Online Shoppers Purchasing Intention Dataset Analysis

Github: https://github.com/wendyhuai/Online-Shoppers-Analysis

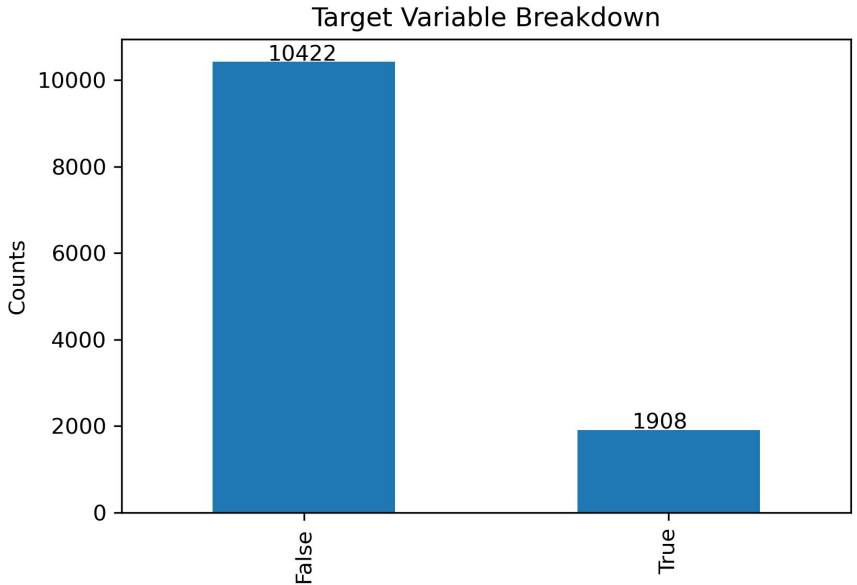
Zuxuan Huai
Brown University
Oct 14, 2020

What factors will classify a website visit as a visit with purchase made or one without?

- E-commerce essential to retail industry
- Classification Problem
- 12330 rows and 18 columns
- 17 Predictor variables
  - 9 Numerical
  - 8 Categorical
- Assume IID
- No Missing Values, No Group Structure

Target Variable Breakdown

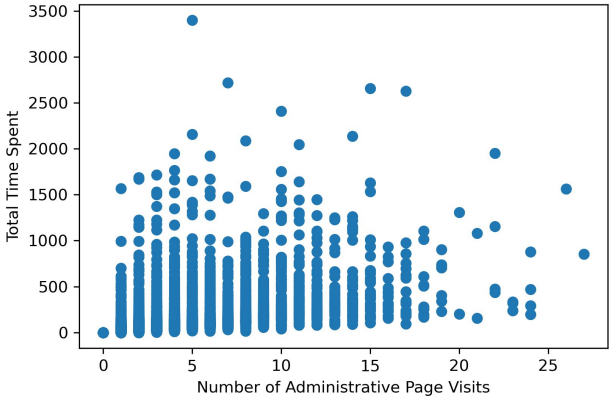
- True / False in Revenue Column



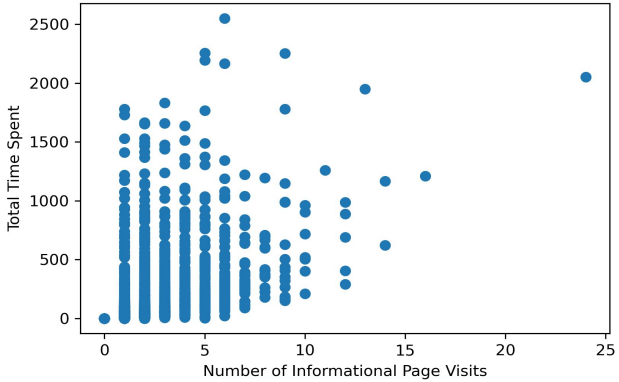
Predictor Variables - Numerical

- Number of Visits: Administrative, Informational, Product-related
- Total time spent on a type of page

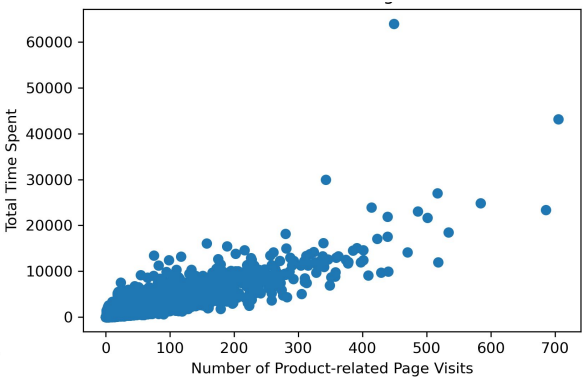
Administrative



Informational

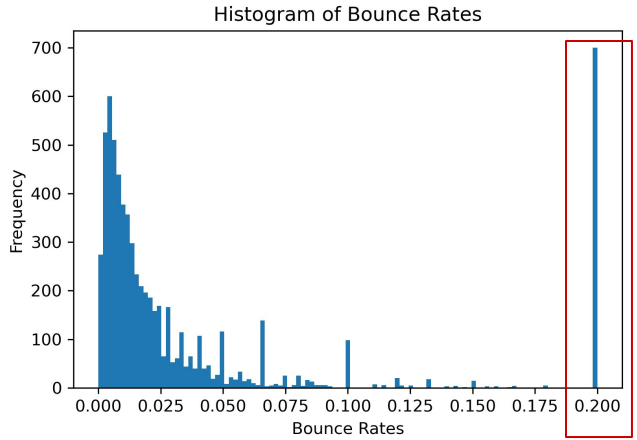


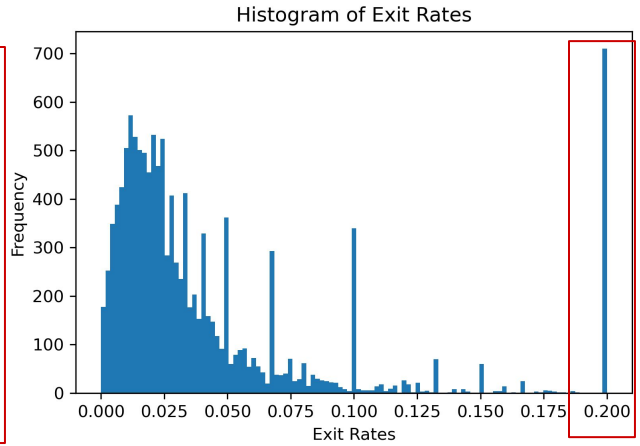
Product - related

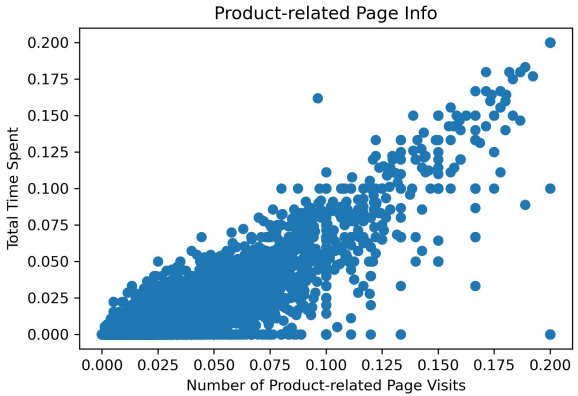


Predictor Variables - Numerical

- **The average rate of the pages visited by the user:**
  - **Bounce Rates:** the percentage of visitors who enter the site from that page and then leave the site
  - **Exit Rates:** for all pageviews to the page, the percentage that were the last in the session

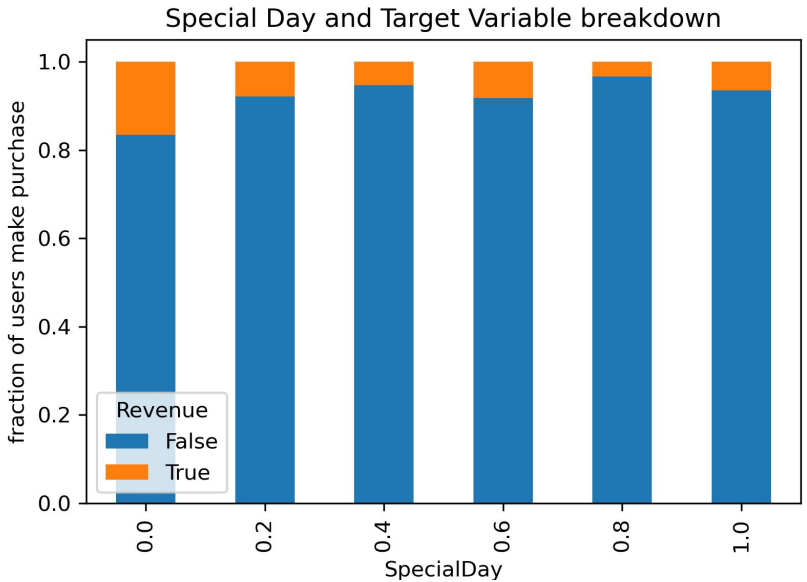






Predictor Variables

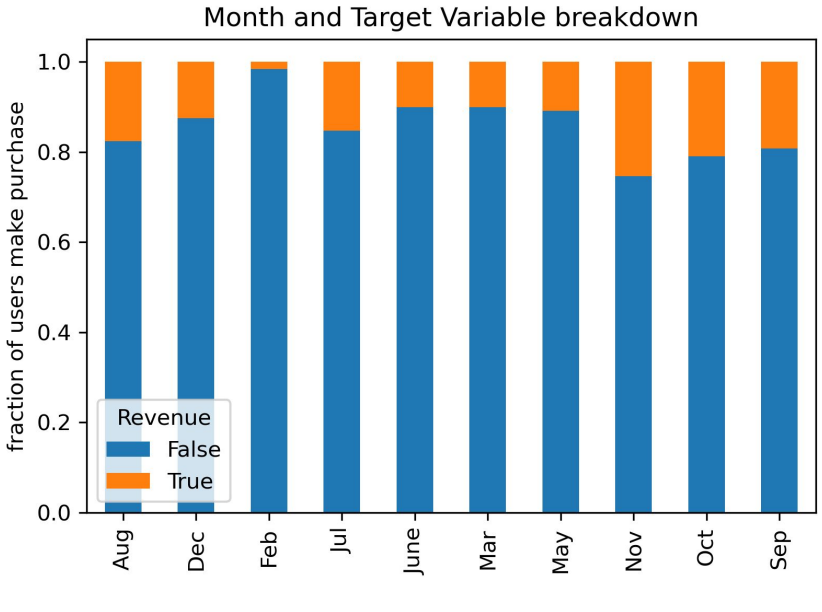
- **Special Day:** numerical [0, 0.2, 0.4, 0.6, 0.8, 1] closeness of the site visiting time to a specific special day

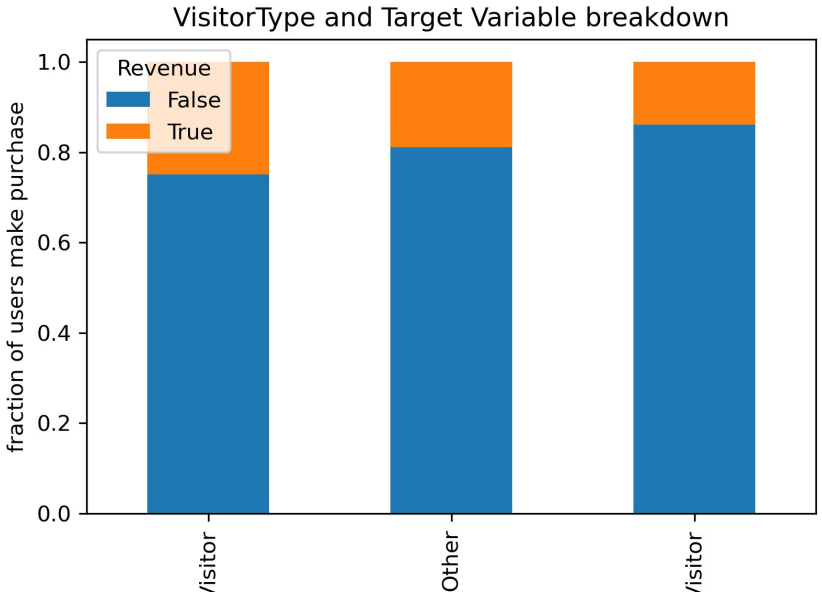


Predictor Variables - Categorical

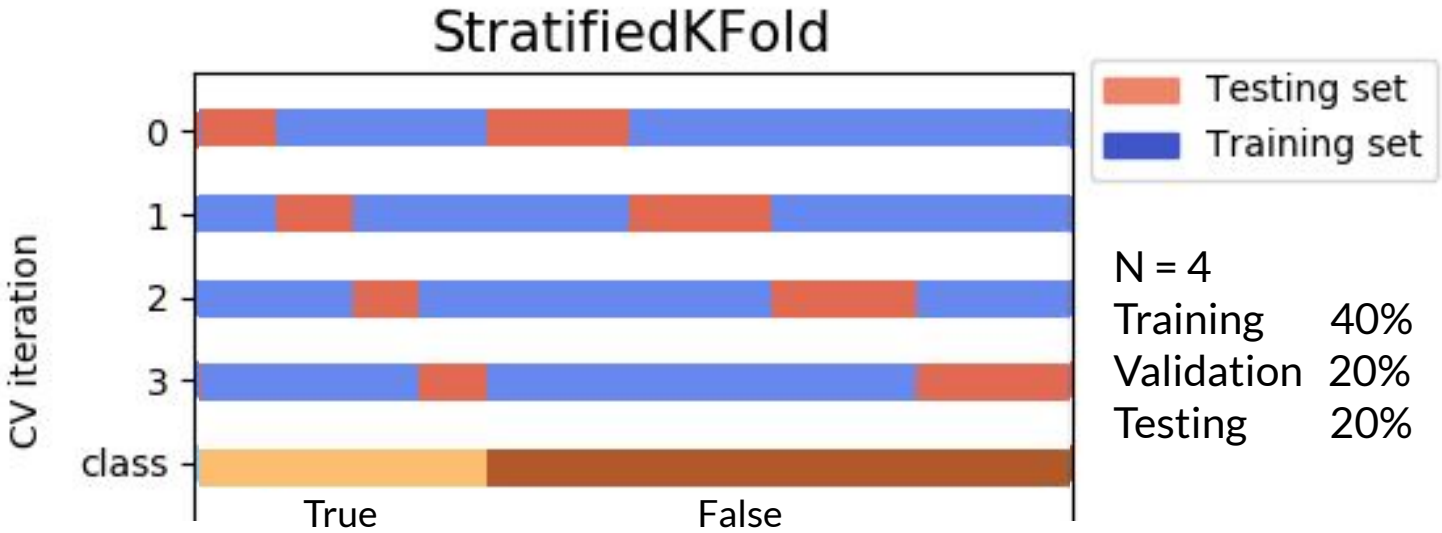
- Month: 10 unique months, excluding Jan and Apr
- Operating Systems: 8 categories
- Browser: 13 categories
- Region: 9 categories
- Traffic Type: 20 categories
- Visitor Type: New Visitor, Returning Visitor, or Other
- Weekend: True/False

Month & Visitor Type





Splitting Data



N = 4

Training	40%
Validation	20%
Testing	20%

Preprocessing

- Ordinal Encoder:
  - Special Day
- One Hot Encoder:
  - Month, Operating Systems, Browser, Region, Traffic Type, Visitor Type, Weekend
- Standard Scalar:
  - Administrative, Administrative Duration
  - Informational, Informational Duration
  - Product-related, Product-related Duration
  - Bounce Rates, Exit Rates, Page Values

Reference

Sakar, C.O., Polat, S.O., Katircioglu, M. et al. Neural Comput & Applic (2018).