

## 1 Machine Learning

Different classifiers and Regressors for pre-processing or dimensionality-reduction operation.

### MODEL ——— ACCURACY

Random Forest—— 86.78%

K-Nearest Neighbors—— 84.47%

Linear Discriminant Analysis—— 83.75%

Logistic Regression —— 83.26%

Support Vector Machine —— 82.47%

Decision Tree —— 78.12%

Quadratic Discriminant Analysis —— 69.86%

Bayes —— 69.34%

## 2 Variables

I used following variables: D2011, D2015, D2016, D2021, D2022, D2023, D2026, D2027, D2028, D2029, D2030, age(codebook) There was a disrupt the distribution of data. In order to make data have a uniform or normal distribution, transformers methods were applied.(Gaussian)

Figure 1: Variables Before Gaussian Form

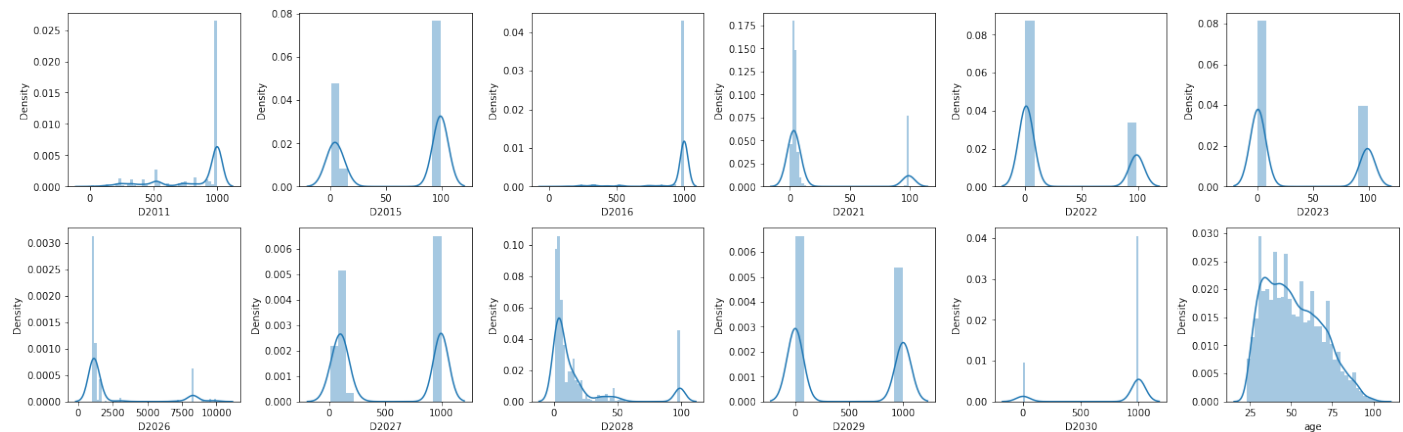


Figure 2: Variables After Gaussian Form

