

Kiértékelés

A megadott adatkészleten keresztül megvizsgáltam a hallgatók költési szokásait, és megnéztem, hogy mi a diákok preferált fizetési módja, hogyan függ össze a demográfiai jellemzőkkel és tanulmányi háttérükkel.

A három különböző gépi tanulási modell (Döntési fa, Random Forest, XGBoost) használatával elemeztem az adatkészletet, és megállapítottam,

hogy az XGBoost modell a legmagasabb pontosságot (35,87%) érte el.

Az eredmények azt mutatják, hogy a diákok preferált fizetési módja összefügg a jövedelmükkel, a tanulóéveikkel és a szakukkal.

- A magasabb jövedelemmel rendelkező diákok valószínűbben használnak mobil- vagy kártyás fizetést, míg az alacsonyabb jövedelemmel rendelkező diákok valószínűbben használnak készpénzt. A felsőbb éves diákok is hajlamosabbak a készpénzmentes fizetési módokat választani, mint az alsóbb évesek. Ezenkívül a művészeti és humán tudományok hallgatói gyakrabban használnak készpénzt, mint a STEM-területeken tanuló diákok.

Fontos megjegyezni, hogy ezek az eredmények csak a megadott adatkészletre vonatkoznak, és további kutatásokra van szükség a diákok költési szokásainak teljesebb megértéséhez, úgy gondolom hogy nagyobb adatkészlet esetén lehetett volna jobb eredményt elérni. Mivel a gépi tanulási modellek pontossága az adatkészlet minőségétől és a modell paramétereitől is függ.

Összességében ez a beadandó bemutatja a diákok költési szokásait és a preferált fizetési módok befolyásoló tényezőket.

Ezeket az eredményeket felhasználhatjuk edukációs célra, például a diákok pénzügyi tudatosságának javítására és a pénzügyi szolgáltatások fejlesztésére, amik jobban megfelelnek a diákok igényeinek.

Saját vélemény: Nagyobb adatkészlettel javíthatnánk a pontosságot

Ebben a beadandóban 1000 diák költési szokásaira vonatkozó adatokat használtunk a hallgatók preferált fizetési módjának feltárására. Bár az XGBoost modell 35,87%-os pontosságot ért el, **a nagyobb adatkészlettel még pontosabb eredményeket érhetnénk el.**

Nagyobb adatkészlettel:

- **Jobban reprezentálhatnánk a diákok populációját:** Jelenleg az adatkészletünk nem feltétlenül tükrözi a teljes diákpopulációt, ami torzításokhoz vezethet az eredményekben. Egy nagyobb adatkészlet pontosabb képet adna a diákok költési szokásairól.
- **Javíthatnánk a gépi tanulási modelleket:** A modellek jobban működnek, ha nagyobb és változatosabb adatkészleten tanítják. Ez azt jelenti, hogy a modellek pontosabban jósolhatnák a jövőben a diákok költési szokásait.

Összefoglalva, a nagyobb adatkészlettel **elkerülhető lenne a torzítás, és feltárhatóak lennének a ritkább minták és összefüggések** a diákok költési szokásai között.