

Predict Clicked Ads Customer Classification by using Machine Learning



Created by:

Haolia

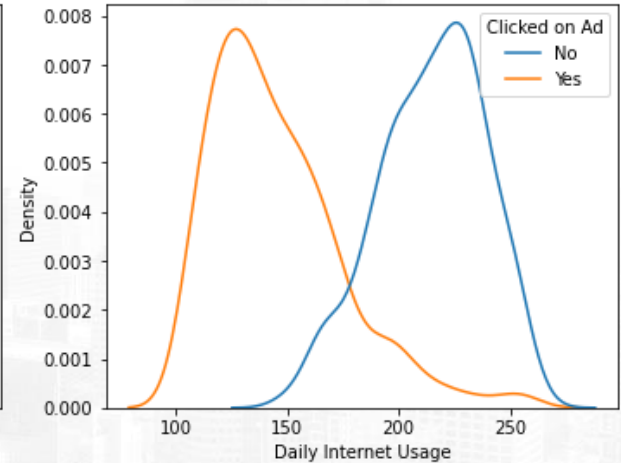
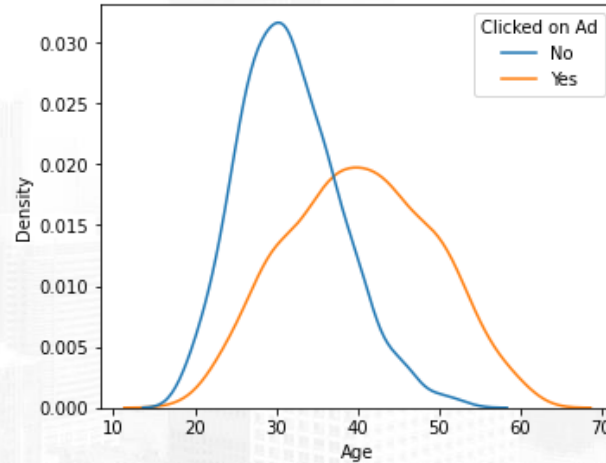
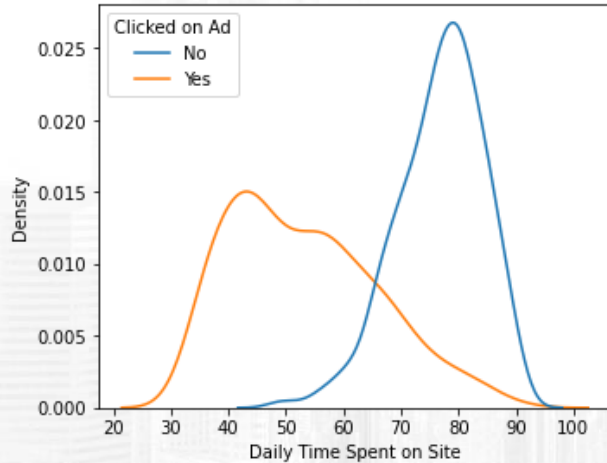
haoliaaul@gmail.com

<https://www.linkedin.com/in/haolia/>

“Geophysical engineering graduates who diligently want to fulfil a role where intellectual, integrity, and curiosity are highly valued. Motivated, able to research, design, implement new features and learn various software. Skill handling problems with unique ways to develop innovative solutions. Proficient using Python, SQL, Tableau and other statistical tools for data multi purposes. Looking for opportunities in data analyst, data science, data engineer and Business Intelligence. ”

“Sebuah perusahaan di Indonesia ingin mengetahui efektifitas sebuah iklan yang mereka tayangkan, hal ini penting bagi perusahaan agar dapat mengetahui seberapa besar ketercapainnya iklan yang dipasarkan sehingga dapat menarik customers untuk melihat iklan.

Dengan mengolah data historical advertisement serta menemukan insight serta pola yang terjadi, maka dapat membantu perusahaan dalam menentukan target marketing, fokus case ini adalah membuat model machine learning classification yang berfungsi menentukan target customers yang tepat ”



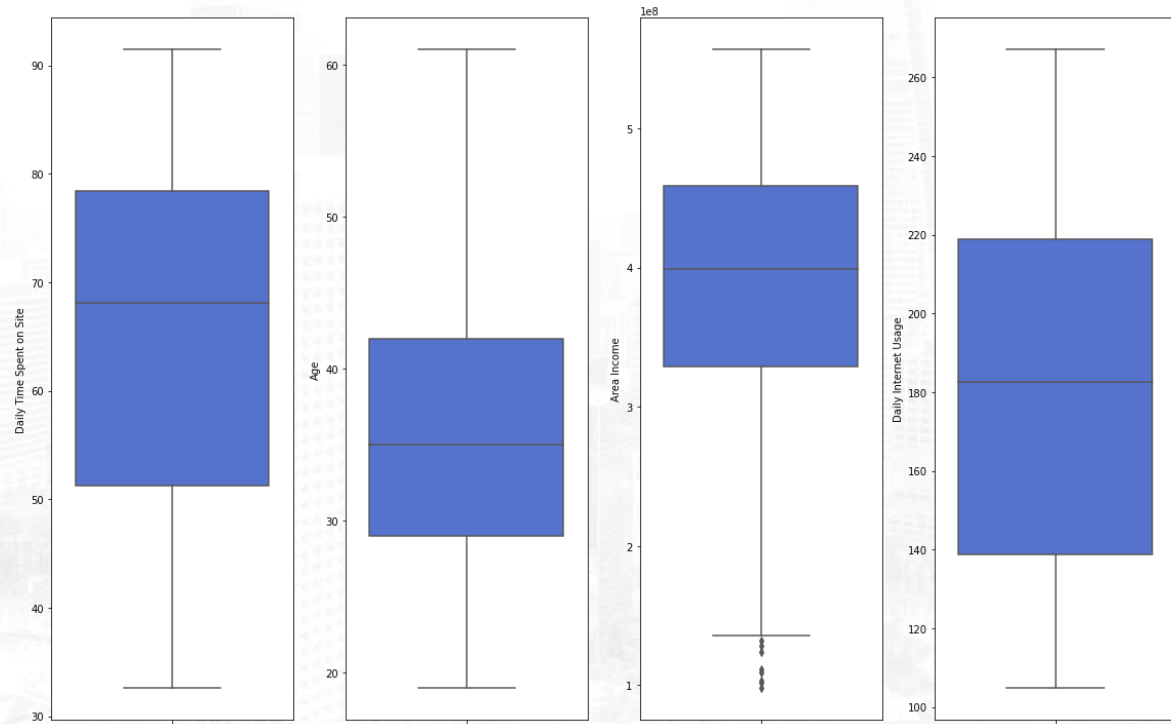
> Observasi :

* user yang mengklik ads adalah user dengan `Daily Time Spent on Site` sekitar 45- 60 menit, sedangkan user yang tidak mengklik ads adalah user dengan `Daily Time Spent on Side` sekitar 72-82 menit.

* user yang mengklik ads rata- rata berusia(`age`) 35-48 tahun. Sedangkan user yang tidak mengklik ads sebagian besar ada pada usia(`age`) 28-35 tahun.

* User dengan `Daily Internet Usage` sekitar 100-160 cenderung mengklik ads. Sedangkan, user dengan `Daily Internet Usage` sekitar 200-250 cenderung tidak mengklik ads.

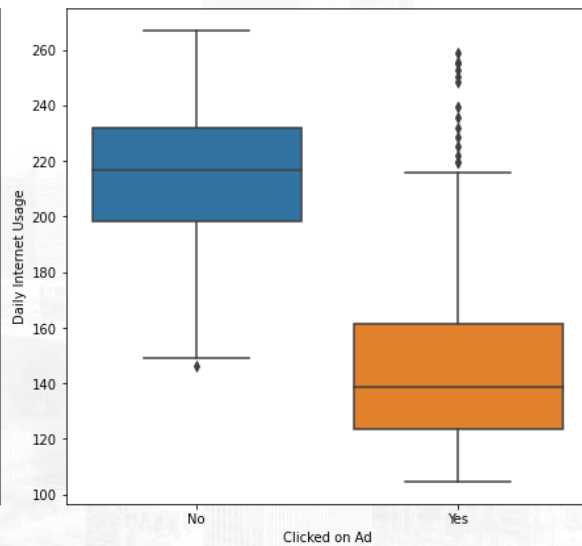
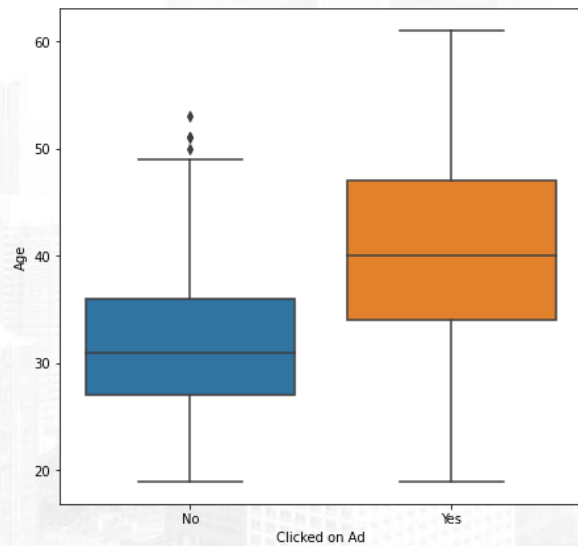
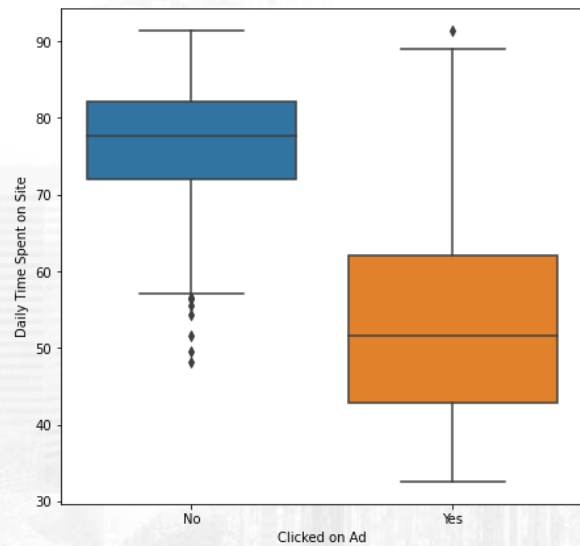
Univariate Analysis Numerical Features



> Observasi:

- Outliers hanya pada feature `Area Income`
- Feature `Dayli Time spent on Site` memiliki distribusi positif skew
- Feature `Age` memiliki distribusi negatif skew
- Feature `Area Income` memiliki distribusi positif skew
- Feature Daily Internet Usage distribusinya mendekati normal

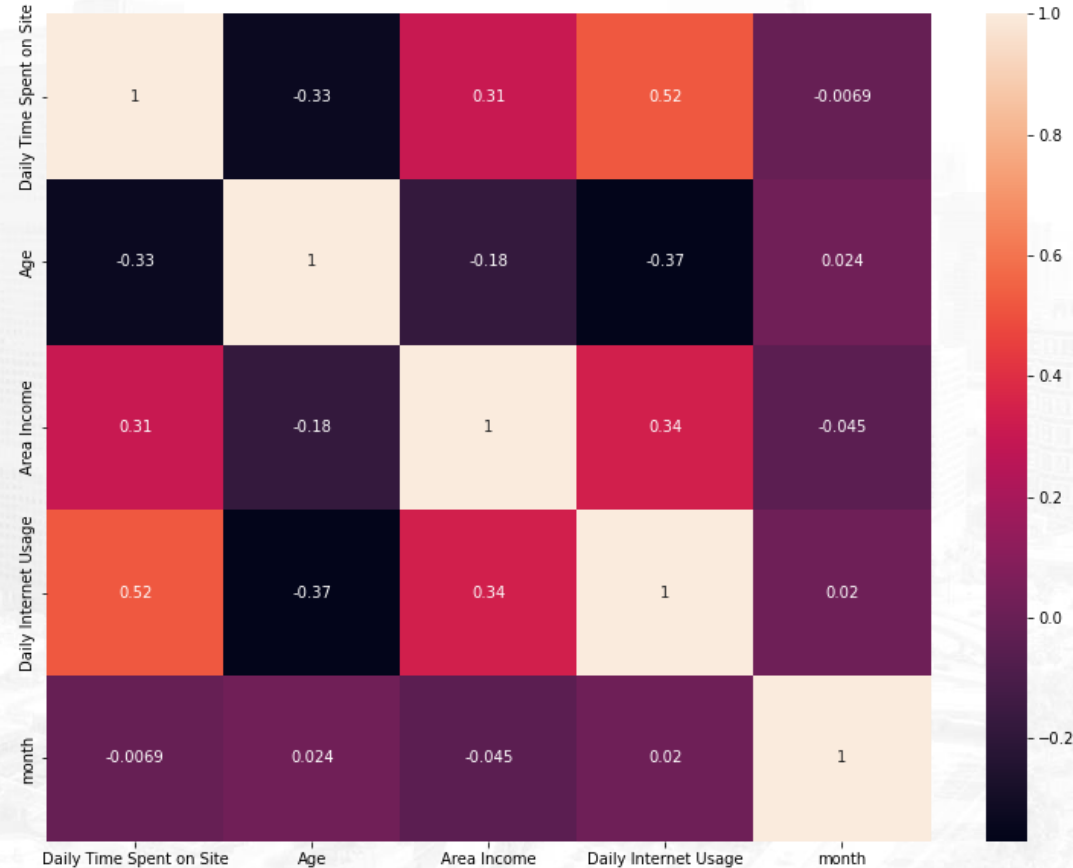
Untuk selengkapnya, dapat melihat jupyter notebook disini



Bivariate Analysis Numerical Features

> Observasi :

- * Feature `Age` berkorelasi negatif lemah dengan feature `Daily Time Spent on Site` dan `Daily Internet Usage`
- * Feature `Area Income` berkorelasi positif dengan feature `Daily Time Spent on Site` dan `Daily Internet Usage` dan berkorelasi negatif dengan feature `Age`



Untuk selengkapnya, dapat melihat jupyter notebook disini