

Mining and Visualization2203588

by Bo Dong

Submission date: 10-Mar-2023 04:28PM (UTC+0000)

Submission ID: 199030551

File name: 42017_Bo_Dong_Mining_and_Visualization2203588_482892_1825917068.docx (13.05K)

Word count: 286

Character count: 1626

Online Shoppers Purchasing Intention Dataset Data Set

The Origins of the Data

An online retailer data.

The Dataset

The dataset consists of feature vectors belonging to 12,330 sessions. The dataset was formed so that each session would belong to a different user in a 1-year period to avoid any tendency to a specific campaign, special day, user profile, or period. Of the 12,330 sessions in the dataset, 84.5% (10,422) were negative class samples that did not end with shopping, and the rest (1908) were positive class samples ending with shopping.

Attribute Information

1.Numerical features

- Administrative
- Administrative duration
- Informational
- Informational duration
- Product related
- Product related duration
- Bounce rate
- Page value
- Special day

2.Categorical features

- OperatingSystems
- Browser
- Region
- TrafficType
- VisitorType
- Weekend
- Month
- Revenue

Suitability for Analysis

The dataset is large enough to allow the application of various data mining techniques to discover interesting connections between individual attributes. High data quality. The logic between attribute relationships is simple and can correctly judge the prediction results of data mining techniques.

Potential Research Questions

The project aims to address the following some questions:

- Can data mining be used to predict visitors' purchase intent?
- Can data mining be used to predict the purchasing power of visitors in different regions?
- Can data mining predict how often visitors will buy over time?

Researcher

1.C. Okan Sakar

Department of Computer Engineering, Faculty of Engineering and Natural Sciences,
Bahcesehir University, 34349 Besiktas, Istanbul, Turkey

2.Yomi Kastro

Inveon Information Technologies Consultancy and Trade, 34335 Istanbul, Turkey

References

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

Note: Due to network reasons, I can't connect to YouTube

FINAL GRADE

GENERAL COMMENTS

Instructor

36/100

PAGE 1

PAGE 2

GRADING FORM: DATA MINING/ VISUALISATION

BO DONG

36

KNOWLEDGE AND UNDERSTANDING

Dataset Suitability (What features of the dataset lend themselves to data mining? Has it been mined previously? Why does the student feel that they have the contextual background to work with this data?)



Feature vectors needed to be explained in greater detail. Has this dataset been mined previously ? Any background knowledge that would support the reasons for suggested analysis?

10

COGNITIVE AND INTELLECTUAL

Objectives (What does the student plan to investigate/expect to find? Are these SMART?)

-  Proposed questions are valid and a form of prediction modelling may best aid in learning more about problem as indicated in submission.

8

PRACTICAL APPLICATION SKILLS

Dataset Description (Where did the data come from? What does the data describe? Is the sample large enough? What data formats and types are involved?)

-  Limited information offered concerning origin of dataset. Data types are not explained nor has submission indicated if chosen dataset is useful for indicated analysis.

15

TRANSFERABLE SKILLS FOR LIFE

Presentation/Slides (Clarity, readability, flow, accuracy, completeness, performance)

-  Overall submission is organised and is readable but does not offer sufficient information and clarity for meeting this criteria in full.

3