

**SLOVENSKÁ TECHNICKÁ UNIVERZITA V BRATISLAVE
FAKULTA ELEKTROTECHNIKY A INFORMATIKY**

Evidenčné číslo: FEI-16605-115297

**ODPORÚČACÍ SYSTÉM S VYUŽITÍM UMELEJ
INTELIGENCIE
BAKALÁRSKA PRÁCA**

SLOVENSKÁ TECHNICKÁ UNIVERZITA V BRATISLAVE
FAKULTA ELEKTROTECHNIKY A INFORMATIKY

Evidenčné číslo: FEI-16605-115297

ODPORÚČACÍ SYSTÉM S VYUŽITÍM UMELEJ
INTELIGENCIE
BAKALÁRSKA PRÁCA

Študijný program :	Aplikovaná informatika
Číslo študijného odboru:	2511
Názov študijného odboru:	9.2.9 Aplikovaná informatika
Školiace pracovisko:	Ústav informatiky a matematiky
Vedúci záverečnej práce:	Ing. Marek Repka, PhD.



ZADANIE BAKALÁRSKEJ PRÁCE

Študent: **Filip Berčík**
ID študenta: 115297
Študijný program: aplikovaná informatika
Študijný odbor: informatika
Vedúci práce: Ing. Marek Repka, PhD.
Vedúci pracoviska: doc. Ing. Milan Vojvoda, PhD.
Miesto vypracovania: Ústav informatiky a matematiky

Názov práce: **Odporúčací systém s využitím umelej inteligencie**

Jazyk, v ktorom sa práca vypracuje: slovenský jazyk

Špecifikácia zadania:

Umelá inteligencia (AI) mení spôsob fungovania ľudstva. Cieľom práce je nájsť a použiť tooly pre umelú inteligenciu a implementovať reálny prípad využitia AI – odporúčací systém, ktorý bude viesť odfiltrovať pre používateľa nezaujímavé položky.

Úlohy:

1. Preskúmať teoretické základy odporúčacích systémov a ich implementácie v umelej inteligencii.
2. Identifikovať a vyhodnotiť existujúce frameworky, nástroje a metódy vhodné pre implementáciu odporúčacieho systému.
3. Realizovať testovanie vybraných nástrojov a metód na základe ich efektivity a schopnosti riešenia konkrétnych úloh. Navrhnuť architektúru a funkcionality odporúčacieho systému s využitím umelej inteligencie. Vyvinúť a implementovať odporúčací systém v súlade s navrhnutými špecifikáciami.
4. Zhodnotiť efektívnosť implementovaného odporúčacieho systému pomocou definovanej metodiky.
5. Podrobne zdokumentovať celý proces od analýzy problematiky a výberu nástrojov až po implementáciu a vyhodnotenie výsledkov.

Zoznam odbornej literatúry:

1. Cop., Aggarwal, C. C. (2016). Recommender Systems. Cham Springer International Publishing Imprint: Springer.
2. Literatúra: Ricci, F., Lior Rokach, & Bracha Shapira. (2015). Recommender systems handbook. Springer

Termín odovzdania bakalárskej práce: 31. 05. 2024

Dátum schválenia zadania bakalárskej práce: 20. 05. 2024

Zadanie bakalárskej práce schválil: prof. Dr. rer. nat. Martin Drozda – garant študijného programu

SÚHRN

SLOVENSKÁ TECHNICKÁ UNIVERZITA V BRATISLAVE
FAKULTA ELEKTROTECHNIKY A INFORMATIKY

Študijný program :	Aplikovaná informatika
Vyberte typ práce	Bakalárska práca
Autor:	Filip Berčík
Vedúci záverečnej práce:	Ing. Marek Repka, PhD.
Miesto a rok predloženia práce:	Bratislava 2024

Táto práca sa zaoberá problematikou, konceptom a dôležitosťou odporúčacích systémov využívajúcich umelú inteligenciu v súčasnom digitálnom svete, kde je pre používateľov neustále generované veľké množstvo obsahu. Následne je preskúmaný teoretický základ, ktorý obsahuje dôležité princípy, techniky a metódy používané v odporúčacích systémoch. V práci je ďalej predstavený návrh a implementácia vlastného odporúčacieho systému, ktorý využíva moderné programovacie jazyky, technológie, frameworky a knižnice ako Python a TensorFlow. Následne je navrhnutý model vyhodnotený pomocou metrík ako presnosť a stratovosť a sú analyzované dosiahnuté výsledky a efektívnosť systému. V záverečnej časti práce sú diskutované možné vylepšenia a budúce smerovanie. Výsledok práce je odporúčací systém pre obrázky s webovým rozhraním, kde odporúčania generuje model umelej inteligencie.

Kľúčové slová: odporúčací systém, umelá inteligencia, personalizované odporúčanie

ABSTRACT

SLOVAK UNIVERSITY OF TECHNOLOGY IN BRATISLAVA
FACULTY OF ELECTRICAL ENGINEERING AND INFORMATION
TECHNOLOGY

Study Programme:	Applied Informatics
Bachelor Thesis:	Recommendation system based on Artificial Intelligence
Author:	Filip Berčík
Supervisor:	Ing. Marek Repka, PhD.
Place and year of submission:	Bratislava 2024

This work focuses on the issues, concepts, and importance of recommendation systems based on artificial intelligence in today's digital world, where a large amount of content is being constantly generated for users. The theoretical basis is examined, covering key principles, methods, and techniques employed in recommendation systems. The work presents the design and implementation of a recommendation system, utilizing modern programming languages, technologies, frameworks, and libraries such as Python and TensorFlow. Then the proposed model is evaluated using metrics such as accuracy and loss, analyzing the achieved results and system effectiveness. In the concluding part of the work, possible improvements and future directions are discussed. The outcome of the work is an image recommendation system with a web interface, where recommendations are generated by an artificial intelligence model.

Key words: recommendation system, artificial intelligence, personalized recommendations

Vyhlásenie autora

Podpísaný Filip Berčík čestne vyhlasujem, že som Bakalársku prácu Odporúčací systém s využitím umelej inteligencie vypracoval na základe poznatkov získaných počas štúdia a informácií z dostupnej literatúry uvedenej v práci.

Uvedenú prácu som vypracoval pod vedením Ing. Marek Repka, PhD. .

V Bratislave dňa 05.06.2024

.....

podpis autora

Pod'akovanie

Týmto by som sa chcel poďakovať môjmu vedúcemu bakalárskej práce Ing. Marekovi Repkovi, PhD. za jeho ústretové a poučné vedenie, odborné rady a nápady počas vypracovania mojej bakalárskej práce.

Obsah

Úvod	1	
1	Odporúčacie systémy	2
1.1	História odporúčacích systémov	3
1.2	Využitie odporúčacích systémov	3
1.2.1	Netflix	4
1.2.2	Amazon	5
1.3	Ciele odporúčacích systémov	5
1.4	Typy odporúčacích systémov	6
1.4.1	Kolaboratívne	6
1.4.2	Obsahové	7
1.4.3	Demografické	8
1.4.4	Community-based	8
1.4.5	Knowledge-based	8
1.4.6	Hybridné	9
1.5	Problémy odporúčacích systémov	10
1.6	Umelá inteligencia v odporúčacích systémoch	11
1.6.1	Hlboké učenie	11
1.6.2	Transferované učenie	12
1.6.3	Aktívne učenie	13
1.6.4	Reinforcement learning	13
1.6.5	Evolučné algoritmy	13
1.6.6	Spracovanie prirodzeného jazyka	13
1.6.7	Fuzzy techniky	14
1.6.8	Počítačové videnie	14
2	Ciele práce	15
3	Návrh systému	16
3.1	Špecifikácia požiadaviek	16
3.2	Klasifikácia položiek	16

3.3	Umelá inteligencia	17
3.4	Vytvorenie vzorových používateľov	17
3.5	Odporúčanie položiek	17
4	Využitie nástroje a technológie.....	18
4.1	Python	18
4.1.1	TensorFlow	18
4.1.2	Flask.....	18
4.1.3	Numpy	18
4.1.4	Pandas	19
4.2	HTML, CSS, JavaScript	19
4.3	Docker.....	19
5	Implementácia	20
5.1	realistic, modern, minimalistic, happy modely	20
5.1.1	categorizeData.py	21
5.2	show model.....	21
5.3	Webové rozhranie	24
5.3.1	Docker.....	24
5.3.2	Flask server.....	24
6	Vyhodnotenie	25
6.1	Metodika vyhodnotenia	25
6.2	Evaluácia modelu.....	25
6.3	Vylepšenia	28
Záver	29
Zoznam použitej literatúry	30
Prílohy	I
Príloha A: Technická dokumentácia	II

Zoznam obrázkov a tabuliek

Obrázok 1 Logický proces fungovania odporúčacieho systému	4
Obrázok 2 Oblasti a techniky umelej inteligencie.....	11
Obrázok 3 Diagram implementovaného systému.....	22
Obrázok 4 Trénovacie a validačné straty a presnosti	26
Obrázok 5 Vývoj RMSE a MAE počas trénovaniak.....	26
Obrázok 6 RMSE metrika používateľov v trénovacích a testovacích dátach obsahového odporúčania	27
Obrázok 7 MAE metrika používateľov v trénovacích a testovacích dátach obsahového odporúčania	27
Obrázok 8 RMSE metrika používateľov pri hybridnom odporúčaním	28
Obrázok 9 MAE metrika používateľov pri hybridnom odporúčaním	28

Zoznam skratiek a značiek

RMSE – Root Mean Squared Error

MAE – Mean Absolute Error

AUC - Area Under the Curve

CSS - Cascading Style Sheets

HTML - HyperText Markup Language

ReLU – Rectified Linear Unit

Úvod

Odporúčacie systémy sú neoddeliteľnou súčasťou našej každodennej interakcie s technológiou. S narastajúcim množstvom dostupného obsahu sa stávajú kľúčovým nástrojom na efektívne filtrovanie a zobrazovanie obsahu podľa osobných preferencií a záujmov. Tieto systémy zohrávajú kľúčovú úlohu v rôznych odvetviach, ako sú e-commerce platformy, sociálne médiá, streamovacie služby a mnoho ďalších.

Naša práca sa sústreďuje na posúdenie a implementáciu moderných prístupov k odporúčacím systémom, ktoré využívajú pokročilé technológie umelej inteligencie. Medzi tieto moderné prístupy patrí napríklad hlboké učenie, ktoré umožňuje modelom naučiť sa komplexné vzťahy a vzory v dátach. Teoretický základ tejto práce sa zaoberá dôležitými princípmi, technikami a metódami, ktoré sú základom pre pochopenie fungovania odporúčacích systémov. Hlavným cieľom tejto práce je navrhnúť a implementovať vlastný odporúčací systém, ktorý bude schopný generovať personalizované odporúčania pomocou modelu umelej inteligencie. Tento systém by mal byť robustný a efektívny a mal by byť schopný adaptovať sa na rôzne typy obsahu a preferencií používateľov.

V rámci vyhodnocovania navrhnutého modelu používame rôzne metriky, dostupné na evaluáciu umelej inteligencie, na posúdenie jeho výkonnosti a účinnosti. Na základe dosiahnutých výsledkov analyzujeme efektivitu systému a identifikujeme prípadné oblasti vylepšenia.

1 Odporúčacie systémy

Odporúčacie systémy sú softvérové nástroje a techniky poskytujúce užitočné odporúčania pre používateľa. Odporúčania sa týkajú všetkých oblastí, kde je viac možností ako je možné realisticky preskúmať, ako napríklad čo kúpiť, akú hudbu počúvať alebo aké filmy a seriály si ako ďalšie pozrieť. Položka (anglicky item) je všeobecný termín používaný na označenie toho, čo systém odporúča používateľom. Odporúčacie systémy sa zvyčajne zameriavajú na konkrétny typ položiek (napríklad spomínané filmy, hudba alebo správy) a jeho návrh, užívateľské rozhranie a odporúčacie techniky používané na generovanie odporúčaní sú prispôsobené tak, aby poskytovali čo najužitejšie a najefektívnejšie odporúčania pre daný typ položiek (1).

V každodennom živote sa spoliehame na odporúčania od ľudí v našom okolí, či už to sú recenzie v časopisoch alebo novinách, dotazníky alebo obyčajné konverzácie. Odporúčacie systémy pomáhajú tomuto prirodzenému sociálnemu procesu a dopĺňajú ho. V typickom odporúčacom systéme ľudia poskytujú odporúčania ako vstupy, ktoré systém potom agreguje a smeruje k vhodným príjemcom. V niektorých prípadoch je hlavnou transformáciou agregácia; v iných prípadoch hodnota systému spočíva v jeho schopnosti vytvárať dobré zhody medzi odporúčateľmi a tými, ktorí hľadajú odporúčania (2).

Najjednoduchším príkladom personalizovaných odporúčaní sú usporiadané zoznamy položiek. Pri tomto usporiadaní sa odporúčacie systémy snažia predpovedať, aké sú najvhodnejšie produkty alebo služby na základe preferencií a obmedzení používateľa. Na to aby plnili túto funkciu, zbierajú od používateľov dáta, ktoré predstavujú ich preferencie. Tie sú buď explicitne vyjadrené, napríklad ako hodnotenia produktov alebo označenia položky že sa vám páči, alebo sú odvodené z interpretačných používateľských akcií. Napríklad odporúčací systém môže považovať navštívenie konkrétnej stránky produktu za implicitný znak preferencie pre položky zobrazené na tejto stránke (1).

Oproti klasickému informačnému systému, kde sa výsledky berú len ako jednoduché zhody s vyhľadávaným výrazom, odporúčania, ktoré poskytuje odporúčací systém sú považované za cennejšie. Tieto odporúčania by mali byť vnímané ako možnosti, ktoré stoja za zváženie (3).

Medzi kľúčové princípy odporúčacích systémov patrí kontextualizácia odporúčaní. To znamená, že odporúčania sú vyhodnocované v kontexte aktuálnej situácie alebo potreby používateľa, čo zvyšuje presnosť odporúčaní. (4)

1.1 História odporúčacích systémov

Odporúčacie systémy sa začali rozvíjať v 90. rokoch 20. storočia, keď sa začali objavovať prvé experimenty s technikami na odporúčanie produktov a obsahu. Jedným z prvých významných prác na túto tému bol článok od Elaine Rich z roku 1979, ktorý sa zaoberal odporúčaním kníh založeným na stereotypoch používateľov (5).

V 90. rokoch vznikli prvé komerčné odporúčacie systémy, ako napríklad systém Tapestry od Xeroxa PARC (6), ktorý bol zameraný na odporúčanie webových stránok na základe hodnotení používateľov. Ďalším dôležitým krokom bola v roku 1994 publikácia práce GroupLens (7), kde boli popísané techniky na kolektívne filtrovanie a odporúčanie obsahu na základe preferencií skupiny používateľov.

1.2 Využitie odporúčacích systémov

Moderné odporúčacie systémy sú nástrojom mnohých online platforiem, ktoré sa snažia zlepšiť užívateľskú skúsenosť a zvýšiť angažovanosť používateľov. Odporúčacie systémy umožňujú používateľom objavovať relevantný obsah alebo produkty na základe ich preferencií a správania. (1)

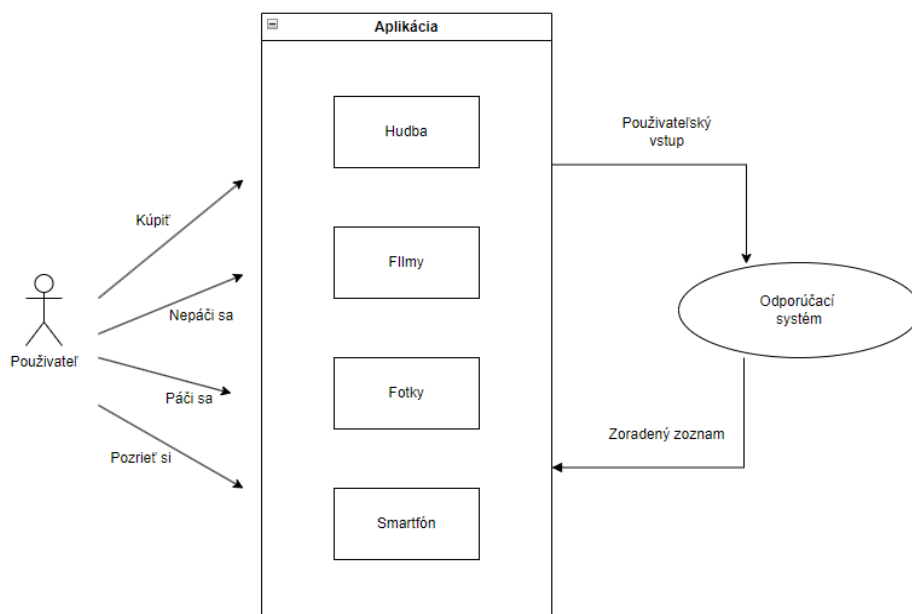
Najzakladanejšie aplikácie odporúčacích systémov v online prostredí definujeme podľa (1) takto:

- Zábava - odporúčania pre filmy, hudbu a hry.
- Obsah - personalizované noviny, odporúčania dokumentov, prác a webových stránok, e-learningové aplikácie a filtre pre e-maily.
- E-obchod - odporúčania pre zákazníkov e-shopov, s produktami všetkých odvetví (napríklad eBay¹ alebo Etsy² (8))

¹ <https://www.ebay.com>

² <https://www.etsy.com>

- Služby - odporúčania cestovných služieb, odporúčanie odborníkov na konzultácie, odporúčanie domov na prenájom (napríklad Airbnb (9)) alebo služby na hľadanie partnera.



Obrázok 1 Logický proces fungovania odporúčacieho systému (45)

1.2.1 Netflix³

Netflix je známy svojím využitím odporúčacích systémov na personalizáciu obsahu pre svojich používateľov. Spoločnosť investuje do výskumu a vývoja nových algoritmov, ktoré analyzujú sledovacie preferencie, hodnotenia a spätnú väzbu používateľov s cieľom generovať odporúčania, ktoré najlepšie zodpovedajú individuálnym záujmom a preferenciám (10) (11).

Netflix pravidelne uskutočňuje experimenty s odporúčajacími systémami, aby zlepšil presnosť a efektivitu svojich algoritmov. Jedným z takýchto experimentov bol Netflix Prize, ktorý začal v roku 2006 a ponúkol odmenu milión dolárov pre tím, ktorý vytvoril algoritmus s najlepšou presnosťou v odporúčaní filmov a seriálov (12).

Netflix Prize viedla k tomu, že mnohí výskumníci začali intenzívne pracovať na vývoji nových techník a metód strojového učenia a výsledky ukázali, že ich prístupy a techniky viedli k významnému zlepšeniu presnosti odporúčaní. Tento výskum poskytol cenné

³ <https://www.netflix.com/sk/>

poznatky nielen pre Netflix, ale aj pre celú komunitu strojového učenia a umelého inteligencie, a posilnil význam odporúčacích systémov v digitálnom priemysle (13).

1.2.2 Amazon⁴

Amazon.com rozsiahlo využíva algoritmy odporúčania na personalizáciu svojej webovej stránky podľa záujmov každého zákazníka. Existujúce algoritmy odporúčania sa nedokážu prispôbiť na stovky miliónov zákazníkov a produktov, preto Amazon vyvinul vlastný algoritmus, ktorý nazvali „item-to-item collaborative filtering“. Ten v reálnom čase produkuje kvalitné odporúčania aj s veľkým množstvom dát (14).

Item-to-item collaborative filtering porovnáva každú zakúpenú a ohodnotenú položku používateľa s podobnými položkami a potom tieto podobné položky kombinuje do zoznamu odporúčaní. Na určenie, ktoré položky sa budú porovnávať, algoritmus vytvára tabuľku podobných položiek tým, že hľadá položky, ktoré zákazníci zvyknú kupovať spolu (14).

Výskum odporúčacích systémov na platforme Amazon je dôležitou súčasťou stratégie spoločnosti zameranej na zlepšenie užívateľskej skúsenosti a zvýšenie konverzie. Úspech týchto systémov je dôkazom ich významu v online obchodovaní.

1.3 Ciele odporúčacích systémov

Pre dosiahnutie primárneho cieľa odporúčacích systémov a to zvýšenia príjmov platforiem, ktoré ich využívajú, sú operačné a technické ciele odporúčacích systémov podľa (15) nasledovné:

- Relevantnosť: Základným cieľom odporúčacieho systému je odporúčať položky, ktoré sú relevantné pre aktuálneho používateľa. Hoci je relevantnosť najdôležitejším operačným cieľom odporúčacieho systému, sama osebe nie je postačujúca. Preto (15) popisuje niekoľko sekundárnych cieľov, ktoré majú vplyv na úspešnosť celého systému.
- Novota: Pre výkonnosť odporúčacieho systému je dôležité aby odporúčané položky boli pre používateľa nové. Nebolo by užitočné odporúčať tie isté položky aj po ich prezretí.

⁴ <https://www.amazon.com>

- “Prijemné prekvapenie” (anglicky serendipity): Ďalšia vec čo môže negatívne ovplyvňovať účinnosť odporúčacieho systému je odporúčanie iba očakávaných odporúčaní, čím sa znižuje diverzita. Narozdiel od novoty, kde sú položky iba nové, položky v takomto systéme sú pre používateľa prekvapivé.
- Zvyšovanie rozmanitosti odporúčaní: Odporúčacie systémy zvyčajne navrhujú zoznam najlepších k položiek. Keď sú všetky tieto odporúčané položky veľmi podobné, zvyšuje sa riziko, že používateľovi sa nepáči žiadna z týchto položiek. Rozmanitosť zabezpečuje, aby tento prípad nenastal a používateľ sa nenudil opakovaným odporúčaním podobných položiek.

1.4 Typy odporúčacích systémov

Odporúčacie systémy môžeme rozoznávať a pomenovávať podľa klasického rozdelenia (16) (1), kde tieto odporúčacie stratégie delíme na 5 kategórii.

1.4.1 Kolaboratívne

Kolaboratívne filtrovanie sa považuje za najpopulárnejšiu a najširšie používanú techniku v odporúčacích systémoch (1). Táto technika je založená na princípe zdieľania preferencií medzi používateľmi a využíva ich podobnosť pri odporúčaní obsahu. Existujú dva hlavné prístupy ku kolaboratívne filtrovaniu: user-based a item-based (17).

User-based filtrovanie je založené na podobnosti medzi používateľmi. Tento postup vytvára profil pre každého používateľa na základe jeho interakcií s obsahom a odporúča sa obsah, ktorý sa páčil iným používateľom s podobnými profilmi (15). Item-based prístup sa zameriava na podobnosť medzi jednotlivými položkami. Pri tejto metóde sa analyzujú interakcie používateľov s jednotlivými položkami a odporúčajú sa položky, ktoré majú podobné interakcie (18).

Na implementáciu kolaboratívneho filtrovania sa používajú rôzne algoritmy a techniky. Medzi najznámejšie patrí algoritmus tzv. „najbližší sused“ (anglicky nearest neighbour), ktorý identifikuje podobných používateľov alebo položky, a metódy založené na faktoroch, ktoré identifikujú skryté vzory v dátach. Tieto techniky sa stále vyvíjajú a zdokonaľujú, aby poskytovali presnejšie a efektívnejšie odporúčania (17).

Silnou stránkou kolaboratívneho filtrovania je schopnosť generovať personalizované odporúčania v situáciách, kde obsah položiek nie je jednoducho opísateľný pomocou štruktúrovaných metadát (16).

Jednou z najväčších nevýhod je problém tzv. „studeného štartu“ (anglicky cold start), kedy noví používatelia alebo nové položky nemajú dostatočnú históriu interakcií na generovanie presných odporúčaní (14). Tento problém môže viesť k zníženej presnosti odporúčaní pre nových používateľov, čo je problém pri nasadení odporúčacích systémov v nových prostrediach. Ďalšou nevýhodou môže byť potreba veľkého množstva dát na efektívne tréningovanie a správne fungovanie kolaboratívneho filtrovania, čo môže byť náročné pre menšie platformy s obmedzeným množstvom údajov.

1.4.2 Obsahové

Odporúčacie systémy s obsahovým filtrom (anglicky content-based) sú založené na analýze vlastností samotných položiek a preferenciách používateľa. Tieto systémy využívajú obsahové charakteristiky položiek, ako sú napríklad kľúčové slová, kategórie alebo popisy, na odporúčanie položiek s podobným obsahom (19).

Využívajú rôzne metódy, ktoré umožňujú systémom porozumieť obsahu položiek a identifikovať ich relevanciu voči preferenciám používateľa, ako napríklad spracovanie textu, extrakcia príznakov a tvorba modelov.

Content-based systémy majú pár výhod ale aj nevýhod oproti kolaboratívnym. Medzi základné výhody podľa (20) patria nasledovné.

- Uvádzajú popis položiek v podobe vlastností, čo nie je možné v kolaboratívnych systémoch.
- Nie je prítomný problém studeného štartu pri nových položkách, keďže nie je potrebná história interakcií medzi položkou a používateľmi.
- Nie sú zložité na používanie a na vytváranie.

Stále však majú problém studeného štartu pri nových používateľoch a zvyknú mať problémy s odporúčaním príliš podobných položiek (20), čím sa narušujú základné operačné ciele novota a diverzita odporúčaných položiek.

1.4.3 Demografické

Demografické filtrovanie v odporúčacích systémoch je založené na analýze demografických údajov používateľov, ako sú napríklad vek, pohlavie, miesto pobytu, alebo sociálno-ekonomický status. Tento prístup umožňuje odporúčať obsah alebo produkty, ktoré sú relevantné pre konkrétne demografické skupiny (3).

To znamená, že pre rôzne demografické kategórie by mali byť generované odlišné odporúčania. Mnoho webových stránok používa jednoduché personalizačné techniky založené na demografických údajoch. Napríklad, používatelia sú presmerovaní na konkrétne webové stránky na základe ich jazyka alebo krajiny (1).

1.4.4 Community-based

Systémy založené na komunite (anglicky community-based recommender systems) sú navrhnuté na generovanie odporúčaní na základe preferencií a správania sa celej komunity používateľov. Tieto systémy využívajú agregované údaje o správaní sa používateľov a preferenciách, aby odvodili odporúčania pre jednotlivých používateľov na základe podobnosti s ostatnými členmi komunity (1).

1.4.5 Knowledge-based

Systémy založené na znalostiach (anglicky knowledge-based recommender systems) sú navrhnuté na odporúčanie obsahu alebo produktov na základe znalostí o preferenciách a potrebách používateľov. Tieto systémy sa spoliehajú na explicitné modelovanie preferencií a znalostí používateľov a následné porovnávanie týchto znalostí s charakteristikami dostupných položiek. (1)

Knowledge-based systémy majú jednu výhodu, a to, že sú mimoriadne užitočné pri odporúčaní položiek, ktoré sú zriedkavo kupované. Položky, ktoré by patrili do tejto kategórie sú: nehnuteľnosti, automobily, finančné služby alebo drahé luxusné tovary. V takých prípadoch nemusia byť dostupné dostatočné hodnotenia pre proces odporúčania pri odporúčacích systémoch ako sú obsahové alebo kolaboratívne. Ďalšie problémy pri odporúčaní takýchto typov položiek je, že môžu mať rôzne typy podrobných možností a je ťažké získať dostatočný počet hodnotení pre konkrétnu inštanciu danej položky (15). A presne v takýchto prípadoch knowledge-based systémy vynikajú.

Knowledge-based systémy zvyknú fungovať lepšie ako ostatné na začiatku ich nasadenia, ale ak nie sú dopĺňané inými komponentami schopnými sa učiť, môžu byť prekonané metódami, ktoré dokážu využívať históriu interakcii používateľa ako napríklad kolaboratívne filtrovanie (1).

1.4.6 Hybridné

Hybridné odporúčacie systémy sú založené na kombinácií doteraz spomenutých metód odporúčania. Kombináciou viacerých metód sa odporúčací systém snaží vyhnúť nedostatkom a nevýhodám samostatných systémov, čím zvyšuje výkon a presnosť odporúčaní (1) (16).

Pri návrhu hybridných odporúčacích systémov je dôležité zvážiť rôzne faktory, ako je napríklad komplexnosť používateľských preferencií, dostupnosť dát, výpočtová náročnosť a pod. Tieto faktory môžu ovplyvniť výber a implementáciu konkrétnych techník a algoritmov v hybridnom systéme (21).

Najčastejšie sa kolaboratívne filtrovanie kombinuje s ďalšou technikou v pokuse o vyhnutie sa problému chladného štartu. V práci (16) uvádzajú niektoré kombinácie metód v hybridných systémoch:

1. Váňované (anglicky weighted): Skóre niekoľkých techník odporúčania sú kombinované spolu na vytvorenie jedného odporúčania.
2. Prepínanie (anglicky switching): Systém prepína medzi technikami odporúčania v závislosti na aktuálnej situácii.
3. Zmiešané: Odporúčania z niekoľkých rôznych odporúčacích systémov sú prezentované súčasne.
4. Kombinácia vstupov: Vstupy z rôznych zdrojov odporúčacích údajov sú spolu zahrnuté do jedného algoritmu odporúčania.
5. Kaskáda: Jeden odporúčací systém upravuje odporúčania z iného.
6. Rozšírenie vstupov: Výstup z jednej metódy sa používa ako vstup pre inú.
7. Meta-úroveň: Model naučený jedným odporúčacím systémom sa používa ako vstup pre iný.

1.5 Problémy odporúčacích systémov

Cold-start je problém, ktorý sa vyskytuje v prípade nových používateľov alebo nových položiek, pre ktoré neexistuje dostatočne veľa historických dát na generovanie presných odporúčaní (22). Tento problém môže mať za následok nepresné alebo nevhodné odporúčania pre nových používateľov alebo položky.

Diversity problem je problém, ktorý nastáva, keď odporúčací systém odporúča príliš perfektné výsledky. Používateľ pravdepodobnejšie nájde položku, ktorú hľadá v zozname odporúčaní s istým stupňom diverzity (1). Diverzita je zvlášť dôležitá na začiatku vyhľadávania, keď používateľ ešte len zisťuje, aké sú všetky možnosti a v prípade, že by chýbala by mohol byť používateľ nespokojný.

Škálovateľnosť: Táto schopnosť hodnotí efektívne fungovanie systému pri náraste objemu informácií. Odporúčací systém by mal byť schopný nezmenene odporúčať položky používateľom, aj keď sa zvýši počet používateľov alebo položiek. Dosiahnutie tohto cieľa si vyžaduje väčšie množstvo výpočtov, čo môže byť nákladné (23).

Problém prílišnej špecializácie: Odporúčané položky používateľom sú založené výhradne na tých, ktoré už máme zaznamenané alebo definované v používateľskom profile, bez skúmania nových položiek a ďalších dostupných možností (24).

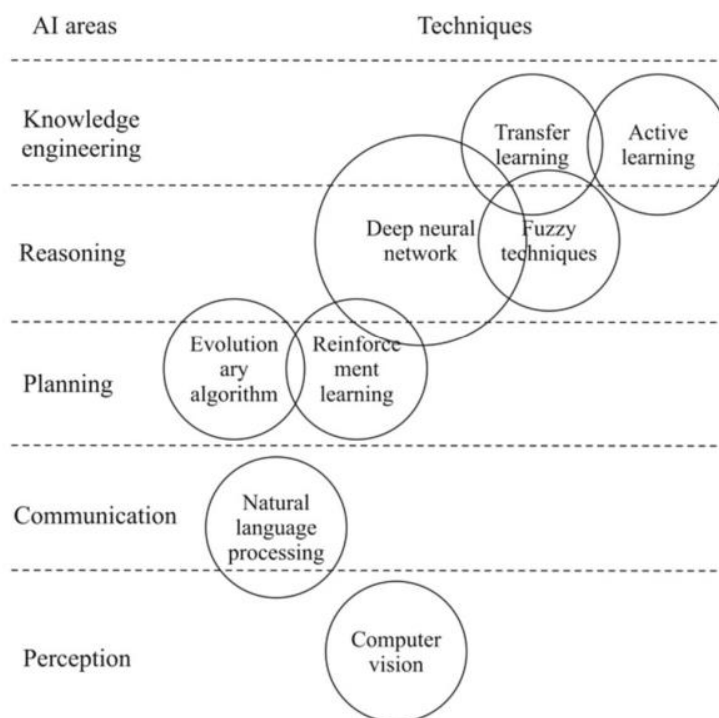
Súkromie: Odporúčacie systémy zhromažďujú veľké množstvo osobných údajov o používateľoch, čo môže viesť k riziku zneužitia alebo úniku dát. Zabezpečenie súkromia a dôvernosti osobných údajov sa stáva kľúčovým faktorom pri návrhu odporúčacích systémov. Ak chceme, aby odporúčacie systémy odporúčali položky zodpovedajúce záujmom používateľov, musíme poznať niektoré informácie o používateľských dátach. Používatelia musia vedieť, aké informácie sú potrebné na odporúčanie položiek a ako sa uplatňujú. (25).

Zhoršovacie útoky (anglicky shilling attacks): Tento jav sa vyskytuje, ak sa nečestný používateľ alebo konkurent dostane do systému a začne poskytovať nepravdivé hodnotenia určitým položkám, či už s cieľom zvýšiť ich popularitu, alebo naopak, ju znížiť (26).

„Sivá ovečka“ (anglicky Grey sheep): Vyskytuje sa v systémoch kolaboratívneho filtrovania, kde názory používateľa sa nezhodujú s žiadnou skupinou a preto nie je schopný získať výhody odporúčaní (27).

1.6 Umeľá inteligencia v odporúčacích systémoch

V tejto podkapitole si predstavíme niekoľko dôležitých metód implementácie umelej inteligencie v odporúčacích systémoch. Tieto metódy boli predstavené a opísané v práci (28).



Obrázok 2 Oblasti a techniky umelej inteligencie (28)

1.6.1 Hlboké učenie

Hlboké učenie (anglicky deep learning) je technika umelej inteligencie, ktorá umožňuje systémom lepšie pochopiť zložité vzory v dátach a zlepšiť presnosť ich odporúčaní. Tu sú niektoré kľúčové techniky hlbokého učenia v odporúčacích systémoch, ktoré popisuje práca (29):

- Viacvrstvový perceptrón predstavuje typ doprednej neurónovej siete, ktorá obsahuje skryté vrstvy umiestnené medzi vstupnými a výstupnými vrstvami. V tejto architektúre môže perceptrón využiť rôzne aktivačné funkcie a nemusí nevyhnutne slúžiť iba ako binárny klasifikátor.
- Autoenkóder je neupravený model, ktorý sa pokúša rekonštruovať svoje vstupné dáta vo výstupnej vrstve. Všeobecne sa ako významná reprezentácia vstupných dát používa tzv. bottleneck vrstva, umiestnená uprostred architektúry.

- Konvolučná neurónová sieť je špeciálny typ doprednej neurónovej siete, ktorá obsahuje konvolučné vrstvy a operácie zhukovania. Dokáže zachytiť globálne aj lokálne vlastnosti dát, čo výrazne prispieva k zlepšeniu efektivity a presnosti. Táto architektúra sa osvedčuje najmä pri spracovaní dát s mriežkovou štruktúrou.
- Rekurentná neurónová sieť je vhodná na modelovanie sekvenčných dát. Na rozdiel od doprednej neurónovej siete v rekurentných neurónových sieťach sa nachádzajú slučky a pamäte, ktoré udržiavajú informácie o predchádzajúcich výpočtoch.
- Deep Reinforcement Learning (slovensky hlboké zosilňovacie učenie) operuje na princípe experimentovania a učenia sa z vlastných chýb. Tento prístup zahŕňa niekoľko kľúčových komponentov, medzi ktoré patria agenti, prostredie, stavy, akcie a odmeny. Je kombináciou hlbokých neurónových sietí s technikami zosilňovacieho učenia a umožňuje extrahovať informácie z rozsiahlych dát a vytvoriť efektívne reprezentácie bez potreby manuálneho navrhovania vlastností alebo heuristik na určitú oblasť.

1.6.2 Transferované učenie

Základná technika transferovaného učenia (anglicky transfer learning) je prenášanie poznatkov z oblasti s dostatočným množstvom údajov do oblasti s obmedzenými údajmi. Jeho cieľom je využiť poznatky z jednej alebo viacerých zdrojových domén na pomoc pri riešení úloh v cieľovej doméne (28). Transferované učenie rozdeľujeme do troch hlavných kategórií (30):

- Induktívny transfer learning sa uplatňuje v prípade, keď sa cieľová úloha odlišuje od zdrojovej.
- Transduktívny transfer learning sa využíva v prípade, keď sú zdrojová a cieľová úloha rovnaké, ale domény sú odlišné.
- Unsupervised transfer learning (slovensky učenie bez učiteľa) sa používa v prípade, keď sú úlohy v oboch doménach unsupervised.

1.6.3 Aktívne učenie

Aktívne učenie je selektívny výber trénovacích údajov, ktorý umožňuje strojovému učeniu dosahovať lepšie výsledky s menším množstvom informácií. Systémy s aktívnou stratégiou učenia požadujú od používateľov označenie alebo popísanie neoznačených položiek (31). Vzhľadom na nákladnosť a časovú náročnosť označovania a niekedy aj jeho nemožnosť sa aktívne učenie uplatňuje v mnohých oblastiach umelej inteligencie ako je napríklad rozpoznávanie reči, informačné vyhľadávanie a výpočtová biológia a je obzvlášť vhodné pre online systémy (28).

1.6.4 Reinforcement learning

Reinforcement learning sa zakladá na princípe trial-and-error, pri ktorom agent postupne objavuje, ktoré akcie vedú k lepším výsledkom s cieľom maximalizovať odmenu. Agent sa učí na základe svojich skúseností a adaptuje svoje rozhodovanie na základe spätnej väzby z prostredia (32). V odporúčacích systémoch sa agent snaží zvoliť najlepšie odporúčania pre jednotlivých používateľov na základe jeho interakcií.

1.6.5 Evolučné algoritmy

Evolučné algoritmy tvoria triedu algoritmov vyhľadávania založených na populácii inšpirovaných prírodou pre globálnu optimalizáciu. Začína sa s počiatočnou populáciou, známou ako rodičovská populácia, ktorá predstavuje súbor kandidátnych riešení problému, v tomto prípade odporúčanie položiek, ktorý treba vyriešiť. Nové riešenia, nazývané potomstvo, sa generujú aplikovaním genetických operátorov, ako sú kríženie a mutácia, na rodičovské jedince. Tento proces pokračuje, kým nie sú splnené určité ukončovacie podmienky (28).

1.6.6 Spracovanie prirodzeného jazyka

Spracovanie prirodzeného jazyka sa zaoberá schopnosťou počítačov analyzovať, porozumieť a generovať ľudský jazyk. Vďaka pokroku v algoritmoch strojového učenia, ako sú rekurentné neurónové siete a transformátory, sa spracovanie prirodzeného jazyka stáva stále účinnejším nástrojom na porozumenie textu, extrakciu informácií, preklad jazyka a dokonca aj generovanie textu. V kontexte odporúčacích systémoch, môžu byť využité na

analýzu obsahu, vyhľadávanie a kategorizáciu informácií, čo prispieva k lepšiemu porozumeniu potrieb a preferencií používateľov (33).

1.6.7 Fuzzy techniky

Fuzzy techniky môžu byť použité na modelovanie reálnych konceptov, ktoré nemôžu byť presne reprezentované; preto sú široko používané v oblasti umelej inteligencie. Vlastnosti položiek a používateľské správanie v reálnych odporúčacích systémoch sú zvyčajne subjektívne, neúplné a nejasné. Fuzzy množiny a fuzzy relácie preto ponúkajú efektívny spôsob riešenia problémov s neistotou informácií a môžu byť tiež použité v odporúčacích systémoch (34) (28).

1.6.8 Počítačové videnie

Počítačové videnie predstavuje oblasť umelej inteligencie, ktorá sa zaoberá schopnosťou počítačov analyzovať, porozumieť a interpretovať vizuálne informácie, či už ide o statické obrázky, videá alebo dokonca reálne scény. Táto disciplína využíva rôzne algoritmy a techniky, ako napríklad konvolučné neurónové siete, na rozpoznávanie vzorov, detekciu objektov, segmentáciu obrázkov a mnoho ďalších úloh, čo je dôležité aj v kontexte odporúčacích systémov (35).

2 Ciele práce

Hlavným cieľom tejto práce je navrhnutie a implementácia odporúčacieho systému s využitím umelej inteligencie a vytvorenie webovej aplikácie, ktorá poskytne užívateľom možnosť otestovať odporúčania vygenerované odporúčacím systémom v reálnom prostredí. Motiváciou bolo vytvoriť odporúčací systém pre e-shop s umením, ktorý by dokázal na základe používateľských interakcií odporúčať relevantné položky, bez potreby filtrácie alebo vyhľadávania.

V predošlej kapitole bolo našim prvým krokom preskúmanie teoretických základov odporúčacích systémov. Tam sme sa dozvedeli o základných pojmoch a výzvach pri vytváraní takéhoto systému.

Následne sa hlavným cieľom je navrhnuť a implementovať odporúčací systém, ktorý je schopný odporúčať umelecké obrazy na základe analýzy ich vizuálnych charakteristík. To bude vyžadovať využitie pokročilých techník strojového učenia, najmä v oblasti počítačového videnia a extrakcie príznakov.

Na návrh a implementáciu backendu sme zvolili programovací jazyk Python s príslušnými knižnicami, podporujúce prácu s dátami a umelú inteligenciu. Na zobrazenie stránky sme použili kombináciu značkovacieho jazyka HTML, CSS a skriptovacieho programovacieho jazyka JavaScript.

Zvolili sme jednoduché webové rozhranie na znázornenie funkcionality nášho odporúčacieho systému. Na domovskej stránke nájdeme obrazy, fotky a abstraktné umenie vo forme mriežky a jednoduché intuitívne ovládanie, kde je možné označiť jednotlivé položky „srdcom“, ktoré signalizuje odporúčaciemu systému záujem o položky podobného typu.

V nasledujúcich kapitolách si bližšie znázorníme funkcionálne a nefunkcionálne požiadavky na fungovanie odporúčacieho systému a jeho komponentov. Ďalej bude nasledovať podrobné vysvetlenie implementácie odporúčacieho systému a jeho evaluácia spolu so záverom.

Celkovo cieľom tejto práce je poskytnúť prínosný príspevok k oblasti odporúčacích systémov a umelej inteligencie, ako aj demonštrovať praktické využitie týchto technológií vo vytvorení odporúčacieho systému pre umelecké obrazy.

3 Návrh systému

V tejto časti sa zaoberáme prístupom a metodológiou, ktorú sme využili na implementovanie nášho odporúčacieho systému.

3.1 Špecifikácia požiadaviek

Na začiatok je dôležité definovať funkcionálne a nefunkcionálne požiadavky na našu aplikáciu pred samotným návrhom a implementáciou. Okrem základnej funkcionality odporúčania položiek podľa interakcií používateľa medzi požiadavky patria:

- Možnosť vymazať svoje dáta
- Možnosť pozrieť si položky, s ktorými sme interagovali
- Možnosť vybrať si akou odporúčacou metódou je odporúčanie vykonávané
- Intuitívnosť
- Responzivnosť
- Jednoduchosť používania

Samozrejme, k požiadavkám môžeme pripojiť aj poznatky, ktoré sme sa naučili v podkapitole 1.3 Ciele odporúčacích systémov, kde sme zistili dôležitosť jednotlivých vlastností odporúčacích systémov ako novota, rozmanitosť a relevantnosť.

3.2 Klasifikácia položiek

Rozhodli sme sa, že odporúčací systém, ktorý vyvineme bude vytvárať zoradené zoznamy odporúčaní na základe vizuálnych artefaktov, čo znamená na základe vizuálnej podobnosti. To dosiahneme tak, že budeme jednotlivé položky umenia zaraďovať pomocou obrázkovej klasifikácie do vybraných kategórií. Takto dosiahneme istú abstrakciu, ktorú môžeme ovplyvňovať podľa počtu kategórií, do ktorých obrázky budeme zaraďovať.

V našom prípade použijeme 4 kategórie, ktoré zahrňujú, čo najväčšie spektrum obrazov. Tieto kategórie sú nasledovné, či je obraz minimalistický alebo je komplexný, či je obraz abstraktný alebo je realistický, či je obraz melancholický alebo veselý a nakoniec či je obraz moderný alebo klasicistický. Kombináciou týchto kategórií by sme mali zachytiť relatívne veľkú vzorku obrazov a zároveň pomáhame odporúčaciemu systému s diverzitou a rýchlosťou generovania odporúčaní.

3.3 Umelá inteligencia

Na spracovanie dát sme využili niekoľko knižníc v Pythone. V projekte je potrebné dáta z jednotlivých modelov zjednotiť do tabuliek vhodných pre spracovanie umelou inteligenciou. Pri správnom tréovaní modelov je kľúčové mať dostatočný objem dát. Preto sme sa snažili získať, čo najväčšiu databázu dát, ktoré sme následne spracovali a použili na tréovanie našich modelov. Normalizácia dát je tiež veľmi dôležitá, pretože zabezpečuje, že dáta majú rovnakú škálu a sú správne porovnateľné. To zlepšuje výkonnosť modelov a znižuje riziko chýb pri odporúčaní.

Celkovo sme venovali veľkú pozornosť implementácii umelej inteligencie do nášho projektu. Dôkladné spracovanie a normalizácia dát sú kľúčové pre dosiahnutie úspešných výsledkov

3.4 Vytvorenie vzorových používateľov

Vytvorenie vzorových používateľov bolo pre nás dôležité ako stratégia na minimalizáciu problému studeného štartu. Tieto prototypy používateľov sme navrhli na základe stereotypov potenciálnych zákazníkov e-shopu. Tieto prototypy zahŕňajú rôzne typy používateľov s rôznymi preferenciami..

Vytvorenie týchto vzorových používateľov nám umožňuje získať prieskumné dáta o tom, ako sa jednotliví typy používateľov správajú a čo preferujú pri prezeraní. Tieto informácie potom využívame na vylepšenie našich algoritmov odporúčania. Konkrétne používame tieto prototypy pre potreby kolaboratívneho filtrovania a hybridného odporúčania, ktoré kombinujú kolaboratívne metódy odporúčania s inými technikami, ako napríklad obsahové.

3.5 Odporúčanie položiek

Odporúčanie položiek prebieha v našom projekte jednoducho, používateľ postupne prezerá položky a označuje tie, ktoré sa mu páčia, to si web aplikácia ukladá do pamäte. Pri stlačení tlačidla odporučiť, vezme a pošle tieto interakcie nášmu modelu, ktorý vytrénuje model podľa aktuálneho výberu obľúbených obrázkov a vytvorí zoradený personalizovaný zoznam odporúčaní. Tieto odporúčania sú potom na webe zobrazené a proces sa môže zopakovať aj s inými metódami odporúčania.

4 Využitie nástroje a technológie

Tu si popíšeme využité nástroje, frameworky a programovacie jazyky a uvedieme, prečo sme si vybrali tieto technológie pre daný problém.

4.1 Python⁵

Python je všeobecný programovací jazyk, ktorý sme použili na vývoj našej webovej aplikácie, prácu s umelou inteligenciou a komunikáciou s frontendom. Vďaka širokej škále knižníc a frameworkov a jednoduchosti používania a implementovania umelej inteligencie sme si ho vybrali ako hlavný programovací jazyk nášho projektu. V nasledujúcich kapitolách si predstavíme niektoré knižnice a frameworky a opíšeme ich hlavnú funkcionality.

4.1.1 TensorFlow⁶

TensorFlow je open-source knižnica pre strojové učenie a hlboké učenie vyvinutá spoločnosťou Google. Pomocou TensorFlow sme implementovali a trénovali neurónové siete na analýzu obrázkov a generovanie personalizovaných odporúčaní pre našich používateľov.

4.1.2 Flask⁷

Flask je mikro webový framework pre Python, ktorý sme použili na vytvorenie backendovej časti našej aplikácie. S jeho pomocou sme definovali API endpointy, ktoré umožňujú komunikáciu medzi frontendom a backendom a spracovávali HTTP požiadavky a odpovede.

4.1.3 Numpy

NumPy je základnou knižnicou pre vedecké výpočty v jazyku Python. Poskytuje výkonné nástroje na prácu s viacrozmernými poliami a maticami, ako aj množstvo matematických funkcií na efektívne spracovanie a analýzu dát. NumPy je preto kľúčovým nástrojom pre prácu s dátami v oblasti strojového učenia a analýzy dát.

⁵ <https://www.python.org>

⁶ <https://www.tensorflow.org>

⁷ <https://flask.palletsprojects.com/en/3.0.x/>

4.1.4 Pandas

Pandas je výkonná knižnica pre analýzu a manipuláciu s dátami v jazyku Python. Vybrali sme si ho na prácu s tabuľkovými dátami, kde umožňuje načítanie, manipuláciu, filtrovanie a agregáciu dát, ako aj ich vizualizáciu.

4.2 HTML, CSS, JavaScript

HTML, CSS a JavaScript sú základné technológie pre vývoj webových aplikácií. Použili sme ich na vytvorenie používateľského rozhrania nášho projektu. HTML sme použili na štruktúrovanie obsahu, CSS na jeho štýlovanie a JavaScript na interaktivitu a dynamické správanie sa webovej stránky. Pomocou týchto technológií sme dokázali vytvoriť prostredie pre jednoduché testovanie a ukážku fungovania nášho odporúčacieho systému s viacerými typmi odporúčania..

4.3 Docker⁸

Docker je platforma pre kontajnerizáciu aplikácií, ktorú sme použili na zabezpečenie konzistentného a izolovaného prostredia pre vývoj, testovanie a nasadenie našej aplikácie. Kontajnerizáciou sme zabalili všetky potrebné závislosti našej aplikácie do jedného kontajnera a umožnili ich jednoducho nasadiť na rôzne platformy bez ohľadu na prostredie.

⁸ <https://www.docker.com>

5 Implementácia

V kapitole implementácia si popíšeme podrobné fungovanie jednotlivých skriptov, súborov a modelov umelej inteligencie. Takisto si popíšeme aj špecifické prvky komunikácie frontendu a backendu a logiku celej aplikácie od výberu tréningových dát modelov až po samotné personalizované odporúčanie.

5.1 realistic, modern, minimalistic, happy models

Ako prvé bolo potrebné získať položky, teda obrázky, na ktorých by mohli modely trénovať klasifikovanie do jednotlivých kategórií. Obrázky sme získali z dvoch datasetov zo stránky Kaggle⁹, ktorá sa venuje datasetom pre strojové učenie a umelú inteligenciu. Prvý dataset obsahoval obrazy klasického umenia (36). Je to zbierka malieb od 50 najvplyvnejších maliarov histórie, ktorý bol pôvodne určený na účely klasifikácie obrazov ku maliarom a obsahuje 8355 obrazov. Druhý dataset je zbierka abstraktného umenia (37), obsahuje 2782 súborov s abstraktným umením a pôvodne bol určený na vytváranie originálnych obrazov podľa tohto datasetu. Tretí dataset sme zahrnuli pre nedostatok realistických položiek (38), ten posluží ako vzor toho, ako by mali vyzeráť realistické obrazy alebo fotky. Obsahuje 4300 prirodzených fotiek krajiny a prírody.

Obrázky, ktoré sme získali, sme následne rozdelili do 8 kategórií tak ako sme navrhli, aby obsahovali dostatočné množstvo obrazov, ktoré sú pozitívne aj negatívne. Kategórie sú nasledovné: realističnosť, modernosť, minimalizmus a veselosť. V každom modeli sme následne spracovali dáta a rozdelili sme ich na tréningové, validačné a testovacie.

Použili sme normalizačnú vrstvu, ktorá vezme každý pixel a vydelí ho hodnotou 255 čím nadobudne hodnotu z intervalu $[0,1]$. Táto metóda pomáha efektívnosti a stabilnosti tréningu modelu. Po normalizovaní dát sme zadefinovali sekvenčný model, v sekvenčnom modeli sme používali nasledujúce vrstvy:

- Conv2D(): Konvolučná vrstva s 16, 32 a 64 filtrami a veľkosťou jadra 3×3 . Vybrali sme si aktivačnú funkciu ReLU (funkcia jednej reálnej premennej), ktorá sa používa na zvýšenie nelinearity výstupu.

⁹ <https://www.kaggle.com>

- `MaxPooling2D()`: Vrstva Max Pooling, ktorá redukuje veľkosť priestoru funkcií získaných z predchádzajúcej konvolučnej vrstvy.
- `Dropout(0.2)`: Vrstva Dropout, ktorá deaktivuje 20% jednotiek výstupnej vrstvy počas tréningu, aby sa predišlo pretrénovaniu.
- `Flatten()`: Vrstva na prevedenie výstupu konvolučných vrstiev na jednorozmerný vektor, aby mohol byť použitý v dense vrstvách.
- `Dense(128, activation='relu')`: Dense vrstva s 128 neurónmi a ReLU ako aktivačnou funkciou.
- `Dense(num_classes, name="outputs")`: Dense vrstva s počtom neurónov rovným počtu tried výstupného datasetu. Používa sa lineárna aktivačná funkcia pre viac tried.

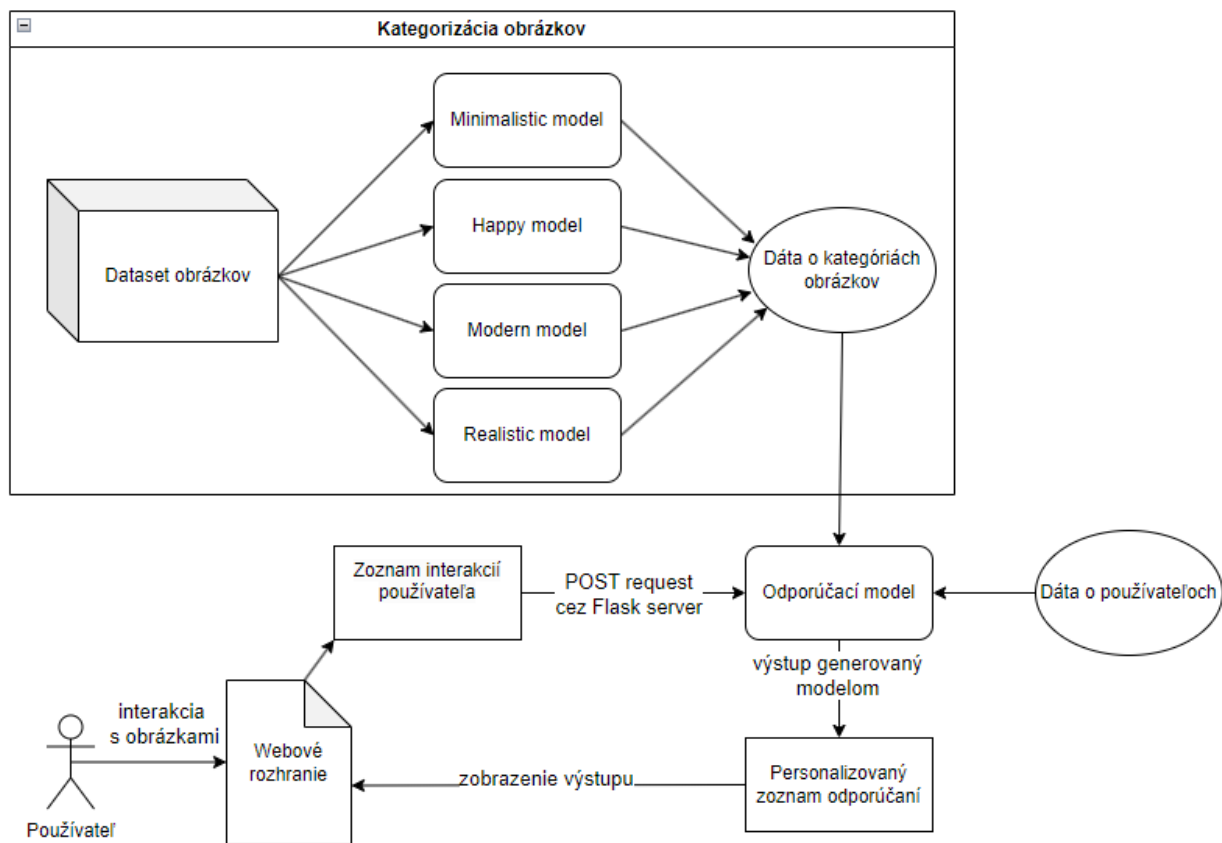
Tento model používa konvolučné vrstvy na extrakciu rôznych príznakov z obrázkov a dense vrstvy na klasifikáciu týchto príznakov do určených kategórií. Ďalším krokom je model natrénovať, kde sme použili optimizer adam. Adam (adaptívne momentové odhadovanie) je algoritmus pre optimalizačnú techniku gradientového zostupu. Táto metóda je veľmi efektívna pri práci s veľkými problémami, ktoré zahŕňajú veľa údajov alebo parametrov (39).

5.1.1 categorizeData.py

Tento jednoduchý skript načíta všetky natrénované modely a pomocou nich spraví predpovede pre každý obrázok z vybraného priečinka za pomoci Python knižníc numpy a pandas. Takúto maticu výsledkov s 4 stĺpcami a N riadkami, kde N je počet obrázkov v priečinku, potom uloží do excel súboru podľa výberu. Tento skript je potrebný spustiť pri akejkoľvek zmene priečinka s obrázkami. Takýmto spôsobom podávame informácie hlavnému odporúčaciemu modelu o tom ako obrazy vyzerajú, a teda ak sa používateľovi páči obraz A, tak sa mu bude páčiť obraz B, ktorý je zaradený v podobných kategóriách.

5.2 show model

Toto je model, ktorý generuje odporúčania na základe zoznamu obľúbených obrázkov, ktoré získal z webovej aplikácie pomocou POST requestu. Súbor show.py zahŕňa však viac a to hlavne spracovanie dát z excel súborov, spracovanie výstupu modelu ale aj následnú evaluáciu ktorej sa budeme venovať v ďalšej kapitole.



Obrázok 3 Diagram implementovaného systému

Najprv si predstavíme súbory a dáta, ktoré potrebujeme na samotné odporúčanie vo funkcií `getRecommendations()`. Zahŕňajú excel súbory a obrázky, ktoré sú položky odporúčania:

- Priečinok `base` obsahuje všetky obrázky, ktoré predstavujú položky v odporúčaní. Táto databáza kombinuje testovacie dáta zo všetkých kategorizačných modelov, čiže obsahuje obrázky z týchto datasetov (37) (38) (36). Okrem týchto vzoriek z testovacích dát boli pridané obrázky z datasetu Wiki Art (40) zo stránky Roboflow¹⁰ a obrazy vedúceho práce (41). Dokopy je počet týchto obrázkov 516 a reprezentujú rozmanitú vzorku abstraktného, moderného ale aj klasicistického umenia.

¹⁰ <https://roboflow.com>

- baseUsers.xlsx. Toto je excel súbor s informáciami o interakciách používateľov, každý stĺpec patrí jednému používateľovi a čísla zaznamenávajú index obrázkov v priečinku base.
- baseCategorized.xlsx. Tu sa ukladajú dáta a informácie do ktorých kategórií patria jednotlivé obrázky.

Po tom ako máme všetky tieto časti pripravené, spracujeme dáta zo spomínaných excel súborov pomocou Python knižníc. Pri odporúčacích systémoch často vzniká problém, kedy máme málo položiek, s ktorými používateľ interagoval vzhľadom na celkový počet položiek, preto sme sa rozhodli použiť váhy pre naše dáta, kde náš model bude brať do úvahy interakciu s položkou viac ako žiadnu interakciu. Takýmto spôsobom sa dáta vyrovnajú a zvýši sa efektivita nášho odporúčacieho systému.

Sekvenčný model sa skladá z dense vrstiev s 64, 32 a 16 neurónmi a ReLU ako aktivačnou funkciou. Posledná vrstva má jeden neurón a aktivačnú funkciu sigmoid. Táto vrstva sa používa na binárnu klasifikáciu, kde výstup je hodnota v rozsahu (0,1), čo predstavuje pravdepodobnosť prítomnosti alebo neprítomnosti daného javu, v našom prípade či sa položka páči alebo nepáči používateľovi. Po testovaní najlepšie výsledky dosahovalo tréningovanie s 50 generáciami a s 10% dát vyčlenených na validáciu.

Akou metódou sa bude odporúčať udáva premenná mode, ktorá je posielaná spolu s používateľskými interakciami v POST requeste. Pri metóde obsahového odporúčania sa jednoducho spracujú poslané dáta a vygeneruje sa zoznam odporúčaní. V prípade, že je zvolené kolaboratívne filtrovanie, nájdú sa pomocou pomocnej funkcie najbližší dvaja používatelia. Tých hľadáme pomocou porovnania vektorov kosínovou podobnosťou. Následne odporúčané položky sú kombináciou položiek nájdených najpodobnejších používateľov.

V prípade vybratia hybridného odporúčania sa nájdú najbližší používatelia, tak ako pri kolaboratívnej metóde ale so zmenou, po ich nájdení sa interakcie týchto používateľov skombinujú s interakciami aktívneho používateľa, natrénuje sa podľa nich model a vygeneruje sa zoznam odporúčaní, ktorý sa pred odoslaním pomocou odpovede naspať na webové rozhranie zoradí.

5.3 Webové rozhranie

Webové rozhranie poskytuje užívateľom pohodlný a interaktívny spôsob prístupu k odporúčaniam. Je postavené na webových technológiách, ktorý sme spomenuli v predošlej kapitole. Základná funkcionality webovej aplikácie je prezeranie si obrázkov, označovanie položiek, ktoré sa používateľovi páčia a vyberanie si metódy odporúčania.

5.3.1 Docker

Pre zabezpečenie stability a jednoduchého nasadenia sme implementovali webové rozhranie v Docker kontajneri. Tento prístup umožňuje ľahké spravovanie a aktualizácie aplikácie, zatiaľ čo zabezpečuje, že všetci používatelia majú prístup k rovnakému prostrediu a funkčnostiam.

5.3.2 Flask server

Komunikácia medzi frontendom a backendom je realizovaná pomocou POST requestov na API endpoint, ktoré sú spracované Flask serverom a tým riadi tok dát medzi užívateľským rozhraním a modelmi umelej inteligencie.

6 Vyhodnotenie

V nasledujúcej časti vyhodnotíme úspešnosť odporúčacieho systému, porovnáme typy odporúčaní, ktoré sme na stránke použili a predstavíme možné návrhy a zlepšenia do budúcnosti pre náš odporúčací systém s umelou inteligenciou.

6.1 Metodika vyhodnotenia

Pri vyhodnocovaní efektivity modelu umelej inteligencie existuje množstvo metodík a techník, ktoré umožňujú posúdenie jeho účinnosti. Dve z najbežnejších metrík používaných na hodnotenie výkonu modelov sú Root Mean Squared Error a Mean Absolute Error.

RMSE je používaná na meranie priemernej veľkosti chýb medzi predpovedanými a vstupnými hodnotami. Táto hodnota je výsledkom odmocnenia priemeru druhých mocnín týchto chýb. Čím nižšia hodnota RMSE, tým presnejší je model v predikcii hodnôt.

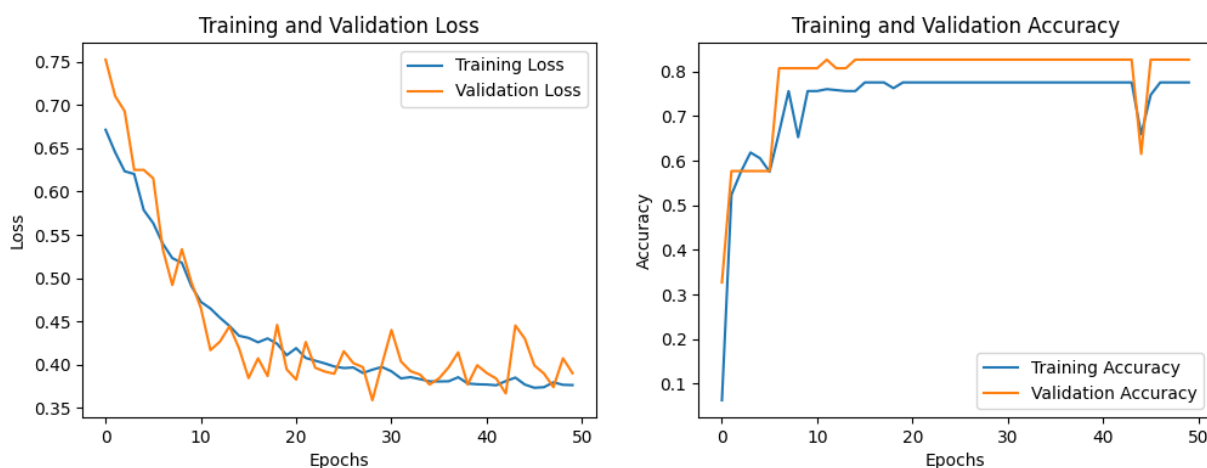
MAE je priemerná hodnota absolútnych hodnôt chýb medzi predpovedanými a pozorovanými hodnotami. MAE poskytuje priamu informáciu o veľkosti chýb, nezávisle od ich smeru.

Okrem RMSE a MAE existuje mnoho ďalších metrík, ktoré môžeme použiť na vyhodnotenie výkonnosti modelov odporúčacích systémov. Medzi ne patria napríklad Precision, Recall, F1 skóre alebo AUC.

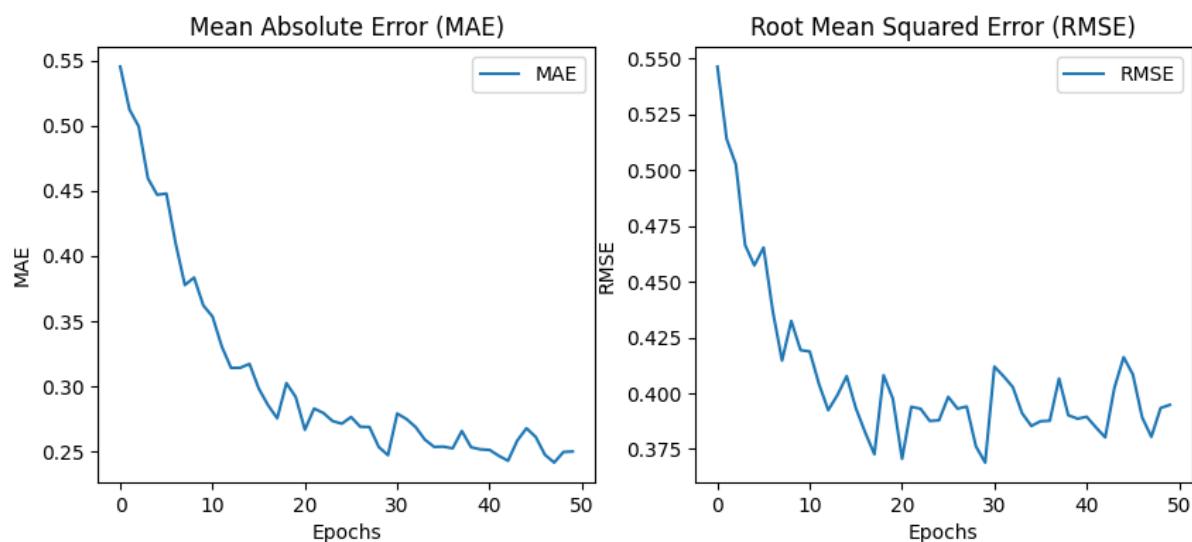
Je tiež dôležité poznamenať, že správne vyhodnotenie modelov odporúčacích systémov nie je len o jednotlivých metrikách, ale aj o ich kontextuálnom porovnaní a interpretácii. Niektoré metriky môžu byť vhodnejšie pre konkrétne scenáre alebo požiadavky, a preto je dôležité brať do úvahy celkový kontext a cieľovú oblasť použitia odporúčacieho systému.

6.2 Evaluácia modelu

Evaluáciu modelu sme vykonali v `showEval.py`, čo je upravená verzia Python skriptu `show.py`, kde je náš odporúčací model. Naším hlavným cieľom bolo dostať, čo najnižšie a najkonzistentnejšie hodnoty RMSE a MAE a porovnať hybridné a obsahové odporúčanie.



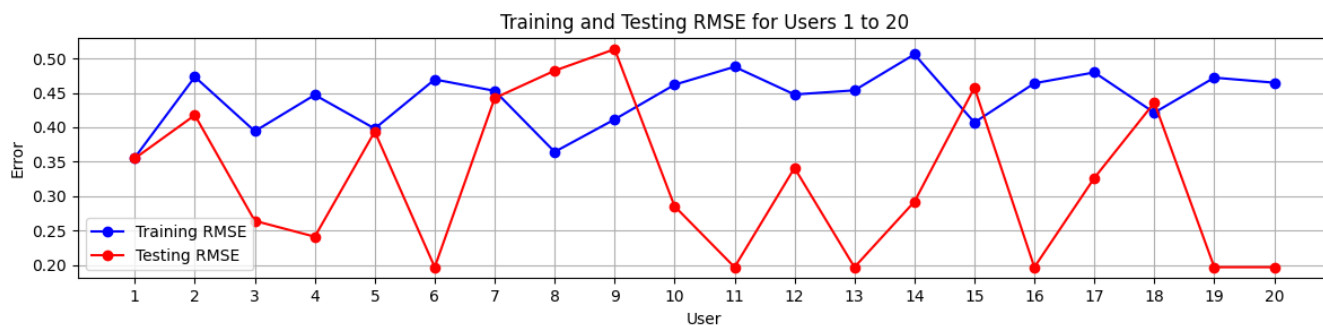
Obrázok 4 Trénovacie a validačné straty a presnosti



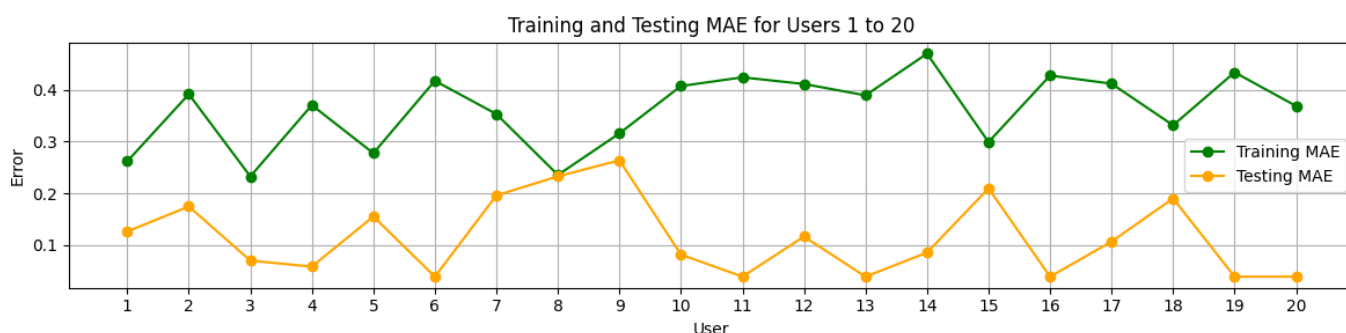
Obrázok 5 Vývoj RMSE a MAE počas tréovania

Najprv sa pozrieme na obsahové odporúčanie. Na grafoch na obrázkoch 4 a 5 môžeme vidieť postupné zlepšovanie všetkých metrík až po stagnáciu okolo generácie 50, kde sme sa rozhodli, že model je dostatočne natrénovaný.

Ako hlavnú časť evaluácie sme sa rozhodli zobrať našich 20 vzorových používateľov, ktorým sme vypočítali RMSE a MAE v trénovacích a testovacích dátach. Grafy na obrázkoch 6 a 7 ukazujú tieto hodnoty kde x os je index používateľa a y os je chyba modelu. Rozdiel chýb v RMSE je $\text{max.} - \text{min.} = 0.51 - 0.2 = 0.31$, a pre MSE to je $\text{max.} - \text{min.} = 0.48 - 0.04 = 0.44$.

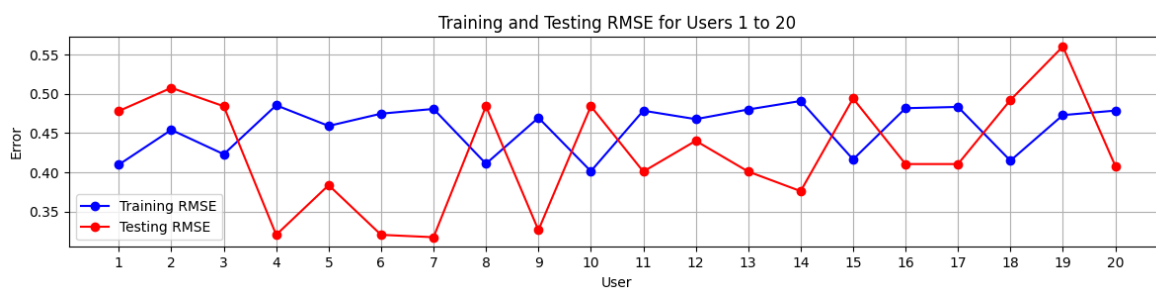


Obrázok 6 RMSE metrika používateľov v tréningových a testovacích dátach obsahového odporúčania

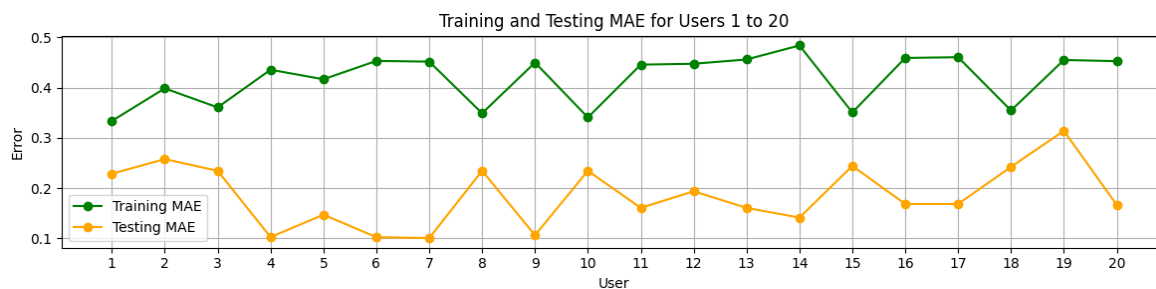


Obrázok 7 MAE metrika používateľov v tréningových a testovacích dátach obsahového odporúčania

Na porovnanie sme rovnaký postup zopakovali s hybridným odporúčaním, ktoré najprv nájde seba najpodobnejších používateľov zo systému a následne model trénuje aj s interakciami nájdených najpodobnejších používateľov ako keby boli jeho vlastné. Tak ako na predošlých obrázkoch grafy na obrázkoch 8 a 9 ukazujú hodnoty RMSE a MAE. Rozdiel chýb v tomto prípade pre RMSE je $\text{max.} - \text{min.} = 0.57 - 0.32 = 0.25$, a pre MSE to je $\text{max.} - \text{min.} = 0.49 - 0.1 = 0.39$. Ako môžeme vidieť rozdiel týchto hodnôt je nižší ako pri čisto obsahovom odporúčaní a naznačuje, že pridanie kolaboratívneho filtrovania do procesu odporúčania zvyšuje jeho presnosť predikcií. Model tak má schopnosť dobre generalizovať a správne predikovať hodnoty aj pre nové, nevidené dáta. Ďalej môže naznačovať, že dáta, na ktorých sme trénovali a testovali náš model, majú podobné charakteristiky a distribúciu, čo znamená, že model je schopný efektívne pracovať v reálnom prostredí s podobnými dátami.



Obrázok 8 RMSE metrika používateľov pri hybridnom odporúčaní



Obrázok 9 MAE metrika používateľov pri hybridnom odporúčaní

6.3 Vylepšenia

Najväčší problém pri riešení tejto práce predstavoval nedostatok dát a nedostatočné množstvo interakcií používateľov. Ako hlavné vylepšenie by bolo preto nasadenie odporúčacieho systému do reálneho prostredia, kde by mohol získavať reálne dáta od používateľov a iteratívnym vývojom sa zlepšovať. To by poskytlo aj dostatočné dáta pre zlepšenie evaluácie a vyhodnocovania nášho modelu pre odporúčanie.

Ďalšie možné vylepšenie predstavuje zavedenie viacerých kategórií pre opis dát, vďaka ktorému by obsahové filtrovanie malo viac možností prispôbiť sa jednotlivým požiadavkám každého používateľa systému. Kategórie, ktoré by mohli byť pridané sú napríklad autor obrázku, krajina pôvodu a čas vzniku. Pri zaznamenávaní interakcií by mohli pomôcť aj informácie o používateľoch ako vek alebo národnosť a branie do úvahy, kedy interakcia vznikla.

Záver

Naším cieľom bolo posúdiť moderné prístupy k odporúčacím systémom, ktoré využívajú pokročilé technológie umelej inteligencie, a zhodnotiť ich použiteľnosť v reálnom prostredí. Podarilo sa nám navrhnúť a implementovať odporúčací systém s obsahovým, kolaboratívnym a hybridným odporúčaním, za pomoci generovania odporúčaní modelom umelej inteligencie.

V rámci hodnotenia sme sa zaoberali najmä metrikami ako RMSE a MAE, ktoré nám umožnili posúdiť presnosť našich modelov. Pri porovnávaní rôznych typov odporúčaní sme zistili, že hybridné odporúčanie, kombinujúce obsahové a kolaboratívne filtrovanie, dosahuje presnejšie výsledky než čisto obsahové odporúčanie. Tento výsledok naznačuje schopnosť nášho modelu generalizovať a efektívne predikovať hodnoty aj pre nové, nevidené dáta.

Avšak, hlavným problémom, s ktorým sme sa stretli, bola obmedzená dostupnosť dát a nedostatok interakcií používateľov. Aby sme mohli náš odporúčací systém ďalej zdokonaľovať, je nevyhnutné nasadiť ho do reálneho prostredia, kde by mohol získavať reálne dáta od používateľov a iteratívnym vývojom sa zlepšovať. Taktiež je dôležité zvážiť zavedenie viacerých kategórií pre opis dát a zohľadnenie ďalších informácií o používateľoch, čo by umožnilo lepšie prispôsobenie odporúčaní individuálnym preferenciám a požiadavkám používateľov.

Celkovo sme preukázali, že využitie moderných prístupov k odporúčacím systémom, založených na umelej inteligencii, má veľký potenciál vytvoriť personalizované a presné odporúčania pre používateľov. Naša práca prináša príspevok k lepšiemu porozumeniu a využitiu týchto technológií v digitálnom svete.

Zoznam použitej literatúry

1. **Ricci, F., Rokach, L., & Shapira, B.** *Introduction to Recommender Systems Handbook. Recommender Systems Handbook*, 1–35. s.l. : Springer US, 2010. 978-0-387-85819-7.
2. **Resnick, P., & Varian, H. R.** Recommender Systems. *Communications of the ACM*. 1997, Zv. 40, 3.
3. **Burke, R.** Hybrid Web Recommender Systems. *Lecture Notes in Computer Science*, 377–408. 2007, DOI: 10.1007/978-3-540-72079-9_12.
4. **Adomavicius, G., Mobasher, B., Ricci, F., & Tuzhilin, A.** Context-Aware Recommender Systems. *AI Magazine*, 32(3), 67. 2011, doi:10.1609/aimag.v32i3.2364.
5. **Rich, E.** User Modeling via Stereotypes*. *Cognitive Science*, 3(4), 329–354. 1979, doi:10.1207/s15516709cog0304_3.
6. **Goldberg, D., Nichols, D., Oki, B. M., & Terry, D.** Using collaborative filtering to weave an information tapestry. *Communications of the ACM*, 35(12), 61–70. 1992, doi:10.1145/138859.138867.
7. **Resnick, P., Iacovou, N., Suchak, M., Bergstrom, P., & Riedl, J.** GroupLens. *Proceedings of the 1994 ACM Conference on Computer Supported Cooperative Work - CSCW '94*. 1994, doi:10.1145/192844.192905.
8. **Hall, Robert.** Personalized Recommendations at Etsy. <https://www.etsy.com>. [Online] 17. 11 2014. [Dátum: 19. 4 2024.] <https://www.etsy.com/codeascraft/personalized-recommendations-at-etsy>.
9. **The Airbnb Tech Blog.** How Airbnb uses Machine Learning to Detect Host Preferences. <https://medium.com>. [Online] 14. 4 2015. [Dátum: 19. 4 2024.] <https://medium.com/airbnb-engineering/how-airbnb-uses-machine-learning-to-detect-host-preferences-18ce07150fa3>.
10. **Xavier Amatriain, Justin Basilico.** Netflix Recommendations: Beyond the 5 stars (Part 1). <https://netflixtechblog.com>. [Online] Netflix TechBlog, 6. 4 2012. [Dátum: 19. 4 2024.] <https://netflixtechblog.com/netflix-recommendations-beyond-the-5-stars-part-1-55838468f429>.
11. —. Netflix Recommendations: Beyond the 5 stars (Part 2). <https://netflixtechblog.com>. [Online] Netflix TechBlog, 20. 6 2012. [Dátum: 19. 4 2024.] <https://netflixtechblog.com/netflix-recommendations-beyond-the-5-stars-part-2-d9b96aa399f5>.
12. **Bell, R. M., & Koren, Y.** Lessons from the Netflix prize challenge. *ACM SIGKDD Explorations Newsletter*, 9(2), 75. 2007, doi:10.1145/1345448.1345465.

13. **Bell, R. M., Koren, Y., & Volinsky, C.** All Together Now: A Perspective on the Netflix Prize. *CHANCE*, 23(1), 24–29. 2010, doi:10.1080/09332480.2010.10739787 .
14. **Linden, G., Smith, B., & York, J.** Amazon.com recommendations: item-to-item collaborative filtering. *IEEE Internet Computing*, 7(1), 76–80. 2003, doi:10.1109/mic.2003.1167344 .
15. **Aggarwal, Charu C.** *An Introduction to Recommender Systems. Recommender Systems: The Textbook*, 1–28. s.l. : Springer, 2016. 978-3-319-29657-9.
16. **Burke, R.** Hybrid Recommender Systems: Survey and Experiments. *User Modeling and User-Adapted Interaction*, 12(4), 331–370. 2002, doi:10.1023/a:1021240730564 .
17. **Desrosiers, C., & Karypis, G.** *A Comprehensive Survey of Neighborhood-based Recommendation Methods. Recommender Systems Handbook*, 107–144. s.l. : Springer, 2010. doi:10.1007/978-0-387-85820-3_4 .
18. **Sarwar, B., Karypis, G., Konstan, J., & Reidl, J.** Item-based collaborative filtering recommendation algorithms. *Proceedings of the Tenth International Conference on World Wide Web - WWW '01*. 2001, doi:10.1145/371920.372071 .
19. **Lops, P., de Gemmis, M., & Semeraro, G.** Content-based Recommender Systems: State of the Art and Trends. *Recommender Systems Handbook*, 73–105. 2010, doi:10.1007/978-0-387-85820-3_3 .
20. **Aggarwal, Charu C.** *Content-Based Recommender Systems, Recommender Systems: The Textbook*, 139-166. s.l. : Springer, 2016. 978-3-319-29657-9.
21. **Su, X., & Khoshgoftaar, T. M.** A Survey of Collaborative Filtering Techniques. *Advances in Artificial Intelligence*, 1–19. 2009, doi:10.1155/2009/421425 .
22. **Masthoff, J.** Group Recommender Systems: Combining Individual Models. *Recommender Systems Handbook*, 677–702. 2010, doi:10.1007/978-0-387-85820-3_21 .
23. **Mohamed, M. H., Khafagy, M. H., & Ibrahim, M. H.** Recommender Systems Challenges and Solutions Survey. *2019 International Conference on Innovative Trends in Computer Engineering (ITCE)*. 2019, doi:10.1109/itce.2019.8646645 .
24. **Gera, Lalita Sharma and Anju.** A Survey of Recommendation System: Research Challenges. *International Journal of Engineering Trends and Technology (IJETT)*. vol 4(5) 2013, pp.1989-1992.
25. **Mahmoud, H., Hegazy, A., & Khafagy, M. H.** An approach for big data security based on Hadoop distributed file system. *International Conference on Innovative Trends in Computer Engineering (ITCE)*. 2018, doi:10.1109/itce.2018.8316608 .

26. **Khusro, S., Ali, Z., & Ullah, I.** Recommender Systems: Issues, Challenges, and Research Opportunities. *Information Science and Applications (ICISA)*. 2016, Zv. 1179–1189., doi:10.1007/978-981-10-0557-2_112 .
27. **Shah Khusro, Zafar Ali and Irfan Ullah.** Recommender Systems: Issues, Challenges, and Research Opportunities. *Information Science and Applications (ICISA)*. 2016, pp.1179–1189.
28. **Zhang, Q., Lu, J., & Jin, Y.** Artificial intelligence in recommender systems. *Complex & Intelligent Systems*, 7(1), 439–457. 2020, doi:10.1007/s40747-020-00212-w .
29. **Zhang, S., Yao, L., Sun, A., & Tay, Y.** Deep Learning Based Recommender System. *ACM Computing Surveys*, 52(1), 1–38. 2019, doi:10.1145/3285029 .
30. **Lu J, Behbood V, Hao P, Zuo H, Xue S, Zhang G.** Transfer learning using computational intelligence: a survey. *Knowl Based Syst* 80:14–23. 2015.
31. **Aghdam HH, Gonzalez-Garcia A, van de Weijer J, López AM.** Active learning for deep detection neural networks. *Proceedings of the IEEE International Conference on Computer Vision, 2019*, pp 3672–3680. 2019.
32. **Sutton RS, Barto AG.** Reinforcement learning: An introduction. 2011, MIT Press, Cambridge.
33. **Chowdhary, K. R.** *Natural Language Processing. Fundamentals of Artificial Intelligence*. Berlin : Springer, 2020. pp 603–649, doi:10.1007/978-81-322-3972-7 .
34. **Chung, F., & Rhee, H.** Uncertain Fuzzy Clustering: Insights and Recommendations. *IEEE Computational Intelligence Magazine*, 2(1), 44–56. 2007, doi:10.1109/mci.2007.357193 .
35. **Szeliski, Richard.** *Computer Vision: Algorithms and Application*. s.l. : Springer, 2010. 9781848829343,1848829345.
36. **Icaro.** Best Artworks of All Time. [Online] 2019. [Datum: 8. 1 2024.] <https://www.kaggle.com/datasets/ikarus777/best-artworks-of-all-time/data>.
37. **Boulé, Bryan.** Abstract Art Gallery. [Online] 2022. [Datum: 8. 1 2024.] <https://www.kaggle.com/datasets/bryanb/abstract-art-gallery>.
38. **Rougetet, Arnaud.** Landscape Pictures. [Online] 2020. [Datum: 8. 1 2024.] https://www.kaggle.com/datasets/arnaud58/landscape-pictures?select=00000000_%283%29.jpg.
39. **Vishwakarma, Neha.** What is Adam Optimizer? *Analytics Vidhya*. [Online] 28. 4 2024. [Datum: 13. 5 2024.] <https://www.analyticsvidhya.com/blog/2023/09/what-is-adam-optimizer/>.
40. **Dataset, Art.** Wiki Art Computer Vision Project. [Online] 3 2022. [Datum: 22. 4 2024.]

41. **Repka, Marek.** Obrazy. [Online] 2024. [Dátum: 11. 4 2024.]
42. **Adomavicius, G., & Tuzhilin, A.** Toward the next generation of recommender systems: a survey of the state-of-the-art and possible extensions. *IEEE Transactions on Knowledge and Data Engineering*, 17(6), 734–749. 2005, doi:10.1109/tkde.2005.99 .
43. **Koren, Y.** Collaborative filtering with temporal dynamics. *Proceedings of the 15th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining - KDD '09*. 2009, doi:10.1145/1557019.1557072 .
44. **Balázs Hidasi, Alexandros Karatzoglou, Linas Baltrunas, Domonkos Tikk.** *Session-based Recommendations with Recurrent Neural Networks*. 2015.
<https://doi.org/10.48550/arXiv.1511.06939>.
45. **Ferreira, D., Silva, S., Abelha, A., & Machado, J.** Recommendation System Using Autoencoders. *Applied Sciences*, 10(16), 5510. . 2020, doi:10.3390/app10165510 .

Prílohy

Príloha A: Technická dokumentácia.....	II
----------------------------------------	----

Príloha A: Technická dokumentácia

Repozitár je verejne dostupný na GitHub¹¹. V súbore README.md je technická dokumentácia a je tam popísaný postup pre spustenie aplikácie lokálne a taktiež postup na stiahnutie kontajneizovanej aplikácie z DockerHub¹² a jej spustenie.

¹¹ <https://github.com/xbercikfilip/bcpraca>

¹² <https://hub.docker.com/r/xbercikf/bcpraca>