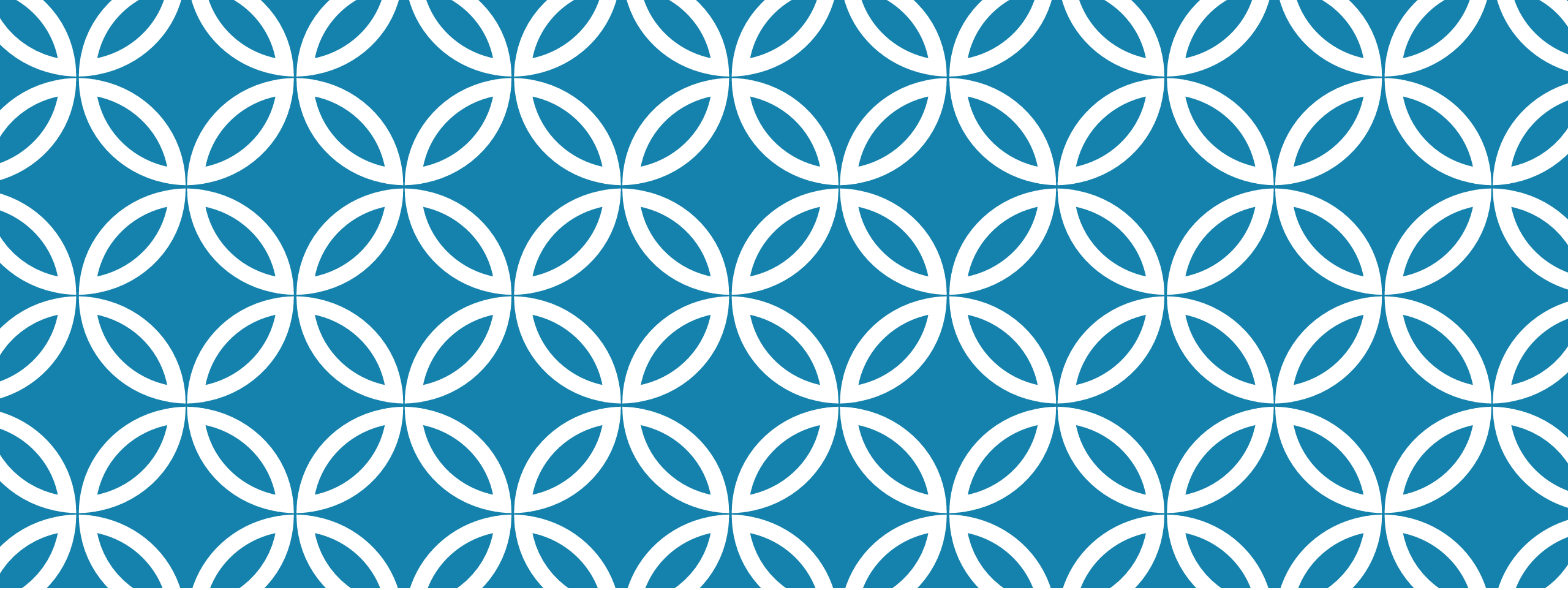


CARDIOVASCULAR DISEASE PREDICTION

Presented by: Waad Alanazi

ABSTRACT :

Cardiovascular diseases (CVDs) are disorders of heart and blood vessels. This is the leading cause of deaths worldwide. Early detection and diagnosis can help the patients. Machine learning can be used to create the predictive model using cardiovascular diseases risk factors like cholesterol level, glucose level and blood pressure

Goal - To predict the presence or absence of cardiovascular disease (CVD) using the patient examination results.

1- DATA UNDERSTANDING

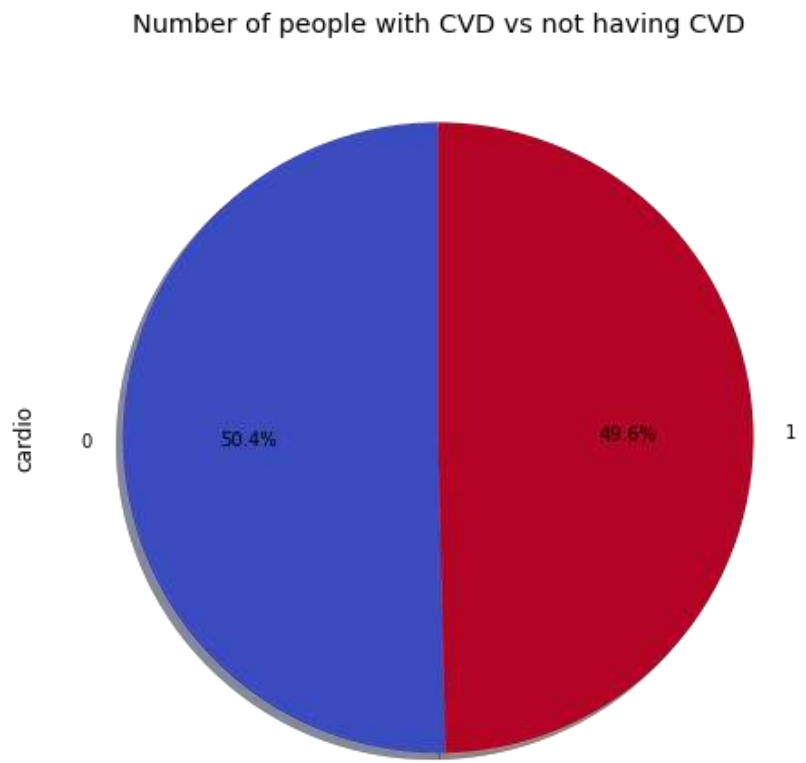
The dataset consists of 70,000 records of patients data in 13 features .

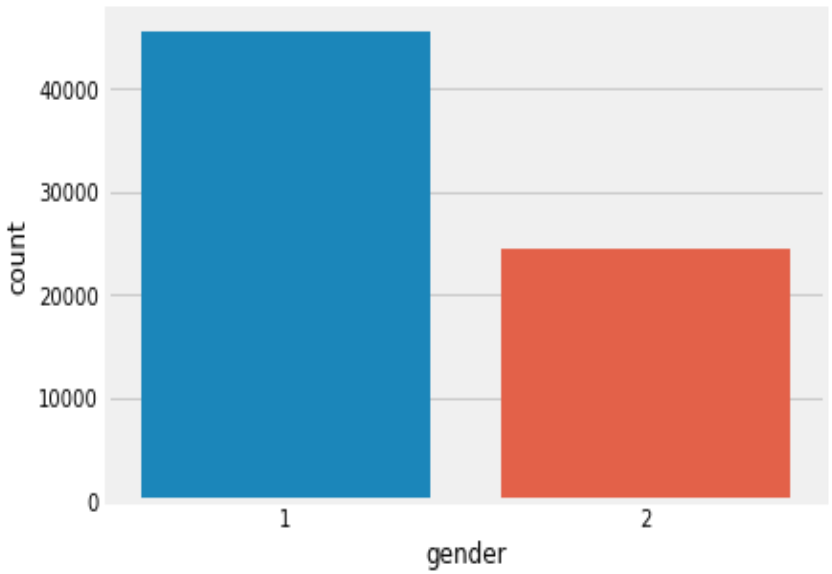
There are 3 types of input features:

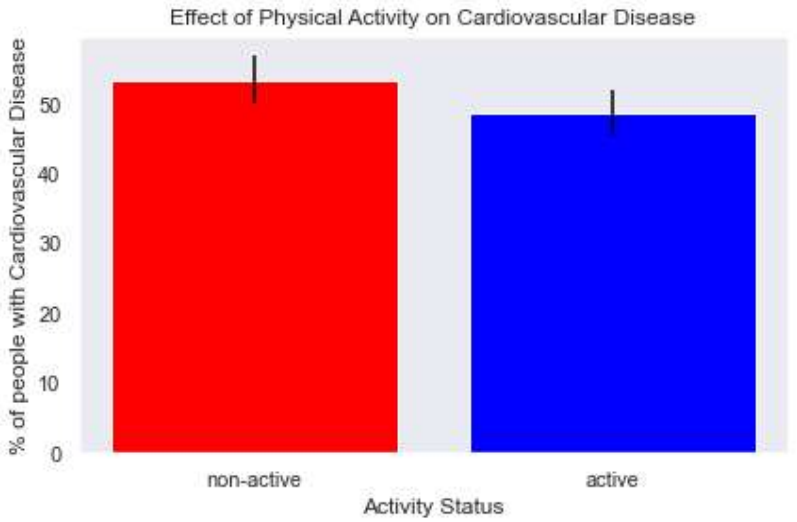
1. *Objective* : factual information.
2. *Examination* : results of medical examination.
3. *Subjective* :information given by the patient.

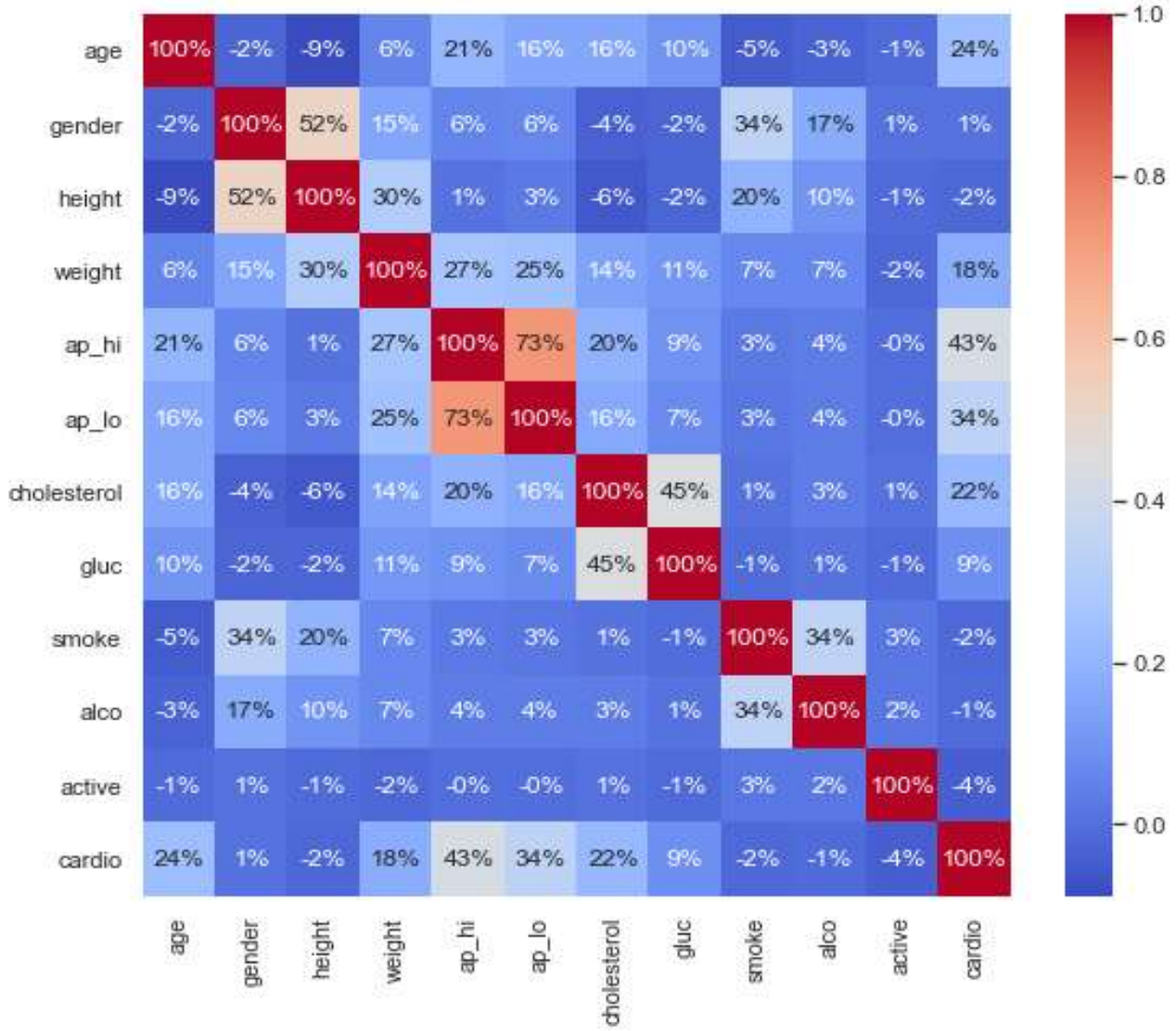
Source of data is kaggle machine learning competitions

2- (EDA & VISUALIZATIONS)







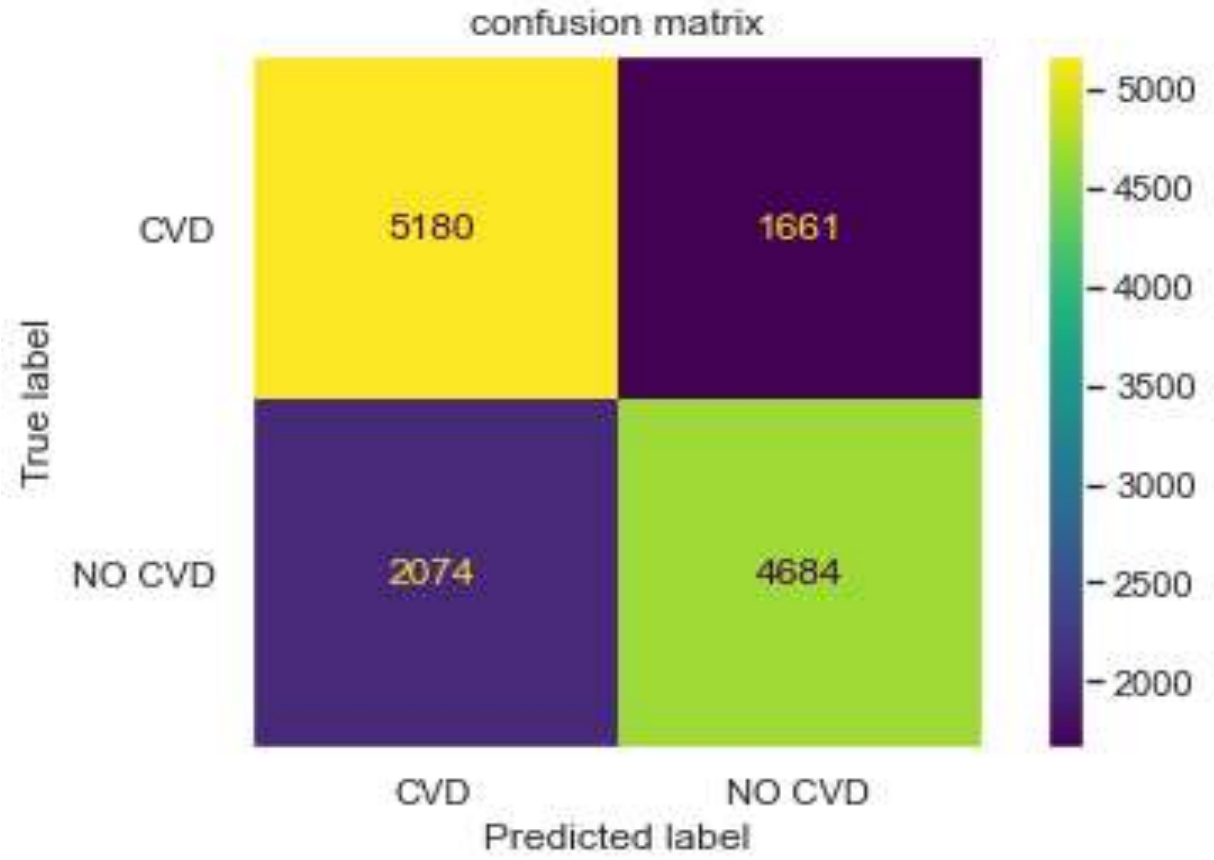


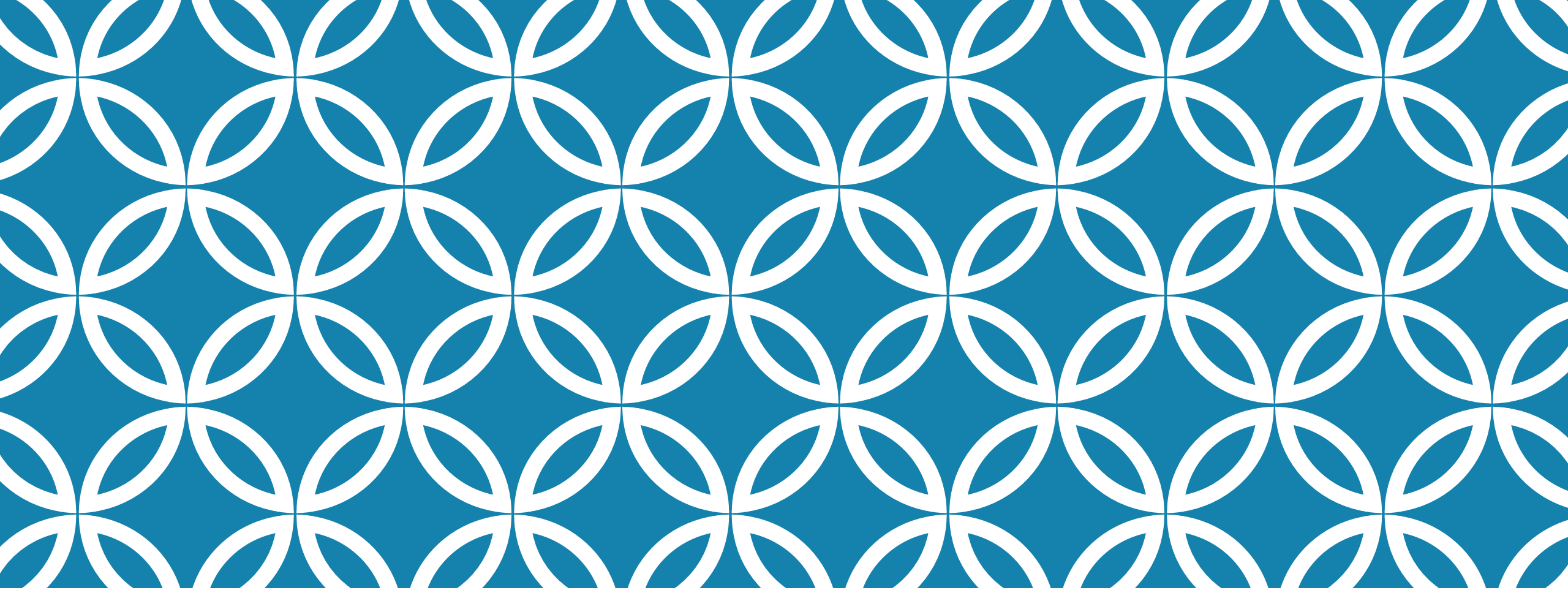
LOOKING INTO THE CORRELATION HEATMAP, AND IT SEEMS LIKE OUR TARGET HAS HIGHER CORRELATION WITH AP_HI, AP_LO ,AGE, WEIGHT,CHOLESTROL.
SOME VARIABLES HAVE HIGHER CORRELATION BETWEEN EACH OTHER.

MODELING

	Classifier	Accuracy with a CV=10	Accuracy after using GridSearchCV()
1	LogisticRegression()	0.709831605265093	0.7167653547905648
2	KNeighborsClassifier()	0.7032150619035672	0.7238982957365837
3	DecisionTreeClassifier()	0.729232791586998	0.7303330416865011
4	MLPClassifier()	0.7140105681631381	0.7180704914635292
5	RandomForestClassifier()	0.7152606308886255	0.7314728286705753

CONFUSION MATRIX FOR RANDOM FOREST CLASSIFIER





THANK YOU FOR LISTENING