Vilniaus Universitetas

Duomenų aibės aprašymas

Darbą atliko:

Vainius Gataveckas, Matas Gaulia, Dovydas Martinkus

Duomenų Mokslas

3 kursas 2 gr.

# Duomenų aibė

Online Shoppers Purchasing Intention duomenų aibė

Duomenų aibės šaltinis: UCI Machine Learning Repository

Nuoroda per internetą: <https://archive.ics.uci.edu/ml/datasets/Online+Shoppers+Purchasing+Intention+Dataset>

Duomenų aibę sudaro duomenys apie 12330 objektų su 18 požymių:

„Administrative“ (kiekybinis, diskretus, santykių skalė) – vartotojo aplankytų šio tipo puslapių skaičius

„Administrive\_Duration“ (kiekybinis, tolydusis, santykių skalė) - laikas, praleistas šios kategorijos puslapiuose

„Informational“ (kiekybinis, diskretus, santykių skalė) – vartotojo aplankytų šio tipo puslapių skaičius

„Informational\_Duration“ (kiekybinis, tolydusis, santykių skalė) – laikas, praleistas šios kategorijos puslapiuose

„ProductRelated“ (kiekybinis, diskretus, santykių skalė) - vartotojo aplankytų šio tipo puslapių skaičius

„ProductRelated\_Duration“ (kiekybinis, tolydusis, santykių skalė) - laikas, praleistas šios kategorijos puslapiuose

„BounceRates“ (kiekybinis, tolydusis, santykių skalė) – procentas vartotojų, kurie palieka svetainę per šį puslapį prieš tai neatlikę jokių papildomų veiksmų

„ExitRates“ (kiekybinis, tolydusis, santykių skalė) – procentas vartotojų, kurie palieka svetainę per šį puslapį

„PageValues“ (kiekybinis, tolydusis, santykių skalė) – Google Analytics Page Value reikšmė

„SpecialDay“ (kiekybinis, tolydusis, santykių skalė) – Artumas iki ypatingų dienų arba švenčių

„Month“ (kategorinis, nominalus) – mėnesis, kada buvo peržiūrėtas puslapis

„OperatingSystems“ (kategorinis, nominalus) - skaičius, identifikuojantis vartotojo naudotą operacinę sistemą

„Browser“ (kategorinis, nominalus) - skaičius, identifikuojantis vartotojo naudotą naršyklę

„Region“ (kategorinis, nominalus) - skaičius, identifikuojantis vartotojo regioną

„TrafficType“ (kategorinis, nominalus) - skaičius, identifikuojantis vartotojo srauto tipą

„VisitorType“ (kategorinis, nominalus) - skaičius, identifikuojantis ar vartotojas yra naujas ar sugrįžtantis

„Weekend“ (kategorinis, nominalus) - skaičius, identifikuojantis ar vartotojo sesija įvyko savaitgalį

„Revenue“ (kategorinis, nominalus) - skaičius, identifikuojantis vartotojo regioną

Alternatvya:

Automobile duomenų aibė

Duomenų aibės šaltinis: UCI Machine Learning Repository

Nuoroda per internatą: <https://archive.ics.uci.edu/ml/datasets/automobile>

Duomenų aibę sudaro duomenys apie 204 objektų su 26 požymių:

Alternatyva:

Default of Credit Card Clients duomenų aibė

Duomenų aibės šaltinis: UCI Machine Learning Repository

Nuoroda per internatą: <https://archive.ics.uci.edu/ml/datasets/default+of+credit+card+clients>

Duomenų aibę sudaro duomenys apie 30001 objektų su 25 požymių:

Alternatyva:

https://archive.ics.uci.edu/ml/datasets/Polish+companies+bankruptcy+data