

Методе евалуације система за препоруку

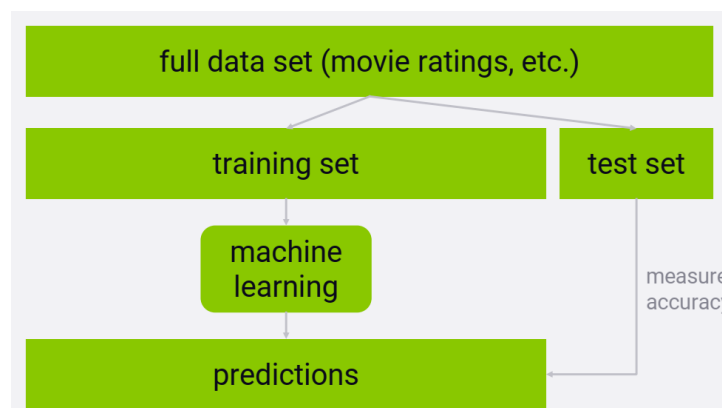
Прави увид у ефикасност система за препоруку се може добити тек извођењем експеримента, где се корисницима, у контролисаним условима, представљају препоруке и прати њихова реакција. Такви тестови се називају А/Б тестови и они нису реално изводљиви за велики број компанија, јер захтевају превише новца, ресурса и времена. Према томе, потребно је окренути се неким другим техникама евалуације, од којих су неке карактеристичне и за остале области машинског учења.

Ниједан од тих поступака евалуације не треба изводити изоловано, већ добијене резултате анализирати холистички и на основу њих извући одређене закључке.

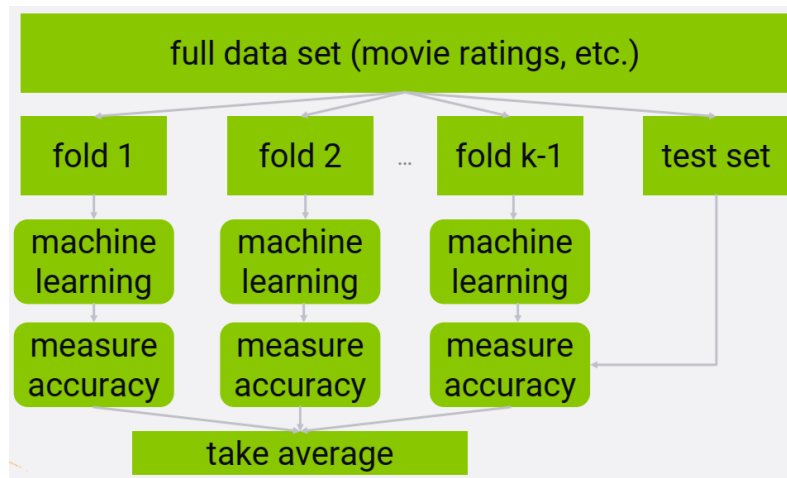
Пре него што се започне детаљнија анализа ових метода, треба истаћи јасну разлику између **предвиђања рејтинга** и **Тор-Н листе**. Наиме, функција предвиђања се односи на предвиђање оцене којом би неки корисник вероватно оценио одређени производ. Са друге стране, као што је речено, *Тор-Н* листа представља списак одређеног броја производа за које се сматра да би били погодни за корисника система, а са којима раније није био упознат. **Битно је истаћи да корисника не интересује способност система да тачно предвиди оцену коју би он дао неком артиклу, већ да добије листу за њега интересантних производа.** Дакле, погрешно је превелику пажњу посветити развоју способности система да изузетно прецизно обавља предикцију рејтинга. Иако се производи могу рангирати у *Тор-Н* листу на основу овог параметра, тачна предикција оцене не гарантује увек квалитетну *Тор-Н* листу. Стога, способност предвиђања рејтинга није небитан, али ни превише значајан показатељ ваљаности система, јер се приликом евалуације искључиво мора посматрати заједно за метрикама које се односе на *Тор-Н* листу.

Припрема података за евалуацију

1. Подела скупа на два скупа: скуп за тренирање података и скуп за тестирање креираног модела.



2. k-fold cross-validation



Методе за евалуацију система за препоруку

У складу са претходном елаборацијом, разликују се две групе метода за евалуацију система за препоруку:

1. Методе евалуације усмерене на предикцију рејтинга и
2. Методе евалуације усмерене на *Top-N* листу.

Методе евалуације усмерене на предикцију рејтинга

У ову групу се убрајају просечна апсолутна грешка (**Mean Absolute Error - MAE**) и корен просека квадрата грешака (**Root Mean Square Error - RMSE**). Наведене метрике су распрострањене и у осталим областима машинског учења.

а) *MAE* се рачуна по следећој формули:

$$\frac{\sum_{i=1}^n |y_i - x_i|}{n} \quad (1)$$

где је n број рејтинга у тестном скупу¹, y_i предикција рејтинга, x_i стварни рејтинг.

Предвиђени рејтинг	Актуелни рејтинг	Грешка
5	3	2
4	1	3
5	4	1
1	1	0

¹ Претпоставља се да је читалац упознат са основним појмовима машинског учења као што су тестни скуп, тренинг скуп итд.

$$MAE = (2 + 3 + 1 + 0)/4 = 1.5$$

b) RMSE се рачуна на следећи начин:

$$\sqrt{\frac{\sum_{i=1}^n (y_i - x_i)^2}{n}} \quad (2)$$

Предвиђени рејтинг	Актуелни рејтинг	Грешка ²
5	3	4
4	1	9
5	4	1
1	1	0

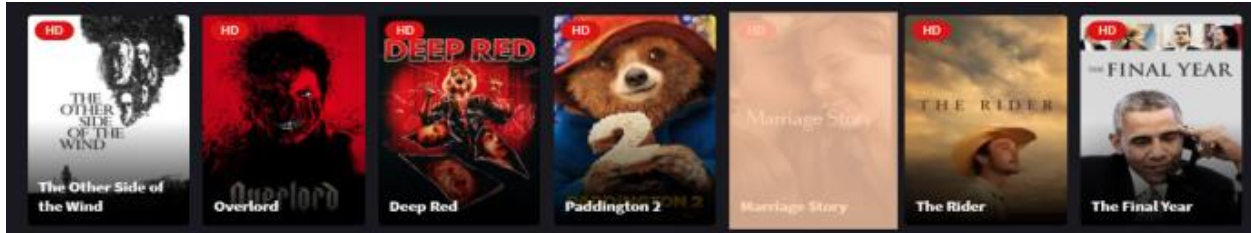
$$RMSE = \sqrt{(4 + 9 + 1 + 0)/4} = 1.87$$

Метод евалуације усмерене на *Top-N* листу

У ову групу се убрајају:

- стопа поготка (*Hit Rate*)
 - просечна реципрочна стопа поготка (*average reciprocal hit rate - ARHR*),
 - кумулативна стопа поготка (*cumulative hit rate - CHR*),
 - *rating hit rate - rHR*
- покривеност (*coverage*),
- разноврсност (*diversity*) и
- степен популарности (*novelty*).

Поступак евалуације се спроводи применом неке од познатих метода валидације у машинском учењу (*k-fold cross-validation*, *leave-one-out validation* итд). На овом месту биће објашњена *leave-one-out* валидација, јер се најчешће примењује при евалуацији *Top-N* листе. Наиме, овим поступком се по један насумично одабран рејтинг сваког корисника смешта у тестни скуп, а остали подаци чине тренинг скуп.



Hit rate

У случају leave-one-out валидације, под „поготком“ (*hit*) се подразумева појава изостављеног артикла из тренинг скупа у *Top-N* листи. У наставку је наведена формула којом се рачуна стопа поготка.

$$\frac{hits}{users} \quad (3)$$

где је *hits* укупан број погодака (јесте у *TopN* листи или није), а *users* укупан број корисника за које се генеришу *Top-N* листе.

Установљено је да корисници највише пажње обрађају на првих неколико препорука на листи. У том смислу, битна је и позиција поготка на *Top-N* листи. Узимајући све ово у обзир, настала је модификована верзија стопе поготка која се назива **просечна реципрочна стопа поготка (average reciprocal hit rate)**.

$$ARHR = \frac{\sum_{i=1}^n \frac{1}{rank_i}}{users} \quad (4)$$

где је *rank* позиција поготка на листи препорука, а $n = users$ број корисника. Ако на листи погодак заузима позицију 3, онда се у формули користи $1/3$.

Кумулативна стопа поготка (cumulative hit rate) је заправо идентична стопи поготка са додатним захтевом да производ има оцену која је већа од одређене вредности. Ова метрика наглашава значај могућности система да препоручи производе који би се допали корисницима.

Hit rank	Predicted rating
4	5.0
2	3.0
1	5.0
10	2.0

Још један начин сагледавања стопе поготка је уобличен метриком која се зове **rating hit rate**. Наведена метрика представља дистрибуцију погодака по рејтинзима. Дакле, за сваку вредност рејтинга из тестног скупа рачуна се однос броја погодака према укупном броју производа са том оценом у тестном скупу.

Нека се у тестном скупу налазе k различитих вредности рејтинга (r_1, r_2, \dots, r_k) . *Rating hit rate* се може представити на следећи начин:

$$\begin{array}{l} r_1 \frac{hits_{r_1}}{test_{r_1}} \\ r_2 \frac{hits_{r_2}}{test_{r_2}} \\ \vdots \\ r_k \frac{hits_{r_k}}{test_{r_k}} \end{array} \quad (5)$$

где је $hits_{r_i}$ број погодака са оценом r_i , а $test_{r_i}$ број појављивања оцене r_i у тестном скупу, $i = \overline{1 \dots k}$.

rating	hit rate
5.0	0.001
4.0	0.004
3.0	0.030
2.0	0.001
1.0	0.0005

Покривеност (Coverage)

Покривеност подразумева број различитих препорука које је систем у стању да креира. Може се посебно дефинисати за каталог (производе) и кориснике.

Покривеност каталога се дефинише као проценат $\langle user, item \rangle$ парова које је систем у стању да предвиди у односи на укупан број потенцијалних парова. Покривеност каталога треба посматрати са посебном пажњом, јер нам ова мера говори колико брзо ће нови производ почети да се појављује у препорукама. На пример, ако се појави нова књига на Амазону, она се неће појављивати у препорукама све док је неколико људи не купи и тако успостави шаблон/везу са куповином других производа. Све док ти обрасци не буду постојали, та нова књига ће смањивати метрику покривености на Амазону.

Најједноставнија мера за покривеност каталога је проценат/удео свих оних производа који потенцијално могу бити препоручени. Ова мере се лако може израчунати на основу алгорита и улазних података на следећи начин:

$$Coverage_{item} = \frac{n}{I} \times 100,$$

где је I укупан број производа у каталогу, а n укупан број производа које систем може да препоручи.

У пракси се покривеност каталога мери на скупу од више препорука (сесија), на пример испитивањем препорука враћених корисницима у одређеном временском периоду. Нека је са I_L^j означен скуп производа садржаних у листи L која је враћена од стране j -th препоруке која је добијена у посматраном временском оквиру. Нека је N укупан број

преорука добијен током овог периода и нека је I скуп свих доступних производа у каталогу.

$$Coverage_{catalog} = \frac{|\bigcup_{i=1}^N I_i^L|}{|N|}$$

Уколико се дозволи преорука само оних производа за које је предикција рејтинга већа од неке унапред задате вредности (*Higher Quality Threshold*), може се десити да Top-N листа остане празна. У овом случају ће се највероватније повећати тачност на штету покривености. Дакле, треба пажљиво одредити границу која ће омогућити компромисно решење између проблема празне листе или листе која је пуна производа који се не допадају кориснику.

Друга мера покривености каталога је **разноликост продаје**, која мери колико корисници бирају неједнако различите производе када се користи одређени систем за преоруку. Ако се за производ i корисник одлучује са вероватноћом $p(i)$ тада је *Gini index* дат формулом:

$$G = \frac{1}{n-1} \sum_{j=1}^n (2j-n-1)p(i_j)$$

где i_1, i_2, \dots, i_n чине листу производа који су сортирани у растућем поретку према $p(i)$ вредности. Индекс узима вредност 0 када се сви производи бирају једнако често, и вредност 1 када се само један производ увек бира. Индекс се додатно може дефинисати и за број пута за који је сваки производ препоручен.

Друга мера **расподеле неједнакости производа** је Шенонова ентропија:

$$H = - \sum_{i=1}^n p(i) \log p(i).$$

Ентропија узима вредност 0 када се увек бира или препоручује само један производ, а $\log(n)$ када је n производа једнако заступљено.

Када је реч о корисницима, **корисничка покривеност** може бити измерена као доступност корисничких профила који су потребни да би се преорука направила. На пример, у случају сарадничког филтрирања ово може бити мера за број производа које корисник мора да оцени пре него што добије преоруку. Ова мера се може израчунати као:

$$Coverage_{user} = \frac{u}{U} \times 100,$$

где је U укупан број корисника, а u укупан број корисника који се појављује у препорукама система.

Коментар: На покривеност каталога неће у толикој мери утицати непопуларне ставке које нису оцењене/препоручене, јер најчеће оне представљају мали део оцена корисника. Исто тако, на покривеност корисника ће се утицати у већој мери када корисници са ниском потрошњом немају препоруке у поређењу са покривеношћу из каталога или интеракције са корисницима.

Разноврсност (*Diversity*)

Разноврсност представља степен различитости производа које систем за препоруке смешта у *Top-N* листе. Пример ниског степена разноврсности јесте *Top-N* листа са књигама из едиције коју тренутно корисник чита, тј. листа препоручених производа не садржи књиге од других аутора, али са сличном тематиком. Ова метрика можда делује као превише „субјективна“, али је мерљива. Наиме, могуће је креирати матрицу сличности између производа из *Top-N* листе, а затим на основу исте израчунати просечну вредност сличности два производа. Формула за рачунање разноврсности се дефинише помоћу следећег израза:

$$1 - S, \quad (6)$$

где је S просечна сличност два производа за препоруку. Битно је напоменути да се метрика за сличност у формули може разликовати од оне која се користи у самом систему за препоруку.

Висока вредност ове метрике није сама по себи увек нужно добра ствар. Необично велики степен разноврсности у листама за препоруку се може постићи препоручивањем потпуно насумично одабраних артикла и много је чешће знак лошег, него доброг квалитета система за препоруку. Као што је и речено, све метрике треба посматрати у контексту осталих, па тако ни ова није изузетак.

Ако систем за препоруку путовања даје 5 препорука за исту локацију, при чему се само разликују хотели, такав систем није толико користан као онај који препоручује више локација.

Новитет (*Novelty*)

Новитет неког дела информација се уопштено односи на то колико је тај део различит у односу на „оно што је претходно видео“ корисник или заједница. Новитет се јавља када систем за препоруку сугерише активном кориснику непознати производ који је могао и самостално да открије. Посматрајмо систем за препоруку филмова који предлаже филмове у режији омиљеног режисера корисника. Ако систем препоручи филм са којим корисник није био упознат, филм ће бити нов, али не и случајан. Са друге стране, вероватније је да ће филм младог, не баш популарног редитеља бити **изненађујући** (*serendipitous*), а такође и нов. Међутим, треба имати у виду да у многим случајевима корисници неће пријавити све предмете које су користили у прошлости.

Многи аутори препознају три дефиниције новитета у системима за препоруку:

1. Нови производ за систем. Производ који је скорије додат и којем корисници нису приступали. Производ је нов за систем, а тиме и за сваког корисника у систему.
2. Заборављени производ. Када корисник заборави на неки производ који је раније користио/куповао/приступао. Понављање заборављених предмета у будућим препорукама производи повећану разноликост (*diversity*) и емоционално узбуђење.

3. Непознати производ. Купац никада није видео ни чуо за тај производ. Ставка је позната систему, али непозната кориснику.

Појам *новитета* се може просматрати и кроз појам *степен популарности* производа (просечан ранг популарности за препоручене производе). У овом случају се мери колико су популарни производи који се препоручују.

Сваки корисник жели да као препоруку види бар неколико производа који су му у извесном смислу блиски. На тај начин се стиче поверење у систем. Ако би се препоручивале само ствари за које никада није чуо, корисник може стећи утисак да га систем не „познаје“ у довољној мери. Напослетку, популарни производи нису популарни без разлога. Наиме, они пријају огромном сегменту популације, па је логично очекивати да би такође биле подесне и за оне који их још увек нису оценили. Ако се у препорукама не појављују популарне ставке, то је најчешће знак лошег система. Такође, треба правилно одмерити присуство популарних и блиских производа кориснику и оних за које корисник никада није чуо и које би могле бити изненађујуће пријатно искуство (*novel or serendipitous items*). Дакле, препорука популарних, а кориснику блиских производа, успоставља поверење корисника у систем. Са друге стране, мало необичнији производи омогућавају кориснику да истражи читав спектар производа који бу му се можда свидели - феномен дугог репа.

Израчунавање ове метрике (преузето из књиге Francesco Ricci et al, *Recommender systems Handbook*):

While we can obviously measure novelty in a user study, by asking users whether they were already familiar with a recommended item, we can also gain some understanding of a system's novelty through offline experiments. For such an experiment we could split the data set on time, i.e. hide all the user purchases that occurred after a specific point in time. In addition, we can hide some purchases that occurred prior to that time, simulating the items that the user has purchased and is hence familiar with, but did not report their purchase to the system. When recommending, the system is rewarded for each item that was recommended and purchased after the split time, but would be punished for each item that was recommended but purchased prior to the split time.

To implement the above procedure we must carefully model the hiding process such that it would resemble the true preference discovery process that occurs in the real system. In some cases the set of purchased items is not a uniform sample of the set of all items the user is familiar with, and such bias should be acknowledged and handled if possible. For example, if we believe that the user will report more purchases of unique items, and less purchases of popular or common items, then the hiding process should tend to hide more popular items.

In using this measure of novelty, it is important to control for accuracy, as irrelevant recommendations may be new to the user, but still worthless. One approach would be to consider novelty only among the relevant items [79].

Another method for evaluating novel recommendations uses the above assumption that popular items are less likely to be novel. Thus, novelty can be taken into account by using an accuracy metric where the system does not get the same credit for correctly predicting popular

items as it does when it correctly predicts nonpopular items [65]. Ziegler et al. [80] and Celma and Herrera [12] also give accuracy measures that take popularity into account.

Осетљивост

Churn означава степен осетљивости система на промену укуса корисника, тачније, колико се често систем мења када се преференце корисника мењају. Неопходно је водити рачуна о томе да ли остављање оцене о неком производу одмах треба да утиче на мењање листе за препоруку или не. Такође се поставља и питање да ли је потребно мењати нешто у листи са препорукама уколико корисник не показује интересовање за препоручене производе². Промени у *Top-N* листи најчешће претходи поновно тренирање читавог система и моделирање преференци корисника, што је изузетно временски захтевно, а и изискује додатно ангажовање рачунарских ресурса.

Responsiveness – колико брзо промене корисничког понашања утичу на систем.

Случајно откривање (*Serendipity*)

Случајно откривање је мера колико је успешна или релевантна препорука изненађујућа. Предложена је од стране професора Joseph Konstan, University of Minnesota и професора Michael Ekstrand, Texas State University. Груби превод речи *serendipity* је „случајност и развој догађаја на срећан или користан начин“. Сама мера је нешто тежа за израчунавање, јер се базира на степену изненађења и релевантности производа за корисника.

У литератури постоји више различитих дефиниција ове метрике. На примр, Herlocker et al. описују *serendipity* препоруке као оне препоруке које помажу кориснику да пронађе изненађујући занимљиве предмете које можда сам ни би открио. Shani and Gunawardana (2011) наводе да случајност укључује позитиван емоционални одговор корисника на нове предмете и мери колико су изненађујуће ове препоруке.

Према овим дефиницијама, случајност у системима за препоруку карактерише занимљивост предмета и изненађење за кориснике који добију неочекиване сугестије.

Већина дефиниција за случајно откривање укључује три компоненте: релевантност (*relevance*), новитет (*novelty*) и неочекиваност (*unexpectedness*). Сваки од ових појмова има више различитих дефиниција, што резултира у осам различитих дефиниција за *serendipity*



Занимљивости:

- a) Pasteur principle, i.e. “chance favors only the prepared mind”, meaning that sudden flashes of insight do not just happen, but they are the product of preparation.

² Да ли корисник тражи детаљније податке о понуђеном артиклу (кликом на линк, додавање у корпу итд.)

Recommender systems could implement the “prepared mind” paradigm by applying information about user preferences in different contexts. For example, if the system knows that a user is interested in science-fiction movies, it might exploit that information when the user is looking for a hotel as well, and suggest the Hilton in Las Vegas because it hosts a Star Trek flight simulator.

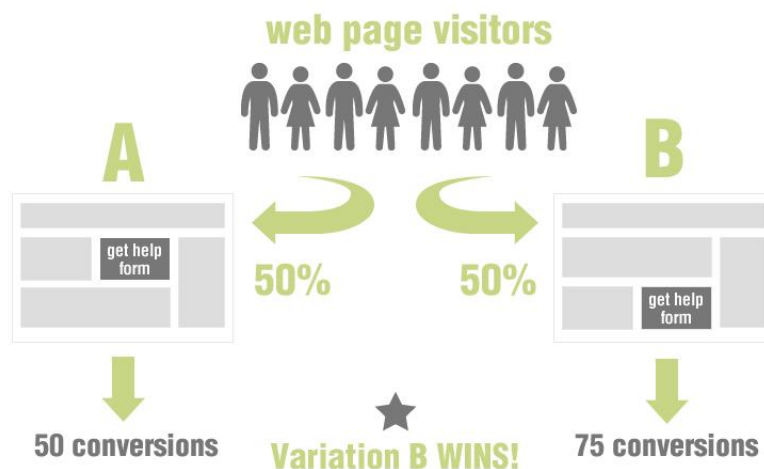
- b) The term serendipity was coined by Horace Walpole by referencing a Persian fairytale, “The Three Princess of Serendip,” in 1758. In the fairytale, the three princes of the country Serendip ventured out to explore the world and made many unexpected discoveries on their way.¹ In his letter, Horace Walpole mentioned that the princes were “always making discoveries, by accidents & sagacity, of things which they were not in quest of”.

Онлајн А/Б тестирање

A/B testing (also known as bucket testing or split-run testing) is a user experience research methodology. A/B tests consist of a randomized experiment with two variants, A and B. It includes application of statistical hypothesis testing or “two-sample hypothesis testing” as used in the field of statistics. A/B testing is a way to compare two versions of a single variable, typically by testing a subject's response to variant A against variant B, and determining which of the two variants is more effective.

A/B tests are useful for understanding user engagement and satisfaction of online features, such as a new feature or product. Large social media sites like *LinkedIn*, *Facebook*, and *Instagram* use A/B testing to make user experiences more successful and as a way to streamline their services.

Извор: Wikipedia



Није нам битно само да погодимо рејтинг који ће корисник дати производу, већ како ће и на тај производ корисник да одреагује. **Netflix награда!**



RMSE грешка није била довољна за добар систем за препоруку. Пословни циљеви нису били испуњени.

Пример развојног пута једног А/Б теста: <https://vwo.com/ab-testing/>

ЗАКЉУЧАК: За крај овог дела још једном треба истаћи значај А/Б тестова и подсетити да су они једини реални показатељи ефикасности изграђеног система за препоруку. Дакле, фокусирање на једну од претходно наведених метрика није довољно. То и потврђује прича о Нетфликс (*Netflix*) награди од милион долара за тим који успе да прецизност предвиђања рејтинга (RMSE) њиховог система за препоруку подигне за 10%. Иако су најбољем тиму исплатити награду, њихово решење нису имплементирали трајно, јер су увидели да повећање прецизности није имало значајнијег утицаја на остварење пословних циљева и повећање профита. Врло често се дешава да изузетно прецизни системи не генеришу очекивано квалитетне листе за препоруку, што је и најбитније за остварење пословних циљева. Истражујући управо овај проблем, *YouTube* је у својој студији названој „*The surrogate problem*“ навео: „Више је уметности него науке у одабиру проблема замене у препорукама“. Под овим се подразумевало да се не може увек користити прецизност (*accuracy metric*) као замена за квалитетну препоруку приликом евалуације система.

ДОМАЋИ: *The Explore-Exploit dilemma* (<https://medium.com/data-science-for-everyone/the-explore-exploit-dilemma-436cb1edff0d>)



ЗАНИМЉИВОСТИ: *The history of Amazon's recommendation algorithm*

<https://www.amazon.science/the-history-of-amazons-recommendation-algorithm>

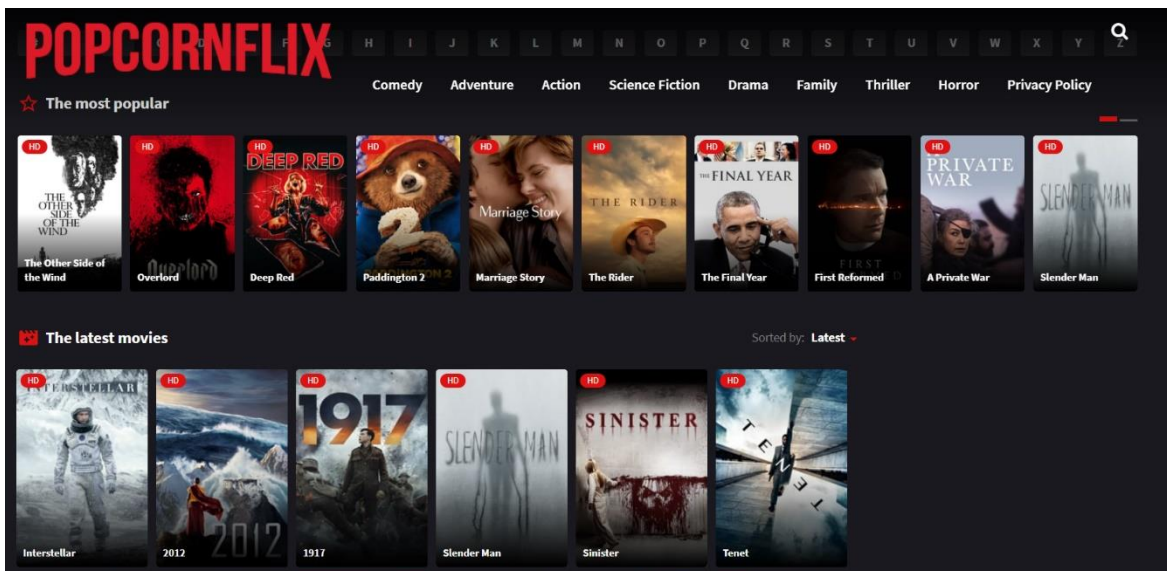
Проблеми система за препоруку

У наставку се даје преглед проблема система за препоруку:

- Хладан старт (*Cold start*)
- Листа „неадекватних” производа (*Stoplist*)
- „Бочица филтера” (*Filter bubble*)

Хладан старт

Хладан старт (*Cold start*) представља један од најпознатијих проблема и мањкавости система за препоруку. Огледа се у немогућности генерисања препорука за новог корисника система. Будући да корисник није оценио ниједан производ, систем нема довољно података, тј. није га довољно добро „упознао” да би креирао персонализоване препоруке. Слично важи и приликом додавања новог производа у каталог.



У даљем тексту се наводе потенцијална решења за хладан старт у случају појаве новог корисника, као и за случај појаве новог производа.

Решења за хладан старт у случају новог корисника:

- **Употреба имплицитних података** – Заснива се на праћењу понашања новог корисника при употреби система. Уколико се задржи на страници са одређеним производом, врло разуман потез би био препорука сличних производа.
- **Употреба колачића (*cookies*)** – Уколико се корисник није представио систему кроз било који вид аутентификације, најбоље је проверити колачиће. На основу

садржаја ових текстуалних фајлова, може се утврдити да ли је можда у питању постојећи корисник, који се можда, овог пута, није пријавио на систем. Приликом примене ове методе треба бити обазрив и поштовати све правне норме о приватности корисника и складу са њима поступати.

- **Употреба IP адресе** – Систем има увид у IP адресе корисника. На основу њих се може приближно одредити локација са које се корисник повезао на систем. Уколико се одлучи за такав приступ, могу се препоручивати најпопуларнији производи у делу света са кога је корисник приступио систему.
- **Препорука најпродаванијих производа или производа на акцији** – Најпродаванији производи нису најпродаванији без разлога, тако да препорука ових производа и није толико лоше решење. Људи врло често купују производе на акцији иако им и нису баш неопходни. Стога и овакви производи заслужују своје место у *Top-N* листи.
- **Интервјуисање корисника** – Овај приступ подразумева да се новом кориснику, након пријаве на систем, понуди да оцени најпопуларније производе. На тај начин, систем може да прикупи довољно података за персонализовање препорука. Међутим, реалност је да већина корисника не жели да буде узнемиравана приликом онлајн деловања. Због тога се овај приступ врло ретко имплементира у савремене системе за препоруку.

Решења за хладан старт у случају појаве новог производа у каталогу:

- **Не предузимати никакве акције** – Иако ово решење делује тривијално, у пракси се показало као врло ефикасно. Наиме, *Top-N* листа није једини начин да се корисници упознају са новим производом. Један од начина је да се производ појави у групи акцијских производа. Такође, произвођачи спроводе различите маркетиншке акције у циљу представљања најновијих производа, надајући се да ће се корисници претрагом система упознати детаљније о њиховим својствима. Како корисници почну да остављају утиске о новим производима, систем ће почети да их препоручује.
- **Користити својства производа (*content-based attributes*)** – На основу својства новог производа, креирати матрицу сличности са осталим производима, што је и принцип рада система заснованих на садржају.
- **Насумично одабран нови производ** – Подесити систем тако да у *Top-N* листу убаци насумично одабран нови производ.

Листа „неадекватних“ производа (Stoplist)

Приликом генерисања *Top-N* листе треба избегавати производе који на одређени начин могу бити погрешно протумачени. Кључне речи које могу указати на такве ситуације се налазе у наслову, опису или називу категорија. Све такве термине треба сместити у стоп-листу (*stoplist*).

Неке од таквих категорија су:

- Садржај за одрасле,
- Производи који садрже разне вулгарне речи у наслову, опису итд.
- Теме које су једноставно неприхватљиве (нпр. нацизам),
- Тероризам и политички екстремизам,
- Психоактивне супстанце,
- Религијски екстремизми.

Посебну пажњу треба обратити на производе (нпр. књиге) који говоре о неким тешким болестима и третманима. Људи такве књиге читају у тешким периодима живота. Уколико неко близак премине од последица болести, последња ствар коју желимо је да систем поново препоручи сличне књиге таквим особама. Зато је најбоље решење да се овакви артикли потпуно изоставе из тренинг скупа.

Дакле, врло је битно градити стоп-листу на начин који ће омогућити брзо прилагођавање окружењу у коме се послује, његовим формалним и неформалним нормама, као и тренутној глобалној друштвено-економској ситуацији.

„Бочица филтера“ (*Filter bubble*)

„Бочица филтера“ представља стање интелектуалне изолације које може произаћи из персонализованих претрага када алгоритам веб странице селективно погоди које информације жели да види корисник.³

Поменути феномен није само проблем система за препоруку, већ и друштвених мрежа. Наиме, системи корисника излажу искључиво садржају који је у складу са његовим интересима, држећи га на тај начин у одређеној врсти балона његових постојећих интересовања.

Како би се овај проблем превазишао, уведена је метрика разноликости (*diversity*), о којој је већ било речи. Поред производа који су у складу са афинитетима корисника, *Top-N* листа треба да садржи и неки мало другачији производ. Како већина производа из листе одговара интересовању корисника, они заиста верује да их систем „познаје“ добро

³ Преузето са Википедије.

(поверење корисника - *user trust*). Због таквог утиска су спремни да истраже и мало другачије објекте из листе за препоруку. Ради одржавања поверења у систем, неопходно је кориснику дати могућност да открије зашто му је препоручен баш тај производ (транспарентност), тј. на који начин је систем дошао до закључка да би му се можда допао баш тај артикал.

На основу претходне елаборације се лако закључује да је транспарентност веома битан фактор у стицању поверења у систем од стране корисника, док је разноликост важна за избегавање свеprisутнијег феномена на интернету – *filter bubble*. Обе ствари нису толико лако изводљиве, али је такође јасно да никако не смеју бити занемарене приликом имплементације система за препоруку.