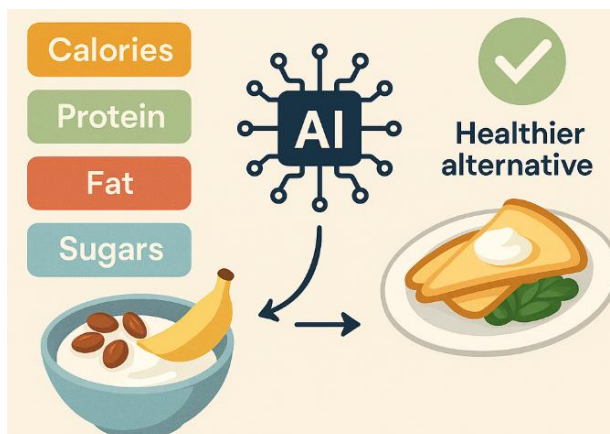


Úloha 1:

Predikcia energetickej hodnoty jedál a tvorba zdravšej alternatívy pomocou AI

V dnešnom svete je informácií o jedle a výžive viac než kedykoľvek predtým. Na sociálnych sieťach sa každý deň objavujú nové tipy, zázračné diéty a odporúčania, ktoré si často odporujú. Napriek tomu počet ľudí so zlou životosprávou, obezitou či nedostatkom pohybu stále rastie. Čoraz jasnejšie vidíme, že najväčším problémom nie je nedostatok dát, ale nedostatok schopnosti tieto dáta správne analyzovať a využiť.



A práve tu prichádzajú na scénu aj metódy umelej inteligencie. Dnes dokážu pomáhať v medicíne, autonómnych vozidlách aj energetike. Prečo by nám teda nemohla pomôcť aj pri jednom z najosobnejších rozhodnutí, čo si dám dnes na tanier?

V tejto úlohe budete pracovať s rozsiahlym nutričným datasetom FoodData Central, ktorý obsahuje tisíce potravín s ich nutričnými hodnotami (kalórie, bielkoviny, tuky, sacharidy a ďalšie).

Vašou úlohou je vytvoriť systém založený na metódach umelej inteligencie (strojovom učení), ktorý:

- predikuje energetickú hodnotu jedla, teda celkovú kalorickú hodnotu jeho ingrediencií.
- Navrhne zdravšiu alternatívu jedla, s nižším obsahom kalórií, nižším obsahom sacharidov a vyšším podielom bielkovín.

Príprava dát:

- Stiahnite si zip súbor [nutričného datasetu](#) zo stránky FoodData Central a načítajte do svojho programu údaje z tabuliek tohto datasetu **food.csv**, **nutrient.csv**, **food_nutrient.csv** (alebo využite notebook s načítaním dát sprístupnený v Edupage).
- Z dát odstráňte riadky s neúplnými alebo extrémne chýbajúcimi hodnotami (napr. potraviny bez uvedených kcal, bielkovín alebo cukrov).
- Vytvorte nové pomocné atribúty (features), napr.:
 - pomer (bielkoviny / kalórie),
 - pomer (sacharidy / kalórie),
 - pomer (tuky / kalórie),
 - kategória potraviny (napr. obilniny, mliečne výrobky, mäso, zelenina) ak je dostupná.

Predikcia kalórií jedla:

- Každé jedlo pozostáva z niekoľkých ingrediencií, každá ingrediencia je potravina z datasetu a množstva v gramoch. Napríklad vstup do programu má tvar:
150 g apples, 30 g lettuce, 50 g bananas
Pričom názvy všetkých dostupných potravín sú v tabuľke **food.csv**.
- Spočítajte „skutočnú“ energetickú hodnotu jedla na vstupe: pre každú ingredienciu nájdite kcal / 100 g z datasetu, prepočítajte podľa gramáže, sčítajte (údaje o kalóriách sú v tabuľke **nutrient.csv** a **food_nutrient.csv**).

-
- Vytvorte vlastný dataset s minimálne 50 jedlami (každé jedlo má aspoň 3 a maximálne 10 ingrediencií). Pre každé jedlo si uchovajte: ingrediencie + množstvá + skutočná energetická hodnota jedla (kalórie) kcal.
 - Rozdeľte dáta: tréningová časť modelu (napr. 80 %) a testovacia (20 %).
 - Natrénujte aspoň dva modely strojového učenia (napr. lineárna regresia, random forest, neurónové siete) na predikciu skutočnej energetickej hodnoty jedla (kalórií) na základe Vami vytvoreného datasetu.
 - Vyhodnoťte ich výkonnosť pomocou metrík ako MAE (Mean Absolute Error) a RMSE (Root Mean Square Error).
 - Vytvorte graf, ktorý porovnáva skutočné vs. predpovedané kalorické hodnoty pre testovaciu sadu.

Návrh zdravšej alternatívy jedla:

- Zvoľte tri jedlá, ktoré ste už použili v predchádzajúcej časti alebo úplne nové (napr. populárne slovenské jedlá).
- Pre každé jedlo navrhните naprogramujte alternatívu, ktorá:
 - má minimálne o 25 % menej kalórií než pôvodné jedlo,
 - má nižší obsah sacharidov,
 - má vyšší podiel bielkovín než pôvodné jedlo.
- Pri návrhu môžete:
 - nahradiť ingrediencie zdravšími variantmi (napr. biela múka vs. celozrnná, tvarohový krém vs. nízkotučný tvaroh, smotanový jogurt vs. nízkotučný jogurt),
 - upraviť množstvá ingrediencií (napr. menej oleja),
 - pridať ingredienciu s vysokým obsahom bielkovín (napr. tofu, grécky jogurt).
- Vytvorte tabuľku porovnania (pôvodné jedlo vs. alternatíva) s údajmi: kalórie, sacharidy, bielkoviny, tuky, percentuálna zmena.

Čo odovzdať:

1. Vytvorený kód so spracovaním oboch úloh (**.py** alebo **.ipynb**),
2. Dokumentácia (súbor **README** alebo komentáre v kóde) vysvetľujúca kroky, ktoré ste použili vo svojom riešení,
3. Prípadne ďalšie materiály, ktoré uznáte za vhodné.
4. Odovzdávate **jeden .zip súbor**, ktorý má v názve Vaše priezvisko a nepresiahne veľkosť 70MB. V tomto súbore sú zahrnuté všetky vyššie uvedené súbory.

Hodnotenie:

0-100 bodov, prihliada sa na:

- kvalitu výsledku,
 - nápaditosť riešenia,
 - zrozumiteľnosť a prínos komentára k riešeniu.
-