Fairness of Measuring Happiness

Yi Lin

03/24/2021

1 Introduction

To analyze the constituents' happiness, we help build up a linear model that predicts the happiness score and identifies such areas fairly. The fairness of the model is critical and prediction should be without any bias, any sensitive information related to ethnicity. Due to focus on the measure of fairness, we need to build up to three models, including Race Aware Model (Prediction Model I), Race Blind Model (Prediction Model II), and Fair Model.

To measure fairness, we applied the correlation between both independent predictors and predicted outcomes as proxy. This task aims to make sure the happiness predictions are not biased towards the ethnic group variables. We also define totalGroup1 and totalGroup2 as our protected variables that yield biases to be observed in the outcome. We also applied data standardization, which coverts data in a common format, helping us process and analyze three models.

For each model, there are 4 scatterplots of the valence (ground truth valence or predictive valence) for the degree (percent_bachelorPlus) and the income (households) conditioned on either totalGroup1 (Group1) or totalGroup2 (Group2).

2 Ground Truth Model

The Ground Truth Model analyzes both correlations and scatterplot of the ground truth valence for the degree (percent_bachelorPlus) or the income (households) with conditioned on either totalGroup1 or totalGroup2.

Based on the correlation (Figure 1), it shows that the correlations between target variable (mean_valence) and totalGroup1, totalGroup2, degree, income are -0.361, 0.335, 0.457, 0.311 respectively. Notice that totalGroup1 has the only negative relationship to the target variable.

Figure 1: Correlation of Ground Truth Model

	meanvalence	totalGroup1	totalGroup2	percent_bachelorPlus	households_meanIncome
meanvalence	1.000000	-0.361185	0.335692	0.457267	0.311013
totalGroup1	-0.361185	1.000000	-0.503950	-0.729305	-0.507290
totalGroup2	0.335692	-0.503950	1.000000	0.726215	0.585684
percent_bachelorPlus	0.457267	-0.729305	0.726215	1.000000	0.752844
households_meanIncome	0.311013	-0.507290	0.585684	0.752844	1.000000

Based on the scatterplots, we cannot state that there is a bias to one ethnic group since the darker spots are evenly distributed around the threshold above or below.

Figure 2: Ground Truth Model: Valence & Degree (Group1)

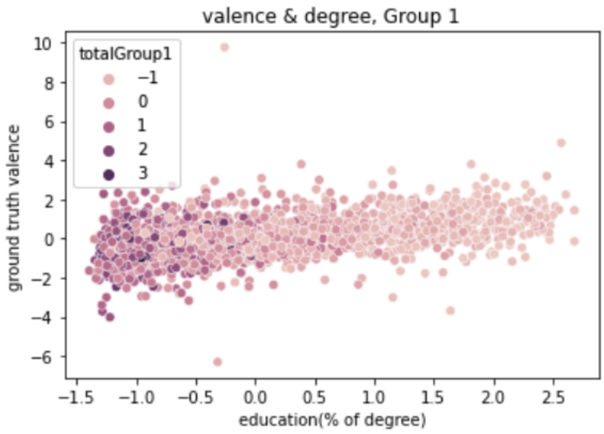


Figure 3: Ground Truth Model: Valence & Degree (Group2)

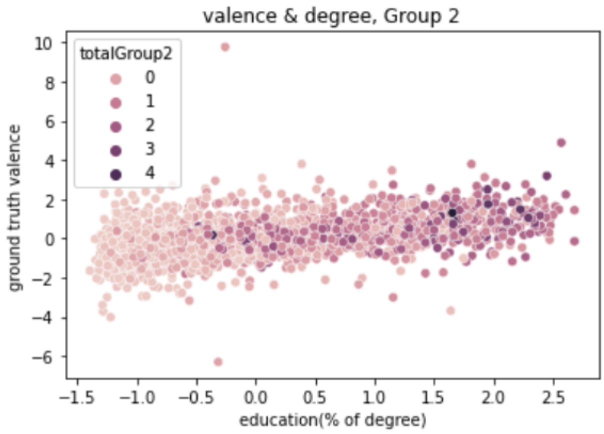


Figure 4: Ground Truth Model: Valence & Income (Group1)

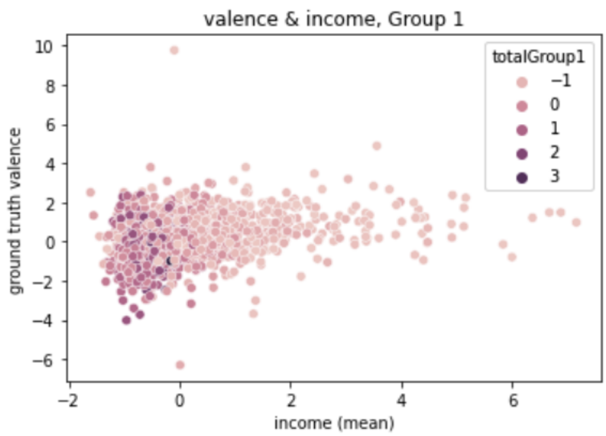
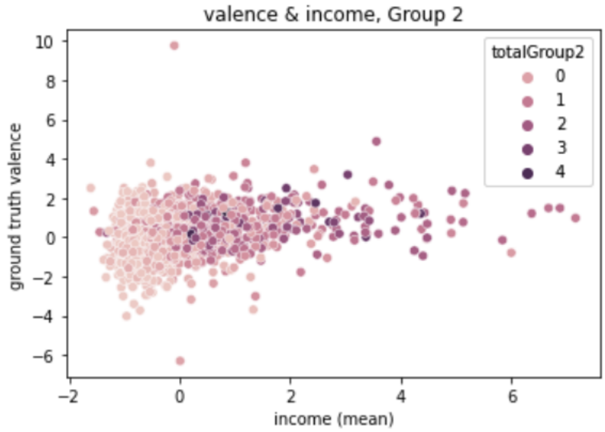


Figure 5: Ground Truth Model: Valence & Income (Group2)



Two predictive models (Prediction Model I and Prediction Model II) are considered fair if the correlations and scatterplot of predicted results (valence_predict) V.S. independent variables conditioned on ethnic group variables (totalGroup1 and totalGroup2) which are similar to that of the Ground Truth Model.

3 Ethnic Group Aware Model (Prediction Model I)

The Prediction Model I uses all predictors (totalGroup1, totalGroup2, degree and income) to build a linear regression model. I applied the ordinary least square (OLS) via statsmodels package to predict the entire dataset's valence. To analyze the predicted valence's fairness, I compared both correlations and scatterplots of the Prediction Model I with the Ground Truth Model. To state whether there is bias to either ethnic Group 1 or Group 2, we standardized the predicted valence threshold set to -0.2520924.

The correlation (Figure 6) shows that the correlation between predicted variable (valence_predict) and totalGroup1, totalGroup2, degree, income are -0.782, 0.727, 0.991, 0.674, respectively. Notice that both Group1 and Group2 have a highly negative and positive relationship with the predicted variable. Also, degree has the strongest positive relationship with the predicted variable (0.991), which implies that as the percentage of degree increases, the predicted valence significantly increases. Assume alpha is 0.05 since the P-values of income (0.014) and degree (0.000) are smaller than 0.05, demonstrating that both income and degree are statistically significant.

Figure 6: Correlation of Prediction Model I

	valence_predict	totalGroup1	totalGroup2	percent_bachelorPlus	households_meanIncome
valence_predict	1.000000	-0.782482	0.727253	0.990636	0.673788
totalGroup1	-0.782482	1.000000	-0.503950	-0.729305	-0.507290
totalGroup2	0.727253	-0.503950	1.000000	0.726215	0.585684
percent_bachelorPlus	0.990636	-0.729305	0.726215	1.000000	0.752844
households_meanIncome	0.673788	-0.507290	0.585684	0.752844	1.000000

Based on the scatterplots with threshold, those scatterplots clearly show the bias to one of the two groups. Compared with Figure 7 and Figure 8, we notice that most ethnic group 1 is clustering below the threshold while the majority of ethnic group 2 is clustering above the threshold in degree. This demonstrates that most ethnic group 1 has a relatively smaller percentage of a degree corresponding to a lower predicted valence score. In comparison, that of ethnic groups has a somewhat larger percentage of degree corresponding to higher predicted valence scores. Compared with Figure 9 and Figure 10, we also found a similar situation in income. The majority of ethnic group 1 is clustering below the threshold, while most ethnic group 2 is clustering above the threshold. This indicates that most ethnic group 1 has relatively low income corresponding to lower predicted valence score. In comparison, that of ethnic groups has relatively high income corresponding to higher predicted valence score. Compare with Prediction Model I and Ground Truth Model, these observations imply that the social class of ethnic group2 is relatively higher than that of ethnic group1 according to both income and degree. Also, the observations imply that ethnic group 2 has relatively high valence in terms of income and degree.

Figure 7: Model I: Predicted Valence & Degree (Group1)

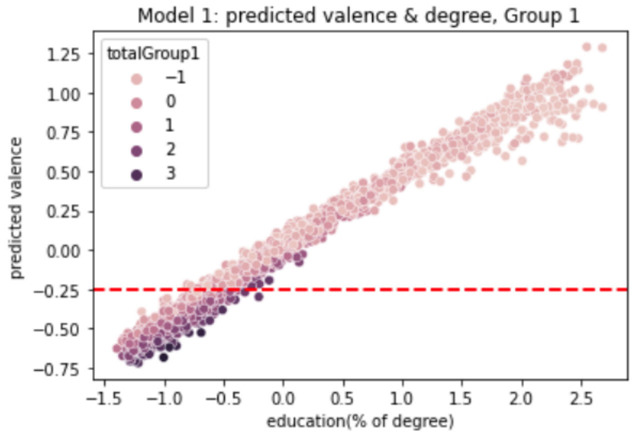


Figure 8: Model I: Predicted Valence & Degree (Group2)

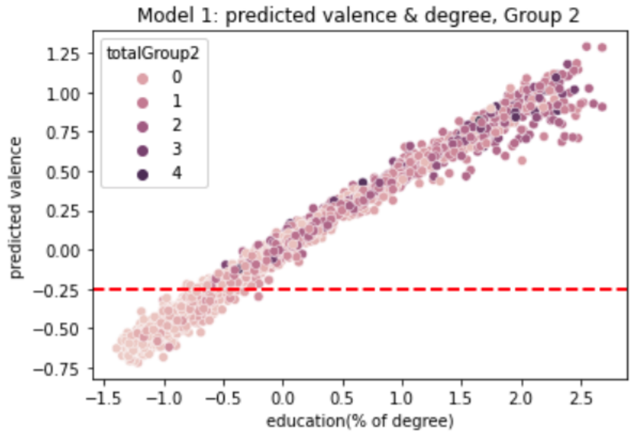
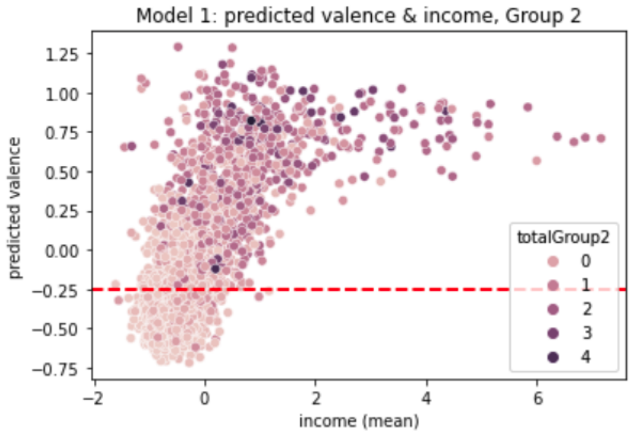


Figure 9: Model I: Predicted Valence & Income (Group1)

Figure 10: Model I: Predicted Valence & Income (Group2)



4 Ethnic Group Blind Model (Prediction Model II)

The Prediction Model II also uses predictors without using totalGroup 1 and totalGroup 2 to build a linear regression model. I also applied the same method as previous Prediction Model II to predict the entire dataset's valence.

Based on the correlation (Figure 11), it shows that the correlation between predicted variable (valence_predict) and totalGroup1, totalGroup2, degree, income are -0.732, 0.715, 0.994, 0.676 respectively. Similar to Prediction Model I, both Group1 and Group2 have a highly negative and positive correlation with predicted variables. The degree has the strongest positive relationship with the predicted variable (0.994), which implies that as the percentage of degree increases, the predicted valence significantly increases. Assume alpha is 0.05, both income and degree has statistically significant because the P-value of income (0.009) and degree (0.000) are smaller than 0.05.

Figure 11: Correlation of Prediction Model II

	valence_predict	totalGroup1	totalGroup2	percent_bachelorPlus	households_meanIncome
valence_predict	1.000000	-0.731863	0.715330	0.993957	0.676047
totalGroup1	-0.731863	1.000000	-0.503950	-0.729305	-0.507290
totalGroup2	0.715330	-0.503950	1.000000	0.726215	0.585684
percent_bachelorPlus	0.993957	-0.729305	0.726215	1.000000	0.752844
households_meanIncome	0.676047	-0.507290	0.585684	0.752844	1.000000

Based on the scatterplot, those scatterplots also show that hiding these protected variables does not change the bias to one of the two groups. Compared with Figure 12 and Figure 13, we have a similar image that most ethnic group 1 has a larger percentage of degree as the clustered darker spots are lying the scale on the larger score of education. And the majority of ethnic group 2 tend to have a smaller percentage of degree as the clustered darker spots lying the scale on the low score of education. Compared with Figure 14 and Figure 15, the majority of ethnic group 2 tends to have a higher income as the clustered darker spots generally greater than -0.25, while ethnic group 1 tends to have lower income as the clustered darker spots generally lower than -0.25. Compare with Prediction Model II and Ground Truth Model, even we don't have protected variables (totalGroup1 and totalGroup2), the hiding protected variables do not change the bias identified in the previous model (Prediction Model I) change here. These observations show similar results from the Prediction Model I as well.

Figure 12: Model II: Predicted Valence & Degree (Group1)

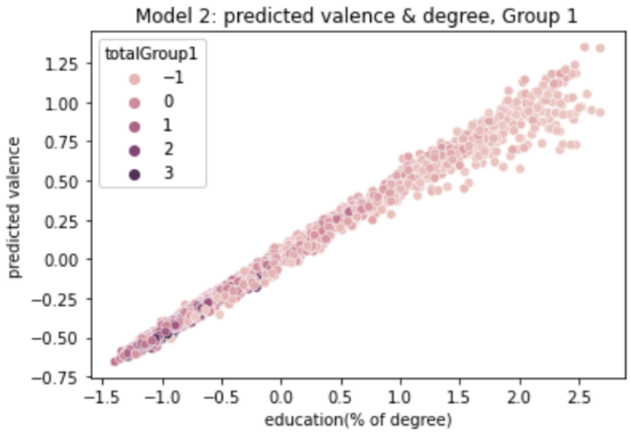


Figure 13: Model II: Predicted Valence & Degree (Group2)

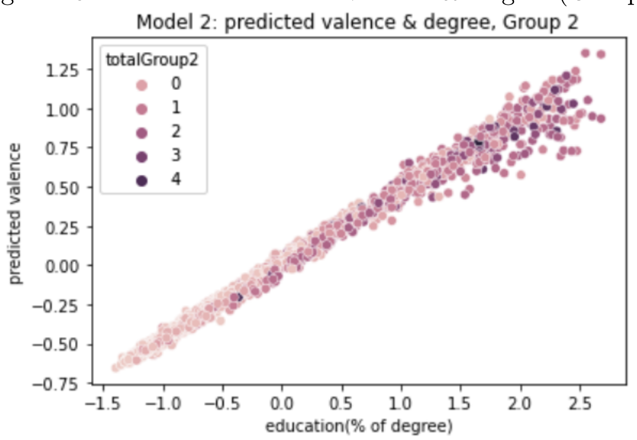


Figure 14: Model II: Predicted Valence & Income (Group1)

Figure 15: Model II: Predicted Valence & Income (Group2)

5 Fair Model

The Fair Model, the unpredictive model, uses a helper function to debias the outcome variable. I also applied the same method as Ground Truth Model. I applied the *gen_latent_fast* to generate the unbiased valence as my target variable.

The correlation (Figure 16) clearly shows that the Fair model with helper function does debias the outcome variable. The correlation between target variable (valence_unbiased) and totalGroup1, totalGroup2, degree, income are -0.000, 0.000, 0.131, 0.065 respectively. There are weak relationships between unbiased valence and the predictors. Notice that there are no correlations between unbiased valence and protected variables (totalGroup1 and totalGroup2) due to correlations equal to 0. And unbiased valence and degree has relative weak relationship (corr= 0.131), unbiased valence and degree has weak relationship (corr= 0.065).

Figure 16: Correlation of Fair Model

	valence_unbiased_fast	totalGroup1	totalGroup2	percent_bachelorPlus	households_meanIncome
valence_unbiased_fast	1.000000	-0.000000	0.000000	0.131062	0.065334
totalGroup1	-0.000000	1.000000	-0.503950	-0.729305	-0.507290
totalGroup2	0.000000	-0.503950	1.000000	0.726215	0.585684
percent_bachelorPlus	0.131062	-0.729305	0.726215	1.000000	0.752844
households_meanIncome	0.065334	-0.507290	0.585684	0.752844	1.000000

Compared to the Ground Truth Model, those scatterplots of the Fair Model do slightly change to debias the outcome variable (valence) since the Fair Model's linear regressions are generally flatter than those in the Ground Truth Model. And we cannot state that there is a bias to one ethnic group since the darker spots are evenly distributed around the threshold above or below.

Figure 17: Fair Model: Valence & Degree (Group1)

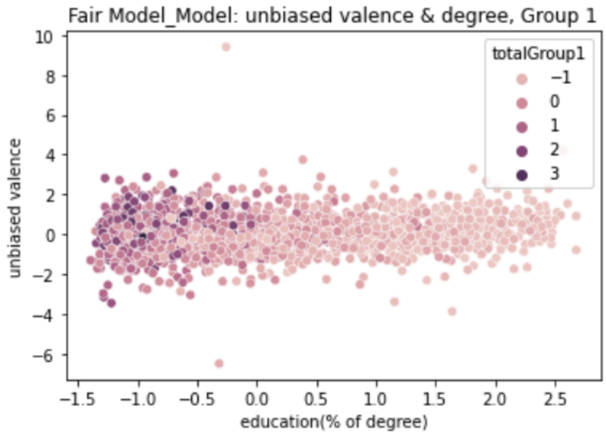


Figure 18: Fair Model: Valence & Degree (Group2)

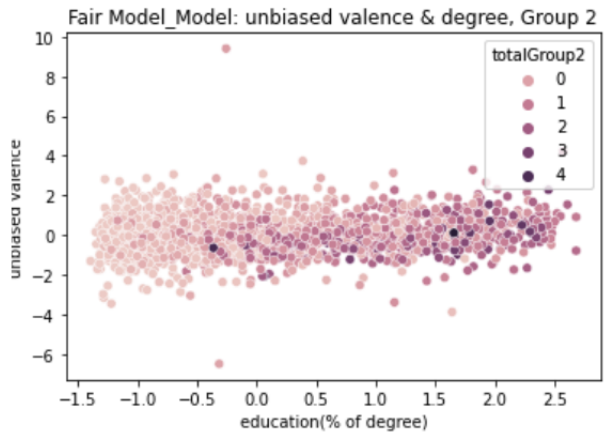


Figure 19: Fair Model: Valence & Income (Group1)

Figure 20: Fair Model: Valence & Income (Group2)