Text Analysis of Applicants for Personality Classification Using Multinomial Naïve Bayes and Decision Tree

Nanda Yonda Hutama[1], Kemas Muslim Lhaksmana[2*], Isman Kurniawan[3]
[1,2,3] School of Computing, Telkom University
[1,2,3] Jl. Telekomunikasi, Terusan Buah Batu, Bandung 40257, West Java, Indonesia
*Corresponding email: kemasmuslim@telkomuniversity.ac.id

Abstract — In talent recruitment, companies seek talents who fit with their corporate culture among many applicants. To conduct this process, companies carry out psychological tests to assess the fitness of applicants' personalities with their corporate culture. However, psychological testing requires cost and much effort. Thus, an automated system necessary to assist companies in the talent recruitment process that can classify personalities through text and to reduce the effort needed. This research is conducted based on the personality traits according to the corporate culture in Telkom Indonesia. The data used is text data that has been labeled, pre-processed, and feature selected. The clean text data is used to create a classification model using multinomial Naïve Bayes and Decision Trees. There are six models built based on three work cultures. The Decision Tree achieves an accuracy of 33%, 66%, 80%, while multinomial Naïve Bayes with an accuracy of 83%, 50%, 60%, which resulted in better performance.

Keywords – personality classification, multinomial Naïve Bayes, Decision Tree, text mining

I. INTRODUCTION

The recruitment process is essential for a company because the quality of employees influences the overall company's performance. Therefore, companies are very selective in finding promising applicants [1]. However, with many applicants in every recruitment process, companies have difficulty getting employees who meet their criteria. Moreover, the recruitment process also requires enormous resources from the process, costs, and human resources [2]. The applicant's personality can be an essential factor for companies to determine whether they can work well or not. According to N. R. Ngatirin, Z. Zainol, and T. L. C. Yoong [3], personality represents a combination of features and qualities that build individual characteristics. Personality traits can be used to understand human behavior regarding many things, including how they work with their environment. Psychological testing is one of the efforts that can be done to determine the applicants' personality but requires a lot of time and cost in its implementation. Therefore, we need a system that can classify the applicants' personality and reduce the time and cost required. Personality classification is done by classifying applicant text using one of the text mining methods. The data used for personality classification is interview verbatim, which has been converted from interview recordings into text.

Multinomial Naïve Bayes and Decision Tree are the classification methods used in this study. Multinomial Naïve Bayes is used because each feature stands individually so that its implementation can produce excellent performance. Multinomial Naïve Bayes can also predict with less time, so it will reduce the time spent to classify the personality and acceptance of applicants to the company. The text mining will have more than a hundred features to naive calculate with the label. Therefore, multinomial Bayes will be suitable to be implemented in this research. Decision Tree is used because it is easy to apply, the classification process can be easily understood. and the learning speed is quite fast. Decision Tree is also suitable for this research that gives output categorical; more than that, we can calculate or see the most relevant word that can

associate with personality traits by visualizing the tree itself.

Research conducted by Rintaspon Bhannarai and Chartchai Doungsa-ard states that personality tests can be performed to predict agile people suitable for an agile methodology based on the theory of big five personality traits. With k-nearest neighbor as the classification technique used, the study gave an accuracy of 65.71% with its best k as 1 [4].

Research conducted by Harshal Chaudhari, Nalini Yadhav, and Yash Shukla predicts what types of jobs are suitable for each individual based on their resumes. The classification is done using the Naïve Bayes Classifier by calculating the likelihood of candidates getting the tasks based on the tokens calculated from the resumes. Tokens are calculated based on academic grades, hobbies, professional experiences, projects, publications, awards, etc. The expected result from this study is that the created model can help, companies select candidates and employ competent applicants in the right position [5].

To classify personalities using text data, personality theory that can be used as a basis for the personality classification is needed. For example, the Big Five Personality Traits theory is a personality theory that categorizes people into five different characteristics [6]. The five characteristics are Openness to Experience, Conscientiousness, Extraversion, Agreeableness, and Neuroticism. But in this study, the personality traits that support the applicant's personality classification process are Conscientiousness and Agreeableness. The dataset used is the text data that needs to be done through the cleansing process first, so that the data can be used for the applicant's personality classification process. The result of the classification of applicants is whether the applicant is suitable or not for the company based on personality theory. Several limitations arise in this study. That is the difficulty of collecting the required dataset. Because the dataset needed is the data result of interviews between applicants and companies. The problem lies in the code of ethics of psychology that forbids anyone other than the parties involved to find out the contents of the interview. There are many procedures needed to be followed to collect the interview data from the company, whereas making the interview data manually requires a significant amount of money to pay volunteers to be interviewed and psychologists to identify the personality of the volunteer. The result is that the data collected is limited.

The objective of this research is to classify applicants' personalities based on existing text data. The steps taken are processing the text data that will be used to classify the personality. The features used in the classification process are a collection of words that have been filtered beforehand. Then, the accuracies of the two methods are calculated to find out which method is better to apply for the classification and whether the system built is good enough to be implemented or not.

II. RESEARCH METHOD

A. Dataset

The dataset used in this study is data in the form of text obtained from interviews stored in audio and then converted into text through speech to text. There are nine assessments in the dataset that are not related to each other; the evaluation comes from the work culture that exists in the company with the associated personality theory, which is Conscientiousness and Agreeableness. The work culture used for three assessment criteria is integrity, enthusiast, and speed. Each of these assessments is labeled with two classes: 1 is to declare the applicant does not fit the company and 2 is to state that the applicant is suitable for the company. In this study, only three assessments are used because there are only three assessments balanced between the proportion of class 1 and class 2. The existing dataset is 53 in total, with a comparison between training data and test data at 90% for training data and 10% for test data.

The scoring process or how the applicant's personality can be classified is defined based on criteria from each work culture used in this study. Based on that criteria, scoring is defined as whether the applicants are suitable or not in the company. Using the number as explained before, 1 is for "not suitable" and 2 is for "suitable". An expert in the psychology major does the scoring process.

Table 1. Score Explanation

Score	Explanation
1	Not Suitable
2	Suitable

Table 1 is the score that represents the "suitable" and "not suitable" label.

Table 2. Personality Evaluation

No	Work Culture	Criteria
1	Integrity	-Honest
		-Behave positively
		-Comply with company rules and social norms
		-Not manipulative
		-Did not hide a mistake
		-Appreciate
		-Abusing the authority or facilities of the company

2	Enthusiast	- Excited - Happy - Seriously working to be the best
3	Speed	-Mental pioneer -Actions as directed -Quickly respond to business opportunities -The accuracy of product explanation -The Speed of providing QCD services (quality, cost, delivery) -Initiative -Agile to serve customers -Agile in business

Table 3. Personality Rubric

Label	Integrity	Enthusiasm	Speed
Suitable	Applicants can effectively articulate honest attitudes, positive behavior, and professional ethics that can positively affect the work environment.	Applicants effectively articulate a passion for sincerity at work, and a desire to do the best for themselves and the company.	Applicants effectively articulate agile attitudes and proactively provide the best for the company.
Not Suitable	Applicants cannot or less effectively articulate honest attitudes, positive behavior, and professional ethics that can positively affect the work environment.	Applicants cannot or less effectively articulate a passion for sincerity at work, and a desire to do the best for themselves and the company.	Applicants cannot or less effectively articulate an agile and proactive attitude to provide the best for the company.

Table 2 and Table 3 are the criteria used to determine to score for the applicant's personality, whether suitable or not in the company.

Table 4. Integrity Dataset

No	Name	Interview Content	Integrity
1	Talent A	balik lagi seperti yang tadi sih kondisinya jadi ya seadanya aja gitu maksimalin seadanya dan sebisanya gitu (*Back again as a condition before so just as it maximizes as it is and what I can*)	1
2	Talent B	biasanya yang menyebabkan buntu itu ada dua ada karena capek atau enggak tahu kalau misalnya capek istirahat kalau emang nggak tahu saya akan nyari informasi entah itu Mencari latar atau bertanya langsung ke orang-orang yang saya pikirkan mereka lebih paham (*usually, two things cause deadlock whether caused by tired or didn't know if tired then take a break if didn't know I will find information by searching for background or ask directly to people that I think more understand*)	1
3	Talent C	sebenernya yang saya percaya sampai sekarang ini nggak pernah yang ada namanya kebohongan untuk kebaikan dan yang saya tangkap dari pernyataan Bapak tadi itu merupakan yang kayak terpaksa itu aku bahkan untuk kebaikan itu sendiri tidak ada yang kata terpaksa untuk melakukan kebohongan tersebut dan bagaimana menyikapinya dengan saya tidak melakukannya karena itu hal buruk itu dapat berdampak baik di awal dan di jangka panjang itu dapat berdampak sangatlah buruk dan sangatlah jelek untuk individual diri sendiri maupun untuk perusahaan tempat saya bekerja nanti (*actually I believe that up to now there has never been such a thing as a lie for good and what I have captured from your statement earlier is that it is like being forced to me even for the goodness itself there is no word forced to do such lies and how to react to me is not doing it because it's a bad thing it can have a good effect in the beginning and in the long run it can have a very bad impact and is very bad for the individual myself or for the company where I work later*)	2

74

Table 4 is an example of the dataset that classifies an applicant's personality with one of the work cultures: integrity.

Table 5. Enthusiast Dataset

No	Name	Interview Content	Enthusiast
1	Talent A	saya sih lebih cenderung pengalaman saya dalam berorganisasi lah ya Ini mah yang sudah saya optimalkan paling saya selalu apabila diberi kepercayaan yang memaksimalkan kepercayaan itu itu karena menurut saya tanggung jawab itu sangat penting gitu (*I am more inclined to my experience in organizing, right? This is the thing that I have optimized, I always give the most trust when maximizing that trust, because I think responsibility is very important.*)	2
2	Talent B	kalau secara aturan yang aturan aturan yang tertulis seperti itu belum pernah melanggar tapi kalau misalnya dari segi budaya bisa di organisasi yang saya ikuti itu ada sedikit ada budaya yang ada pada budaya kemudian karena Suatu kondisi yang itu lumayan urgent ya istilahnya penting itu sayang melanggar sedikit tapi itu dengan izin kau yang lebih paham gitu ya Tergantung kepentingan yang ngomong itu penting banget dan terus pelanggaran itu mas istilahnya pelanggaran itu tuh nggak mempengaruhi ke yang lainnya juga kayak gitu (*if the rules are written rules like that have never been violated but if for example in terms of culture in the organization that I follow there is a little culture that exists in the culture then because A condition that is quite urgent yes the term is important it is unfortunately breaking a little but That is with your permission, you know better. Depending on the interests of the conversation, it is very important and then the violation, mas, the term violation does not affect the others as well.*)	1
3	Talent C	untuk yang upayakan sampai setelah lulus ini ya ikutlah dan mengajar di laboratorium mengajar di apa itu namanya Pelajaran di menjadi asisten laboratorium jadi asisten praktikum dan juga mengikuti pengabdian masyarakat dan lagi juga ada update paper ke dalam kalender nasional di sini juga kan sebelum presentasi dan memperoleh nilai IP yang cukup baik dikarenakan saya belajar dengan baik itu selama ini (*for those who are working until after graduating, yes, come and teach in the teaching laboratory, what is it called? Learning in being a laboratory assistant, becoming a practicum assistant and also following community service, and again there is also an update paper into the national calendar here, right before the presentation and obtaining IP scores are quite good because I studied well so far*)	2

Table 5 is an example of the dataset that classifies an applicant's personality with one of the work cultures: enthusiast.

Table 6. Speed Dataset

No	Name	Interview Content	Speed
1	Talent A	pernah sebenarnya itu solusinya jika dari saya saya melakukan sebanyak mungkin dulu berarti kita mengoptimalkan sisa waktu jadi ya sebanyak mungkin kita mau atau tidak karena sudah dihapus seperti itu jadi maksimalkan saja (*I have ever experienced it actually that's the solution if from me I did as much as possible first means we optimize the remaining time so yes as much as possible we want it or not because it's been deleted like that so just maximize it*)	2

75

No	Name	Interview Content	Speed
2	Talent B	Iya pernah mungkin kalau dibilang lumayan sering itu lumayan sering agak misal di harinya atau beberapa hari sebelumnya kayak gitu karena memang kebetulan ketika kuliah lumayan banyak (*Yes, it was once possible if you said it was quite often it was quite often quite a bit on the day or a few days before like that because it was a coincidence when there were quite a lot of lectures*)	1
3	Talent C	pasti pernah ya kalau misalnya kita dalam ada tugas yang seperti itu mengatakan ada mungkin seperti risoles yang kurang atau mungkin kayak kita dapat butuh sesuatu ternyata sesuatu hilang atau sesuatu kita tidak punya dan kita pas kita beli itu kita nggak apa kita nggak punya waktu lagi untuk membelinya karena udah mepet ya itu sebenarnya salah satu untuk mengatasinya itu Ya kita melakukan pekerjaan tersebut jauh-jauh hari jadi kalau misal kesalahan-kesalahan tersebut itu yang memang kita bisa temui akan bisa kita atasi didalam jauh-jauh hari jadi kita mempunyai persiapan lebih untuk mengerjakannya dan juga mempunyai waktu yang lebih juga untuk mempersiapkan jika memang ada sesuatu yang dibutuhkan yang kita tidak punya i (*there must have been yes if for example we were in a task like that saying there might be something like lacking risoles or maybe like we could need something it turned out that something was missing or something we didn't have and we were right when we bought it we didn't what we didn't have time to buy it because it's already tight yes it's actually one to overcome it Yes we do the work well in advance so if for example those mistakes that we can actually meet we can overcome in advance so we have more preparation to do it and also have more time also to prepare if there really is something needed that we don't have*)	2

Table 6 is an example of the dataset that classifies an applicant's personality with one of the work cultures: speed.

Table 4, Table 5, and Table 6 above show a lot of speech that contains informal speech for the Indonesian language. It is because the recorded audio is based on spontaneous interviews, resulting in colloquial words that the interviewee directly said. The applicant talks whether formal, semi-formal, or informal depends on the atmosphere or situation created by the interviewer.

B. Propose Method

a) Personality Traits

According to B. Y. Pratama and R. Sarno [7], personality is a combination of individual characteristics and behavior in overcoming various situations. Besides, personality also influences interactions with other people and the environment. Personality can be used as an assessment for employee recruitment, career consulting, relationships, and health. Attitudes and behavior can be explained by personality traits. Personality traits are beneficial for knowing individual psychological differences, psychological similarities between individuals, and identification of human nature [8]. It has a strong relation to the personal level, interpersonal level, life, and work decision [9], Personality traits affect leadership [10], ways of learning as well as work performance [11], academic ability and motivation [12]. So, when the company knows the applicant's personality, the company can put the applicant into a position that suits their personality best. By placing applicants in a suitable position, the company will increase its overall performance. Psychology offers personality tests to find out the personality traits of individuals [13]. An example of a personality test is the Big Five Inventory (BFI) [14].

The Big Five Personality traits or the Five-Factor Model is a personality theory that categorizes people into five different characteristics [6]. The five factors are Openness to Experience, Conscientiousness, Extraversion, Agreeableness, Neuroticism. Each of these factors has its features, namely Openness to Experience, with an active imagination, sensitive to feelings, and intellectual curiosity. Conscientiousness has characteristics that are neat, efficient, disciplined, conscientious, people who have these traits tend to be hard-working and reliable. Extraversion has features that are easy to socialize with other people and the environment; people who have these characteristics tend to enjoy spending time with many people by partying, doing community activities, and public demonstrations. People with these characteristics also tend to be better at doing work in groups. Agreeableness has features that are friendliness, sympathetic, cooperative, and caring; people who have these characteristics tend to prioritize shared interests rather than their interests. Neuroticism has characteristics that are anxious, worried, afraid, grumpy, frustrated, jealous, guilty, depressed, and lonely. People with these

76

characteristics tend to be moody, exaggerated responses to typical situations and consider a little problem as a huge problem.

b) Text Mining

Text mining is a method of taking information from unstructured text data [15]. Steps taken to gain knowledge in the text is usually by determining a specific pattern. Before determining the pattern, there are preprocessing stages in text mining that makes the data in the text to be the only data related to the information. This stage includes tokenization, normalization, and weighting of each word.

According to D. L. Olson and D.Delen [16], there are several techniques used in text mining: information retrieval, information extraction, text categorization, text summarizing, and text clustering. Text categorization will be used to collect the text, process it, and analyze it to determine the topic of the text. In this study, the personality possessed by each applicant is determined based on the text data of the applicant.

c) Multinomial Naïve Bayes

Naïve Bayes is a classification with probability and statistical methods. For each class selected, calculate the possibility based on the condition that the selected class is valid, where the vector is the object information [16]. Multinomial Naïve Bayes is a modified form of Naïve Bayes. With the same approach using probabilities, Multinomial Naïve Bayes is designed for document text by calculating the frequency of words [17]. By calculating the likelihood of a word's appearance in text data, then the category of the text can be determined. The formula used to calculate Multinomial Naïve Bayes is as follows:

$$P(X|c) = \log \frac{N_c}{N} + \sum_{i=1}^{n} \log \frac{t_i + \propto}{\sum_{i=1}^{n} t + \propto} \qquad (1)$$

Where:

$P(X|c)$: Probability of document X in class c
N_c : Number of documents in class c
N : Number of documents
t_i : Words that have value
$\sum_{i=1}^{n} t$: Words that have value in class c
$\propto$: Parameter smoothing

d) Decision Tree

Decision Tree is a classification method by making a tree structure like a flow chart where each node represents a feature or attribute as a classification criterion and a leaf node as a result of its classification [18]. Determination of features at each branch node in the Decision Tree is calculated based on the Gini index to determine which features or attributes most influence the classification

process. The formula for calculating the value of a feature based on the Gini index is as follows:

$$G = \sum_{K=1}^{K} P(1 - P) \qquad (2)$$

Where:

G: The gini value of a feature
K: Number of classes in attribute
P: The percentage of classes that appear in the attribute

e) Term Weighting TF-IDF

TF IDF is a technique in NLP (Natural Language Processing) used to extract essential keywords in a document. The TF-IDF algorithm has two working steps:

• Term Frequency

Calculates the frequency of a word based on the number of times the word appears in the text, the more often the word appears, the higher the value of the word.

• Inverse Document Frequency

Count the unique words that exist in a text, by separating the words that appear only in one text with the words that appear in every text. To calculate the value of TF-IDF is as follows:

$$tfidf_t = f_{t,d} \times log \frac{N}{df_t} \qquad (3)$$

Where:

$tfidf_t$: The value of the word t
$f_{t,d}$: The appearance of the word t in document d
N : Number of documents
df_t : The number of documents containing the word t

C. *System Development*

In this research, several steps need to be done to build a classification model, preprocessing data to clean the dataset for the training phase and validation phase, training the model classification and test the model for the validation phase.

a) Preprocessing

Because the dataset is not well-structured for classification yet, it is necessary to cleanse the data by removing punctuation, affixing, and changing every word in the text into a basic word. The process carried out is case-folding, remove punctuation, stopword removal, stemming, and tokenization.
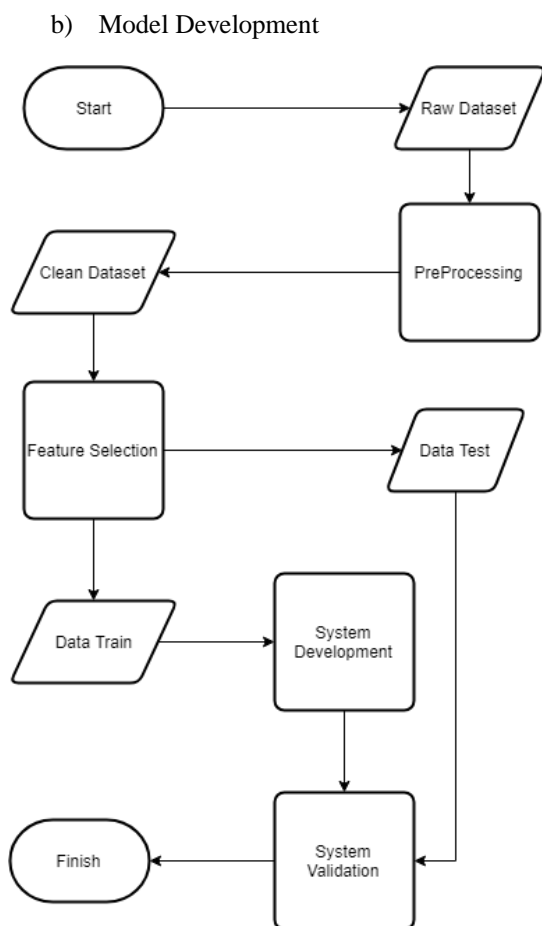
b) Model Development

Fig.1. Model Development Flowchart

Figure 1 is the flowchart that illustrates how the system's flow to classify an applicant's personality worked.

After processing the text, a feature selection is performed on the preprocessed text data using TF-IDF to calculate the frequency of words in the text data. After successfully giving value on the words, every word that passes the feature selection will be used to build a system model using Multinomial Naïve Bayes and Decision Tree. Before entering that stage, the dataset is divided into train data and test data for system development with a 90% ratio for train data and 10% for test data. The data train is used to build a classification model using the Multinomial Naïve Bayes and Decision Tree. Before entering that stage, the dataset is divided into train data and test data for system development with a 90% ratio for train data and 10% for test data. The learning process in the Decision Tree is done by calculating the feasibility of the feature in the dataset using the Gini index. After each feature gets its Gini value, the feature with the largest Gini value will be used as the root node. Other features that have smaller Gini values will act as branches from the root node, and the class in each assessment category act as its leaf node. Each preprocessing text data will enter each node in the Decision Tree,

starting from the root, branch node, and leaf node. If the text data can pass through the existing node until it reaches the leaf node, the text data class can be defined. Each assessment has its respective classification model because it has no connection with each other, so there are three classification models built in this study from each method.

c) Validation

After the system has been successfully created, it is necessary to validate the process to prove whether the system produced satisfactory results or not. Results can be evaluated by trying to test the existing test data. The parameter used to validate the system is Accuracy, recall, precision, and F-1 measure. Then the evaluation is done using the parameter equation as follows:

$$Accuracy = \frac{TP + TN}{TP + FP + FN + TN} \quad (4)$$

$$Precision = \frac{TP}{(TP + FP)} \quad (5)$$

$$Recall = \frac{TP}{(TP + FN)} \quad (6)$$

$$F-1\ Score = \frac{2 \times (Recall \times Precission)}{Recall + Precission} \quad (7)$$

Where:

$TP = True\ Positive$
$TN = True\ Negative$
$FP = False\ Positive$
$FN = False\ Negative$

With these parameters, it can be seen how much performance that the system has. After that, we can see how well the system performed by using the confusion matrix in Table 7,

Table 7. Confusion Matrix

	Actual Positive	**Actual Negative**
Predicted Positive	True Positive	False Positive
Predicted Negative	False Negative	True Negative

- True Positive (TP) is the value when the predicted data shows suitable personality, and the actual data shows a suitable personality
- True Negative (TN) is a value when the predicted data shows not suitable personality, and the actual data shows not suitable personality
- False Positive (FP) is the value when the predicted data shows suitable personality, and the actual data shows not suitable personality

- False Negative (FN) is a value when the predicted data shows not suitable personality, and the actual data shows a suitable personality

III. RESULT

After the construction of the classification model and data, validation is done. The following are the accuracy data of the classification model results made from the Multinomial Naïve Bayes and Decision Tree:

Table 8. Comparison of Method Accuracy

	Accuracy	
	Multinomial NB	Dec Tree
Integrity	83.4%	33.33%
Enthusiast	50%	66.66%
Speed	60%	80%

Table 8 shows a comparison of how well the performance of each classification model was built using the Multinomial Naïve Bayes and Decision Tree. The validation method is also performed by calculating the confusion matrix of each model built along with the recall, precision, and f-1 score

Table 9. Confusion Matrix of Classification Model

	Confusion Matrix			
	Multinomial NB		Decision Tree	
Integrity	1	1	1	1
	0	4	3	1
Enthusiast	0	3	1	2
	0	3	0	3
Speed	0	2	2	0
	0	3	1	2

Table 9 is the confusion matrix result from the three models that have been built with two methods, from confusion matrix we can calculate recall, precision, and an f-1 score of the model so we can know how well the model performed so far.

Table 10. Validation Table of The Multinomial Naïve Bayes

	Multinomial Naïve Bayes					
	Integrity		Enthusiast		Speed	
Class	Not Suit-able	Suit-able	Not Suit-able	Suit-able	Not Suit-able	Suit-able
Precision	1.00	0.80	0	0.50	0	0.60
Recall	0.50	1.00	0	1.00	0	1.00
F-1 Score	0.67	0.89	0	0.67	0	0.75

Table 11. Validation Table of The Decision Tree

	Decision Tree					
	Integrity		Enthusiast		Speed	
Class	Not Suit-able	Suit-able	Not Suit-able	Suit-able	Not Suit-able	Suit-able
Precision	0.25	0.50	1.00	0.60	0.67	1.00
Recall	0.50	0.25	0.33	1.00	1.00	0.67
F-1 Score	0.33	0.33	0.50	0.75	0.80	0.80

Table 10 and 11 is the validation result of the model that implements the multinomial Naïve Bayes method and Decision Tree method. From these tables, we can know the precision, recall, and f-1 score from each label at both methods. To compare two methods of how well they predict positive classes correctly, we can use the f-1 score. From the tables, multinomial Naïve Bayes has the f-1 score for integrity, enthusiast, and speed is 0.89, 0.67, and 0.75, respectively, and Decision Tree has 0.33, 0.75, and 0.80, respectively.
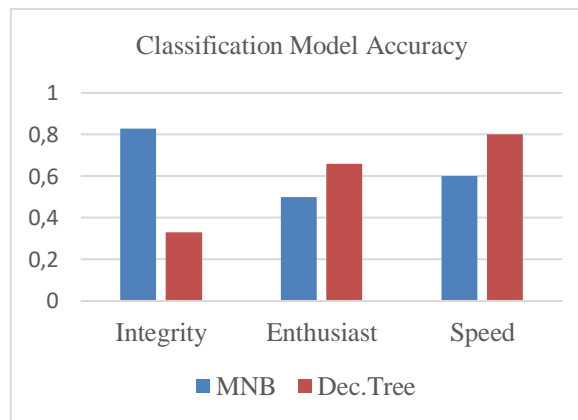

Fig.2. Comparison Graph Accuracy Method

Figure 2 illustrates the comparison of the accuracy from both method by using a graph. We can see that two models from the Decision Tree perform higher than multinomial Naïve Bayes.

IV. DISCUSSION

Based on the test results that have been shown previously, Multinomial Naïve Bayes can provide better performance than the Decision Tree of the three models. However, the results can be said, not optimal. Even though from accuracy and f-1 score, we can see that two models from Decision Tree can perform better than multinomial Naïve Bayes. But if we see it based on what method that more stable, multinomial Naïve Bayes has better performance. Because all models from multinomial Naïve Bayes can have accuracy 50% or higher, on the other hand, there is one model from the Decision Tree that has accuracy, only 33%. Many factors can affect the performance of the built classification model. One of them is from the dataset

itself because the dataset collected to create and test the classification model is interview data recorded into audio and converted into text data using speech to text. Inside the dataset, there are many data records whose sentences cannot be understood because of the words contained in them. This is most likely due to missed speech to text in converting the interviewee's words into text. Here are a few words or sentences that were rated wrong by researchers:

Table 12. Missed Sentence

Before Pre-Processing	After Pre-processing
jujur Tidak tidak jangan galau jangan memanipulasi angkat itu Nggak cuman ditelepon aja harusnya ya Kita sebagai pribadi sebagai pribadi yang yang lebih kalau itu general setiap orang lagi itu ya Mau di mana (*Truth no no don't be indecisive don't manipulate lift that not just trough calling supposedly yes we as an individual as an individual that that more if that's general everyone again that yes want where*)	jujur galau manipulasi angkat nggak cuman telepon aja pribadi pribadi general (*Truth indecisive manipulation lift not just through calling individual individual general*)
saya akan selalu berusaha untuk mengedepankan keuntungan untuk organisasi atau perusahaansaya akan selalu berusaha untuk mengedepankan keuntungan untuk organisasi atau perusahaan (*I will always prioritize the benefits for mycompany or the organization I will always try to prioritize the benefits for the company or the organization*)	usaha depan untung organisasi perusahaansaya usaha depan untung organisasi usaha (*The future effort mycompany's organization* *benefit the future effort for business* *organization's benefit*)

Table 12 shows the sentence that considered mistakenly converted. This can affect the performance of the classification model. Some words have the same meaning, but speech to text miss converting it. Thus, the system considers the word as a new feature, resulting in performance's reduction of the classification model. As in the table, the word "Perusahaan (company)" which is the basic word, became "Usaha (business)". Because of the missed conversion, there is a word that becomes "perusahaansaya (my company)" so that the basic word remains "perusahaansaya (my company)," this makes the word "Perusahaan (company)" in "perusahaansaya (my company)" valued as a new feature and not the same as the word "perusahaan (company)" from before.

In addition to the strange sentence or word factors above, the classification model's performance is also influenced by the stopword list. In the construction of personality classification models, libraries and lists for stopwords that are used are derived from literature. Instead of omitting them from the documents, some of stopwords in literature are being kept for giving broader information on personality classification.

Another factor also comes from the comparison of label classes that exist in each recording in each assessment. The following is the data comparison of label classes at each evaluation:

Table 13. Comparison of The Number of Label Classes

Assessments Criteria	Comparation of Label	
	Not Suitable	Suitable
Integrity	23	28
Enthusiast	20	31
Speed	20	28

Based on the data table above, there are only three appraisal data in which the number of label classes is quite balanced, namely in Integrity, Enthusiast, and Speed. This can affect the performance of the classification model because the system can accurately classify the model if the system truly has enough data for each label class. In addition to the comparison of biased label classes, the number of existing records is inadequate. It can be said that this test can have an excellent performance on several assessments because it only has two classes of each evaluation. If there is only one class with a sufficient amount of training data to be able to guess well in that class, then the recorded data, which was not included in the class classification, will be included in the other class. Still, if there are three classes, then the performance of the system will decrease.

V. CONCLUSION

Based on the research that has been done for the classification of applicants' personalities using the Multinomial Naïve Bayes and The Decision Tree, it can be concluded that the Multinomial Naïve Bayes provides better performance than the Decision Tree. Even so, these results cannot be said to be good because if the averaged results between the three models of the two methods, the multinomial Naïve Bayes produce an accuracy of 64.3% and the Decision Tree produce an accuracy of 59.67%, this is due to insufficient data available for the construction of the classification model. The preprocessing stage is also an essential factor to improve the performance of the model made, therefore even though this proposed classification model can classify the personality of the applicants based on the suitable with the work culture

of the company. Still, this classification model cannot be implemented yet on the actual condition.

For related research in the future, it is deemed necessary to have a special stopword list for the construction of personality classification models so that the results obtained can be more accurate. Besides, the dataset collection also needs to be done more, and the dataset has a balanced comparison of label classes.

ACKNOWLEDGMENT

REFERENCES

[1] F. Ahmed, M. Anannya, T. Rahman, and R. T. Khan, "Automated CV processing along with psychometric analysis in job recruiting process," 2nd Int. Conf. Electr. Eng. Inf. Commun. Technol. iCEEiCT 2015, no. May, pp. 21–23, 2015, doi: 10.1109/ICEEICT.2015.7307521.

[2] S. Sawleshwarkar, N. Rangnani, V. Mariwalla, and A. Halbe, "Simplified Recruitment Model Using Text-Mining on Psychometric and Aptitude Tests," Proc. 2nd Int. Conf. Electron. Commun. Aerosp. Technol. ICECA 2018, no. Iceca, pp. 586–589, 2018, doi: 10.1109/ICECA.2018.8474769.

[3] N. R. Ngatirin, Z. Zainol, and T. L. C. Yoong, "A comparative study of different classifiers for automatic personality prediction," Proc. - 6th IEEE Int. Conf. Control Syst. Comput. Eng. ICCSCE 2016, no. November, pp. 435–440, 2017, doi: 10.1109/ICCSCE.2016.7893613.

[4] I. Conference, I. Technology, R. Bhannarai, and C. Doungsa, "Person Identification Through Personality Test and k - NN Classification," pp. 215–219, 2016.

[5] H. Chaudhari, N. Yadav, and Y. Shukla, "A predictive analysis on job recruitment," pp. 6 (5 .)-6 (5 .), 2018, doi: 10.1049/cp.2016.1474.

[6] S. Dang, "A Review of Text Mining Techniques Associated with Various Application Areas," Int. J. Sci. Res., vol. 4, no. February, pp. 2461–2466, 2015.

[7] B. Y. Pratama and R. Sarno, "Personality classification based on Twitter text using Naïve Bayes, KNN and SVM," Proc. 2015 Int. Conf. Data Softw. Eng. ICODSE 2015, pp. 170–174, 2016, doi: 10.1109/ICODSE.2015.7436992.

[8] M. Fallahnezhad, M. Vali, and M. Khalili, "Automatic Personality Recognition from reading text speech," 2017 25th Iran. Conf. Electr. Eng. ICEE 2017, no. C, pp. 18–23, 2017, 10.1109/IranianCEE.2017.7985447.

[9] D. J. Ozer and V. Benet-Martínez, "Personality and the Prediction of Consequential Outcomes," Annu. Rev. Psychol., 2006, doi: 10.1146/annurev.psych.57.102904.190127.

[10] R. Hogan, G. J. Curphy, and J. Hogan, "What we know about leadership: Effectiveness and personality.," Am. Psychol., 1994, doi: 10.1037/0003-066x.49.6.493.

[11] A. Furnham, C. J. Jackson, and T. Miller, "Personality, learning style and work performance," Pers. Individ. Dif., 1999, doi: 10.1016/S0191- 8869(99)00053-7.

[12] M. Komarraju and S. J. Karau, "The relationship between the big five personality traits and academic motivation," Pers. Individ. Dif., 2005, doi: 10.1016/j.paid.2005.02.013.

[13] R. M. Kaplan and D. P. Saccuzzo, Psychological Testing seventh edition. 2012.

[14] S. John, O. P., & Srivastava, "Big Five Inventory (Bfi)," Handb. Personal. Theory Res., 1999, doi: 10.1525/fq.1998.51.4.04a00260.

[15] S. Sathya and N. Rajendran, "A Review on Text Mining Techniques," Int. J. Comput. Sci. Trends Technol., vol. 3, no. 5, pp. 274–284, 2013, [Online]. Available: www.ijcstjournal.org.

[16] D. L. Olson and D. Delen, Advanced data mining techniques. 2008.

[17] P. P. Surya, L. V Seetha, and B. Subbulakshmi, "Analysis of user emotions and opinion using Multinomial Naïve Bayes Classifier," 2019 3rd Int. Conf. Electron. Commun. Aerosp. Technol., pp. 410–415, 2019, doi: 10.1109/iceca.2019.8822096.

[18] H. Sharma and S. Kumar, "A Survey on Decision Tree Algorithms of Classification in Data Mining," Int. J. Sci. Res., vol. 5, no. 4, pp. 2094–2097, 2016, doi: 10.21275/v5i4.nov162954.