

# Personality Type Based on Myers-Briggs Type Indicator with Text Posting Style by using Traditional and Deep Learning

Sakdipat Ontoum<sup>\*†</sup>, Jonathan H. Chan<sup>\*†</sup>

<sup>\*</sup>Computer Science Program, School of Information Technology

<sup>†</sup>Innovative Cognitive Computing (IC2) Research Center

King Mongkut's University of Technology Thonburi

Bangmod, Thung-Khru, Bangkok, Thailand

**Abstract**—The term “personality” may be expressed in terms of the individual differences in characteristics pattern of thinking, feeling, and behavior. This work presents several machine learning techniques including “Naive Bayes”, “Support Vector Machines”, and “Recurrent Neural Networks” to predict people personality from text based on “Myers-Briggs Type Indicator (MBTI)”. Furthermore, this project applies “CRISP-DM”, which stands for “Cross-Industry Standard Process for Data Mining”, to guide the learning process. Since, “CRISP-DM” is kind of iterative development, we have adopted it with agile methodology, which is a rapid iterative software development method, in order to reduce the development cycle to be minimal.

**Index Terms**—Machine Learning; Naive Bayes; Support Vector Machines; Recurrent Neural Networks; Myers-Briggs Type Indicator (MBTI); Personality Prediction; CRISP-DM; Agile Methodology

## I. INTRODUCTION

The “Myers-Briggs Type Indicator (MBTI)” is an approach that reflects and identifies a person’s character. According to [2], a person’s personality is described as a group of characteristics that describe the likelihood of a person’s unique behavior, feelings, and thoughts. These characteristics of a person evolve and may differ in different contexts. In other words, personality is a collection of the traits that contribute to an individual’s distinct personality. Nowadays, many different types of personality models are utilized to characterize personalities, such as the “Big Five Personality Traits model” [3], “VIA Classification of Character Strengths” [4], “Myers-Briggs Type Indicator (MBTI)” model [5,6], and “Jung’s Theory of Personality Type models”. It has been discovered that “MBTI” is more powerful because it has a broader application in a variety of disciplines. However, it has a few reliability and validity concerns. We chose the “MBTI” personality model for this study because of its popularity and the potential for utilizing it in a variety of fields.

Machine learning is a method that researchers frequently employ to predict personality. They can utilize machine learning to learn personality patterns because it can be utilized to learn historical data and predict future data [1, 30]. In psychology science, such an application exists as a tool for

personality evaluation and prediction [30]. The development of machine learning algorithm-based personality prediction technologies has become more important to businesses and recruiters for choosing the best candidates [30].

Nowadays, both natural language processing and social science communities have expressed an increasing interest in automatic personality prediction from social networks [9,30]. Social networks can also be utilized to investigate people’s personalities. A social networking post can be made in a variety of ways, including by utilizing an image, a URL link, and music. Unlike traditional personality tests, which have mostly been utilized in human resources management, counseling, and clinical psychology. The application of automatic personality prediction from social networks is much larger. Dating websites and social media marketing are two examples [10, 31].

In this paper, personality type prediction from the text by utilizing the “Myers-Briggs Type Indicator” on the “MBTI” dataset to predict candidates’ personality types [12]. The proposed method employs machine learning to create a classifier that utilizes text as input to predict the author’s “MBTI” personality type. The remainder of this paper is structured as follows: Section 2 discusses related works; Section 3 outlines the research methodology; Section 4 presents results and discussion; and Section 5 concludes the work.

## II. RELATED WORKS

### A. “Machine Learning”

“Machine learning” is used to train machines on effective handling of data. Sometimes it is difficult for human to interpret a pattern or extract information from vast amount of data. Machine learning is utilized in this case [13]. Generally, machine learning is grouped into three types. For instance, “supervised learning”, “unsupervised learning”, and “reinforcement learning” [32,33]. “Supervised learning” and “Unsupervised learning” are the most widely utilized and widely accepted methods [32]. The supervised machine learning algorithms are those that require assistance from a human. The training and testing datasets are separate, without overlap between them.

An output variable must be predicted or classified in the training dataset. It is expected that an algorithm would learn patterns from training data and then apply them to test data to predict or classify. “K-Nearest Neighbor (KNN)”, “Logistic Regression”, and “Stochastic Gradient Descent” are examples of classification algorithms.

Furthermore, unsupervised learning algorithms only extract a few features from the data. Then when unseen data is introduced, it utilizes previously learned features to predict the class of the new data. “K-Means”, “Mean Shift”, and “K models” are just a few examples [14]. Meanwhile, “reinforcement learning” is employed when the task at hand entails making a series of decisions that lead to a final reward. An artificial agent is rewarded or punished for the actions it takes during the training process. The goal of the learning process is to maximize total reward. “Q-Learning” and the “Markov Decision Process” are two examples of reinforcement learning algorithms.

#### B. “Personality Prediction System from Facebook Users” [10]

“Facebook” has utilized Personality Prediction Systems for many years to predict a user’s personality based on their “Facebook” functionalities [10,30,31]. For users’ personalities, “Facebook” utilized the “Big Five Personality Traits model” [31]. This model would be utilized to discover “Conscientiousness”, “Extraversion”, “Agreeableness”, “Openness”, and “Neuroticism” [10]. The researchers utilized two dataset collections in this study to predict the personalities of the users [30]. The first dataset is made up of data samples from the “myPersonality” project, while the second was built by hand [30]. Before proceeding to the next stage, English-language texts are adjusted in the preprocessing stage. The removal of URLs, symbols, names, and spaces, as well as the lowering of the case, stemming, and removal of stop words, are all part of the preprocessing steps. Slang and non-standard words in Bahasa Melayu data are replaced manually in a separate preprocessing phase before those of the texts are translated to English.

In this study, traditional machine learning algorithms and deep learning were utilized to run a series of tests to predict candidates’ personality types for specifying job positions with maximum accuracy in the classification process. Traditional machine learning algorithms include “Support Vector Machines”, “Gradient Boosting”, “Naive Bayes”, “Linear Discriminant Analysis (LDA)”, and “Logistic Regression” [30]. Deep learning implementations, on the other hand, typically employ four architectures: “Multi-Layer Perceptrons (MLP)”, “Long Short-Term Memory (LSTM)”, and “1-Dimensional Convolutional Neural Networks (1D CNN)” [30].

Experimental algorithms of “deep learning” indicated that the “MLP architecture” had the highest average accuracy in the “myPersonality” dataset [10,30], while the “1D CNN+LSTM” architectures had the highest accuracy in the gathered dataset [10,30]. To summarize, “deep learning” algorithms can be utilized to improve dataset accuracy, even for low-accuracy traits.

#### C. “Myers–Briggs Type Indicator (MBTI)”

Research in the personality field has long piqued the interest of psychologists, and one such study was conducted on the “Myers–Briggs Type Indicator” by a psychiatrist named “Carl Jung”. Then, “Katharine Briggs” and “Isabel Myers Briggs” created the “Myers–Briggs Type Indicator” for testing personality in the 1920s, based on “Jung’s theory of psychological types” [5,8,30]. This model instrument has 16 personality types represented by a “personality types key” as shown in Fig. 1 [7]. In the “MBTI” system, for example, people classified as “INTPs” prefer “Introversion (I)”, “Intuition (N)”, “Thinking (T)”, and “Perception (P)” personality traits. We can classify the needs or behaviors of individuals according to labels, and then the machine can learn the patterns.

As shown in Fig. 2, the preferences in four dimensions are indicated by combining the 16 personality types. Each dimension corresponds to two distinct personalities. The four dimensions are “Introversion (I) – Extraversion (E)”, “Intuition (N) – Sensation (S)”, “Feeling (F) – Thinking (T)”, and “Perception (P) – Judgment (J)” [8].

Text or related data can be utilized to extract a variety of personality-related features. We can determine a user’s “MBTI” type by analyzing their posts on the social platform utilizing “Term Frequency-Inverse Document Frequency (TF-IDF)” to detect and quantify a person’s most frequently utilized words.



Fig. 1. A chart with descriptions of each Myers–Briggs personality type [24]

#### D. “Recurrent Neural Networks”

“Recurrent Neural Networks” are a subset of feed-forward neural networks that can send information along with time

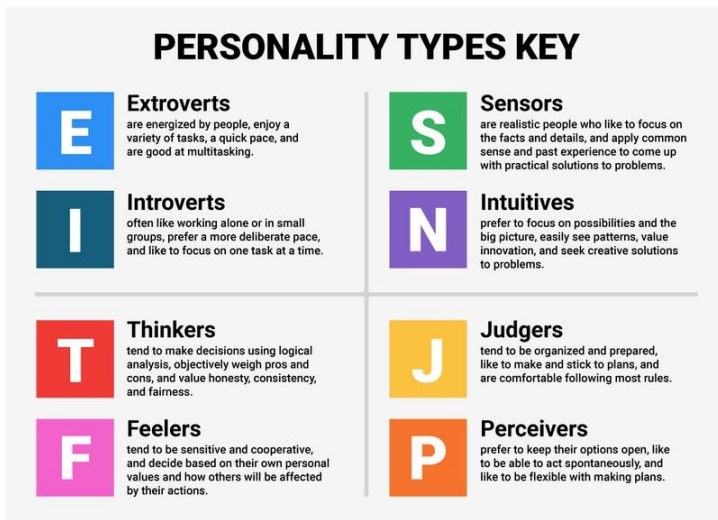


Fig. 2. Key Personality Types [25]

steps. They are a diverse family of models that can perform nearly arbitrary computations [19,20]. It has been demonstrated that finite-sized “Recurrent Neural Networks” with sigmoidal activation functions can simulate a universal Turing machine in a well-known result [20]. In practice, the ability to model temporal dependencies makes “Recurrent Neural Networks” particularly well suited to tasks where the input and/or output consists of non-independent sequences of points [20,21].

#### Advantages

- “Recurrent Neural Networks” can store records in such a way that each pattern is assumed to be dependent on previous ones.
- “Recurrent Neural Networks” are also utilized in combination with convolutional layers to broaden the effective pixel neighborhood [41][42].

#### Disadvantages

- “Recurrent Neural Networks” mostly have problems with exploding and vanishing gradients.
- It is hard to train “Recurrent Neural Networks”.
- If tanh or ReLU is utilized as an activation function, it cannot handle very long sequences.

#### E. “Support Vector Machine” Classifiers

The “Support Vector Machines” Classifier is a supervised machine learning algorithm useful for “classification” and “regression tasks” [17,34,36]. However, it is mostly utilized in classification tasks [36]. Each data point is represented in N-dimensional space by the “Support Vector Machines” algorithm, and each feature is represented in a specific coordinate [36]. Then, it can classify the data by locating the hyperplane that best separates the two classes [18,27,34].

#### Advantages

- When there is a clear margin between classes, the “Support Vector Machines” perform reasonably well [35,39].

- In high-dimensional spaces, “Support Vector Machines” are more effective.
- “Support Vector Machines” are effective when the number of dimensions exceeds the number of samples [35,39].
- “Support Vector Machines” utilize a small amount of memory.

#### Disadvantages

- The “Support Vector Machines” algorithm usually perform poorly for large datasets.
- If the dataset contains noise, overlap classes, and “Support Vector Machines” do not perform well [35].
- The “Support Vector Machines” may perform poorly when the number of features in each data point outstrips the sample size of the training data [35].

#### F. “Naive Bayes” Classifiers

The Bayes theorem is the foundation of the supervised machine learning algorithm on “Naive Bayes”. Each feature will contribute independently and equally to the target class or label [15,32]. Furthermore, when the likelihood of a sample belonging to a specific class is increased, Then they stop interacting with one another [32]. The “Naive Bayes” classifier is easy to implement, computationally fast, and effective on large, high-dimensional datasets [16,26,32].

#### Advantages

- “Naive Bayes” can be used to solve multi-class prediction problems quickly and efficiently.
- When the feature independence hypothesis is true, it can outperform other models while requiring substantially less training data.

#### Disadvantages

- “Naive Bayes” presupposes that all characteristics are independent, which is seldom the case. This restricts the algorithm’s usability in real-world use applications [40].
- This approach encounters the “zero-frequency problem” in which it assigns zero probability to a categorical variable whose category is in the test data set [40]. Also, it was not included in the training dataset [40].
- Because its predictions might be inaccurate in some instances, its probability outputs should not be taken too seriously.

### III. RESEARCH METHODOLOGY

#### A. “Cross Industry Standard Process for Data Mining (CRISP-DM)”

A schematic of “Cross Industry Standard Process for Data Mining (CRISP-DM)” is shown in Fig. 3. It is a free model that has become the industry standard in data mining methodology [38]. Because of its industry and tool independence, “CRISP-DM” can give instructions for the structured execution of any project. All scheduled tasks are often divided into six separate periods [11, 38].

- “Business Understanding”: Learn the domain, clarify the goal, and plan [22].



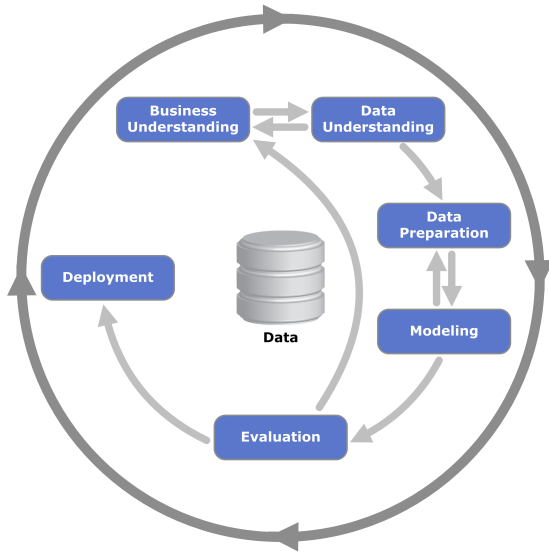


Fig. 3. "CRISP-DM" process framework diagram [23]

- **"Data Understanding"**: Understand the data in all aspects [22].
- **"Data Preparation"**: Preprocessing and feature engineering [22].
- **"Modeling"**: Build the model [22].
- **"Evaluation"**: Evaluate model performance [22].
- **"Deployment"**: Deploy the model to end users [22].

The cycle process serves as a good overall framework for an **end-to-end data science** product [22]. However, when tied to a product cycle, particularly in the case of a modeling product, **things become more complicated; this is where agile data science can come into play** [22].

#### B. "Agile Methodology" with "CRISP-DM"

The agile development process is a development cycle that has been reduced to a bare minimum [22]. The key to agile data science is, to begin with, a **minimal viable model as a baseline, then refine the model through multiple iterations**. We may not build the perfect model during this process, but we will deliver a model with **business value**. In each sprint of our project, we utilize "Agile methodology" and "CRISP-DM" as shown in Fig. 4.

#### C. Data Understanding

##### Dataset

The "MBTI" dataset is shown in Fig. 9, which is publicly available on "Kaggle" [12] and contains 8675 rows of post data. **Each row has two columns, which are the personality type based on "MBTI" and personal social networking posts from "personalitycafe.com"**. Users first complete a questionnaire that determines their "MBTI" type. Then, it allows them to publicly chat or forum with another user. since each user has fifty posts available to them. There are 430,000 data points in total. The data in each row contains the following:

- **Type** (This persons 4 letter MBTI code/type)

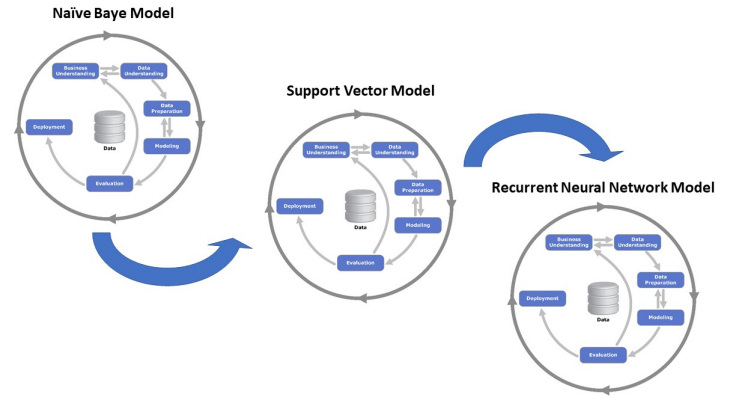


Fig. 4. "CRISP-DM" process framework with Agile Development diagram

- **A section from each of their last 50 posts** (each entry separated by " | | | " (3 pipe characters)) [12].

#### Exploratory Data Analysis

**The objective of the exploratory data analysis was to use various tables and graphs to generate a visual representation of the data** for further investigation. As shown in Fig. 5, the proportionality of "MBTI" type is shown for the general population. There is a significant imbalance in the "Introversion (I) - Extroversion (E)" and "Intuition (N) - Sensing (S)" pairs. On the other hand, "Perception (P) - Thinking (T)" and "Judgment (J)" pairs are quite balanced.

**Then, we categorized the personality type keys into four dimensions** as shown in Fig. 6. In the first category of "Extroversion (E) – Introversion (I)", the distribution of "Extroversion (E)" is significantly bigger than that of "Introversion (I)". In the second category, "Sensation (S) – Intuition (N)", "Sensation (S)" have a significantly higher dispersion than "Intuition (N)". In the third category, "Thinking (T) – Feeling (F)", "Thinking (T)" has a somewhat greater distribution than "Feeling (F)". Lastly, in the fourth category, "Judgment (J) – Perception (P)", "Judgment (J)" is more prevalent than "Perception (P)".

**Then, for the concepts that were used the most frequently by specific classes of the personality dimensions, we constructed word clouds**. They were created by deleting postings that had the most extreme class probability prediction. The size of each term in the word clouds is then proportional to its frequency of appearance in the top posts. We feel that these word clouds reflect some of the unique ways that different "MBTI" utilize language, as shown in Figs. 7 and 8.

#### D. Data Preparation

##### Preprocessing

**To better comprehend the datasets, we included four more columns that were categorized based on respondents' responses to the four dimensions of "MBTI"**: Introversion (I) - "Extroversion (E)", "Intuition (N) - Sensation (S)", "Feeling (F) - Thinking (T)", and "Perception (P) - Judgment (J)". The purpose of the procedure is **to improve the accuracy of "Naive**

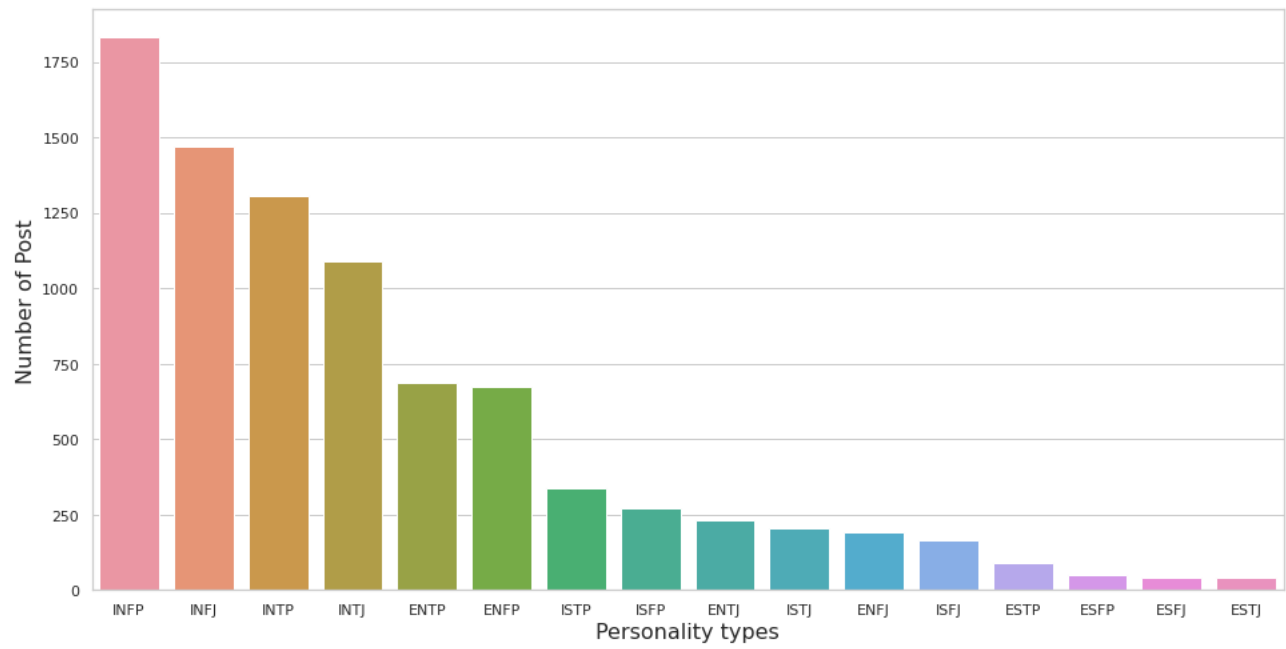


Fig. 5. Proportionality diagram of “MBTI” types

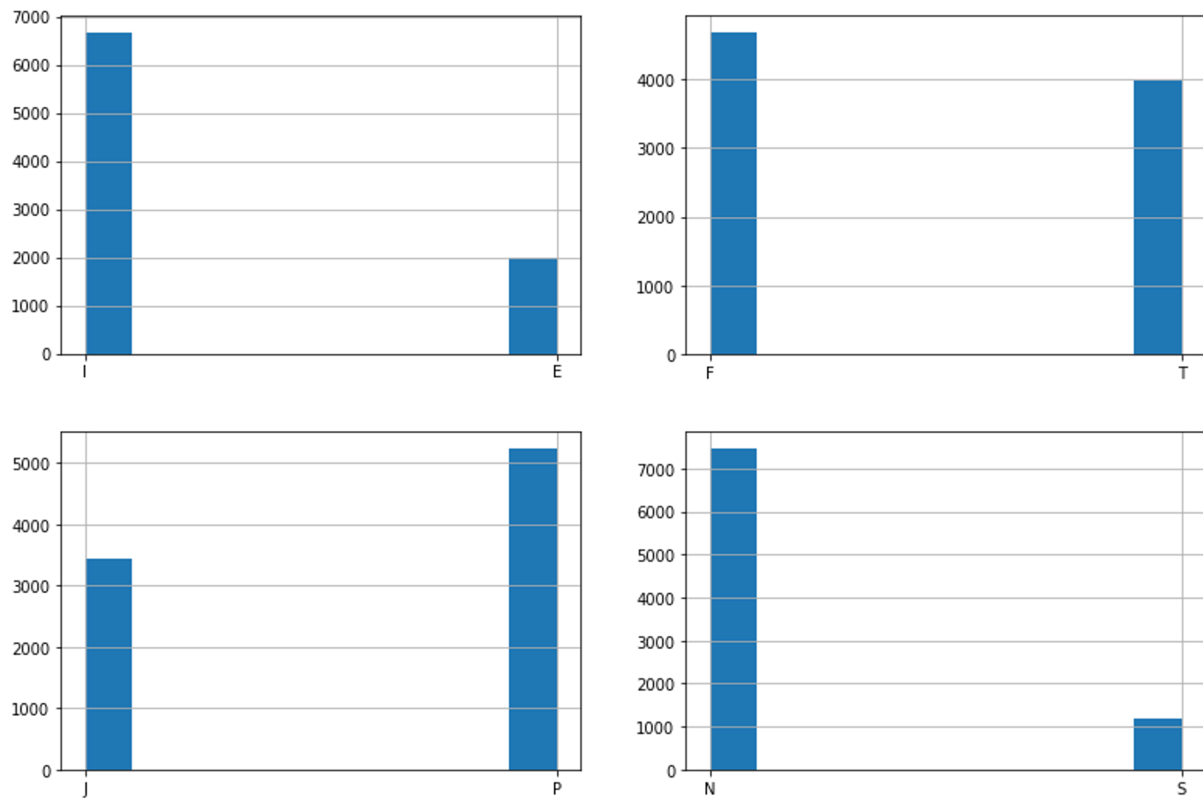


Fig. 6. Proportionality diagram of each “MBTI” type key





Fig. 8. Unique word cloud for individual “MBTI” types

### Modelling on “Naive Bayes” and “Support Vector Machines” classifiers

We utilized “Scikit-learn” to construct the “Naive Bayes” classifier, which is a “multinomial-NB probabilistic” learning method that is commonly utilized in “Natural Language Processing (NLP)”. It computes the likelihood of each tag for a given sample and outputs the tag with the highest likelihood.

Second, we utilized “Support Vector Machines” with regularization parameter with cost with 1, kernel parameter is linear with degree 3, and we also applied gamma with auto to avoid exact match as per the training data set, which tends to cause over-fitting.

### Modelling on “Recurrent Neural Networks”

As shown in Fig. 10, we utilize the embedding word to define the indexes into dense vectors of fixed size with a total length of vocabulary of 256. We decided to utilize “CONV1D”

on “Recurrent Neural Networks” because “CONV1D” moves along a single axis. It makes perfect sense to apply this type of convolution layer to sequential data, such as text. Then, as shown in Fig. 11, to improve the performance of the “Recurrent Neural Networks” model, we added “Bi-directional Long Short-Term Memory (BI-LSTM)” with a size of 64, allowing the “Recurrent Neural Networks” to store sequence information in both directions, backward (future to past) and forward (past to future).

## IV. RESULTS AND DISCUSSION

Table I shows that the “Recurrent Neural Networks” outperform other machine learning models in terms of accuracy in all four personality types. The accuracy of “Recurrent Neural Networks” is significantly higher than that of “Naive Bayes” and “Support Vector Machines” for all categories. On



	type	posts
0	INFJ	'http://www.youtube.com/watch?v=qsXHcwe3krw   ...
1	ENTP	'I'm finding the lack of me in these posts ver...
2	INTP	'Good one _____ https://www.youtube.com/wat...
3	INTJ	'Dear INTP, I enjoyed our conversation the o...
4	ENTJ	'You're fired.   That's another silly misconce...
...	...	...
8670	ISFP	'https://www.youtube.com/watch?v=t8edHB_h908   ...
8671	ENFP	'So...if this thread already exists someplace ...
8672	INTP	'So many questions when i do these things. I ...
8673	INFP	'I am very conflicted right now when it comes ...
8674	INFP	'It has been too long since I have been on per...
8675 rows × 2 columns		

Fig. 9. “MBTI” Personality Type Dataset [12]

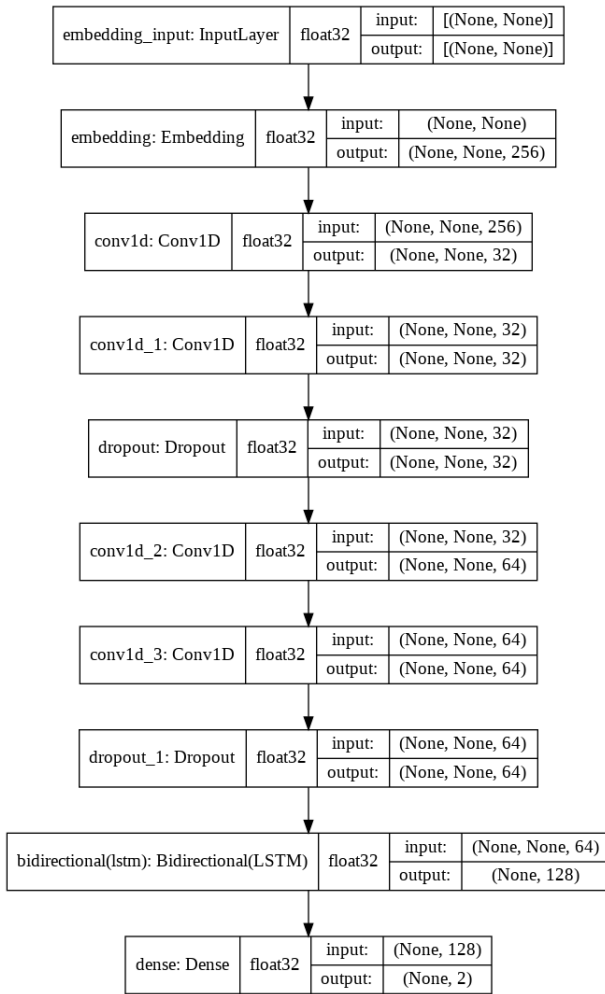


Fig. 10. “Recurrent Neural Networks” Architecture

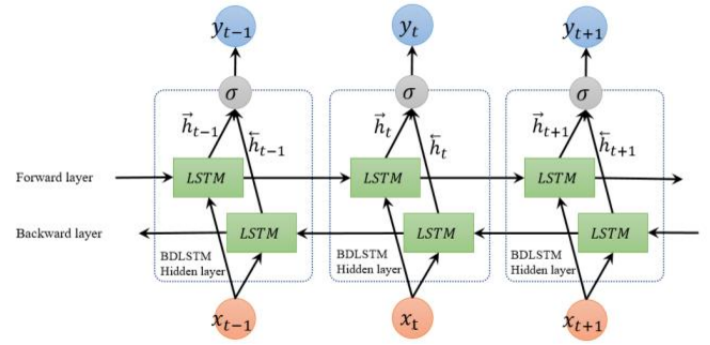


Fig. 11. Bi-directional Long Short-Term Memory (BI-LSTM) Architecture [29]

TABLE I  
ACCURACY OF SUBSET “MBTI” MODEL

Model	Type			
	I/E	N/S	F/T	P/J
“Naive Bayes”	78.28%	86.95%	80.63%	74.78%
“Support Vector Machines”	82.15%	87.32%	80.49%	72.70%
“Recurrent Neural Networks”	83.59%	93.22%	80.00%	77.40%

TABLE II  
OVERALL ACCURACY OF “MBTI” MODEL

Model	Overall Accuracy
“Naive Bayes”	41.03%
“Support Vector Machines”	41.97%
“Recurrent Neural Networks”	49.75%

the other hand, the accuracy of “Naive Bayes” and “Support Vector Machines” for “Perception (P)” and “Judgment (J)” is significantly lower than that of the “Recurrent Neural Networks”. As a result, the “Recurrent Neural Networks” outperform the other three machine learning models on this dataset.

In contrast with the overall accuracy in each model shown in Table II, there appears to be a general weakness in our model’s ability to classify all four “MBTI” dimensions accurately. However, the number that represents perfect classification does not indicate the efficacy of our model’s predictions of overall “MBTI” types.

Since the “Recurrent Neural Networks” outperform, the confusion matrix of the “Recurrent Neural Networks” model is shown in Fig. 12. For each “MBTI” model, the results are mostly positioned as “True Positive” which means they are projected as positive and turn out to be true. which is a false positive, implying that the prediction is positive but incorrect.

Lastly, the classification report of the “Recurrent Neural Networks” is shown in Fig. 13 and Fig. 14. This shows that the F-score or F-measure, which is a weighted average score of the true positive (recall) and precision, is around 76% when



P/J model			N/S model		
Actual \ Predicted	J	P	Actual \ Predicted	N	S
J	764	79	N	1800	83
P	411	915	S	64	222

I/E model			F/T model		
Actual \ Predicted	E	I	Actual \ Predicted	F	T
E	423	95	F	774	359
I	261	1390	T	76	960

Fig. 12. Confusion Matrix on each “MBTI” models with RNN

I/E model				
Classes	Precision	Recall	F1-score	Support
E	0.62	0.82	0.70	518
I	0.94	0.84	0.89	1651
accuracy			0.84	2169
macro avg	0.78	0.83	0.80	2169
weighted avg	0.86	0.84	0.84	2169
F-measure: 0.7952				

F/T model				
Classes	Precision	Recall	F1-score	Support
F	0.91	0.68	0.78	1133
T	0.73	0.93	0.82	1036
accuracy			0.80	2169
macro avg	0.82	0.80	0.80	2169
weighted avg	0.82	0.80	0.80	2169
F-measure: 0.7980				

Fig. 13. Classification Report on each “MBTI” models with RNN (1)

N/S model				
Classes	Precision	Recall	F1-score	Support
N	0.97	0.96	0.96	1883
S	0.84	0.35	0.50	286
accuracy			0.90	2169
macro avg	0.85	0.87	0.86	2169
weighted avg	0.93	0.93	0.93	2169
F-measure: 0.7211				

P/J model				
Classes	Precision	Recall	F1-score	Support
J	0.65	0.91	0.76	843
P	0.92	0.69	0.79	1326
accuracy			0.77	2169
macro avg	0.90	0.89	0.89	2169
weighted avg	0.82	0.77	0.78	2169
F-measure: 0.7730				

Fig. 14. Classification Report on each “MBTI” models with RNN (2)

calculated as the mean of all models.

## V. CONCLUSION

In conclusion, this study was able to predict personality utilizing social network data with a machine learning algorithm that we applied. Also, the “Recurrent Neural Networks” machine learning algorithm is the best model for predicting personality based on “MBTI”. Companies would substantially benefit from this since they would be able to study their candidates’ social network profiles before choosing suitable employees.

### Limitation

The “Myers-Briggs Type Indicator” from the text is only the first tier in developing a personality type model. The only social network utilized for this project was the Personality Cafe forum. Other social networks may provide useful data, allowing the prediction model to be improved. When building teams to produce new software, play sports, or fight crime, a technical position is one of the most important factors to consider when joining a team [30]. In addition, people’s soft skills, such as mindset and personality, as well as their living environment must be considered.

### Future work

We intend to utilize the Generative Pre-trained Transformer 3 (GPT-3) which is an autoregressive language model that leverages deep learning to generate human-like text [19].

## REFERENCES

- [1] Zhang, W. (2016). The Application of Machine Learning Algorithms in Data Mining. DEStech Transactions on Engineering and Technology Research, iect. <https://doi.org/10.12783/dtet/iect2016/3817>.
- [2] Ahmed, F., Campbell, P., Jaffar, A., Alkobaisi, S., & Campbell, J. (2010). Learning & Personality Types: A Case Study of a Software Design Course. Journal of Information Technology Education: Innovations in Practice, 9, 237–252. <https://doi.org/10.28945/1329>.
- [3] de Raad, B., & Mlačić, B. (2015). Big Five Factor Model, Theory and Structure. International Encyclopedia of the Social & Behavioral Sciences, 559–566. <https://doi.org/10.1016/b978-0-08-097086-8.25066-6>.
- [4] Niemiec, R. M. (2012). VIA Character Strengths: Research and Practice (The First 10 Years). Cross-Cultural Advancements in Positive Psychology, 11–29. [https://doi.org/10.1007/978-94-007-4611-4\\_2](https://doi.org/10.1007/978-94-007-4611-4_2).
- [5] Varvel, T., & Adams, S. (2003). A Study Of The Effect Of The Myers Briggs Type Indicator. 2003 Annual Conference Proceedings. Published. <https://doi.org/10.18260/1-2-11800>.
- [6] Boyle, G. J. (1995). Myers-Briggs Type Indicator (MBTI): Some Psychometric Limitations. Australian Psychologist, 30(1), 71–74. <https://doi.org/10.1111/j.1742-9544.1995.tb01750.x>.
- [7] Poursafar, Z., Devi, N.R., & Rodrigues, L.L. (2015). EVALUATION OF MYERS-BRIGGS PERSONALITY TRAITS IN OFFICES AND ITS EFFECTS ON PRODUCTIVITY OF EMPLOYEES: AN EMPIRICAL STUDY. International journal of current research and review, 7, 53–58. [http://ijcrr.com/uploads/411\\_pdf.pdf](http://ijcrr.com/uploads/411_pdf.pdf).
- [8] Mallari, S.D., & Pelayo, J.M. (2017). Myers-Briggs Type Indicator (MBTI) Personality Profiling and General Weighted Average (GWA) of Nursing Students. <https://files.eric.ed.gov/fulltext/ED579286.pdf>.
- [9] Nguyen, D., Doğruöz, A. S., Rosé, C. P., & de Jong, F. (2016). Computational Sociolinguistics: A Survey. Computational Linguistics, 42(3), 537–593. [https://doi.org/10.1162/coli\\_a\\_00258](https://doi.org/10.1162/coli_a_00258).
- [10] Tander, T., Hendro, Suhartono, D., Wongso, R., & Prasetyo, Y. L. (2017). Personality Prediction System from Facebook Users. Procedia Computer Science, 116, 604–611. <https://doi.org/10.1016/j.procs.2017.10.016>.

- [11] Shearer, C. (2000). The CRISP-DM model: the new blueprint for data mining. *Journal of data warehousing*, 5, 13–22. <https://mineraacadados.files.wordpress.com/2012/04/the-crisp-dm-model-the-new-blueprint-for-data-mining-shearer-colin.pdf>.
- [12] (MBTI) Myers-Briggs Personality Type Dataset. (2017, September 22). Kaggle. <https://www.kaggle.com/datasnaek/mbti-type>.
- [13] A, D. (2016). Machine Learning Algorithms : A Review. *International Journal of Computer Science and Information Technologies*. Published. <http://ijcsit.com/docs/Volume%207/vol7issue3/ijcsit2016070332.pdf>.
- [14] Sah, S. (2020). Machine Learning: A Review of Learning Types. *Machine Learning: A Review of Learning Types*. Published. <https://doi.org/10.20944/preprints202007.0230.v1>.
- [15] Misra, S., & Li, H. (2020). Noninvasive fracture characterization based on the classification of sonic wave travel times. *Machine Learning for Subsurface Characterization*, 243–287. <https://doi.org/10.1016/b978-0-12-817736-5.00009-0>.
- [16] SHORT SURVEY ON NAIVE BAYES ALGORITHM. (2017). *International Journal of Advance Engineering and Research Development*, 4(11). <https://doi.org/10.21090/ijaerd.40826>.
- [17] Keeman, V. (2005). Basics of Machine Learning by Support Vector Machines. *Real World Applications of Computational Intelligence*, 49–103. [https://doi.org/10.1007/11364160\\_2](https://doi.org/10.1007/11364160_2).
- [18] Tomar, D., & Agarwal, S. (2013). A survey on Data Mining approaches for Healthcare. *International Journal of Bio-Science and Bio-Technology*, 5(5), 241–266. <https://doi.org/10.14257/ijbsbt.2013.5.5.25>.
- [19] Siegelmann, H. T., & Sontag, E. D. (1991). Turing computability with neural nets. *Applied Mathematics Letters*, 4(6), 77–80. [https://doi.org/10.1016/0893-9659\(91\)90080-f](https://doi.org/10.1016/0893-9659(91)90080-f).
- [20] Lipton, Z. C. (2015, May 29). A Critical Review of Recurrent Neural Networks for Sequence Learning. *ArXiv.Org*. <https://arxiv.org/abs/1506.00019>.
- [21] Sharkawy, A. N. (2020). Principle of Neural Networks and Its Main Types: Review. *Journal of Advances in Applied & Computational Mathematics*, 7(1), 8–19. <https://doi.org/10.15377/2409-5761.2020.07.2>.
- [22] Wu, T. S. D. (2020, February 26). What's The Right Agile Data Science Process For Your Organization? A42 Labs. <https://www.a42labs.io/agile-data-science-1>.
- [23] Jensen, K. (2012, April 12). IBM SPSS Modeler CRISP-DM Guide. *International Business Machines Corporation (IBM)*. <http://public.dhe.ibm.com/software/analytics/spss/documentation/modeler/18.0/en/ModelerCRISPDM.pdf>.
- [24] Beech, J. (2014, January 28). A chart with descriptions of each Myers-Briggs personality type and the four dichotomies central to the theory [Chart]. <https://commons.wikimedia.org/w/index.php?title=File:MyersBriggsTypes.png&oldid=538599539>.
- [25] Tieger, P. D., Barron, B., & Tieger, K. (2021). *Do What You Are: Discover the Perfect Career for You Through the Secrets of Personality Type* (Revised ed.). Little, Brown Spark.
- [26] Zainab Mahmood Fadhil. (2021). Hybrid of K-means clustering and naive Bayes classifier for predicting performance of an employee. *Periodicals of Engineering and Natural Sciences*, 9. <http://dx.doi.org/10.21533/pen.v9i2.1898>.
- [27] Pereira, S., & Joshi, U. (2014). Implementation of Support Vector Machines Technique in Feedback Analysis System. *International Journal of Computer Applications*, 96(17), 24–28. <https://doi.org/10.5120/16887-6906>.
- [28] Qaiser, S., & Ali, R. (2018). Text Mining: Use of TF-IDF to Examine the Relevance of Words to Documents. *International Journal of Computer Applications*, 181(1), 25–29. <https://doi.org/10.5120/ijca2018917395>.
- [29] Cui, Z., Ke, R., Pu, Z., & Wang, Y. (2020). Stacked bidirectional and unidirectional LSTM recurrent neural networks for forecasting networks-wide traffic state with missing values. *Transportation Research Part C: Emerging Technologies*, 118, 102674. <https://doi.org/10.1016/j.trc.2020.102674>.
- [30] Abidin, N. H. Z., Akmal, M., Mohd, N., Nincarean, D., Yusoff, N., Karimah, H., & H. A. (2020). Improving Intelligent Personality Prediction using Myers-Briggs Type Indicator and Random Forest Classifier. *International Journal of Advanced Computer Science and Applications*, 11(11). <https://doi.org/10.14569/ijacsa.2020.0111125>.
- [31] Amirhosseini, M. H., & Kazemian, H. (2020). Machine Learning Approach to Personality Type Prediction Based on the Myers-Briggs Type Indicator. *Multimodal Technologies and Interaction*, 4(1), 9. <https://doi.org/10.3390/mti4010009>.
- [32] Misra, S., Li, H., & He, J. (2019). *Machine Learning for Subsurface Characterization* (1st ed.). Gulf Professional Publishing.
- [33] So, A., Joseph, T., V., John, R. T., Worsley, A., & Asare, S. (2020). *The Data Science Workshop: Learn how you can build machine learning models and create your own real-world data science projects*, 2nd Edition (2nd ed.). Packt Publishing.
- [34] Smys, S., Bestak, R., Palanisamy, R., & Kotuliak, I. (2021). *Computer Networks and Inventive Communication Technologies: Proceedings of Fourth ICCNCT 2021 (Lecture Notes on Data Engineering and Communications Technologies Book 75)*. Springer.
- [35] Dhar, V., Beg, M., Saxena, A., & Sharma, S. (2021). Capillary Suction Timer and machine learning techniques as tools for evaluating the performance of different shale inhibitors used in drilling mud. *Journal of Natural Gas Science and Engineering*, 96, 104301. <https://doi.org/10.1016/j.jngse.2021.104301>.
- [36] Sohana, J. F., Rupa, R. J., & Rahman, M. (2021). Bengali Stop Word Detection Using Different Machine Learning Algorithms. *Algorithms for Intelligent Systems*, 131–142. [https://doi.org/10.1007/978-981-16-0586-4\\_11](https://doi.org/10.1007/978-981-16-0586-4_11).
- [37] Qaiser, S., & Ali, R. (2018). Text Mining: Use of TF-IDF to Examine the Relevance of Words to Documents. *International Journal of Computer Applications*, 181(1), 25–29. <https://doi.org/10.5120/ijca2018917395>.
- [38] Castelo-Branco, F., Reis, J. L., Vieira, J. C., & Cayolla, R. (2019). Business Intelligence and Data Mining to Support Sales in Retail. *Marketing and Smart Technologies*, 406–419. [https://doi.org/10.1007/978-981-15-1564-4\\_38](https://doi.org/10.1007/978-981-15-1564-4_38).
- [39] K, D. (2020, December 26). Top 4 advantages and disadvantages of Support Vector Machine or SVM. *Medium*. <https://dhirajkumarblog.medium.com/top-4-advantages-and-disadvantages-of-support-vector-machine-or-svm-a3c06a2b107>.
- [40] Sweta Bhise , Shrutika Gadekar , Aishwarya Singh Gaur , Simran Bepari, Deepmala Kale, Dr. Shailendra Aswale, 2021, Breast Cancer Detection using Machine Learning Techniques, *INTERNATIONAL JOURNAL OF ENGINEERING RESEARCH & TECHNOLOGY (IJERT)* Volume 10, Issue 07 (July 2021).
- [41] Reddy P, B. S. (2020, May 26). A Chatbot in Python using nltk - The Startup. *Medium*. <https://medium.com/swlh/a-chatbot-in-python-using-nltk-938a37a9eacc>.
- [42] Gudikandula, P. (2021, March 14). Recurrent Neural Networks and LSTM explained - purnasai gudikandula. *Medium*. <https://purnasaigudikandula.medium.com/recurrent-neural-networks-and-lstm-explained-7f51c7f6bbb9>.