


Machine Learning Model Predicting Myers-Briggs personality type from... *social media posts.*


Today, I'm a Director of User Experience at Google, and an aspiring student of Artificial Intelligence at University of Colorado Boulder, but many years ago I studied...
Psychology .



*Young me after graduation with
MA in Cognitive Psychology*

I found on Kaggle the Myers-Briggs personality test results matched with social media posts. And that gave me a thought of creating a model predicting personality type based on users text.

 Search

 MITCHELL J · UPDATED 8 YEARS AGO

(MBTI) Myers-Briggs P

Includes a large number of people's MBTI type and o

Data Card

Code (126)

Discussion (5)

Suggest

About Dataset

Context

The Myers Briggs Type Indicator (or MBTI for short) is a pers
4 axis:

- Introversion (I) – Extroversion (E)
- Intuition (N) – Sensing (S)
- Thinking (T) – Feeling (F)
- Judging (J) – Perceiving (P)

[\(More can be learned about what these mean here\)](#)

So for example, someone who prefers introversion, intuition,
are lots of personality based components that would model

It is one of, if not the, the most popular personality test in the
simple google search reveals all of the different ways the tes
world in terms of its use.

From scientific or psychological perspective it is based on th
model of 8 distinct functions, thought processes or ways of
transformed into several different personality systems to ma

Recently, its use/validity has come into question because of
to as being a very useful tool in a lot of areas, and the purpo
and their style of writing, which overall explores the validity

What is Myers-Briggs (MBTI)?

It's a personality test! MBTI was developed in the 1940s by American Psychologists – Katherine Cook Briggs and Isabel Briggs Meyers. The test identifies 16 personality types which are grouped by four pairs of opposite preferences:

- extraversion (E) or introversion (I),
- sensing (S) or intuition (N),
- thinking (T) or feeling (F),
- and judging (J) or perceiving (P).



Goal

Building a supervised model classifying any text provided as a sign of one of the MBTI personality types.

Research Plan

1. EAD. After loading the data set I'm going to thoroughly evaluate it. I'll pay a lot of attention to how balanced the data set is (number of observations per type).
2. Data processing. I assume that I'll have to process the data to clean it up and prepare it for training of a machine learning model.
3. Vectorization. The text data will have to be vectorized to be used by my supervised model.
4. Model training. Based on the aforementioned analysis I'll decide which supervised model, or models, should I use.
5. Model tuning. I'll analyze the performance of the model and try to tune it.
6. Inference. I'm going to write a little inference system and test the final model on an additional data set – my own social media posts.