The Personality Inventory for DSM-5 (shortPID5)

Ngoc Diem Le
ngocdiem.le@studenti.unipd.it

Sebastian Rojas Ardila
josesebastian.rojasardila@studenti.unipd.it

1. Introduction

Personality Disorders have become an important condition in mainstream psychiatry across the world, which affects around 6% of the world population and the differences between countries show no consistent variation [2]. The Personality Inventory for DSM-5 (PID-5) was developed by the DSM-5 Personality and Personality Disorders Work Group to assess maladaptive traits that describe variation in the presentation of patients with personality disorder diagnoses.

In this experiment, a group of people took the questionnaire contains 25 questions in order to evaluate the individual's personality dysfunction. In general, people with high scores are more vulnerable to anxiety and depression while individuals with low scores tend to be calm and well adjusted. However, the aim of the test is not about finding the one with mental disorder but instead, this group of people was asked to answer twice the questionnaire, firstly with honest attitude and secondly with dishonest attitude when pretending to have mental disorder. The objective here is to build a model to detect a fake response. Faking to a psychological test is often observed whenever an examinee may gain an advantage from it [1]. This is the reason why faking detection plays a viral part in terms of distinguishing a faker and more than that is the questions people tend to fake.

With the advantages of technology and specially machine learning, the task seems like not as challenge as before. This project will approach several machine learning classification and regression algorithms in order to discriminate dishonest responses and reconstruct these answers to honest responses based on the dataset collected by asking the group of people above. Each method has its own advantages and disadvantages, in this dataset, Random Forest classifier (MLP), K-Means and Multi-layer perceptron give the best result with 96% of accuracy among other basic techniques like Logistic Regression, Decision Tree. Meanwhile, for the multi output regressor the task is a little bit more challenging because in this project we use different approaches to do the reconstruction of an honest response based on a dishonest response including libraries such as scikit learn, XGBOOST and a neural network based on Tensorflow, and the different methods have a different accuracy and have a good improvement comparing to the baseline provided by the trivial solution, but still it's hard to find an accuracy around 80% because there is no enough data to do a good generalization.

2. Dataset

The dataset was collected based on the responses of 519 people in 25 questions to evaluate personality and related disorders based on the dimensional trait model (DSM-5). Each question is rated on a 4-point scale (i.e., 0 = very false or often false; 1 = sometimes or somewhat false; 2 = sometimes or somewhat true; 3 = very true or often true). The overall measure has a range of scores from 0 to 75, with higher scores indicating greater overall personality dysfunction. Moreover, 25 questions are grouped into 5 trait domains with 5 questions per domain (Negative Affect, Detachment, Antagonism, Disinhibition, and Psychoticism). Each trait domain ranges in score from 0 to 15, with higher scores indicating greater dysfunction in the specific personality trait domain.

The main dataset contains 2 conditions - Honest and Dishonest because all 519 subjects were asked to answer twice the questionnaire above following different instructions: the first time by answering honestly, while the second time by pretending to have a mental disorder, which means the answers were fake responses.

3. Methodology

3.1. Exploratory Data Analysis

In this part, we generated some plots and tables in order to have a more detailed view of the data.

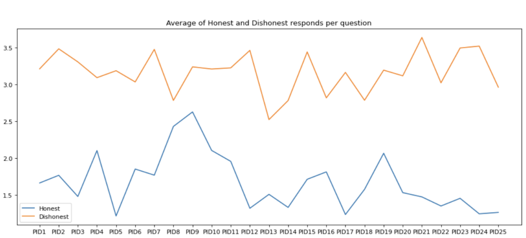

Figure 1. Average responses per question

First of all, the graph 1 shows the average of honest and dishonest responses for each question. Overall, the average of dishonest responses were significantly higher than the honest responses. However, there are some small gaps between both answers for questions PID8 and PID9.

Domain	All		Negative Affect		Detachment		Antagonism		Disinhibition		Psychoticism	
	Mean	STD	Mean	STD	Mean	STD	Mean	STD	Mean	STD	Mean	STD
Honest	1.68	0.31	2.17	0.5	1.67	0.49	1.49	0.49	1.6	0.45	1.45	0.47
Dishonest	3.17	0.53	3.18	0.67	2.8	0.78	3.1	0.83	3.24	0.72	**3.52**	**0.66**

Figure 2. Statistics for each domain

The table 2 shows mean and standard deviation of the data and for each domain based on honest and dishonest responses. As we can see, dishonest responses have considerable higher values compare to honest responses. In particular, Psychoticism had the highest mean in terms of dishonest condition, which shows that people had a tendency to make serious in Psychoticism questions. However, Detachment domain showed the lowest average for dishonest responses.

Domain	Negative Affect	Detachment	Antagonism	Disinhibition	Psychoticism
Unchanged	0.19%	0.19%	**1.16%**	0.39%	0.77%
Increase	**97.88%**	96.92%	94.99%	96.15%	97.11%

Figure 3. Percentage of changes for each domain

As we can witness on the table 3 that among five domains, Antagonism had the highest percentage of people who did not change their answers and Negative Affect had the highest percentage of people who increased their scores for all five questions when changes from honest to dishonest responses, followed by Psychoticism with a slightly lower value.

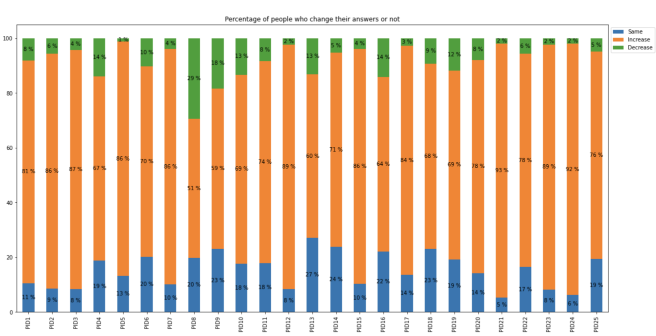

Figure 4. Percentage of changes for each question

The graph 4 illustrates the percentage of changes for each question. Most of the cases people chose to increase their scores when they changed from honest to dishonest responses. But there was still a significant high proportion of people who decreased their score, it was the question PID8 which is ”I worry about almost everything” and it belongs to the Negative Affect domain.

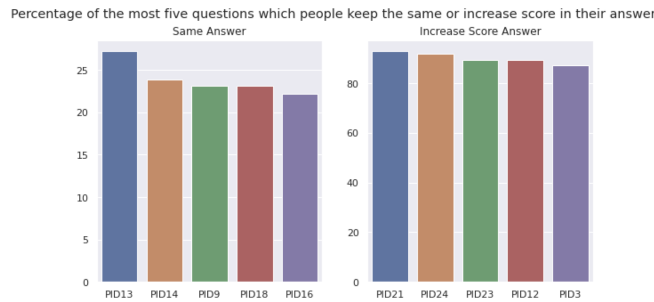

Figure 5. Percentage of the most five questions

The figure 5 shows the percentage of the most five questions in which people kept the same or increased score in their answers. On the same answer graph, the most five questions people had a tendency to keep the same answer are PID13, PID14, PID9, PID18 and PID16. Except the PID9, 4 remaining questions belong to the Detachment domain which somehow can show the less important of Detachment domain with the mental disorder trait. The graph on the right shows the most five questions people have a tendency to increase their score, which are PID21, PID24, PID23, PID12 and PID3. Except the PID3, 4 remaining questions belong to the Psychoticism domain which can be easily understood that when they pretend to have mental disorder, people have a tendency to increase their score in the questions belong to the Psychoticism domain.

Figure 6. Frequency

The frequencies in figure 6 once again show that Psychoticism had the most contrast between honest and dishonest responses. Moreover, there is no any answer with ”Very True or Often True” in honest responses but instead high proportion of ”Sometimes or Somewhat True” and ”Very True or Often True” in dishonest responses.

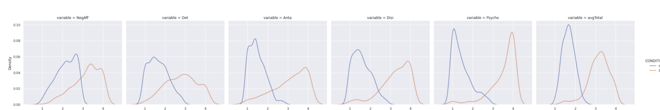

Figure 7. Distribution

Figure 7 shows the distribution plots of the honest and dishonest responses, a significant different belongs to the Psychoticism domain and Antagonism domain. Moreover, only the Psychoticism can effectively classify between honest and dishonest answers, while there is no significant different in the Negative Affect which can be explained that if alone Negative Affect, it's not reliable to distinguish Honest and Dishonest response.

3.2. Classification

For the classification task, some algorithms were used to recognize fake responses. The table 1 shows the accuracy of 6 methods our group used. The highest accuracy we achieved was 96%, belonged to Multi-layer perceptron classifier, K-Means and Random Forest.

Decision Tree	0.94
MLP	**0.96**
Bayesian Rule Classifier	0.92
Logistic Regression	0.94
K-Means	**0.96**
Random Forest	**0.96**

Table 1. Accuracy

Next, we can discuss in more detail about the decision rules, which means to identify the items that maximise the accuracy.

```
training bayesian_rule_list...
Trained RuleListClassifier for detecting Dishonest
=================================================
IF PID21 <= 2.5 and PID25 <= 1.5 THEN probability of Dishonest: 3.9% (2.0%-6.4%)
ELSE IF PID12 > 3.5 THEN probability of Dishonest: 99.6% (98.6%-100.0%)
ELSE IF PID2 > 3.5 THEN probability of Dishonest: 98.2% (93.6%-100.0%)
ELSE IF PID17 > 2.5 and PID5 > 2.5 THEN probability of Dishonest: 94.7% (85.8%-99.3%)
ELSE IF PID20 > 3.5 THEN probability of Dishonest: 90.5% (75.1%-98.8%)
ELSE IF PID18 > 3.5 THEN probability of Dishonest: 80.0% (51.8%-97.2%)
ELSE IF PID24 <= 3.5 and PID24 > 2.5 THEN probability of Dishonest: 33.3% (16.4%-52.9%)
ELSE IF PID24 > 3.5 THEN probability of Dishonest: 87.5% (59.0%-99.6%)
ELSE probability of Dishonest: 4.5% (1.5%-9.0%)
================================================
```

Figure 8. Bayesian Rule

Figure 8 shows the output from Bayesian Rule Classifier. In general, this algorithm builds probabilistic rule lists based on each question. The output has a logical structure that is a sequence of IF - THEN rules. From the output, we have a chance at 99.6% to recognize a fake response if looking only at question PID12 which has the score higher than 3.5, followed by question PID2 with 98.2% accuracy with the same threshold.

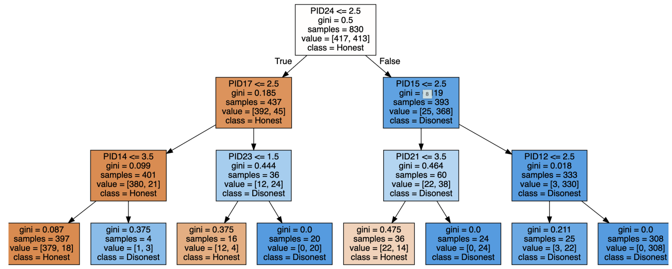

Figure 9. Decision Tree

The graph 9 illustrates the output of decision tree algorithms. At the top is the root node which evaluates the variable that best splits the data, which is the question PID24. Next, in order to find fake responses, question PID15 is used at depth 1. Travelling down the tree branches, question PID21 and question PID12 are used consequently at depth 2.

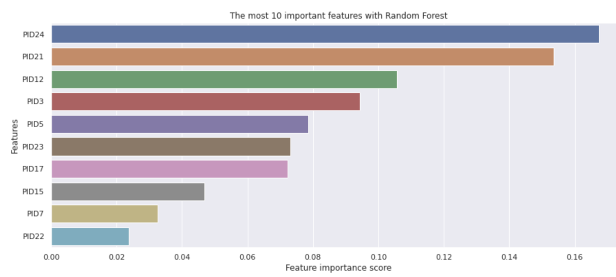

Figure 10. Important features extracted from Random Forest

Finally, in this classification task, we extracted the most 10 important features with Random Forest algorithm. Based on the information in the graph 10, question PID24 which is "Things around me often feel unreal, or more real than usual" belongs to Psychoticism was the most important question, followed by question PID21 "I often have thoughts that make sense to me but that other people say are strange", question PID23 "I often "zone out" and then suddenly come to and realize that a lot of time has passed" and question PID12 "I have seen things that weren't really there". All of these questions are in the Psychoticism domain, which strongly affirm that this domain plays an important role to distinguish between honest and dishonest responses.

In order to optimize the model, we tried to use Principal component analysis to reduce the number of features. In this case, 5 components were chosen to run PCA algorithm. By using the data contains the information of 5 components, we run again Random forest algorithm and received 97% of accuracy, which slightly increased compared to the original data.

3.3. Multi output Regressor

Basically, with the multi output regressor we are doing the forecasting of the honest response of a dishonest person. The input of this kind of regressors we use the 25 responses that the dishonest person gives and at the end we are going to obtain the 25 responses based on the behavior of the honest person. For this task we are going to compare different methods such a trivial solution that is the base line and the benchmark to the other methods like a multioutput regressor based on the library of XGBOOST that use a regression with a cost function of square error, the next one is a solution is a multioutput regressor of scikit learn based on support vector classification and the last one is a neural network with long short-term memory neurons.

The trivial solution use the average of the difference between the honest and the dishonest responses and take that difference to subtract that value to the dishonest responses and giving to us the honest responses. For this method we have an accuracy of 21.01% ,a mean error of 1.14 and a standard deviation of 0.74 comparing the predictions with the real honest responses.

For the solution with the XGBoost the library provide to you the kind of regressor that you want to use, in this case is a multi output and the type of the cost function for this model we used Mean squared error. For this model the accuracy have a good improvement giving to us an accuracy of 67.31% a mean error of 0.43 and standard deviation of 0.36

With the model using Support vector classification we obtain an accuracy of 57% and a mean error of 0.53.

And finally for the last model in which we decide to use a neural network on the framework of tensorflow and keras we decide to use the LSTM neurons to try to have a relation between the memory of the responses and how the people try to respond to this kind of surveys. In this case we needed to change the activation function in order to obtain a number between 1 and 4. After 20 epochs with a batch size of 128 we obtain that the accuracy was 49.6% the and the error was 0.61.

At the end in conclusion with the comparison of the models the one that have better results is the model using XGBoost that have an accuracy of 67.31% and the smallest number of mean error.

4. Experiments

In this part, we randomly created the answers of 25 questions and aim to reconstruct the honest response if they are fake responses. In order to do that, firstly, we used Random forest to classify whether it is a honest or dishonest responses. If it is a honest response, nothing happens. On the other hand, if it is a dishonest response, XGBOOST model is used to reconstruct the honest one.

In order to demonstrate how this work we create a random vector of responses in which domains such as pyschoticism and antagonism have a severe qualification. In this order of ideas the random input vector have the qualification that we can see in the Figure 11.

```
Total raw score 86
The Negative Affection is Moderate
The detachment is Severe
The antagonism is Moderate
The disihibition is Severe
The pyschoticism is Severe
```

Figure 11. Qualification of the domains

Automatically the software will classify this responses between the honest and dishonest, in this case as we can see in the Figure 12, the response is dishonest so the software will use the XGBOOST model of multioutput regressor to give us the honest response of this person

```
Its a dishonest response
Your Honest response should be:
[array([2., 2., 2., 2., 1., 2., 2., 3., 3., 2., 2., 2., 2., 1., 2., 2., 2.,
       2., 2., 2., 2., 2., 1., 2., 1.], dtype=float32)]
```

Figure 12. Final output

5. Conclusion

In conclusion, among five domains, people have a tendency to fake their responses in psychoticism, antagonism questions and the questions PID24 as well as PID21 play as a key role to classify fake responses when the people do this kind of test, specially because we can analyze each domain and see that the gap between a honest an a dishonest in this domains make a great difference and it's easy to classify this in any of the models that we proposed such as the logistic regression, the k-means and the MLP.

On the other hand the multioutput regressor doesn't have the best accuracy in all the models because it is not easy to obtain the correct answers for an honest response because the model need more samples to generalize better the result of the regression and on the other hand maybe we can explore another architectures on the neural network model in order to improve the accuracy, but anyway the models that we proposed accomplish the goal and have a descent response in terms of the mean error.

Finally this kind of test could be automatized in order to provide more information to the person that analyze this kind of data for a single person comparing and detecting more than a personality disorder and do a best diagnosis.

References

[1] Purpura Alberto, Giorgianni Dora, Orrù Graziella, Melis Giulia, and Sartori Giuseppe. Identifying faking to single item responses in personality tests: a new tf-idf based method.

[2] Peter Tyrer, Roger Mulder, Mike Crawford, Giles Newton-Howes, Erik Simonsen, David Ndetei, Nestor Koldobsky, Andrea Fossati, Joseph Mbatia, and Barbara Barrett. Personality disorder: a new global perspective. *World Psychiatry*, 9(1):56, 2010.