

Je odporúčací systém Netflix-u problém ?*

Adam Glogovský

Slovenská technická univerzita v Bratislave
Fakulta informatiky a informačných technológií
xglogovsky@stuba.sk

5. október 2024

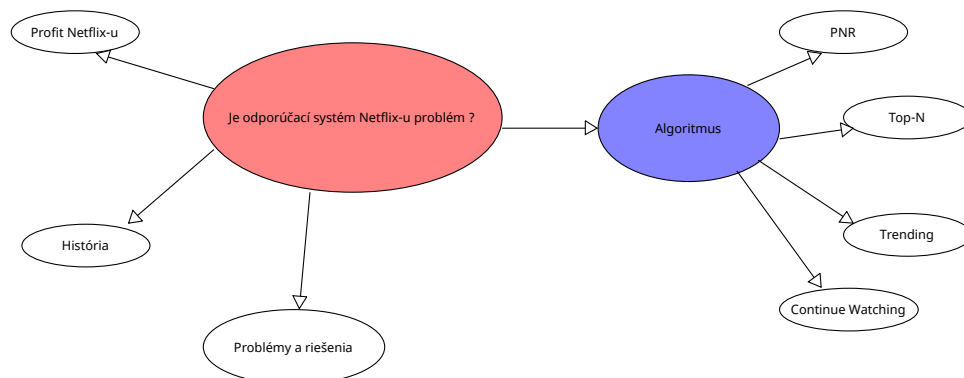
Abstrakt

Touto prácou by som chcel poukázať na to, ako Netflix ako firma spracúva užívateľské informácie a či je to správne. Na začiatku tejto práce objasním problematiku všeobecných odporúčacích systémov. Potom ukážem kde Netflix posiela užívateľské informácie. Ako ich spracúva a ako podľa nich dokáže nám užívateľom spríjemniť fungovanie na platforme. Ak nám to zvyšuje zážitok z používania tejto platformy, čo to teda tú platformu stojí a či je to vo výsledku pre danú platformu efektívne a finančne prospešné. Taktiež sa pokúsím čo navyiac priblížiť samotné spracovanie týchto údajov a poukážem na to, či je správne daným spôsobom spracúvať tieto informácie. [1] Dôležitý taktiež pre platformu je, algoritmus akým spracúvajú dané informácie a samotná infraštruktúra v ktorej to celé funguje. Samozrejme touto prácou chcem zistiť, či kvôli tomuto algoritmu a spôsobu spracovania informácií, Netflix neubližuje menej populárnemu alebo menej "mainstream" obsahu, ktorý je teda ešte o to menej ukazovaný užívateľom. A tiež poukážem na to, či je tento menej populárny obsah iba nefavorizovaný a posúvaný nižšie v hľadaní, alebo je kompletne zahľavený určitým užívateľom. Na konci tejto práce teda budete vedieť ako Netflix spracúva informácie, či je správne spracovávať užívateľské informácie týmto spôsobom, či je to profitabilné pre Netflix a taktiež, či to prospieva užívateľom tejto platformy.

Úvod

Zaujímalo vás niekedy, ako Netflix určuje, čo by sa vám pravdepodobne páčilo keď prechádzate cez ich ponuku filmov a seriálov ? Viac o tejto tématike sa dozviete tu: 3. Alebo ako určuje kde a ako tieto filmy a seriály rozložiť na hlavnú stránku ? Viac o tomto tu:2. Možno vás niekedy zaujímal či je takýto algoritmus profitabilný a prečo vlastne chce Netflix zlepšiť svojim používateľom používanie ich platofrmy ? Odpoveď nájdete tu:4

*Semestrálny projekt v predmete Metódy inžinierskej práce, ak. rok 2024/2025, vedenie: Mgr. Yevheniia Kataieva, PhD.



1 Krátka história

Už v roku 2006 mal Netflix záujem o odporúčacie systémy. [1] Preto, oznámil Cenu Netflixu ako súťaž strojového učenia o to, kto vytvorí aspoň o 10% lepší odporúčací systém ako bol ten ich, Cinematech. Nový odporúčací systém mal byť lepší v konkrétne $RMSE(\text{root mean squared error})$ ¹ pre predpokladané hodnotenie. A zrazu sa veľmi veľa rozdielných firiem snažilo zlepšiť svoje algoritmy.

2 Problem zlatej rybky

Keď majú ľudia príliš veľa vecí, na ktoré sa môžu pozrieť, nepozrú sa na nič. Štúdia spomenutá v [3] zistila, že v priemere typický používateľ stráca pozornosť po približne minúte prezerania si možností, čo si pozrie. No a teda sa Netflix veľmi usiluje aby počas tejto minúty, ukázal užívateľovi čo najlepší výsledok. Určite sa aj vám niekedy stalo, že ste otvorili Netflix hlavnú stránku, chvíľku popozerali a potom Netflix vyplí. To je pre Netflix najhoršia vec, čo môže nastať, pretože vy ste si nič nepozreli a aj napriek tomu, že im platíte mesačne, ak sa toto stane viac krát, určite to predplatné zrušíte. Nikto predsa nechce službu ktorá im neponúka to, čo chcú.

3 Algoritmus

Ak človek nemá žiadne odporúčanie od kamaráta alebo známeho, je odkázaný prehľadávať veľké množstvo filmov a seriálov, o ktorých nevieme či sú dobré alebo nie. Jednoducho si o každom z filmov alebo seriálov dokážeme dohľadať hodnotenia a žánre, ale nikdy nevieme, či sa to bude páčiť práve nám. A preto prichádzajú "na scénu", veľmi potrebné algoritmy ktoré vďaka veľkému množstvu dát dokážu predpovedať ako veľmi sa nám bude daná šou páčiť, ukázať nám hodnotenia ľudí podobnej filmovej chute ako my a dokonca aj predpovedať aké hodnotenie jej dáme. "V Netflixu nieje jeden model ktorý poháňa všetky odporúčania ale skôr set techník ktoré sa navzájom prepájajú v rovnakom celi -

¹RMSE - root mean squared error - odmocnina priemerného rozdielu medzi reálnymi a očakávanými hodnotami [8]

zvýšiť užívateľskú spokojnosť” [7]. No a teda aké všetky algoritmy vlaste používa Netflix na svojej platforme ? Viac sa dozviete v najbližších podkapitolách. . .

3.1 Prispôsobený hodnotič videí

Na hlavnej stránke môžete vidieť naozaj veľa filmov a seriálov. Sú ale rozdelené do riadkov, vždy podľa toho čo ich spája. Pre uvedenie príkladu, riadok komediálnych filmov sa pre nás vytvorí, ak sme si niekedy predtým pozreli nejakú komédiu. Samozrejme sa to odvíja od toho, aký člen domácnosti práve pozerá, tak isto od toho aké hodnotenie dal danej komédii a tak isto ako dlho ju pozeral a koľko takých komédií si pozrel. Na všetkom z tohto pracuje *PVR (Personalised Video Ranker)*² [3]

3.2 Top-N hodnotič videí

Iný typ riadku ktorý je tiež na hlavnej stránke ale ukazuje trošku niečo iné. Jeho cieľ je taký istý - ukázať užívateľovi najlepší obsah. Pre zmenu, to teraz nefiltruje podľa iba určitého žánru ale z celého katalógu nám ukazuje to najlepšie čo bol schopný nájsť pre daného užívateľa a ukáže prvých 10, inokedy viac, odporúčaní, ktoré sú na začiatku "listu" ktoré vytvára tento algoritmus. [1]

3.3 Teraz populárne

Tento algoritmus je veľmi jednoduchý aj naraz personalizuje aj ukazuje krátko-dobo populárne filmy a seriály.

3.4 Pokračujte v pozeraní

Keďže je veľmi dôležité ukladať pokrok pozerania seriálov alebo filmov na pokračovanie, tento riadok je veľmi potrebný.

4 Profitabilita

Keďže 80% sledovania na platforme tvorí sledovanie odporúčaného obsahu práve predtým spomenutými algoritmami, tak to znamená, že fungujú veľmi dobre a, že sa im vyplatia [6]. V podstate bez ohľadu na to, koľko by ich to stálo, vyplatí sa im udržiavať a vyvíjať tieto algoritmy. Keďže za prehrávanie určitého obsahu, či už populárneho a dobre hodnoteného alebo nepopulárneho a menej dobre hodnoteného, platí Netflix prakticky tú istú sumu a teda im je to v podstate jedno, čo užívateľovi zobrazí, jediné na čom záleží je, že si to užívateľ pozrie, je spokojný a príde na platformu znova. [1] Pre jednoduché porovnanie odporúčacieho systému Netflixu s napr. Spotify odporúčacím systémom. Môžeme vidieť veľmi veľký problém v tom, že Spotify uprednostňuje populárnu hudbu a obsah oproti tomu menej populárnemu, pretože z toho má väčší profit a teda je stavaný aj ich odporúčací systém okolo profitu [2]. Keďže každá firma chce mať profit, je to v relatívne v poriadku, ale problém je, že potom sú utláčaní nepopulárni umelci a majú ťažšie vytvoriť zisk. Na rozdiel od Spotify, Netflix

²PVR - personal video ranker - Prispôsobený hodnotič videí [3]

profituje za všetok obsah rovnako, je menej podstatné uprednostňovať populárny obsah. O tom, že Netflix je na tom finančne veľmi dobre a rastie aj naďalej oproti minulým rokom, hovorí aj ich výročná správa o profite [4]. Tento profit sa odvíja aj na hodnote ich akcie, tá bola v 01. novembra 2023 iba 420,19\$ ale tohto roka v ten istý dátum sa vyšplhala na neuveriteľných 756,10\$ za akciu. To je zvýšenie o takmer 80% za jeden rok. [5]

5 Rozloženie hlavnej stránky

Keďže už sme si popísali ako fungujú algoritmy, teraz si môžeme popísať ako personalizovaný obsah rozkladá na stránku. Ako už predtým bolo spomenuté, užívateľ má krátku dobu pozornosti a Netflix preto rozdeľuje tento obsah do riadkov hneď na hlavnú stránku. Treba si ale uvedomiť že Netflix je zväčša používaný v domácnostiach a teda do napr. top 10 sa snaží dávať obsah ktorý si užije napr. otec, mama a ďalší členovia rodiny. Tak isto sa snaží nájsť položku pre celú rodinu. Ak je rodina jednočlenná, tak sa snaží nájsť výber z rozdielnych nálad a záľub daného užívateľa ktoré zisťujú dané algoritmy. [1] Netflix chce aby sme vedeli, ako sa ich systém adaptuje našej chuti. To som si ale ja napríklad sám nevšimol. . . A tak isto sa snaží vybudovať dôveru v tieto odporúčania a tiež sa snaží vysvetliť prečo by sme si to chceli pozrieť. Toto reprezentuje popisok k danému filmu/serialu, tak isto to ukazuje predikciu hodnotenia ake by sme dali danému filmu/seriálu a tiež či si nejaký z našich priateľov pozrel daný obsah.

Literatúra

- [1] X. Amatriain and J. Basilico. Recommender systems in industry: A netflix case study. In *Recommender systems handbook*, pages 385–419. Springer, 2015.
- [2] T. Bachal, A. Kumar, A. Bachal, and V. S. Gutte. Tuning in to personalized music: A spotify api-based hybrid recommendation system integrating content-based and popularity-based approaches. In *2023 2nd International Conference on Futuristic Technologies (INCOFT)*, pages 1–6, 2023.
- [3] C. A. Gomez-Urbe and N. Hunt. The netflix recommender system: Algorithms, business value, and innovation. *ACM Trans. Manage. Inf. Syst.*, 6(4), Dec. 2016.
- [4] Netflix. 2023 annual report. https://s22.q4cdn.com/959853165/files/doc_financials/2023/ar/Netflix-10-K-01262024.pdf. Citované: 03.11.2024.
- [5] Netflix. Value of stock of netflix. <https://ir.netflix.net/stock-information/stock-chart/default.aspx>. Citované: 03.11.2024.
- [6] D. Rastogi, T. S. Parihar, and H. Kumar. A parametric analysis of ava to optimise netflix performance. *International Journal of Information Technology*, 15(5):2687–2694, 2023. <https://link.springer.com/content/pdf/10.1007/s41870-023-01281-z.pdf>.

- [7] H. Steck, L. Baltrunas, E. Elahi, D. Liang, Y. Raimond, and J. Basilico. Deep learning for recommender systems: A netflix case study. *AI Magazine*, 42(3):7–18, Nov. 2021.
- [8] K. Stewart. Mean squared error. <https://www.britannica.com/science/mean-squared-error>. Citované: 03.11.2024.