SCHOOL OF COMPUTER SCIENCE

UNIVERSITY OF ADELAIDE

Student:
Deepesh Kumar Malik
A1804938

Using Machine Learning to predict Myers–Briggs Type Indicator

ABSTRACT

Our aim for this project is to use Machine learning algorithms which will classify people into personality type denoted by Myers-Briggs Type Indicator (MBTI) on the basis of their posts on social media. In this project, author aims at experimenting on dataset available on Kaggle with different embedding techniques like Bag of Words, TF-IDF, and word2Vec, combined with different machine learning models like XGBoost, Support Vector Machines, Logistic Regression and Bert.

* Corresponding author
Email addresses: deepesh.malik89@gmail.com (Deepesh Kumar Malik)

CONTENTS

I. INTRODUCTION

In real world, we often hear our friends or coworkers describe themselves as "I'm an INFJ or I'm an ENTP!" or some other seemingly odd and random four-letter combination. In most cases they're talking about their MBTI test results.

With increased focus and ongoing researches in psychology domain, the very concept of an individual personality is a very strong but indistinct construct. Hence, there is a requirement of a much more conclusive, pragmatic measures of existent models of personality. In this project, author aims to work on the understanding of one of the personality testing model: MBTI (Myers-Briggs Type Indicator). This project focuses on building an NLP algorithm, which will process on the input text from social media posts and classify the users into different MBTI categories by making a predicting on the input text which can then be used as a strong linguistic baseline for Myers-Briggs Type Indicator or a user personality in general. Such a text based classifier can have potentially important applications in psychology domain, there is a connection between the personality type and natural language.

Hence, this prediction will give us a better understanding of any individual at a high level, although these predictions are open to personal interpretation.

Myers–Briggs Type Indicator (MBTI) test was developed in the 1940's with more than 60 years of R&D. Test consists of 90-plus "forced choice" questions where no preference is good or bad, better or worse based on text samples from social media. 4 axis are:

Introversion (I) – Extroversion (E)

Intuition (N) – Sensing (S)

Thinking (T) – Feeling (F)

Judging (J) – Perceiving (P)

The widespread usage of social media enables such a classifier to be trained on huge amount of data to do personality tests, enabling more and more people to see their MBTI personality type, even more quickly and with more reliable results. Nowadays, even the private sector and big corporates companies are showing remarkable interests along with researchers in psychology sector. These employers like to know the personality type of the individuals they plan on hiring, so that they can effectively manage their firms culture. Also, another reason is that when these tests are taken again by an individual in different contexts, it usually generate different classifications since, personality tests done by trained psychologists retest error hovering around 0.5. Hence, such a machine learning classifier can possibly act as a verification system and will help individuals have more confidence about the results of their tests and help them understand their personality and apply their personal strengths. Also, such a classifier is capable of operating on a far larger amount of dataset compared to data present in a single personality test

In MBTI, each individual is corresponding to 4 opposite pairs known as "dichotomies" and typically referred with its letter abbreviation: Extraversion (E) – Introversion (I), Sensing (S) – Intuition (N), Thinking (T) – Feeling (F), and Judging (J) – Perceiving (P) and forming 16 personality types with a combination of 4 of these dichotomies example, INTJ.

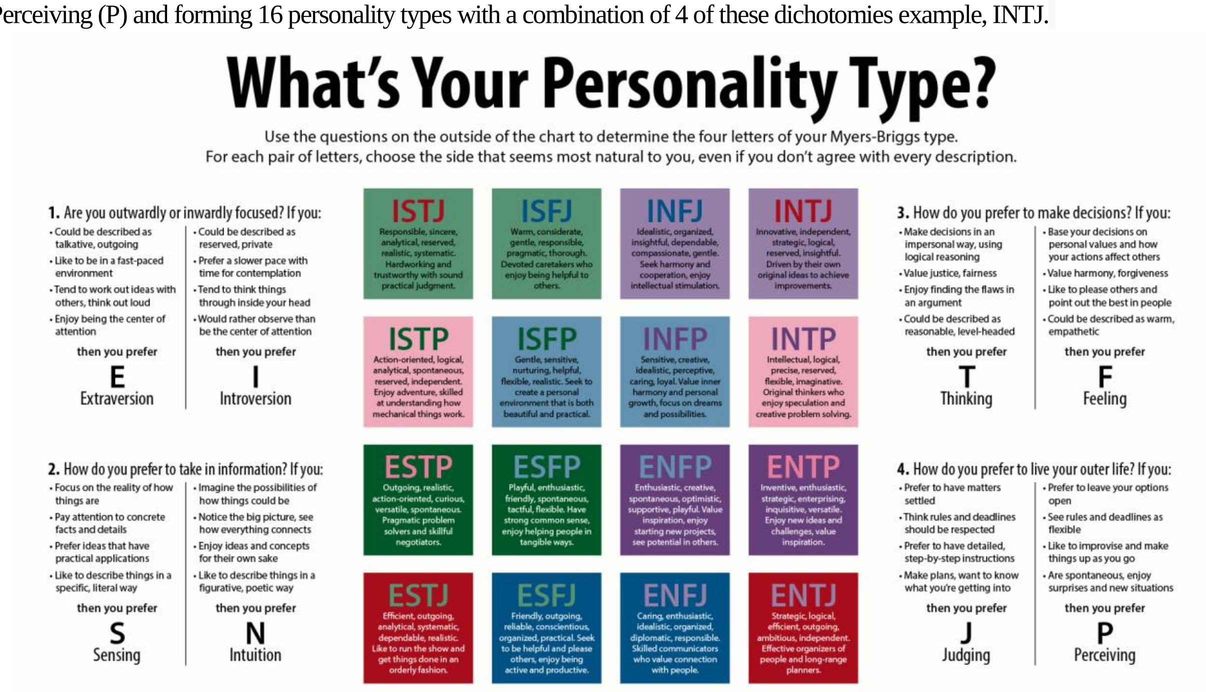


Fig. 1. Chart for each MBTI personality type and 4 dichotomies central to theory

II. LITERATURE REVIEW

There has been increased demand of various text processing methods and models and researches in the area of natural language processing. Various applications of text processing spreads from image recognition, speech recognition, email and spam filtering, fraud detection, stock market trading, sentiment analysis, and language translation and so on.

In 1944, Katherine Cook Briggs and Isabel Briggs Myers published 'Briggs Myers Type Indicator Handbook' aiming to help women find jobs according to their personalities.

There is debate on the validity of Myers-Briggs Type Indicator being a benchmark personality type. As opposed to MBTI test, there's another psychometrics used known as Big Five personality classification system which is a measure of 5 statistically orthogonal personality dimensions: Extraversion, Agreeableness, Openness, Conscientiousness, and Neuroticism which unlike MBTI is derived statistically and provides predictions which predicts on set of features in any individual's life, like marital status, income, education level etc. which are mostly stable in that individuals lifetime.

Pennebaker et al. [1] shows good correlations between 4 Myers-Briggs dimensions with 4 of the Big Five personality traits which shows a good correlation of MBTI personality traits to determine personality. Jonathan Adelstein et al. [16] confirmed that Big five personality domains have neural correlations suggesting that cognitive and affective processing of individuals brain had a different activation pattern on different Big Five personality dimension. Furnham et al. [22] examined at management level factors like personality traits, intelligence and personality disorders.

In other work, Mihai [19] and Champa [23] 3-layer feed forward model on handwritten textual data was done along with textual characters, but due to small sample size was an issue but still provides a proof that machine learning models are efficient in predicting MBTI personality type with good accuracy. Mike K. et al. [21] used bog of words for feature engineering with Naïve Bayes and SVM to achieve good prediction of MBTI personality type.

In other NLP implementations, Liu [17] implemented opinion mining of a block or block of texts, and information from article, reviews from authors or organisations. Balahur et al. [18] said scope should be defined to identify good or bad news from good or bad sentiments.

III. DATA PREPARATION

One advantage of supervised learning is that the data is properly labeled and can be used directly to train our models with some preprocessing done on the texts data available. The data for this project is publicly available on Kaggle [2] which is an online community for data science and machine learning practitioners.

This dataset consists of 8675 rows of data which is acquired by the author of the dataset from PersonalityCafe online forum, where each row consists of 2 columns:

1. Type →MBTI Personality Type e.g. INTJ, ENTP of a user
2. Posts → Upto 50 posts of each user taken from social media (fig. 2. Shows the posts record at row1 for a user)

```
data_set.iloc[0][1]
```

```
"'http://www.youtube.com/watch?v=qsXHcwe3krw|||http://41.media.tumblr.com/tumblr_lfouy03PMA1qa1rooo1_500.jpg|||enfp and intj moments  https://www.youtube.com/watch?v=iz7lE1g4XM4  sportscenter not top te
n plays  https://www.youtube.com/watch?v=uCdfze1etec  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://v
imeo.com/22842206|||Hello ENFJ7. 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 time
s 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 on
e 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 y
our now, current favorite video games? :cool:|||https://www.youtube.com/watch?v=QyPqT8umzmY|||It appears to be too late. :sad:|||There's someone out there for everyone.|||Wait... I thought confidence wa
s 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 alway
s 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.|||ht
tp://www.youtube.com/watch?v=gDhy7rdfm14  I really dig the part from 1:46 to 2:50|||http://www.youtube.com/watch?v=msqXffgh7b8|||Banned because this thread requires it of me.|||Get high in backyard, roa
st and eat marshmellows in backyard while conversing over something intellectual, followed by massages and kisses.|||http://www.youtube.com/watch?v=Mw7eoU3BMbE|||http://www.youtube.com/watch?v=4V2uYORhQ
Ok|||http://www.youtube.com/watch?v=SlVmgFQQ0TI|||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 He
alth class clearly taught you nothing about peer pressure.|||Banned for a whole host of reasons!|||http://www.youtube.com/watch?v=IRcrv41hgz4|||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...|||a pokemon world  an infj society  everyone becomes an optimis
t|||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/pok
emon-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 signa
ture.|||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 wit
h the roaches.|||http://www.youtube.com/watch?v=w8IgImn57aQ|||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 j
ust overloading myself with too...|||I like this person's mentality. He's a confirmed INTJ by the way. http://www.youtube.com/watch?v=hGKLI-GEc6M|||Move to the Denver area and start a new life for mysel
f.'"
```

Fig. 2. First row record showing user posts separated by triple pipes '|||'

To experiment our models and embedding techniques on this dataset, we first need to analyze the data and do some preprocessing on the user posts to handle clean the text data. A null check on the dataset tells us that no row is having null or unfilled data (see fig. 3.) and hence, we can process with further checks on the dataset. To analyze the data first the dataset distribution of posts for each personality type is check as seen in fig. 4.

```
RangeIndex: 8675 entries, 0 to 8674
Data columns (total 2 columns):
 #   Column  Non-Null Count  Dtype
---  ------  --------------  -----
 0   type    8675 non-null   object
 1   posts   8675 non-null   object
dtypes: object(2)
```

Fig. 3. Dataset Info

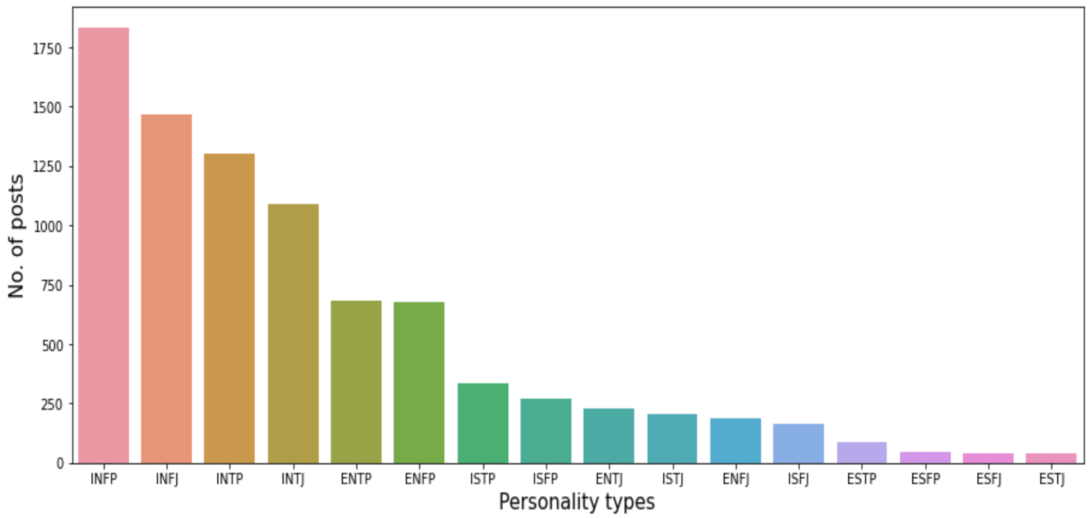


Fig. 4. Distribution of Number of Posts vs. Personality Types

TABLE I
DISTRIBUITION OF POSTS FOR EACH TYPE IN DATASET

INFP	INFJ	INTP	INTJ	ENTP	ENFP	ISTP	ISFP	ENTJ	ISTJ	ENFJ	ISFJ	ESTP	ESFP	ESFJ	ESTJ
1832	1470	1304	1091	685	675	337	271	231	205	190	166	89	48	42	39

From fig.4. And Table I, we can see that the data is not proportionate, as most of the number of users for INFP, INFJ, INTP and INTJ are way high compared to the number of users for personalities like ESTP, ESFP, ESFJ and ESTJ.

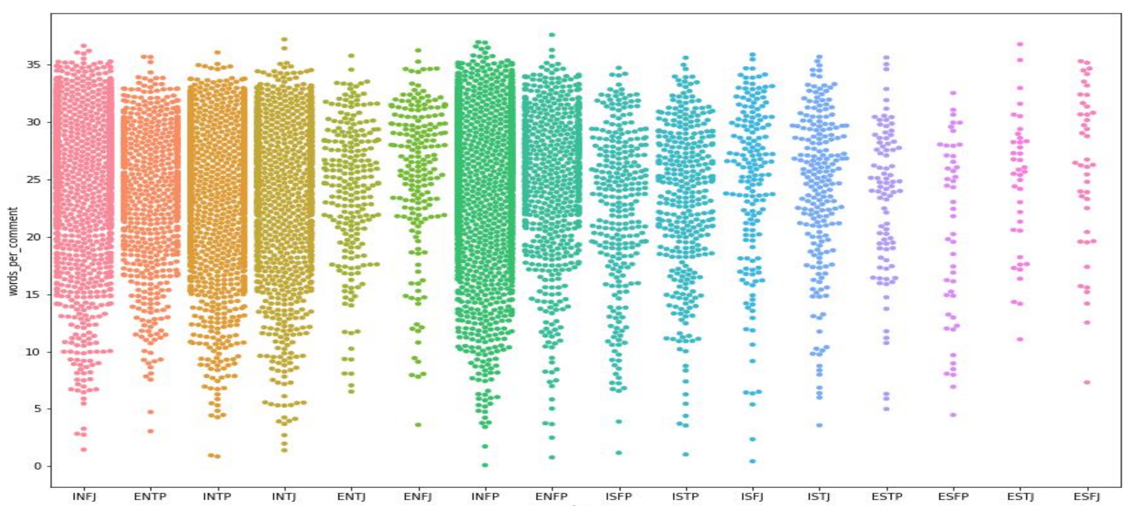


Fig. 5. Words per comment vs. personality type

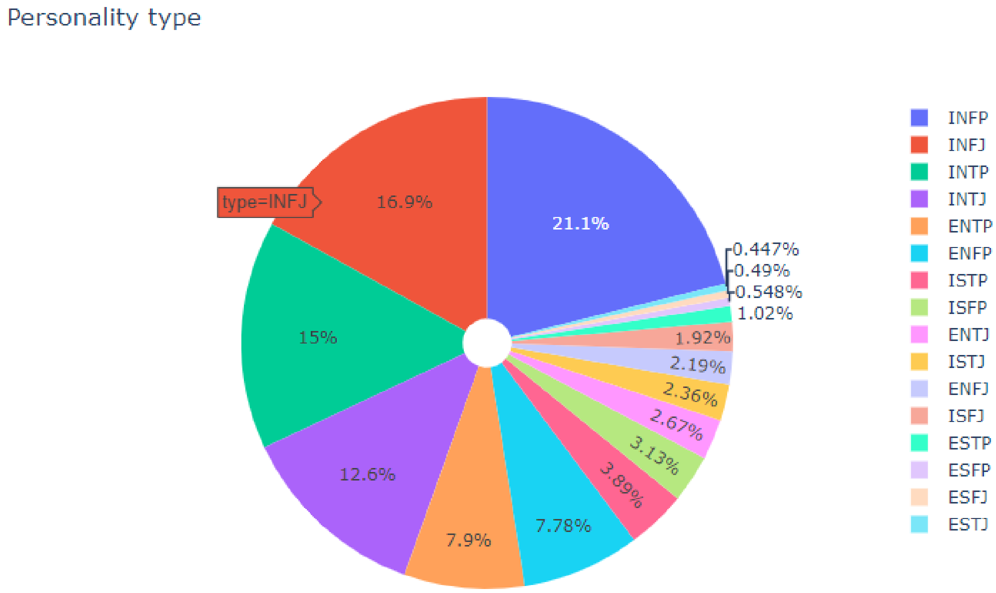


Fig. 6. Data Proportion for each personality type

Fig. 7. Show the distribution of words per comment for each personality type, telling us that most posts on an average have a word count from 15-35 words per post.

Fig. 7. Top 40 words with most count in the dataset

Fig. 7. Gives us an idea of the most common words in the dataset. This gives us an idea about the words, symbols etc. which needs to be handled or removed in the pre-processing stage, as we can see that INFJ and INTP (see fig 8 common words for type = 'INFJ') are couple of the most common words in the user posts and can have major effect in predicting user's personality type.

Fig. 8. Top 40 words with personality type = 'INFJ'

In data cleaning stage, following text cleaning and handling is done:

a) ||| Separators removal → each user post consists of around 50 posts of that user which are separated by triple pipe which needs to be removed from the data
b) Links removal → There are many posts which consists of links to different pages like YouTube etc. which needs to be removed as web links doesn't have major impact on a user personality
c) Punctuations removal → There are punctuation marks and symbols which are removed from the dataset
d) Lower case → the entire dataset is set to lower case to make the data consistent
e) Stop-words removal → English stop words are removed from the dataset
f) Non-words removal → Numbers and symbols are removed from the dataset

Once above preprocessing data cleaning is done, the dataset is ready to apply different tokenization and embedding techniques. To work with different embedding techniques, the dataset needs to be tokenized and the size of the text data needs to be reduced to have minimal words to have all the related words map to the same words, and is done by using nltk library's porterstemmer [3]. Stemming is 'normalization' process which helps reduce the words corpus by reducing the dictionary size by 2-3 folds, hence reducing the input dimension to our machine learning models and hence, making it faster to train and predict the results. Stemming basically reduces the size of our lookup, and hence normalize the sentences. Basically, words are reduced to their root word after removing the verbs and tenses of those words.

After all the steps above, the dataset is ready to apply the planned embedding techniques on it and then to train our models. In this project, the dataset is divided in the ratio of 70-30 for training and testing respectively.

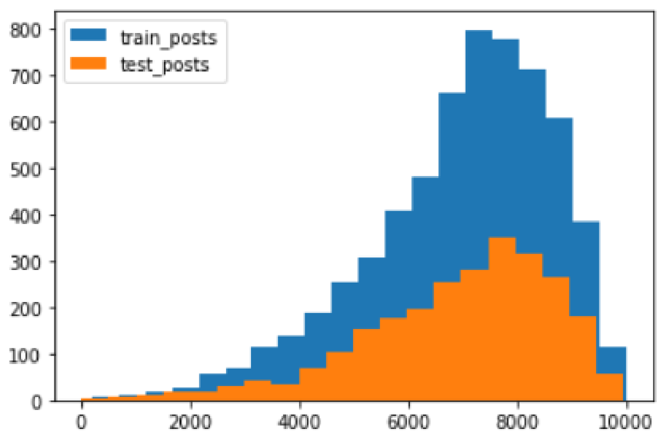


Fig. 9. Length of Train posts vs. Test posts

IV. METHODOLOGY

In methodology, we aim to test embedding like Bag of Words, Tf-IDF, word2Vec and then train these embedding with models like XGBoost, Logistic Regression, and Linear Support Vector Classifier, and compare the results with the predictions made by BERT.

Embedding is representation of text data to vectors of real numbers. When text data is passed to machine learning models, it requires inputs to be numbers and not text, and hence, embedding is required for text analysis. Embedding helps us to limit the length of the input vectors to our models. In this project we will implement different word embedding techniques as below.

4.1) **Embedding techniques:**

4.1.1) **Bag-of-Words →**

Also known as BOW, is one of the most basic and simple model which builds a vocabulary out of corpus of documents and then keep a count of number of words appearing in each document which can be seen in a way where every word in the vocabulary is considered as a feature and the document being denoted as a vector of the same length as vocabulary i.e. a bag of words. This method leads to high dimensionality issue and hence, it requires preprocessing like stemming to reduce the dimensions. For this experiment, the most common and classical method of CountVectorizer is implemented which tokenizes the text while very basic preprocessing is performed. CountVectorizer have parameters as max_features = 1000, min_df = 0.2 and max_df = 0.9
min_df specifies how much importance is will be given to less frequent words in the document.

max_df specifies the importance given to the most frequent words in the document

4.1.2) **Tear frequency-inverse document frequency →**

Famously known as Tf-IDF tells the importance of a word to a document important a word in a vocabulary corpus. For example, we have a collection of N documents. Let's say fij = frequency of word i in document j. Then, term frequency TFij is,

$$TF_{ij} = f_{ij}/max_k f_{kj}$$

Here denominator is the maximum occurrence of a word in that document, making the maximum frequency word in document a TF of 1 and rest others are set in fractions. The IDF of word I appearing in n_i of N documents is,

$$IDF_i = \log_2 N/n_i$$

4.1.3) **Word2Vec →**

This predictive embedding model was developed by Google in 2013, which uses a shallow neural network for vectorization using Continuous Bag of Words (CBOW) and Skip Gram methods by working on next word prediction. In this words are converted to softmax probabilities of a given dimension where if a word will be given same probabilities or they will be assigned same vector even if that word appears in a same context in a sentence.

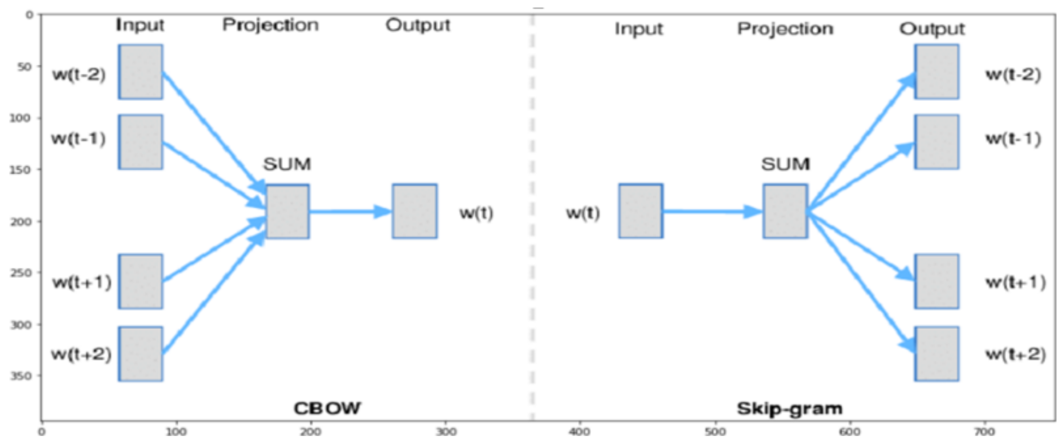


Fig. 10. Word2Vec architecture

4.2) **Models:**

4.2.1) **Logistic Regression →**

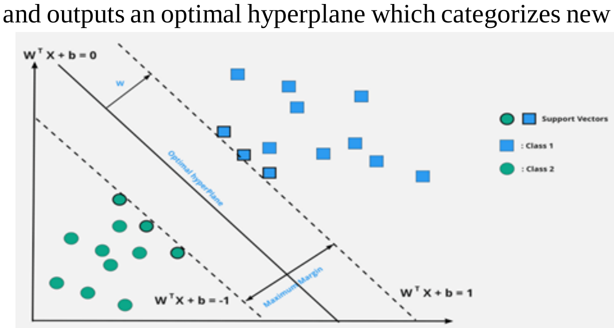
Logistic regression is one of the simplest classification algorithm which is used to assign observations to a set of classes. Logistic regression uses a logistic sigmoid function which helps transforming its output to return a probability value which can then be mapped to two or more discrete classes.

4.2.2) **Extreme Gradient Boosting (XGB) →**

Uses a technique called Gradient Boosting. It builds strong classifier from several weak classifiers in series. XGB provides a parallel tree boosting. Unlike other boosting algorithms where weights of misclassified branches are increased, in Gradient Boosted algorithms the loss function is optimized. XGBoost is an advanced implementation of gradient boosting along with some regularization factors. Boosting controls both the aspects - bias & variance.

4.2.3) **Support Vector Machine (SVM) →**

It's one of the most popular supervised learning algorithms for classification or regression problem which aims to create best decision boundary that can segregate n-dimensional space into classes which enables the system to assign new data point to correct category. This best decision boundary is called a hyperplane. It's a discriminative classifier which intakes training data and outputs an optimal hyperplane which categorizes new examples. This algorithm works



by finding a point closest to the lines for all the classes, which are known as support vectors and the distance between these vectors and the hyperplane is termed as margin and SVM aims at maximizing this margin and give an optimal hyperplane with maximum margin. Extreme points or vectors chosen will help in creation of the hyperplane. Such extreme cases are known support vectors, and hence the name Support Vector Machine.

Fig. 11. SVM classification

4.2.4) **Bidirectional Encoder Representations from Transformers (BERT) →**

For this experiment, a basic BERT implementation has been done. Pre-trained Bert tokenizer – "**bert-large-uncased**" has been used to tokenize on the clean text dataset for training and testing. Sparse Categorical Cross entropy loss function is used as it is considered to be more effective for classification tasks and Adam optimizer is used to train the models as it helps improve the SGD optimizer and combines the advantages of AdaGrad and RMSProp. Bert is considered as it get rid of the decoder from traditional transformer as it has a stack of Transformer encoder blocks. Bert is pre-trained with next sentence prediction and masked language modeling and is a combination of ELMO context embedding and several transformers and assign different vectors for words based on the contexts.

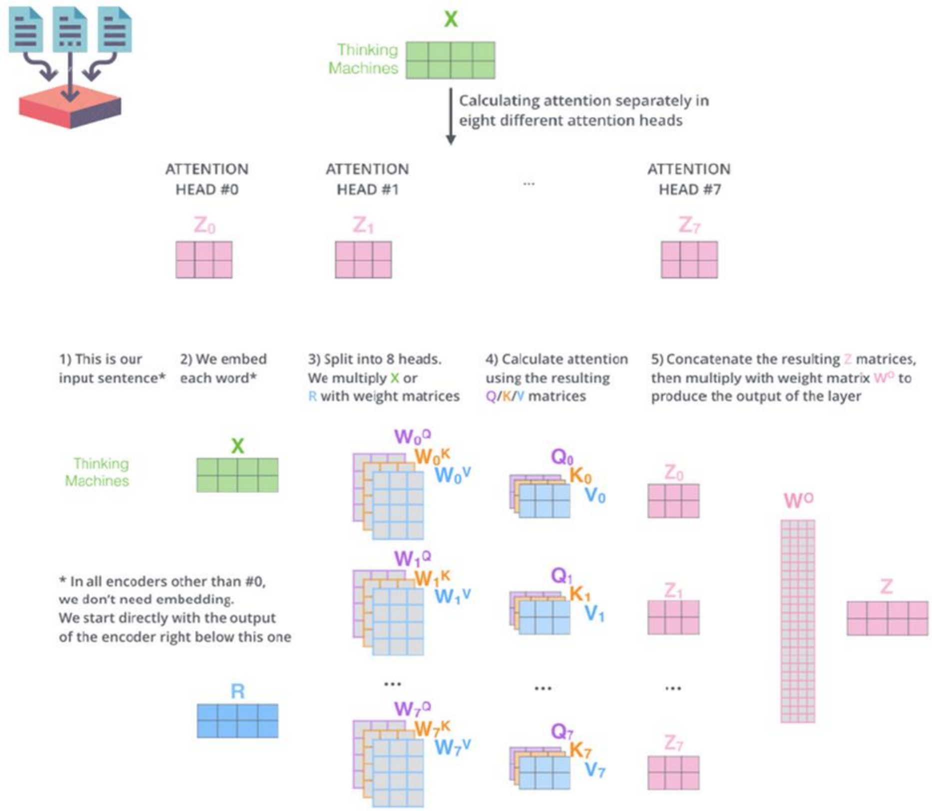


Fig. 12. BERT Model

4.3) **Metrics:**

For this experiment, the metrics used to evaluate the performance of each model is Accuracy. Accuracy is the measure of a classification model's performance and is usually expressed as percentage. It is count of predicted value equal to the true value.

V. Plan vs Progress

The overall plan of this project was as below:

TABLE II
Timeline for tasks

Tasks	Timeline
Topic Selection	Week 1-3
Dataset preparation	Week 4
Data Preparation	Week 5-6
Methodology	Week 7-9
Final Report & Presentation	Week 10-12

Following embedding and models seen in the below table were originally planned for this project:

TABLE III
Planned Embedding and Models Combination

Embedding	Model
BOW	Logistic Regression
TF-IDF	SVM
Word2Vec	XGB
FastText	LSTM
GloVe Elmo	BERT

Following embedding and models seen in the below table were successfully implemented in this project:

TABLE IV
Successful Implementation of Embedding and Models Combination

Embedding	Model
BOW	Logistic Regression
TF-IDF	SVM
Word2Vec	XGB
	BERT

Not all the embedding and models were implemented, due to several issues faces during the development of this project because of the issues mentioned below:

- Started implementation on pytorch but had to switch to Keras for development as found the implementation simpler with keras
- Different embedding and models required different version of tensorflow and keras
- Elmo embedding required various libraries to be downgraded which made other models to fail
- Different embedding required different input size and tokens
- Different runtime required for different models like TPU for BERT, GPU for LSTM
- BERT implementation required many parameters knowledge and high level of hyper-parameter tuning
- BERT takes very long to run for more number of epochs which is an issue when using free TPU allocation
- LSTM didn't have any change in accuracy after first epoch, and couldn't figure out the issue, which could be with the input and output size
- Up-sampling of data didn't enhance accuracy result and need to test with different upsampling techniques

VI. Experiments & Results

Accuracy of classical models and embedding combinations gave the following result, when sorted on the basis of accuracy, see table below. For this experiment, number of features for all the embedding was kept as constant 1000. Word2Vec embedding was set to skip gram model by setting parameter sg = 1 with context window size set to 5.

	Models	Test accuracy
0	Linear Support Vector classifier_TFIDF	0.670343
1	XGBoost Classifier_BOW	0.662466
2	XGBoost Classifier_TFIDF	0.660102
3	logistic regression_TFIDF	0.639228
4	logistic regression	0.550217
5	Linear Support Vector classifier_BOW	0.470264
6	logistic regression_w2v	0.221347
7	Linear Support Vector classifier_w2v	0.217802
8	XGBoost Classifier_w2v	0.183931

Fig. 12. Accuracy of Classical Models with embedding sorted descending as per accuracy

See table 4 for the best performing model and embedding combination on test dataset, experimented in this project.

TABLE IV
Test Accuracy for Model and Embedding combination

Model	Embedding	Test Accuracy %
Linear SVM	BOW	47.0264
	TF-IDF	**67.0343**
	Word2Vec	18.3931
Logistic Regression	BOW	55.0217
	TF-IDF	**63.9228**
	Word2Vec	22.1347
XGB	BOW	**66.2466**
	TF-IDF	**66.0102**
	Word2Vec	18.3931
Bert	Inbuilt - ELMO	**62.103**

VII. CONCLUSION AND DISCUSSION

BOW and TF-IDF gave good results with almost all the models with accuracy ranging between 60-70%. Word2Vec gave the least accuracy with all the models. With minimal hypertuning of parameters Bert performed very well with test accuracy of 62.103% for 10 epochs. Training time of BERT is very high as it has millions of parameters, hence, fine-tuning BERT should give much better results.

This project enable me to learn about various toolsets and libraries and techniques:

- Natural Language Toolkit (NLTK)
- Gensim
- Sklearn
- Tensorflow
- Good understanding of NLP process
- Brief idea of Transformers
- Brief idea of CNNs for text processing
- Different embedding techniques

VIII. FUTURE WORK

This project can be developed or improved by implementing various other models and embedding and process. Some of the work that can be done on this project are:

- Handle imbalance of data by upsampling the dataset
- Predict on some popular comments or essays by celebrities to predict their personality type
- Implement different RNN with all the embedding's
- Try n-grams range for BOW and TF-IDF
- Further hyper tune BERT parameters and try different pre-trained weights
- Try to implement classification method based on individual features, and then predict the overall personality, which may give better results i.e. predicting for
    - Extraversion (E) or Introversion (I)
    - Sensing (S) or Intuition (N)
    - Thinking (T) or Feeling (F)
    - Judging (J) or Perceiving (P)

IX. REFERENCES

[1] James W Pennebaker and Laura A King. Linguistic styles: Language use as an individual difference. Journal of personality and social psychology, 77(6):1296, 1999

[2] Mbti kaggle data set. https://www.kaggle.com/datasnaek/mbti-type

[3] https://riptutorial.com/nltk/topic/8793/stemming

[4] https://medium.com/mlearning-ai/nlp-tokenization-stemming-lemmatization-and-part-of-speech-tagging-9088ac068768

[5] https://www.myersbriggs.org/my-mbti-personality-type/mbti-basics/

[6] https://www.talkspace.com/blog/myers-briggs-personality-test-how-to-take-it/

[7] https://towardsdatascience.com/breaking-bert-down-430461f60efb

[8] https://scikit-learn.org/stable/modules/generated/sklearn.svm.SVC.html

[9] https://xgboost.readthedocs.io/en/latest/index.html

[10] https://docs.paperspace.com/machine-learning/wiki/accuracy-and-loss

[11] Jacob Devlin, Ming-Wei Chang, Kenton Lee, Kristina Toutanova BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding: https://arxiv.org/abs/1810.04805v2

[12] https://www.tensorflow.org/tutorials/text/word2vec

[13] Yun-tao, Z., Ling, G. & Yong-cheng, W. An improved TF-IDF approach for text classification. J. Zheijang Univ.-Sci. A 6, 49–55 (2005). https://doi.org/10.1007/BF02842477

[14] https://scikit-learn.org/stable/modules/generated/sklearn.feature_extraction.text.CountVectorizer.html

[15] https://medium.com/@dhartidhami/understanding-bert-word-embeddings-7dc4d2ea54ca

[16] Jonathan S. Adelstein, Zarrar Shehzad, Maarten Mennes, Colin G. DeYoung, Xi-Nian Zuo, Clare Kelly, Daniel S. Margulies, Aaron Bloomfield, Jeremy R. Gray, F. Xavier Castellanos, and Michael P. Milham. Personality is reflected in the brain's intrinsic functional architecture. PLOS ONE, 6(11):1–12, 11 2011

[17] Liu, B., 2010. Sentiment analysis and subjectivity. *Handbook of natural language processing*, *2*(2010), pp.627-666.

[18] Balahur, A., Steinberger, R., Kabadjov, M., Zavarella, V., Van Der Goot, E., Halkia, M., Pouliquen, B. and Belyaeva, J., 2013. Sentiment analysis in the news. *arXiv preprint arXiv:1309.6202*.

[19] Mihai Gavrilescu. Study on determining the myers-briggs personality type based on individual's handwriting. The 5th IEEE International Conference on E-Health and Bioengineering, 11 2015.

[20] Mayuri P. Kalghatgi, Manjula Ramannavar, and Dr. Nandini S. Sidnal. A neural network approach to personality prediction based on the big-five model. International Journal of Innovative Research in Advanced Engineering, 2015.

[21] Mike Komisin and Curry Guinn. Identifying personality types using document classification methods. Proceedings of the 25th International Florida Artificial Intelligence Research Society Conference, FLAIRS-25, pages 232–237, 01 2012.

[22] Furnham, A., & Crump, J. (2015). The Myers-Briggs Type Indicator (MBTI) and Promotion at Work. Psychology, 6, 1510-1515

[23] Champa HN and Dr. KR Anandakumar . Artificial neural network for human behavior prediction through handwriting analysis. International Journal of Computer Applications, 2010.