

# The Performance of Personality-based Recommender System for Fashion with Demographic Data-based Personality Prediction

Iman Paryudi<sup>1</sup>, Ahmad Ashari<sup>2</sup>, Khabib Mustofa<sup>3</sup>

Department of Computer Science and Electronics, Universitas Gadjah Mada, Yogyakarta, Indonesia<sup>1</sup>

Department of Informatics, Universitas Pancasila, Jakarta, Indonesia<sup>1</sup>

Department of Computer Science and Electronics, Universitas Gadjah Mada, Yogyakarta, Indonesia<sup>2</sup>

Department of Computer Science and Electronics, Universitas Gadjah Mada, Yogyakarta, Indonesia<sup>3</sup>

**Abstract**—Currently, the common method to predict personality implicitly (Implicit Personality Elicitation) is Personality Elicitation from Text (PET). PET predicts personality implicitly based on statuses written on social media. The weakness of this method when applied to a recommender system is the requirement to have minimal one social media account. A user without such qualification cannot use such system. To overcome this shortcoming, a new method to predict personality implicitly based on demographic data is proposed. This proposal is based on findings by previous researchers stating that there is a correlation between demographic data and personality trait. To predict personality based on demographic data, a personality model (rule) is needed. This model correlates demographic data and personality. To apply this model to a recommender system, another model is needed, that is preference model which connects personality and preference. These two models are then applied to a personality-based recommender system for fashion. From performance evaluation, the precision of and user satisfaction to the recommendation is 60.19% and 87.50%, respectively. When compared to precision and user satisfaction of PET-based recommender system (which are 82% and 79%, respectively), the precision of demographic data-based recommender system is lower whereas the satisfaction is higher.

**Keywords**—Implicit personality elicitation; demographic data; personality-based recommender system; personality trait

## I. INTRODUCTION

The first method to be used in a recommender system was content-based filtering which recommend items based on similarity between keywords on item description and on user's profile [1][2]. However, as it turned out, a content-based filtering has several weaknesses, one of which is its inability to distinguish the quality of items. This is because a good quality item will have the same keyword as a bad quality item [1].

Because of the glaring weakness in content-based filtering method, a new method, collaborative filtering, is used. The inability of content-based method to differentiate between different item qualities is solved by collaborative filtering by asking users to rate all the consumed items. This rating data is then used to calculate the rating of all new items [3], then this method will select top N items with highest ratings and then recommend these items along with the estimated ratings [4]. In practice, a rating-based collaborative filtering also has several

weaknesses, one of which is cold start problem or the new user problem [5]. This issue occurs when a recommender system is unable to provide a new user with accurate recommendations, because a new user does not have a record of what items has been consumed and the rating (the user profile is still empty).

To deal with the cold start problem, a user profile must be made available as soon as a new user becomes a member of a recommender system. The trick is that new users must fill in certain data when registering as a new member. Data that can be used in this case is personality trait. Afterward users will be given recommendations that match their personality traits. There are three advantages of using this personality trait [6]. The data pertaining to personality trait can be obtained in two ways, i.e. explicitly (Explicit Personality Elicitation) and implicitly (Implicit Personality Elicitation). The explicit method requires the user to answer a personality trait questionnaire to predict the personality trait. The commonly used personality trait questionnaire is based on the Big Five. As the name implies, Big Five consists of five factors/traits, namely: openness, conscientiousness, extraversion, agreeableness, and intellect. There are many Big Five based questionnaires that are available free of charge ranging from the longest with 504 questions to the shortest with only 10 questions [7].

If a recommender system utilizes the explicit method to obtain a user's personality traits, the user must answer a personality trait questionnaire before becoming a member of the system. Despite the fact that the method can accurately predict a user's personality traits; however, this method is burdensome and time-consuming for the user; therefore, the explicit method is only suitable for use in laboratory studies [8].

To overcome the weaknesses of the explicit personality elicitation method, a researcher may opt to use the implicit personality elicitation method. By using the latter method, a user's personality trait can be predicted, albeit indirectly. The current technique is called the Personality Elicitation from Text (PET). As the name implies, the users' personality traits are predicted from the posts they write, in this case on social media [9][10][11][12][13][14]. However, this method has one obvious weakness when applied to a recommender system, i.e. the user must have at least one active social media account.

In order to cope with the shortcoming of PET, a novel method of implicit personality elicitation is proposed, that is based on demographic data. In this new method, the user's personality is predicted based on demographic data.

To date, demographic data has been applied directly to a recommender system, hence the name Demographic Recommender System. Here the demographic information about the users is used by the classifiers to learn about how to find correlations between certain demographic data with ratings or buying tendencies [15]. However, there has been no research on the use of demographic data to predict human personality traits. The research is useful to overcome the weakness of implicit personality elicitation method, which is based on writing. This is where the gap lies in the personality-based recommender system research, specifically the implicit personality elicitation. This paper presents the result of the work in creating personality model connecting demographic data and personality traits. Next, the model is applied in a personality-based recommender system.

The idea to apply demographic data to predict personality traits come from the results of previous studies which found relationship between personality traits and demographic data. According to [16], one's personality can change or is stable at certain period in the course of his or her life. Except [17], other researchers such as [18] and [19] found that gender also affected personality traits.

Other demographic data such as race/ethnicity/country, hobbies, sport, occupation, zodiac, blood type, and color are known to also affect personality traits. Reference [20] found that persons from different countries have different personality traits. Furthermore, people with different personality traits tend to have different hobbies. An example is a person with a high score in openness is more likely to enjoy something abstract. Therefore, they are most likely are connoisseur of the arts and other forms of culture. Regarding sport, psychologically, a person's preference for a certain type of sport will be supported physically along with preference for a certain movements. Researchers also believe that different personalities will favor different movements.

Researchers also found correlation between personality traits and demographic data such as occupation [21], zodiac [22], blood type [23], and color [24].

The potential benefits of this new method of implicit personality elicitation are that it can be applied to any personality-based recommender system such as in an online shop, library, and travel company. By using this method, the system can give accurate and satisfying recommendations based on the users' demographic data instead of the users' rating history.

A summary of research trends in the recommender system and the proposed method is presented in Fig. 1.

The goal of this research is to create personality and preference models that when applied to a recommender system, then such system:

1) Can be used by any users without the need to have social media account or write status with certain length.

2) Has quite high users' satisfaction.

To achieve the goals, the following research questions must be answered:

1) Which demographic data or combination of demographic data that makes up the best model?

2) When the model is applied to a recommender system, how is the precision of the recommendation and satisfaction to the items recommended?

This paper is structured as follows: the next section, Section II, talk about the methods that are used in processing the data. Results and Discussion is presented in Section III. Section IV concludes the paper.

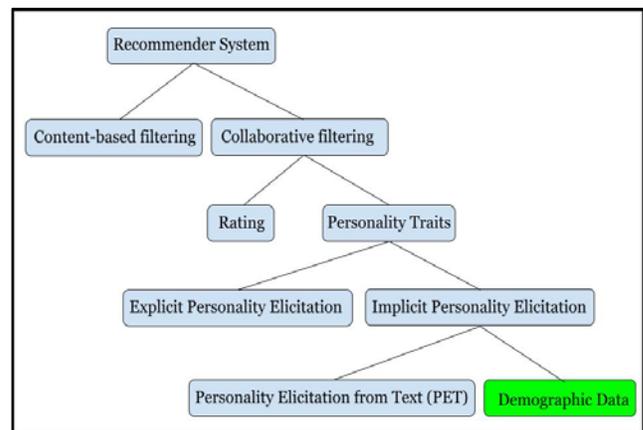


Fig. 1. Research Trends in the Recommender System and the Proposed Method.

## II. METHODS

The detailed survey methods has been presented in [25]. Below is the summary.

A total of 1014 respondents from several cities in Indonesia were involved in the current study. The questionnaire used in this study consists of three parts: i.e. demographic data, personality traits, and preferences.

In this survey, personality questionnaire based on Big Five was used. From a number of Big Five personality trait questionnaires that are available, the Indonesian IPIP 50 questionnaire [26] was chosen. The difference between IPIP 50 and the other Big Five-based questionnaires is that IPIP 50 does not use the term neuroticism; in its place, it uses the term emotional stability which is the opposite of neuroticism. Moreover, IPIP 50 also does not use the term openness; it uses the term intellect instead. To assess the personality traits, respondents were asked to score each question with a score of 1-5 where 1: strongly disagree, 2: disagree, 3: neutral, 4: agree, and 5: strongly agree.

In this survey, the following demographic data were collected: year of birth, marital status, city of residence, sport, occupation, hobbies, ethnicity, favorite color, zodiac, and blood type.

To learn more about preferences, data on respondents' preferences with regard to clothing styles were collected. There

are seven main clothing styles, i.e. Rebellious, Natural, Feminine, Elegant Chic, Dramatic, Creative, and Classic. It should be pointed out that initially this classification of clothing styles was intended for female users; hence, the styles are referred to as feminine and elegant chic. If the terms were applied to a male user, then obviously they would not be referred to as feminine and elegant chic. The term feminine for male users would be substituted with a style that matches the gender, i.e. masculine, whereas, the elegant chic style for male users would simply be referred to as fashionable.

As many as 105 samples of clothes that match the seven clothing styles, or 15 samples for each clothing/fashion style were provided. The respondents were asked to choose samples of clothes they liked. If they liked all the samples then they must choose all and vice versa, if they did not like any of the samples then they did not have to choose any.

After collecting the data, a three steps initial data processing were performed: (1) converted the year of birth into age, (2) calculated the total score of personality traits for each trait, (3) classified the data on the clothes samples selected by the respondents for each clothing style and counted the total number. The number of items selected by the respondent is used to determine the respondent's level of preference for a particular clothing style. If the number of selected items ranges from 1 to 5, then the level of preference is weak. If the number is from 6-10, then the level of preference is moderate. And if the number is greater than that, i.e. between 11 and 15 items, then the level of preference is strong. After that, the three most preferred clothing styles were determined, i.e. the three clothing styles with the highest number of selected items.

While doing the initial processing, the Cronbach alpha value was also calculated to determine the internal consistency of the data. Cronbach alpha was calculated using the following formula:

$$\alpha = \frac{N.c}{v+(N-1).c} \quad (1)$$

where:

$\alpha$  = Cronbach alpha.

N = the number of items.

c = average covariance between item-pairs.

v = average variance.

From the calculation, the following values were obtained: 0.801 for extraversion (good internal consistency), 0.773 for agreeableness (acceptable internal consistency), 0.844 for conscientiousness (good internal consistency), 0.908 for emotional stability (excellent internal consistency), and 0.749 for intellect (acceptable internal consistency).

It should be noted that the questionnaires for the research was made using Google Forms that does not put a limit on how many times a respondent can fill in the questionnaires. Therefore, a check needs to be done to find respondents who fill in the questionnaires more than once. During this check, as many as 25 such respondents were found; therefore, one of the duplicate data was deleted.

In addition to checking for duplicate data, another thing that needs to be checked was the presence of a certain respondent known as a self-enhancer. The presence of a self-enhancer is characterized by a high interscale correlation value, which is the value of the correlation coefficient between attributes. To calculate the interscale correlation, Pearson's correlation coefficient (r) formula was used:

$$r_{xy} = \frac{n(\sum x_i y_i) - (\sum x_i)(\sum y_i)}{\sqrt{[n \sum x_i^2 - (\sum x_i)^2][n \sum y_i^2 - (\sum y_i)^2]}} \quad (2)$$

where:

$r_{xy}$  = correlation between variables x and y.

n = the sample size.

$x_i$  and  $y_i$  = the  $i^{\text{th}}$  sample points.

Since there are five traits, there are 10 correlation coefficients must be calculated such as correlation between extraversion and agreeableness, correlation between extraversion and conscientiousness, etc. All these correlation coefficient values are then averaged to get the average interscale correlation.

The high interscale correlation value can happen because when filling in the questionnaires, self-enhancers tend to rate themselves higher than they should. For that reason, a self-enhancer will show a high level of personality traits in all traits. Therefore, to search for the presence of a self-enhancer, the total scores of personality traits of each respondent must be checked. Respondents who have a maximum score or near to the maximum score on all traits are considered as self-enhancers. In their study, [19] obtained an interscale correlation value of 0.19 and they claimed that such a value indicated that there were not many self-enhancers in the data. However, since a value of 0.38 was obtained, it was assumed that there were quite many self-enhancers in the data. After checking the data, 94 self-enhancers were found. After the data were deleted, the interscale correlation value dropped to 0.24.

When checking the data, data that did not make any sense at all were found. The data were from respondents who rated themselves with a score of 3 on all questions. Therefore, the data were deleted. After deleting the unwanted data, the remaining data is 894.

In building the model, the attribute that act as dependent attribute is level of personality traits. There are two levels: i.e. high and low. The level of personality trait was obtained in the following way: (a) by calculating the average score for each trait, (b) scores that were smaller than averages were designated as low level and scores that were higher than average were labeled as high level.

Additionally, in the modeling, the respondents' age group (categorical type) was used instead of age (numerical type). Therefore, the age data were grouped according to the classification laid down by the Indonesian Ministry of Health [25]. Based on the classification, the respondents were grouped into three groups, i.e. the middle age, adulthood, and adolescents.

The last stage in data processing was to remove some of the attributes that will not be used in the modeling stage. In the modeling, the following attributes were used: blood type, occupation, favorite color, gender, hobby, sport, zodiac group, zodiac component, age group, marital status, ethnicity, intellect level, emotional stability level, conscientiousness level, agreeableness level, extraversion level, preferred clothing style 1, preference level 1, preferred clothing style 2, preference level 2, preferred clothing style 3, and preference level 3.

### III. RESULT AND DISCUSSION

#### A. Modeling

It takes two models to build a personality-based recommender system for fashion; first, a personality model that links the demographic data with personality traits and the other one is a preference model that links the personality traits with a person's preference over fashion. Accordingly, in this stage, the two models were built.

1) *Personality model*: The process to create personality model has been presented in detail in [25]. Below is the summary.

The attributes that were used in the personality trait modeling were blood type, occupation, favorite color, gender, hobby, sport, zodiac group, zodiac component, age group, marital status, ethnicity, intellect level, emotional stability level, conscientiousness level, agreeableness level, and extraversion level. The first eleven attributes are demographic data that serve as independent attributes in the modeling using a decision tree. In addition to using the demographic data individually, a combination of two demographic data (e.g. blood type-occupation, blood type-age group) is used. By combining two demographic data, as many as 54 combinations are obtained, so the total number of demographic data used in the modeling is 65. Hence for each trait, as many as 65 models were created. Meanwhile, the level of personality traits (intellect level, emotional stability level, conscientiousness level, agreeableness level, and extraversion level) were used as the dependent attribute. To evaluate the model, a 10-fold cross validation was used.

Only one model will be used at a later stage. To select the model, the following criteria were used: (1) to make sure that the model can be used by everyone repeatedly at a later date, it has to be made certain that the demographic data in the model will never change, (2) it has to be also made sure that the model are fairly accurate. Based on these criteria, the model

based on age group and gender is chosen [25]. Another reason to choose this model is because previous research found that age and gender has very close relationship to personality traits [19]. Table I presents the model.

2) *Preference model*: As in the personality model, the detailed process in making this preference model has been presented in [27]. The summary is presented.

In the data processing stage, three preferences data and their level (preferred clothing style 1, preference level 1, preferred clothing style 2, preference level 2, preferred clothing style 3, and preference level 3) were selected from each respondent. However, before building the model, a preference data for each clothing style must be created. The data was obtained by combining the respondent's three preferences data into one. Table II shows example preferred data for Natural clothing style that will be used in the modeling. Preferred data for other clothing styles, i.e. Dramatic, Classic, Elegant Chic, Creative, Rebellious, and Feminine were also created.

The data used in building the preference model comprises the levels of the five personality traits and preferences. In the modeling, the class association rule method was used with personality trait's levels as antecedents and preferences as consequent or class.

Association rule is a method to discover a rule that connect items on a transaction. There are at least two measures that are used to identify good rule: support and confidence. If N is the number of transaction, then support of item X, Y is defined as:

$$Support(X) = \frac{Frequency\ of\ X}{N} = P(X) \tag{3}$$

$$Support(X, Y) = \frac{Frequency\ of\ XY}{N} = P(X \cap Y) \tag{4}$$

Meanwhile confidence of  $(X \rightarrow Y)$  is defined as:

$$Confidence(X \rightarrow Y) = \frac{Support(XY)}{Support(X)} = P(Y|X) \tag{5}$$

The personality trait model shown in Table I reveals that the levels of personality traits for extraversion, agreeableness, and conscientiousness are the same for all groups, they are: low for extraversion, high for agreeableness and conscientiousness. Therefore, in this preference modeling, only data with low extraversion, high agreeableness, and high conscientiousness were used. The modeling also only used the data with moderate and strong preference levels. Since it is not possible to recommend men's clothing to women and vice versa; therefore, the men and women were separated.

TABLE I. PERSONALITY MODEL (E: EXTRAVERSION, A: AGREEABLENESS, C: CONSCIENTIOUSNESS, ES: EMOTIONAL STABILITY, I: INTELLECT)

	E	A	C	ES	I	Group
Adolescence Male	Low	High	High	Low	High	1
Adulthood Male	Low	High	High	High	Low	2
Middle Age Male	Low	High	High	High	High	3
Adolescence Female	Low	High	High	Low	Low	4
Adulthood Female	Low	High	High	High	Low	5
Middle Age Female	Low	High	High	High	Low	6

TABLE II. SAMPLE OF PREFERRED DATA ON NATURAL CLOTHING STYLE

Preference	Level
Elegant Chic	Strong
Natural	Weak
Natural	Strong
Natural	Moderate
Natural	Moderate
Rebellious	Strong
Natural	Strong
Natural	Moderate
Natural	Weak
Natural	Moderate
Feminine	Moderate
Natural	Moderate

TABLE III. PREFERENCE MODELING RESULT. THE YELLOW CELLS SHOW THE PREFERRED CLOTHING STYLE OF EACH PERSONALITY GROUP (E = EXTRAVERSION, A = AGREEABLENESS, C = CONSCIENTIOUSNESS, ES = EMOTIONAL STABILITY, I = INTELLECT, CL = CLASSIC, CR = CREATIVE, DR = DRAMATIC, EC = ELEGANT CHIC, FE = FEMININE, NA = NATURAL, RE = REBELLIOUS.)

Group	Personality Traits					Fashion Style						
	E	A	C	ES	I	Cl	Cr	Dr	EC	Fe	Na	Re
1	Low	High	High	Low	High	0,6	0,5	0,5	0,4	0,7	0,8	0,4
2	Low	High	High	High	Low	0,5	0,7	0,7	0,8	0,5	0,8	0,7
3	Low	High	High	High	High	0,6	0,26	0,3	0,7	0,6	0,9	0,3
4	Low	High	High	Low	Low	0,6	0,4	0,4	1	0,3	0,9	0,4
5	Low	High	High	High	Low	0,5	0,3	0,4	1	0,3	0,8	0,3
6	Low	High	High	High	Low	0,5	0,3	0,4	1	0,3	0,8	0,3

TABLE IV. THE PREFERENCE MODEL

Group	Personality Traits					Favorite Fashion Style
1	Low Extraversion	High Agreeableness	High Conscientiousness	Low Emotional Stability	High Intellect	Natural, Feminine
2	Low Extraversion	High Agreeableness	High Conscientiousness	High Emotional Stability	Low Intellect	Elegant Chic, Natural
3	Low Extraversion	High Agreeableness	High Conscientiousness	High Emotional Stability	High Intellect	Natural, Elegant Chic
4	Low Extraversion	High Agreeableness	High Conscientiousness	Low Emotional Stability	Low Intellect	Elegant Chic, Natural
5	Low Extraversion	High Agreeableness	High Conscientiousness	High Emotional Stability	Low Intellect	Elegant Chic, Natural
6	Low Extraversion	High Agreeableness	High Conscientiousness	High Emotional Stability	Low Intellect	Elegant Chic, Natural

Two preferred clothing styles for each personality group are selected based on the highest confidence value. The two fashion styles with the highest confidence value were chosen as the preferred clothing styles (Table III). After selecting the two preferred clothing styles, the preference model can be created and is presented in Table IV.

### B. System Performance

1) *System architecture*: The author in [7] developed a personality-based recommender system in which personality traits were predicted using a questionnaire (explicit method).

Basically, the system consists of two parts, i.e. a part for predicting the personality traits and a part to find the nearest neighbors. Meanwhile, [28] used the personality elicitation from text (implicit method) to predict personality traits in their system. The system they propose basically also consists of two parts, i.e. a part for predicting the personality traits and a part to find the nearest neighbors.

In reference to the two studies above, the recommender system that is built also consists of two parts, i.e. a part for predicting the personality traits and a part to find the nearest neighbors (Fig. 2).

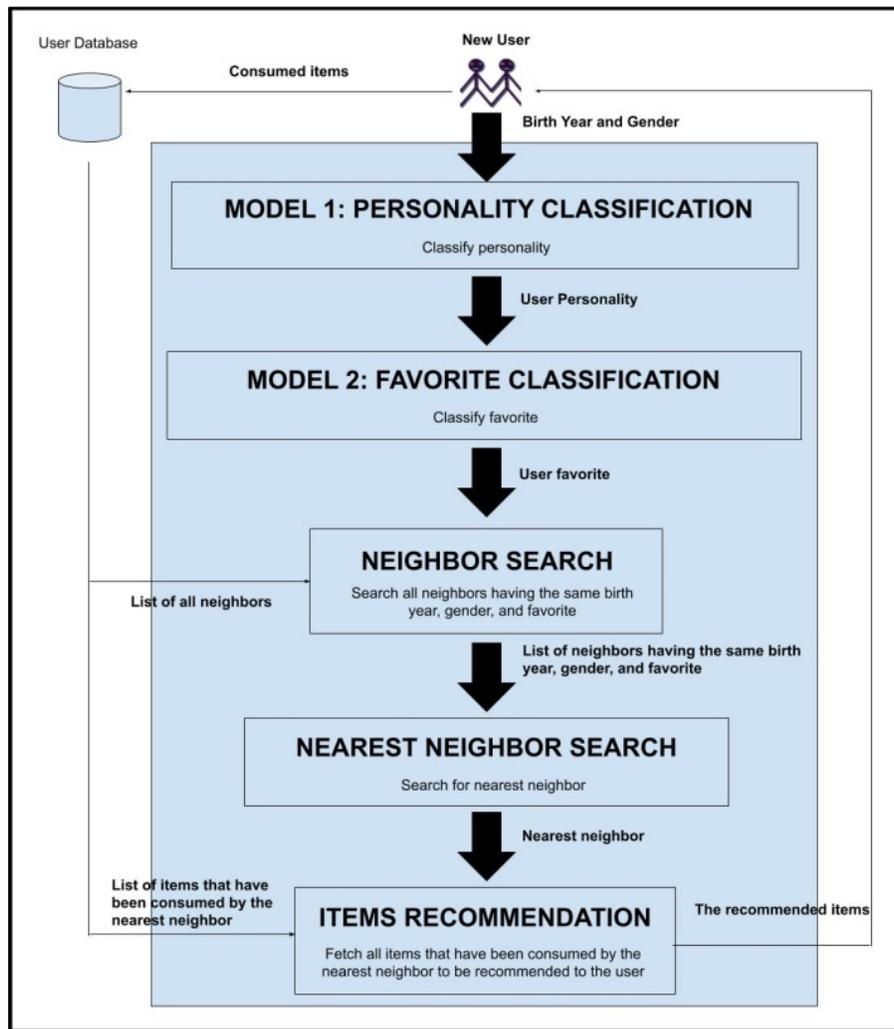


Fig. 2. The Architecture of the Proposed Recommender System.

Before discussing the model, it is necessary to explain what is meant by neighbors—neighbors are respondent's data collected during the data collection stage and modeled at the modeling stage. When the respondents were filling in the questionnaire, they were asked to pick the items that they liked; therefore, this data were treated as data about users who had consumed certain items. These data were stored in the user database as a basis for providing recommendations.

The built system includes a personality model (Model 1) and preference model (Model 2) which was obtained in the previous stage. As mentioned before, the personality model contains rules that link the demographic data to personality traits. Meanwhile, the preference model contains rules that link the personality traits to preferences.

The system starts working when a new user enters the year of birth and gender into the system. Using these data, the system will classify users into certain personality traits (predicting the user's personality traits) based on Model 1 (personality model). These personality traits are passed along to Model 2 (preference model) that will classify user's preferences based on the user's personality traits (predicting user's preferences).

After generating user's preferences, the system will retrieve all neighbors in the user's database in search of all neighbors whose year of birth, gender, and preferences are the same as the user's year of birth, gender, and preferences. The result is a list of neighbors with the same year of birth, gender, and preferences as the user.

By using the filtered list of neighbors, the system will search for the nearest neighbor. It should be explained that in this study, all users are using the model for the first time. Therefore, the user has never consumed any items. Because of that, at this stage, the system looks for the nearest neighbor based only on the same year of birth. If there are several neighbors with the same year of birth, then the system will pick one neighbor.

After the nearest neighbor is obtained, the system will collect all the items that have been consumed by the nearest neighbor. These are the items that will be recommended to the users. After consuming some or all of the recommended items, the new user data will be saved to the user database.

2) *System evaluation:* A number of researchers have used direct method to evaluate the systems they have built. A survey

conducted by [29] involved 21 students as the respondents. In the survey, the respondents were asked about the novelty of the items recommended, the accuracy, and satisfaction with the recommendations given. In another study, [30] also carried out a direct survey to users to find out the respondents' level of satisfaction with the built model.

To evaluate the performance of the proposed recommender system, a direct survey involving 74 respondents was conducted. In the survey, the respondents were asked to interact with the system.

- Relevance of the recommended items. In this experiment, the system recommends a number of items to the user. The system provides a Like button on all recommended items. When checking the recommended items one by one, if the button was pressed, it means that the item is relevant. Then the precision of the relevant items can be calculated. Precision is the percentage of relevant items from all the recommended items.

$$\text{Precision} = \frac{\text{Number of relevant items}}{\text{Number of recommended items}} \quad (6)$$

Here each user has one precision value, whereas the system's precision is the average precision of all users.

- Satisfaction to the recommended items. To find out user's satisfaction, the CSAT method is used. By using the method, user's satisfaction can be gauged by asking them how satisfied they are with the goods or services they used. There are five scales used to assess user's satisfaction, i.e. very dissatisfied, dissatisfied, neutral, satisfied, and very satisfied. The percentage of satisfaction is calculated using the following formula:

$$\text{Satisfaction} = \frac{\text{Number of responses}}{\text{Total number of responses}} \times 100\% \quad (7)$$

Note that the responses used in the formula are only Satisfied and Very Satisfied ones.

In the study, the system will ask the question: "How do you rate the recommended items?" immediately after the respondent finished checking the recommended items. The respondents were supplied with five answers as follows: very dissatisfied, dissatisfied, neutral, satisfied, and very satisfied. The percentage of user satisfaction is only calculated from the responses that give the value of Satisfied and Very Satisfied.

### C. Evaluation Result

From the evaluation, the following facts were obtained:

- 1) Precision of the recommendation was 60.19%.
- 2) Satisfaction to the items recommended was 87.50%.
- 3) There were as many as 17 respondents whose precision less than 50%, nevertheless satisfied or even very satisfied with the recommendation.
- 4) There were as many as 5 respondents whose precision more than 50%, yet dissatisfied (neutral) with the recommendation.

From the facts above, it can be said that a respondent with low precision can still be satisfied with the recommendation. In opposite, a respondent with high precision can be dissatisfied with the recommendation. Therefore, it is hypothesized that satisfaction correlates with level of preference, and not with precision. The statement on satisfaction has no correlation with accuracy has been confirmed by [31] and [32]. Note that precision is one of accuracy metrics used in recommender system besides recall, MAE, and MSE.

User's satisfaction is a psychological condition that can be measured from the user's expectation. A user will satisfy if the products or services offered to them exceeds or at least the same as the expectation. On the contrary, a user does not satisfy if the experience when using the products or services below the expectation [33]. Based on this, the reason why 17 respondents with low precision still satisfied or very satisfied with the recommendation is because they like the items recommended very much (high preference level) despite they only like a few of the many items recommended. In other word, the user's expectations are fulfilled. Meanwhile, in the case of 5 respondents with high precision but dissatisfy; it is because they do not really like the items recommended (medium preference level). As a result, they dissatisfy with the recommendation although they like many items. In other word, the user's expectation still not met.

The author in [29] reported that the precision and satisfaction of PET-based system were 82% and 79%, respectively. Compared to demographic data-based system performance, the precision of PET-based system is better but demographic data-based system is better in satisfaction. In recommendation system, accuracy is important but accuracy alone is not enough. This is because user satisfaction is more important. The next paragraph explains about this.

According to [31] two main tasks of accuracy metrics in recommender system are:

- 1) To measure the accuracy of single prediction. This is called predictive accuracy metrics. This metric calculate how close the predicted rating from the actual rating. Mean absolute error (MAE) and mean squared error (MSE) are used in this metric.
- 2) To evaluate the effectivity of the system in selecting the high quality items from a set of available items. This metric is called decision-support metrics and uses precision and recall to declare the accuracy.

Furthermore, [31] stated that building a recommender system with high accuracy was not enough. This is because the most accurate recommendation based on those metrics above, is sometimes not the useful recommendation for or liked by the users. This causes dissatisfaction to the users. In other word, user satisfaction does not always correlate with high accuracy [32][31]. Knowing the importance of user's satisfaction, [31] and [34] stated that it is not fair to judge a recommender solely from its accuracy. The users must be taken into account since they do not care with the algorithm to increase the recommendation accuracy. They only want to have useful recommendations.

Back to the performance comparison between demographic data-based system and PET-based one where the demographic data-based system has lower accuracy but higher satisfaction. Based on the above explanation, recommender system whose satisfaction is higher is better. In other word, demographic data-based system is better than the PET-based system.

#### IV. CONCLUSION

A research has been carried out to find model connecting demographic data and personality traits. As many as 65 models of each trait were created. From those models, the one based on age group and gender was selected as the working model since it satisfied two criteria. Besides that, previous research also found that age and gender had very close relationship to personality traits.

From the performance evaluation, precision and satisfaction of the demographic data-based recommender system were 60.19% and 87.50% respectively. When compared to PET-based system, demographic data-based system is lower in precision but higher in satisfaction. Other advantage of demographic data-based system compared to PET-based system is there is no obligation to have social media account.

Despite the strength of demographic data-based system compared to PET-based system, this research has some limitations:

- 1) The fashion classification may differ from other classifications. So, when a clothing is classified into a certain clothing style, others may classify it into different clothing style.
- 2) Some of the respondents may be come from low-income communities who do not care with fashion hence the choice of preferred clothing different from other respondents with the same category (e.g. age group – gender).

To test the hypothesis about the relationship between satisfaction and level of preference, another experiment is needed. In the experiment the respondents are asked not only to determine whether they like an item or not, but also the level of preference to the item. One way to obtain the level of preference is by providing five stars on each item. The more stars given by a respondent to an item, the higher the level of preference of the respondent to that item.

#### REFERENCES

[1] U. Shardanand and P. Maes, "Social information filtering: algorithms for automating 'word of mouth,'" *Conf. Hum. Factors Comput. Syst. - Proc.*, vol. 1, pp. 210–217, 1995.

[2] H. Li, F. Cai, and Z. Liao, "Content-based filtering recommendation algorithm using HMM," *Proc. - 4th Int. Conf. Comput. Inf. Sci. ICCIS 2012*, no. March, pp. 275–277, 2012, doi: 10.1109/ICCIS.2012.112.

[3] R. Burke, A. Felfernig, and M. H. Göker, "Recommender systems: An overview," *Assoc. Adv. Artif. Intell.*, 2011.

[4] T. Bogers and A. van den Bosch, "Collaborative and Content-based Filtering for Item Recommendation on Social Bookmarking Websites," in *Proceedings of the ACM RecSys'09 Workshop on Recommender System & the Social Web*, 2009, doi: 10.1007/978-3-642-25694-3.

[5] N. Rubbens, M. Elahi, M. Sugiyama, and D. Kaplan, "Active Learning in Recommender Systems," in *Recommender Systems Handbook*, Second Edi., F. Ricci, L. Rokach, and B. Shapira, Eds. New York: Springer, 2015, pp. 809–846.

[6] I. Paryudi, A. Ashari, and A. M. Tjoa, "Personality Estimation using

Demographic Data in a Personality-based Recommender System: A Proposal," 2019, doi: 10.1145/3366030.3366098.

[7] M. Augusta Silveira Netto Nunes, "Recommender Systems based on Personality Traits," *Université Montpellier II*, 2008.

[8] M. Tkalčić and L. Chen, "Personality and Recommender Systems," in *Recommender Systems Handbook*, Second Edi., F. Ricci, L. Rokach, and B. Shapira, Eds. New York: Springer, 2015, pp. 715–739.

[9] V. Ong, A. D. S. Rahmanto, W. Williem, N. H. Jeremy, D. Suhartono, and E. W. Andangsari, "Personality Modelling of Indonesian Twitter Users with XGBoost Based on the Five Factor Model," *Int. J. Intell. Eng. Syst.*, vol. 14, no. 2, pp. 248–261, 2021, doi: 10.22266/ijies2021.0430.22.

[10] A. V. Kunte and S. Panicker, "Using textual data for Personality Prediction:A Machine Learning Approach," in *2019 4th International Conference on Information Systems and Computer Networks, ISCON 2019*, 2019, no. March, pp. 529–533, doi: 10.1109/ISCON47742.2019.9036220.

[11] H. Christian, D. Suhartono, A. Chowanda, and K. Z. Zamli, "Text based personality prediction from multiple social media data sources using pre-trained language model and model averaging," *J. Big Data*, vol. 8, no. 1, pp. 1–20, 2021, doi: 10.1186/s40537-021-00459-1.

[12] A.-R. Feizi-Derakhshi, M.-R. Feizi-Derakhshi, M. Ramezani, N. Nikzad-Khasmakhi, and M. Asgari-Chenaghlu, "The state-of-the-art in text-based automatic personality prediction," *arXiv*, pp. 1–27, 2021.

[13] Z. Wang, C. H. Wu, Q. B. Li, B. Yan, and K. F. Zheng, "Encoding text information with graph convolutional networks for personality recognition," *Appl. Sci.*, vol. 10, no. 12, 2020, doi: 10.3390/APP10124081.

[14] X. Wang, Y. Sui, K. Zheng, Y. Shi, and S. Cao, "Personality classification of social users based on feature fusion," *Sensors*, vol. 21, no. 20, 2021, doi: 10.3390/s21206758.

[15] C. C. Aggarwal, *Recommender Systems the Textbook*, First Edit. New York: Springer, 2016.

[16] M. A. Harris, C. E. Brett, W. Johnson, and I. J. Deary, "Personality Stability From Age 14 to Age 77 Years," *Psychol. Aging*, vol. 31, no. 8, pp. 862–874, 2016, doi: 10.1037/pag0000133.supp.

[17] J. S. Hyde, "The gender similarities hypothesis," *Am. Psychol.*, vol. 60, no. 6, pp. 581–592, 2005, doi: 10.1037/0003-066X.60.6.581.

[18] M. Vianello, K. Schnabel, N. Sriram, and B. A. Nosek, "Gender Differences in Implicit and Explicit Personality Traits," *Pers. Individ. Dif.*, 2013, doi: 10.2139/ssrn.2249080.

[19] C. J. Soto, O. P. John, S. D. Gosling, and J. Potter, "Age Differences in Personality Traits From 10 to 65: Big Five Domains and Facets in a Large Cross-Sectional Sample," *J. Pers. Soc. Psychol.*, vol. 100, no. 2, pp. 330–348, Feb. 2011, doi: 10.1037/a0021717.

[20] D. P. Schmitt et al., "The geographic distribution of Big Five personality traits: Patterns and profiles of human self-description across 56 nations," *J. Cross. Cult. Psychol.*, vol. 38, no. 2, pp. 173–212, 2007, doi: 10.1177/0022022106297299.

[21] D. Heller, W. Q. E. Perunovic, and D. Reichman, "The future of person-situation integration in the interface between traits and goals: A bottom-up framework," *J. Res. Pers.*, vol. 43, no. 2, pp. 171–178, 2009, doi: 10.1016/j.jrp.2008.12.011.

[22] J. J. F. van Rooij, "Introversion-extraversion: astrology versus psychology," *Pers. Individ. Dif.*, vol. 16, no. 6, pp. 985–988, 1994, doi: https://doi.org/10.1016/0191-8869(94)90243-7.

[23] M. Rogers and A. I. Glendon, "Blood type and personality," *Pers. Individ. Dif.*, vol. 34, pp. 1099–1112, 2003.

[24] R. Miao, "Colour preference," *Birmingham*, 2017.

[25] I. Paryudi, A. Ashari, and K. Mustofa, "Modelling Relationship between Demographic Data and Personality Traits," *Turkish J. Comput. Mat. Educ.*, vol. 12, no. 14, pp. 2247–2255, 2021.

[26] H. Akhtar and S. Azwar, "Indonesian Adaptation and Psychometric Properties Evaluation of the Big Five Personality Inventory: IPIP-BFM-50," *J. Psikol.*, vol. 46, no. 1, pp. 32–44, 2019, doi: 10.22146/jpsi.33571.

[27] I. Paryudi, A. Ashari, and K. Mustofa, "Creating Personality and Preference Models based on Demographic Data for Personality-based Recommender System for Fashion using Decision Tree and Association

- Rule,” Turkish J. Comput. Mat. Educ., vol. 12, no. 14, pp. 5165–5174, 2021, [Online]. Available: <https://turcomat.org/index.php/turkbilmate/article/view/11548>.
- [28] A. Roshchina, J. Cardiff, and P. Rosso, “TWIN: Personality-based Intelligent Recommender System,” J. Intell. Fuzzy Syst., vol. 28, no. 5, pp. 2059–2071, 2015, doi: 10.3233/IFS-141484.
- [29] A. Di Rienzo and A. Neishabouri, “Recommendations with personality traits extracted from text reviews,” in Studies in Computational Intelligence, vol. 616, Springer Verlag, 2016, pp. 355–364.
- [30] R. Hu, “Design and User Issues in Personality-based Recommender Systems,” in RecSys’10, 2010, p. 386.
- [31] C.-N. Ziegler, S. M. McNee, J. A. Konstan, and G. Lausen, “Improving recommendation lists through neighbor diversification,” 2005.
- [32] S. M. McNee et al., “On the Recommending of Citations for Research Papers On the Recommending of Citations for Research Papers,” in Proc. of ACM CSCW 2002, 2002, no. November 2002, doi: 10.1145/587078.587096.
- [33] F. Tjiptono, Service, Quality, and Satisfaction. Yogyakarta: Penerbit Andi, 2016.
- [34] C. Hayes, P. Massa, P. Avesani, and P. Cunningham, “An on-line evaluation framework for recommender systems,” in Workshop on Personalization and Recommendation in E-Commerce, 2002, no. May.