Statistical Analysis of Differences in Personality Among Ages and Genders

Nathan Kardos

The College of New Jersey

Abstract

Personality and measuring the dimensions on which people's temperament varies may seem like an elusive subject to many. However, using the Big Five theory of personality we can create a model from which we can conduct such measures. Using this model, we see that women and men differ the most in two traits, Neuroticism and Agreeableness. While age has a positive correlation with people's expression in traits Conscientiousness and Agreeableness while it has a strong negative correlation with trait Neuroticism. Also, once we have a way to measure personality we can make predictions on gender and age based on one's personality measures. From a data set of 19512 observations ranging from ages 13 to 70, we were able to predict gender with 64.7% accuracy using Linear Discriminant Analysis and Logistic Regression while we predicted age with a mean absolute difference of 7.69 years between actual values and predicted values using Support Vector Machines.

Introduction

Personality has been theorized about from multiple levels of abstraction and detail. All of these levels help us understand the variance among individual experience and behavior. There has been endless debate among personality researchers and practitioners about which model from the immense array of personality scales is the best at describing an individual. Many researchers had hoped that they would invent an overarching taxonomy that would integrate the entirety of personality research and succeed in having the scientific community reach a consensus in personality vernacular. After decades of study, the field started to approach unanimous agreement on a general classification of personality traits, the "Big Five" personality dimensions (John, 1999).

The traits that make up the Big Five are the following: Neuroticism, Extraversion, Agreeableness, Openness/Intellect, Conscientiousness. The defining attribute of the Big Five model is that it does not stem from any particular theorizing of personality traits, rather it is derived from an analysis of the language and terms that people use to describe themselves and each other.

Neuroticism

Neuroticism describes one's propensity to experience negative emotions and responses to perceived threat and punishment such as anxiety, depressions, anger, etc (Weisberg, 2011). Previous research has found women to score higher than men on Neuroticism when measured at the big five level. Women tend to score higher on almost all facets of Neuroticism that are included in the trait. The only facet of Neuroticism that women do not consistently have higher scores in is Anger, or Angry Hostility (Terracciano, 2001). Neuroticism increases for women from childhood to adolescence then decreases throughout adulthood. Males however, tend to gradually decrease in Neuroticism from childhood into adulthood (Soto,2011).

Conscientiousness

Conscientiousness encapsulates traits related to control of impulses, self-discipline, organization, and the ability to demonstrate self-control to follow guidelines and maintain goals. No significant difference between genders has been found in this trait at the Big Five trait level (Weisberg, 2011). Past adolescence, Conscientiousness has also been found to increase with age (Soto, 2011).

Agreeableness

Agreeableness is built from traits relating to altruism and the desire to maintain cooperation and social harmony. Women have consistently been shown to score higher in this trait (Weisberg, 2011). Agreeableness tends to increase with age and has an almost inverse relationship with neuroticism, this trend can especially be observed in women (Soto, 2011).

Extraversion

Extraversion categorizes one's ability to be social, assertive, and experience positive emotion, all of which have been linked to sensitivity to positive emotions (DeYoung, 2009). Gender differences on the big five are very small but the small effect size could be due to strong presence of gender differences in different directions at the facet level. Women score higher on some traits that comprise extraversion such as Warmth, Gregariousness, and positive emotion. On the other hand, men score higher on Assertiveness and Excitement Seeking (Terracciano, 2001).

Openness/Intellect

Openness/Intellect reflects traits related to creativity, intellectual curiosity, and appreciation of aesthetics. This trait broadly encompasses one's ability to process complex stimuli. Similar to Extraversion, there are no significant differences between genders possibly because of the divergent facets of the trait. Women score higher on the facets of Esthetics and Feelings, where men consistent score higher on the facet of Ideas (Terracciano, 2001). Put simply, women are more interested in esthetics while men are more interested in ideas, another good example to why it is useful to break apart the Big Five into smaller categorizations.

In essence, the Big Five theory makes two claims, one stronger and one weaker. The strong claim is that the dimensions on which humans perceive each other, or in other words, the dimensions on which personality varies, can be derived from language and a grouping of correlated terms. Assuming traits are hierarchically organized, the weaker claim is that five dimensions are the highest order factors and encapsulate all of personality variations (Ashton, 2004). Personality is typically centered around 2 levels of traits: the broad Big Five regions, and many more specific traits, called facets, which are grouped inside the Big Five. In this study, we will only examine personality variances to the specificity of the Big Five domains. Also it is important to note the relationships between the facets and their shared variance within the Big Five domain as summarize by;

> If the Big Five constituted the level of the personality hierarchy immediately above the facets, only one factor should be necessary to explain the shared variance of the facets within a given Big Five domain. However, a large behavioral genetic study revealed that *two* distinct factors were necessary to account for the shared genetic variance among the facets within each domain. In a separate study using factor analysis of 15 different facets within each domain, two phenotypic factors similar to the genetic factors were found for each of the Big Five dimensions. (Weisberg, 2011)

These studies point towards each of the Big Five traits possessing two separable, yet correlated, characteristics that are situated below the Big Five domains and above the many facet scales in the personality categorizing hierarchy. These aspects were labeled the following: Volatility and Withdrawal for Neuroticism, Enthusiasm and Assertiveness for Extraversion, Compassion and Politeness for Agreeableness, Intellect and Openness for Openness/Intellect, Industriousness and Orderliness for Conscientiousness (DeYoung, 2007).

In this study we will be examining gender, age, and race differences in personality across the Big Five traits. Correlation between these identity aspects and personality traits have been thoroughly studied with some correlations found. For instance, to summarize what was stated earlier, women have been noted to score higher than men on traits Neuroticism and Agreeableness as well as almost all facets of both traits. Very little or no gender differences have been found for traits Extraversion, Openness/Intellect, and Conscientiousness. However, there is evidence for differences in some facets that comprise those traits (Terracciano, 2001). Gender differences has been a heated contemporary issue with much debate over causes of differences. According to theory developed in the field of evolutionary psychology, the differences between men and women are attributed to evolved concerns over different reproductive issues and parental participation in raising children (Buss, 2016).

Other theories suggest gender norms are shaped by society or culture. However, research shows that gender differences in personality are more pronounced in more developed cultures with less imposed traditional gender roles (Terracciano, 2001). Personality is also heavily varied among age differences. Gender differences in personality are not noticeable at all before late childhood and adolescence is reached. There is also a negative correlation between age and trait Conscientiousness scores between ages ten through twenty and a positive correlation from age twenty to sixty (Soto, 2011). This study will examine any differences in personality traits between genders and ages. We hope to produce similar results that have been demonstrated in previous research. However, we will ignore any investigation into why differences do or do not exist among our classes.

Methods

Analysis of personality has been widespread practice since the early 20th century. In this analysis of personality, we will try to build a model to predict gender and age the best we can from large selection of personality surveys. The software used was R-Studio Version 1.0.136 (R 3.3.2). The data set has 19719 observations. The description of the variables showed that there were some missing values for gender and some entries for age were obviously entered incorrectly. The data was trimmed to remove these observations, a total of 207, since it made up an insignificant portion of the data. The remaining number of observation was 19512 of which 7583 were men and 11929 were women.

The data set was sourced from http://personality-testing.info/_rawdata/BIG5.zip. Users were also asked to confirm that their answers were accurate and could be used for research. Participants who did not were not recorded. From initially looking at the questionnaire it was evident that some items were phrased in a sense where if the participants marked a high score (4 or 5) it would mean they would have a greater expression of that trait, while for other items the exact opposite was true. We will refer to the items phrased with a positive correlation to trait expressiveness as positively phrased items and the ones with a negative correlation as negatively phrased items. For example, trait Extraversion has exactly 5 negatively phrased questions and 5 positively phrased questions in this questionnaire. "I am the life of the party" would be a positively phrased item, while "I don't like to draw attention to myself" would be a negatively phrased item. These intuitive findings are backed by research as studies have shown that trait extraversion is linked to the dopaminergic regions of the brain and people that score higher in trait extraversion are more sensitive to social interactions and other exploratory activities (Cohen). By doing simple calculations we could also show that these two questions are inverses of each other. When we added the scores of a positively phrased item with a negatively phrased item from the same trait the sum was a distribution that was centered around a score of 6. The graphs below illustrate this phenomenon:

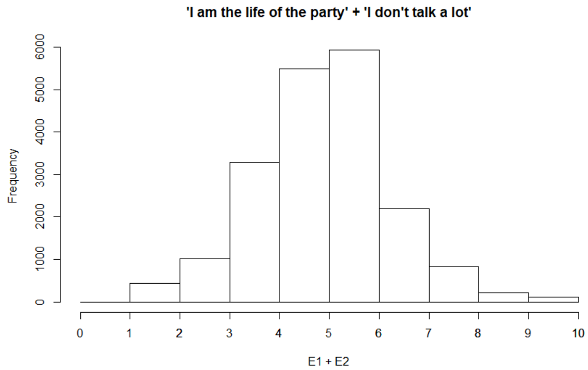

Figure 1, Reported frequency of the sum of a negatively and a positively phrased question related to trait Extraversion

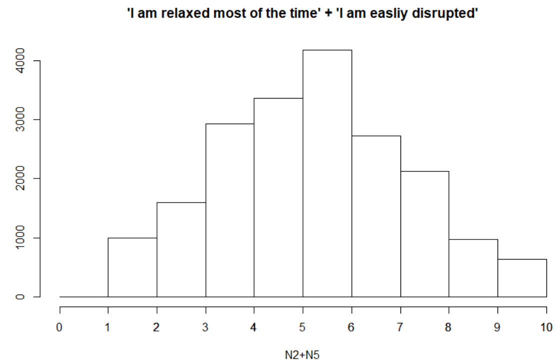

Figure 2, Reported frequency of the sum of a negatively and a positively phrased question related to trait Neuroticism

We found similar histograms to these for almost every trait. Similar histograms for trait Openness and Agreeableness were a little left skewed and were centered around 7 or 8 instead of 6. The questions for these traits were so obviously worded to have negative correlations with their respected traits that we decided to invert the scales even though the histograms weren't as obvious as with the other traits. We inverted the questions by inverting the scales of the negative questions treating 3 as the inversion point. Scores of 1 would now become 5 and scores of 2 would now become 4 and also the reverse.

Our models and plots were created from the following variables:

E	Mean score of questions relating to trait Extraversion
A	Mean score of questions relating to trait Agreeableness
N	Mean score of questions relating to trait Neuroticism
O	Mean score of questions relating to trait Openness
C	Mean score of questions relating to trait Conscientiousness
Evar	Variance in answers relating to trait Extraversion
Avar	Variance in answers relating to trait Agreeableness
Nvar	Variance in answers relating to trait Neuroticism

Ovar	Variance in answers relating to trait Openness
Cvar	Variance in answers relating to trait Conscientiousness

Table 1, Descriptions of independent variables

The first model we decided to build was a logistic regression model using the means and variances we found as independent variables. Logistic regression is a regression model where the dependent variable is categorical, and therefore is an appropriate model for this situation. We generated our logistic regression model with the glm function in R. The logistic function can be written as,

$$F(x) = \frac{1}{1 + e^{-(\beta_0 + \beta_1 x)}} \tag{1}$$

while the inverse of the logistic function (log odds) is,

$$\ln\left(\frac{F(x)}{1 - F(x)}\right) = \beta_0 + \beta_1 x \tag{2}$$

From there we can see how the odds are defined by exponentiating both sides. Generalized linear models use link functions, hence the raw variable coefficients are not easy to interpret. We transformed the log-odds produced by the glm function to odds, so we can use the predict function to make clear statements about our model.

The next model we build was done using Support Vector Machines (SVM). A SVM is a discriminative classifier that implements a separating hyperplane. It is a form of supervised learning, in other words it requires a training set. The algorithm outputs an optimal hyperplane which may be used to classify new data. The SVM model is mapped so that the data points of the separate classes are divided by a gap that is as wide as possible. For linear Support Vector Machines, set of hyperplanes is chosen so $f(x) = (w \cdot x) + b$ where w is the weight vector, x

(3)

is input vector, and b is bias. The equation for the maximum margin separating hyperplane is (Berwick),

$$f(x) = \sum_{i=1}^{m} (\alpha_i y_i x_i \cdot u) + b$$

where only the support vectors will have nonzero α's. The equation above is derived from solving the Lagrangian function with introducing α as new 'slack variables' (Berwick). New examples are then mapped into that same space and predicted to belong to a class depending on which side of the decision boundary they fall on. SVMs are extremely robust and can perform non-linear classifier using what is called the kernel trick which enables the function to operate in a higher dimension. The kernel trick essentially turns any linear model into a non-linear model by replacing its predictors by a kernel function which can be specified by the user. The kernel function is noted as (Weston):

$$K(x_i, x) = \Phi(x_i) \cdot \Phi(x) \tag{4}$$

And the new decision rule is:

$$\sum_{i=1}^{m} \alpha_i \Phi(x_i) \cdot \Phi(x) + b \tag{5}$$

We used the svm function in the e1071 package built for R to create our model. Since SVMs are extremely robust, we used them to predict age as well as gender.

We also implemented Linear Discriminant Analysis (LDA) to predict gender and also understand the relationship between our variables better. LDA is an application of Fisher's

(6)

linear discriminants to find a linear combination of variables that best separates the data into two or more classes. The linear discriminant function is,

$$f_i = \mu_i C^{-1} x_k^T - \frac{1}{2}\mu_i C^{-1} \mu_i^T + \ln(P(i))$$

Where μ_i is the vector mean and C_i is the covariance matrix of group i (Teknomo). This algorithm produces a linear classifier as well as reduces the dimensionality of the data to at most 1 less than the number of classes in the data set, hence LDA is a popular form of dimensionality reduction before further classification is applied. We used the lda function in the MASS package built for R to create our model. Since we only have two classes for gender, the later application is of no use to us.

Finally, we built a linear regression model to predict age with the same dependent variables as before. We used the lm function in R to produce our model.

Results

As we noted before, previous research has shown that men and women generally have the most differentiated scores in Neuroticism and Agreeableness at the Big Five trait level. Taking the mean scores for each trait by gender in our data we see that women and men differ the most in their mean scores of Neuroticism and Agreeableness, keeping in line with previous studies. The mean score for trait Neuroticism was 2.92 and 3.21 for women. The mean score for Agreeableness was 3.66 for men and 3.98 for women. While the next largest difference in mean scores by gender came from Openness/Intellect with a difference in scores of 0.13, less than half of the two largest differences. A two-sample t-test showed all differences in means to be statistically significant at α=0.05. While comparing means between our dependent variables might be a convenient way to compare differences, the variance is lost when we take this approach and it might be important for our model. Therefore, we calculated the variance of the scores between the 10 questions for each individual trait. The histograms of the variances were all right skewed demonstrating that the questions are consistent within traits and measuring the same thing, which is one of our assumptions about the structure of the questions. The following plots are good visualizations for the reader to better understand basic trends in personality between ages and genders:

Figure 3, mean scores of Conscientousness by age and gender (red-female, black-male)

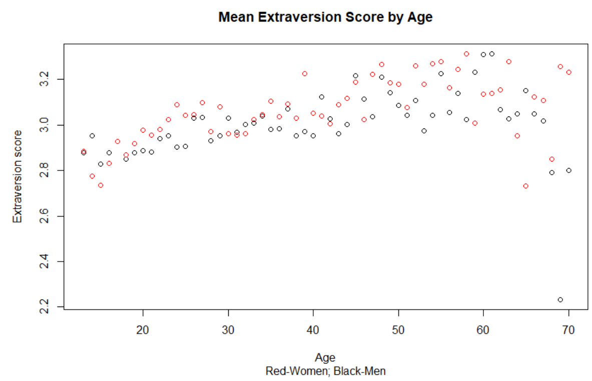

Figure 4, mean scores of Extraversion by age and gender (red-female, black-male)

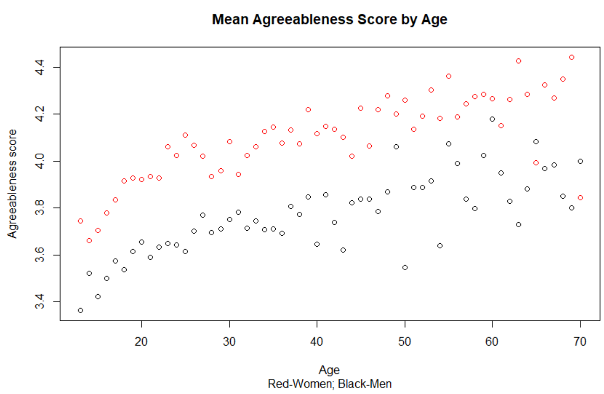

Figure 5, mean scores of Agreeableness by age and gender (red-female, black-male)

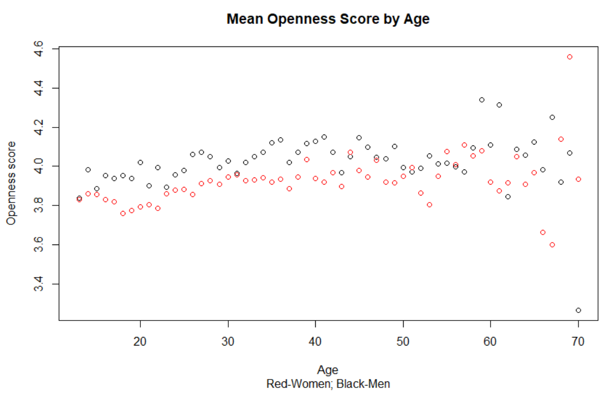

Figure 6, mean scores of Openness by age and gender (red-female, black-male)

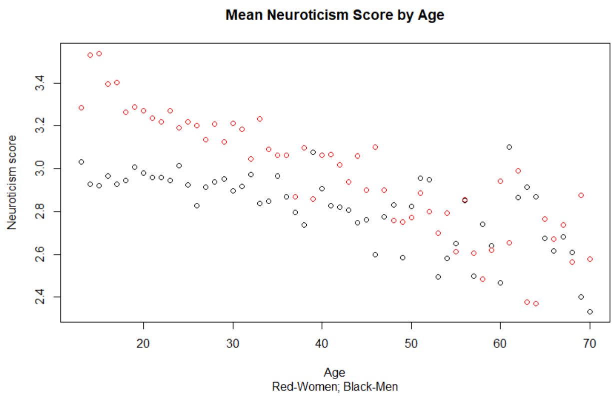

Figure 7, mean scores of Neuroticism by age and gender (red-female, black-male)

Overall, the trends in our data match those in previous research. Figure 3 gives evidence to the positive correlation between age and conscientiousness with little difference among genders. Figure 4 shows the lack of any trend between Extraversion and age or gender at the Big Five trait level. Figure 5 exhibits two clusters in trait Agreeableness which are characterized by gender. It also shows a positive correlation between age and trait Agreeableness. Figure 6 does not give any evidence to any correlation between age and trait Openness, nor to any differences in gender. Figure 7 explicitly shows a strong negative correlation between age and Neuroticism. It can also be seen that scores in Neuroticism between genders converge in mid-adulthood, a trend similar to what has been shown in previous research.

The best model we found using linear regression was,

$$P(\text{female})/P(\text{male})= -2.70+.025(E) +0.527 (N) + 0.724 (A) + 0.147 (C) - 0.419 (O) + 0.041 (Evar) - 0.047 (Nvar) - 0.074 (Avar) -0.0.24 (Cvar) - 0.047 (Ovar) \tag{7}$$

The variables that were significant at α=0.05 were Neuroticism, Agreeableness, Conscientiousness, and Openness. After we built our model and converted the log-odds to odds, we could use the model to analyze how different scores in personality affected the probability of either gender using the predict function in R. For example, we could put forth the question “How much more likely is someone with a mean score of 4 in trait neuroticism to be a woman compared to someone with a mean score of 2?” According to our model, the person with a mean score of 4 in trait Neuroticism has a 73% probability of being a woman while the person with a mean score 2 in trait Neuroticism has a 48% probability of being a woman. We generated effect plots for each of our variables (Figure 8).

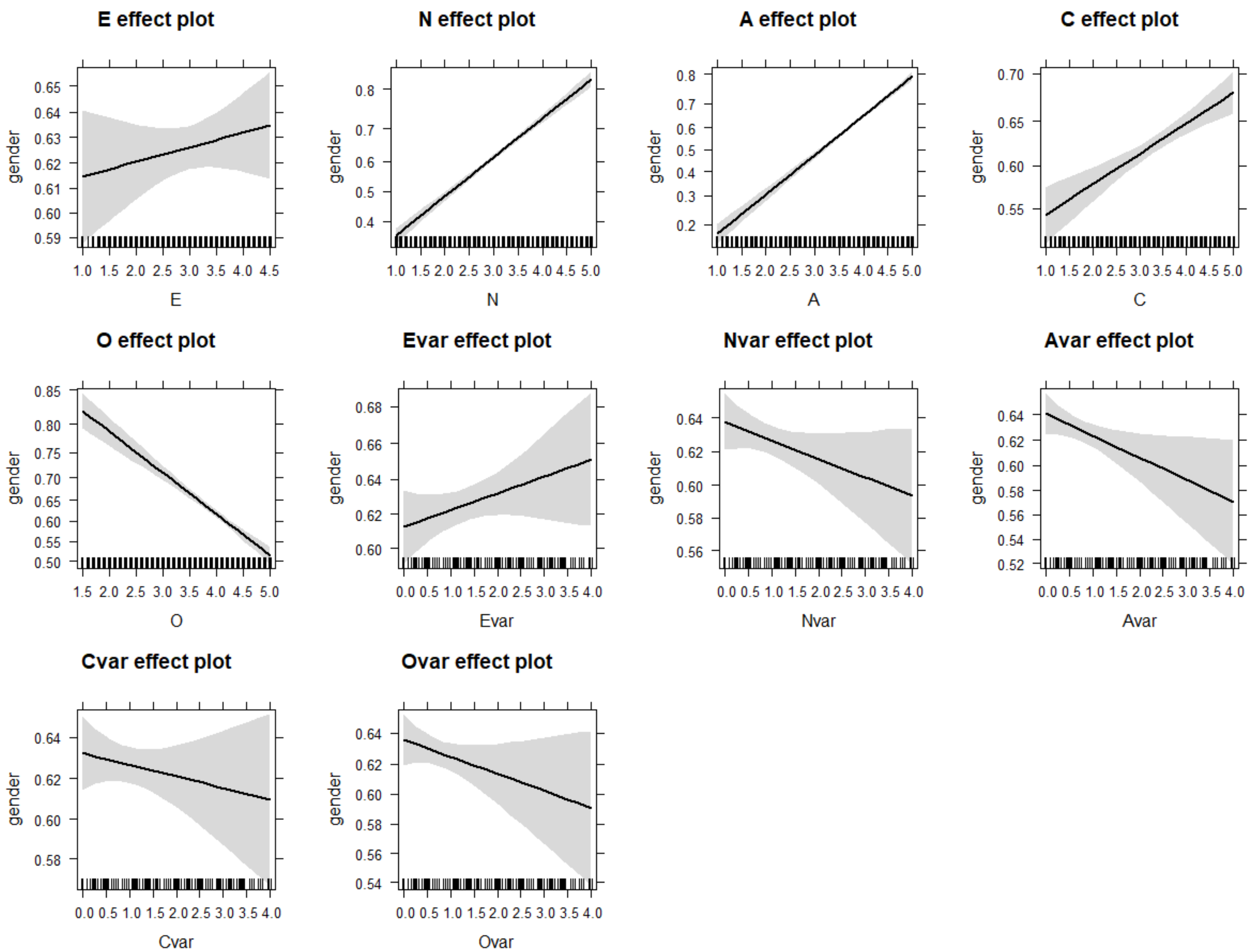

Figure 8, Effect plots produced by transforming log-odds to odds. The shaded regions indicated 95% confidence intervals

Notice how the effect plots point to Neuroticism and Agreeableness having the greatest effects in classifying gender with very small confidence intervals. The effect plots also give more precise measurements to the relationships with gender and other traits that aren't apparent enough to be deduced visually from our previous plots. We used this model to predict the genders in our test set and we managed to do so with an accuracy of 64.7%. Alarmingly our accuracy on predicting the gender of men was 31.8%. This might suggest that we need a data set that is more evenly balanced in terms of the number of men and women represented.

LDA lets us understand our data through a single linear discriminant as well as make predictions on our binary classes. The coefficients of our single linear discriminant can be seen in the table below:

Variables	Linear Discriminant
E	0.03868265
N	0.78104366
A	1.09605522
C	0.21273142
O	-0.62357010
Evar	0.05264580
Nvar	-0.05958539
Avar	-0.10402676
Cvar	-0.04545984
Ovar	-0.07891226

Table 2, coefficients of linear discriminant

The coefficients in table 2, again, show that Agreeableness and Neuroticism are the most influential traits when distinguishing between genders. The discrimination distribution plotted in figure 9 shows that no clear separation exists between men and women in temperament overall.

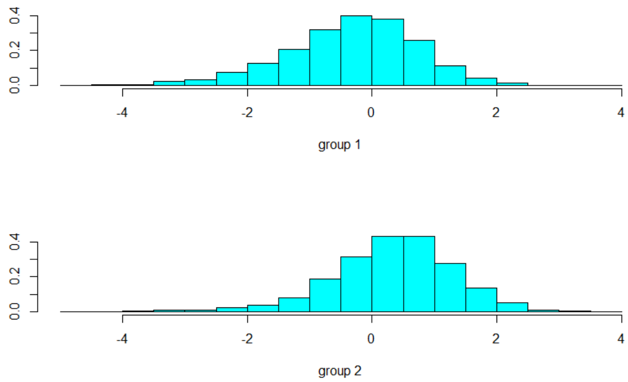

Figure 9, discrimination distribution for the two groups

LDA gave us almost identical results to logistic regression in prediction of our test set with 64.7% prediction accuracy.

Using SVM with a linear kernel, we were able to predict gender in our test set. with 63.6% accuracy. Figure 10 shows the decision boundary in our model of the two most important variables (Agreeableness and Neuroticism).

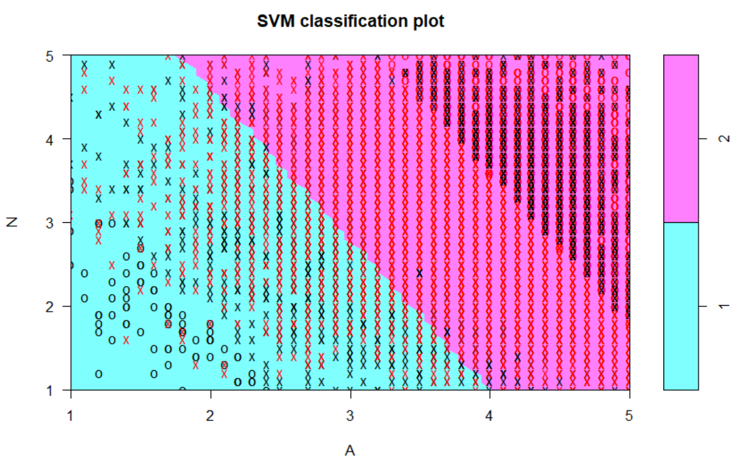

Figure 10, decision boundary using a linear kernel

Using SVM with a radial kernel, we were able to classify our test set with 64.9% accuracy. Figure 11 shows the decision boundary of the two most important variables when using a radial kernel.

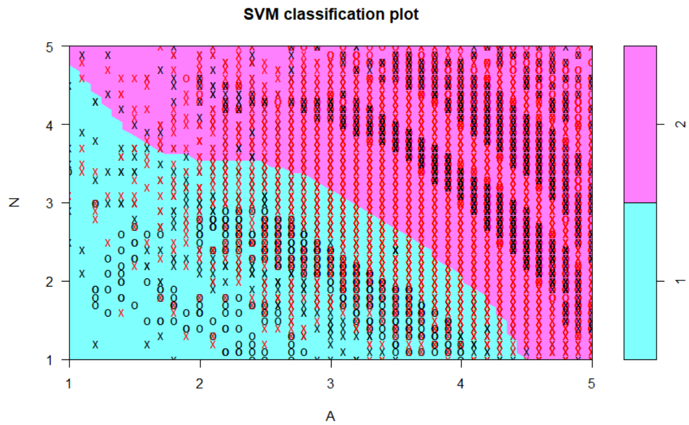

Figure 11, decision boundary using a radial kernel

The first linear regression model we created was,

$$\text{Age}= 12.0115+0.2838(\text{E}) -1.4269(\text{N}) + 1.6662(\text{A}) + 2.8541(\text{C})+ 0.9881 (\text{O}) - 0.5780(\text{Evar}) -2.0965(\text{Nvar}) -0.6552(\text{Avar}) + 0.6961(\text{Cvar})+ 0.9880(\text{Ovar}). \quad (8)$$

Again, Extraversion was not significant at α=0.05. The model's R^2 = 0.1078, F-statistic = 166, and p-value < 2.2e-16.

The Normal Q-Q residual plot shows that our assumption of normally distribute residuals do not hold (Figure 12). We tried a log transformation and our model improved. Our new model was,

$$\log(\text{age})= 1.162304 + 0.005364(\text{E}) -0.020031(\text{N}) + 0.024741(\text{A}) + 0.044657(\text{C}) + 0.014194(\text{O}) -0.009755 (\text{Evar}) -0.032734(\text{Nvar}) -0.008950(\text{Avar}) + 0.011589(\text{Cvar}) + 0.014080(\text{Ovar}). \quad (9)$$

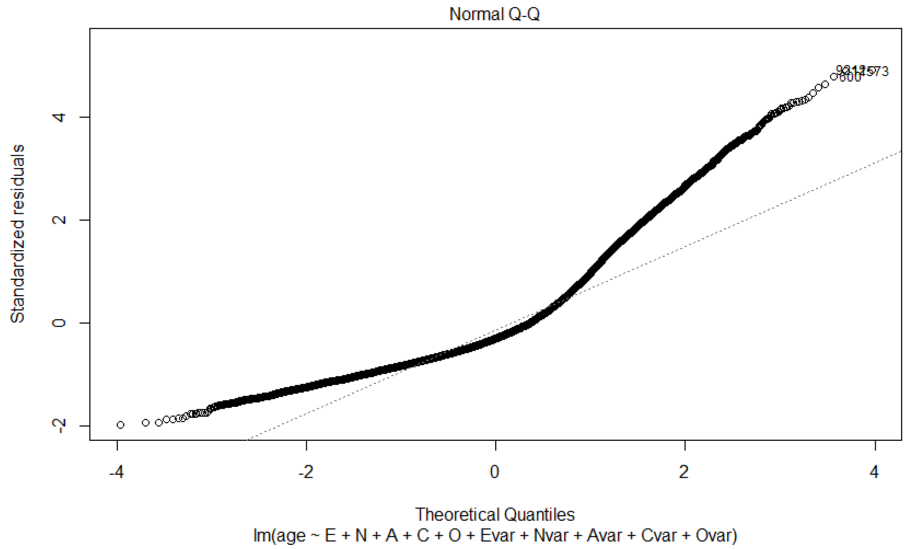

Figure 12, residual plot showing that residuals deviate from normality

All variables were significant at α=.05. Our new Normal Q-Q plot conforms much better to our assumptions about the residuals (Figure 13).

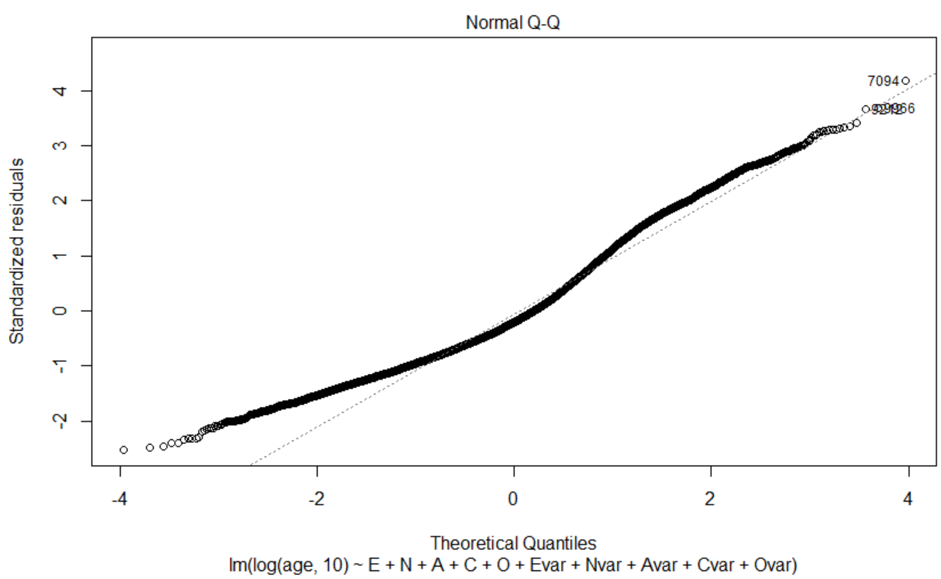

Figure 13, residual plot showing that residuals are closer to being distributed normally than before the transformation. The large gap at the negative values suggest a likely need for age dependent models as earlier suggested

After adjusting our linear regression model so that the residuals would be normal, we were able to predict age in our test set using linear regression with mean absolute difference of 7.73 years of our test set and our predictions. Using SVM to predict age gave us very similar results with a mean absolute difference of 7.69 years.

Discussion

The results in this study aligned with the previous research we cited. However, this analysis only considered gender and age in relationship to personality and only looked at traits at the Big Five level of analysis. There are infinite number of sects that divide people and a finite number that people are interested in. Some of these may include political affiliation, college major, job occupation, marital status, etc. Using analysis similar to that which we have conducted, there may be other trends within these sects. However, a different data set would be required for other interests. For example, as businesses and marketers are known to divide their customers by income classes to be more effective in their marketing strategies, they might also consider analyzing how personality affects their customer's willingness to buy, and plan their market strategies with that information. Personality traits have also been able to predict some significant features of people's lives, namely income, life-satisfaction, social behavior, and many others (Blatný). A personal interest of ours would be helping bewildered freshman choose appropriate majors using personality and I.Q. tests to match them with the field of study they would be most successful in.

References

Ashton,, Michael C., Kibeom Lee, and Lewis R. Goldberg,. "A Hierarchical Analysis of 1,710 English Personality-Descriptive Adjectives." *PsycARTICLES [EBSCO]*. N.p., Nov. 2004. Web. 31 Mar. 2017.

Berwick, R. *An Idiot's Guide to Support Vector Machines (SVMs)*. N.p.: n.p., n.d. PDF.

Blatný, Marek, Katarína Millová, Martin Jelínek, and Terezie Osecká. "Personality Predictors of Successful Development." *PLOS ONE*. Public Library of Science, n.d. Web. 12 May 2017.

Buss, David M. *Evolutionary Psychology: The New Science of the Mind*. 4th ed. London: Routledge, 2016. Print.

Cohen, Michael. "Individual Differences in Extraversion and Dopamine Genetics Predict Neural Reward Responses." *Cognitive Brain Research* 25.3 (2005): 851-61. Web. 12 May 2017.

DeYoung, Colin G., Lena C. Quilty, and Jordan B. Peterson. "Between Facets and Domains: 10 Aspects of the Big Five." *Journal of Personality and Social Psychology* 93.5 (2007): 880-96. *PsycARTICLES [EBSCO]*. Web. 31 Mar. 2017.

DeYoung, C.G., and Gray, J.R. (2009). "Personality neuroscience: explaining individual differences in affect, behavior, and cognition," in *The Cambridge Handbook of Personality Psychology*, eds P. J. Corr and G. Matthews (New York: Cambridge University Press), 323–346.

John, Oliver P., and Sanjay Srivastava. "The Big-Five Trait Taxonomy: History, Measurement, and Theoretical Perspectives." *Handbook of Personality: Theory and Research (2nd Ed.)* (1999): n. pag. Web. 31 Mar. 2017.

Soto, Christopher J., Oliver P. John, Samuel D. Gosling, and Jeff Potter. "Age Differences in Personality Traits from 10 to 65: Big Five Domains and Facets in a Large Cross-sectional Sample." *Journal of Personality and Social Psychology* 100.2 (2011): 330-48. *PsycARTICLES [EBSCO]*. Web. 31 Mar. 2017.

Teknomo, Kardi. "Derivation of Linear Discriminant Analysis Formula." *Linear Discriminant Analysis (LDA) Formula Tutorial*. N.p., 2015. Web. 12 May 2017.

Terracciano, Antonio, Robert R. Mccrae, and Paul Costa. "Gender Differences in Personality Traits across Cultures: Robust and Surprising Findings." *Journal of Personality and Social Psychology* 81.2 (2001): 322-31. *PsycARTICLES [EBSCO]*. Web. 31 Mar. 2017.

Weisberg, Yanna J., Colin G. DeYoung, and Jacob B. Hirsh. "Gender Differences in Personality across the Ten Aspects of the Big Five." *Frontiers*. Frontiers, 14 July 2011. Web. 31 Mar. 2017.

Weston, Jason. *Support Vector Machine (and Statistical Learning Theory) Tutorial*. Princeton: NEC Labs America, n.d. PDF.