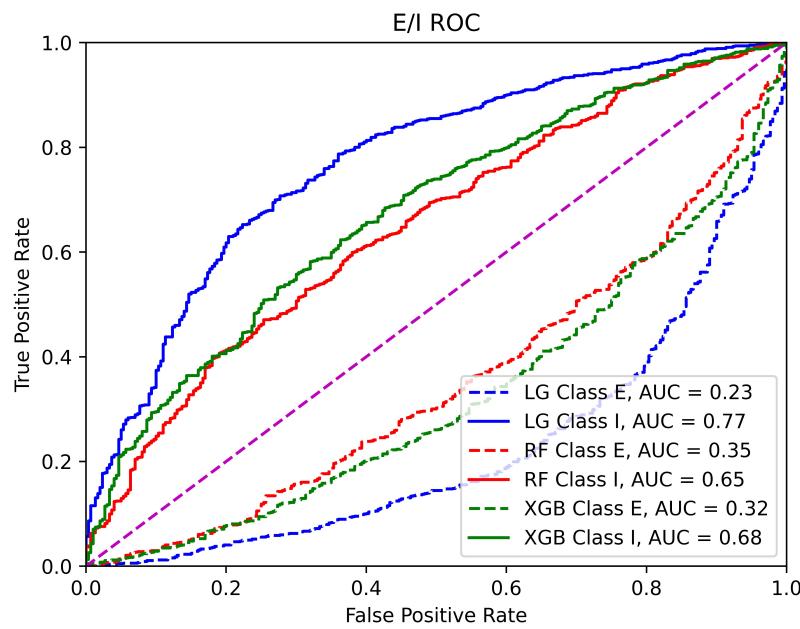
Introversion / Extroversion

| Algorithm | Accuracy | Train Time (s) | Test Time (s) |
|---------------------|---------------|----------------|---------------|
| Logistic Regression | 0.7842 | 1.0319 | 0.0022 |
| Linear SVM | 0.7696 | 0.1627 | 0.0021 |
| Random Forest | 0.7696 | 1.551 | 0.3694 |
| XGBoost | 0.7696 | 38.2684 | 0.1381 |
| Classify all as I | 0.7696 | | |

Intuition / Sensing

| Algorithm | Accuracy | Train Time (s) | Test Time (s) |
|---------------------|---------------|----------------|---------------|
| Logistic Regression | 0.8648 | 2.040 | 0.0026 |
| Linear SVM | 0.8702 | 7.256 | 0.0027 |
| Random Forest | 0.8618 | 1.806 | 0.3587 |
| XGBoost | 0.8679 | 32.20 | 0.1341 |
| Classify all as N | 0.862 | | |

ROC (n=2)



Classification Results (n = 2) cont.

Thinking / Feeling

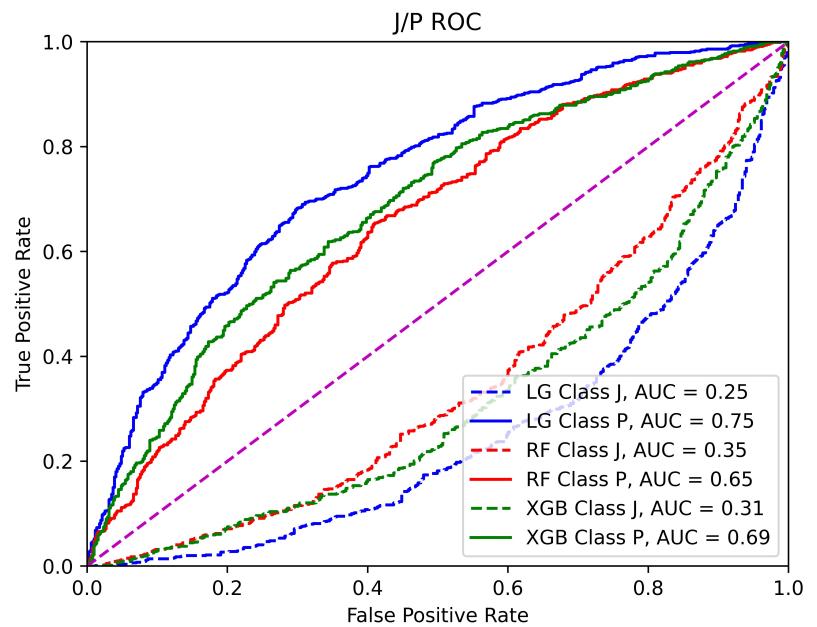
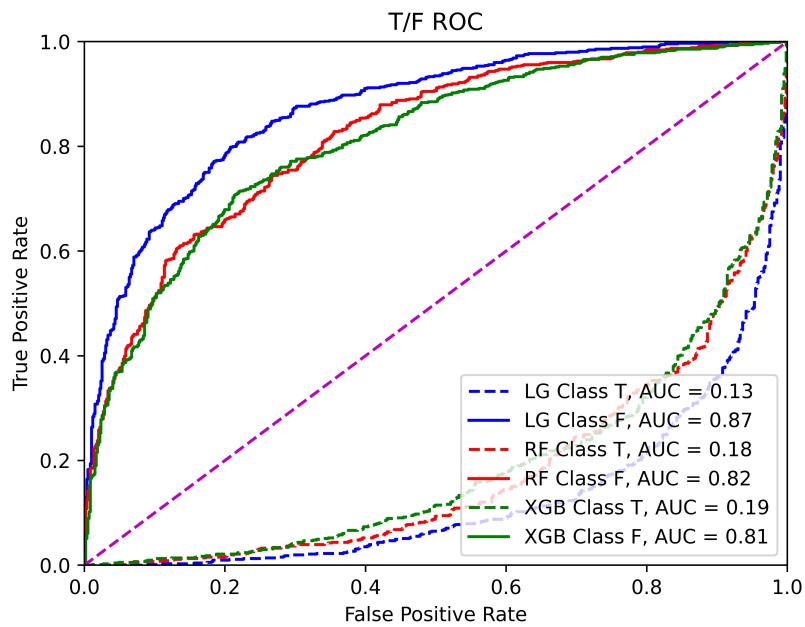
| Algorithm | Accuracy | Train Time (s) | Test Time (s) |
|---------------------|---------------|----------------|---------------|
| Logistic Regression | 0.7926 | 0.9804 | 0.0022 |
| Linear SVM | 0.7811 | 2.0433 | 0.0023 |
| Random Forest | 0.7358 | 190.4 | 3.750 |
| XGBoost | 0.7404 | 26.34 | 0.1297 |

Judging / Perceiving

| Algorithm | Accuracy | Train Time (s) | Test Time (s) |
|---------------------|---------------|----------------|---------------|
| Logistic Regression | 0.6966 | 1.185 | 0.0027 |
| Linear SVM | 0.6935 | 2.709 | 0.0024 |
| Random Forest | 0.6175 | 15.93 | 0.4118 |
| XGBoost | 0.6536 | 44.10 | 0.1456 |

30

ROC (n=2)



Summary

- Tested ML pipelines to generate of Myers-Briggs type using social media posts
- Low classification accuracies (maximum 42%)
- More accurate results than prior work with same dataset (26%)
- More balanced dataset would likely yield a more accurate model

References

1. Anon. MBTI® basics. Retrieved November 5, 2021 from <https://www.myersbriggs.org/my-mbti-personality-type/mbti-basics/>
2. Anon. MBTI Facts. Retrieved November 5, 2021 from <https://www.themyersbriggs.com/en-US/Support/MBTI-Facts>
3. Mitchell J. 2017. (MBTI) myers-briggs personality type dataset. (September 2017). Retrieved November 5, 2021 from <https://www.kaggle.com/datasnaek/mbti-type>
4. Gjurković, M. and Šnajder, J., 2018, June. Reddit: A gold mine for personality prediction. In *Proceedings of the Second Workshop on Computational Modeling of People's Opinions, Personality, and Emotions in Social Media* (pp. 87-97).
5. Bharadwaj, S., Sridhar, S., Choudhary, R. and Srinath, R., 2018, September. persona traits identification based on myers-briggs type indicator (MBTI)-a text classification approach. In *2018 International Conference on Advances in Computing, Communications and Informatics (ICACCI)* (pp. 1076-1082). IEEE.
6. Cui, B. and Qi, C., 2017. persona traits identification based on myers-briggs type indicator (MBTI)-a text classification approach.
7. Khan, A.S., Ahmad, H., Asghar, M.Z., Saddozai, F.K., Arif, A. and Khalid, H.A., 2020. Personality Classification from Online Text using Machine Learning Approach. *International Journal of Advanced Computer Science and Applications*, 11(3).
8. <https://eu.themyersbriggs.com/en/tools/MBTI/Myers-Briggs-history>