REPORT
Predicting Click on Ad

Patrick Gomes de Oliveira

Hello, Mr. Jhon! This report is about the model build to predict Clicks on Ad at your startup.
The model would receive data with the format below:

Daily Time on Site	Age	Area Income	Daily Internet Usage	Clicked on Ad
35.49	47.0	36884.23	170.04	1.0
57.2	42.0	57739.03	110.66	1.0
81.22	53.0	34309.24	223.09	0.0
74.38	40.0	74024.61	220.05	0.0
60.39	45.0	38987.42	108.25	1.0

The graphical analysis (in notebook) was able to show us how separated are the Clickeds given by their features.
With this, we can understand that the variables where the Clicks are better separated:

* Daily Time Spent on Site

* Age

* Daily Internet Usage

* Area income

Chart 1 - Clicked or not clicked by relevants characteristics
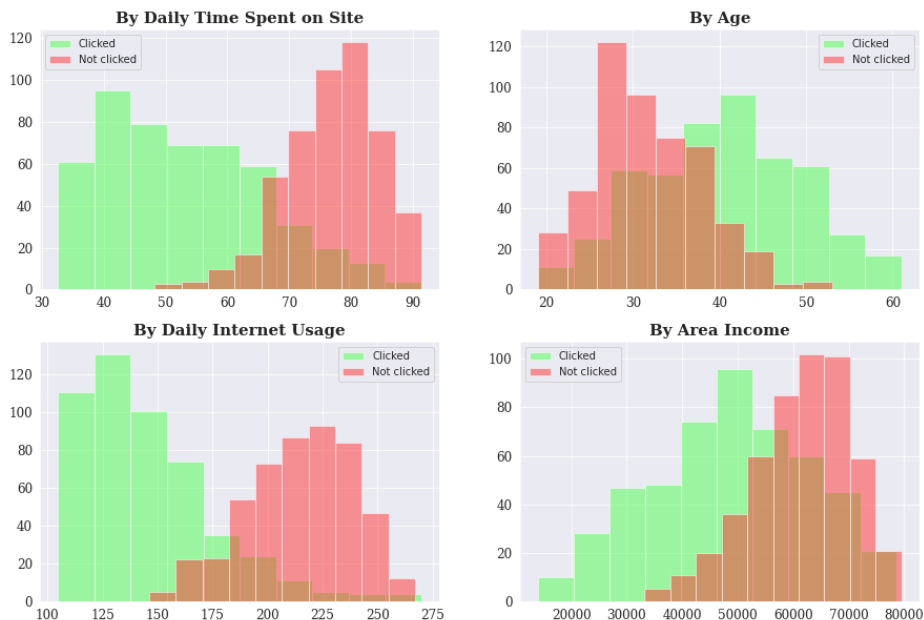
REPORT
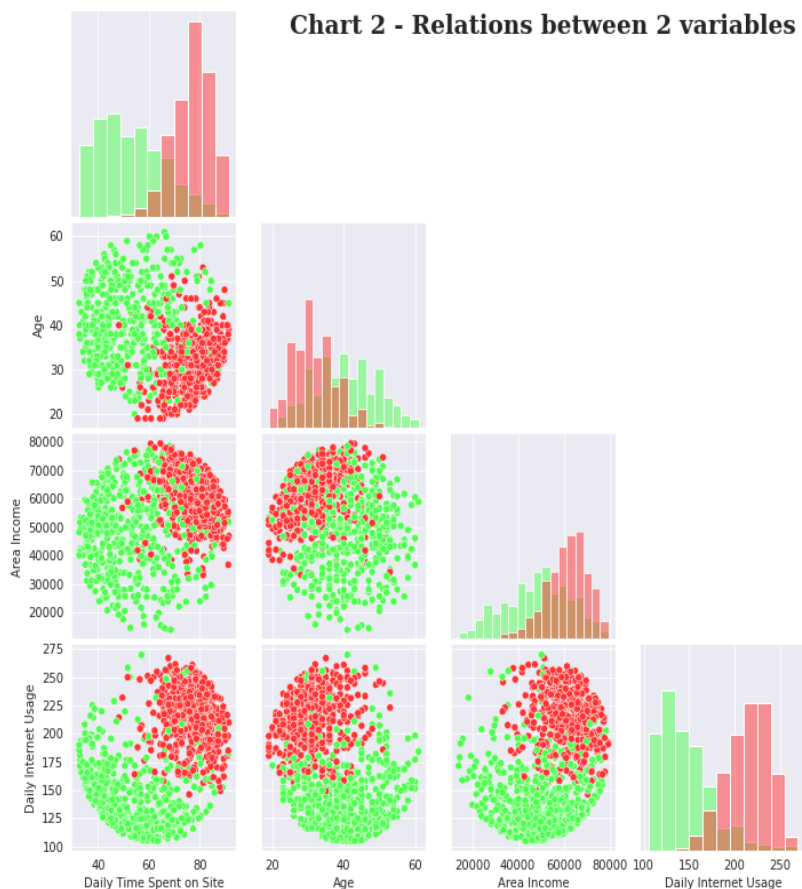# Predicting Click on Ad

Patrick Gomes de Oliveira

Business insights given by univariate analysis:

   * Daily Time Spent on Site: 30 ~ 70 minutes maximize your revenue;

    * Age: 35 ~ 60 years maximize your revenue;

   * Daily Internet Usage: 100 ~ 175 minutes maximize your revenue;

   * Area Income: $20 000 ~ $55 000 maximize your revenue;

 With the univariate analysis we can realize how the Clicked are divided, It facilitates our analysis and model's interpretability.

Bivariate analysis

    Make an bivariate analysis show us a lot of valuous information about the problem and insights to model. For this, we plot the chart below:

Chart 2 - Relations between 2 variables

REPORT
Predicting Click on Ad

Patrick Gomes de Oliveira

Now we unterstand the relation between the different variables, being differencied by "Clicked or Not Clicked".

The next step is the modeling to predict Clicks on Ad

Model

First, was selected the 60% of the data to fitting our model.
Second, was visualize the Confusion Matrix to avaluate the model. We can see below:

Conclusion

Our model was able to predict Clicks on Ad correctly 96.36% of cases.
Was possible and not expensive was able to predict Clicks on Ad using Logistic Regression model,
because your computing costs is low and your precision and interpretability are high