

Cuprins

1. Instrumente specifice algoritmilor de automatizare social media	2
1.1. Problematica generală asociată mediilor de socializare	2
1.2. Stadiul actual al cercetărilor algoritmilor de automatizare în domeniul social media	3
1.3. Motivarea alegerii temei	6
2. Cercetări privind algoritmi de automatizare social media	7
2.1. O privire de ansamblu asupra algoritmilor de automatizare asociați mediilor de socializare	7
2.2. Studiul literaturii dedicate algoritmului YouTube	23
2.3. Cercetări practice privind algoritmi de automatizare prezențați	31
3. Studiul și proiectarea sistemului informatic pentru automatizarea luării deciziei în social media	45
3.1. Tipuri de medii de socializare	46
3.2. Resurse software	46
3.3. Efecte pozitive și negative ale utilizării mediilor de socializare	47
3.4. Proiectarea, realizarea și testarea aplicației	48
4. Concluzii și direcții de cercetare ulterioare	61

1. Instrumente specifice algoritmilor de automatizare social media

Algoritmii de *automatizare* social media sunt un mod de sortare a postărilor în contul utilizatorilor pe baza relevanței, în loc de timpul de publicare. Rețelele de socializare acordă prioritate conținutului pe care un utilizator îl vede mai întâi în contul său, prin probabilitatea că acesta va dori de fapt să îl vadă. Înainte de trecerea la algoritmi, majoritatea fluxurilor de socializare afișau postările în ordine cronologică inversă. Pe scurt, au apărut mai întâi cele mai noi postări din conturile pe care le-a urmărit un utilizator. În mod implicit, algoritmii de automatizare social media determină ce conținut să livreze utilizatorului pe baza comportamentului lui.

De exemplu, Facebook și Twitter ar putea pune pe “peretele” utilizatorului postări ale familiei și ale celor mai apropiați prieteni, deoarece acestea sunt conturile cu care utilizatorul interacționează cel mai des. Sau în ceea ce privește recomandarea videoclipurilor de vizionat pe YouTube, acest lucru se bazează din nou pe comportamentul utilizatorului individual, explorând ceea ce a urmărit în trecut și ceea ce vizionează utilizatorii ca el. Elemente precum categorii, etichete și cuvinte cheie fac parte, de asemenea, din conținutul recomandat în orice rețea dată.

1.1. Problematika generală asociată mediilor de socializare

O mare parte din motivul pentru care algoritmii de automatizare sunt atât de controversați se datorează impactului lor asupra creșterii numărului de utilizatori care sunt abonați pe o anumită platformă de socializare pe un anumit cont.

Algoritmii sunt în continuă evoluție, încercând să rezolve probleme și să ofere cea mai bună experiență de utilizare posibilă. Drept urmare, specialiștii în marketing trebuie să se adapteze în mod constant la ei. Aceasta înseamnă experimentarea constantă a conținutului și schimbarea strategiilor de marketing.

Algoritmii de automatizare nu sunt neapărat perfecți. Există o mulțime de cazuri de algoritmi care aparent „ascund” conținut pe Facebook la întâmplare, în ciuda faptului că au fost optimizați pentru un anumit lucru. Pe de altă parte, există un fenomen bine documentat despre YouTube în legătură cu videoclipuri care au milioane de vizualizări peste noapte deoarece aceste videoclipuri, care par în prima parte aleatorii au fost recomandate utilizatorilor. [21]

1.2. Stadiul actual al cercetărilor algoritmilor de automatizare în domeniul social media

Dintre toate videoclipurile postate pe YouTube, a existat la un moment dat o recomandare pe care algoritmul de automatizare a afișat-o mai des. Videoclipul respectiv este “Bath Song | +More Nursery Rhymes & Kids Songs - Cocomelon (ABCKidTV).” YouTube l-a recomandat de peste 650 de ori printre cele 696.468 de sugestii potrivit unui studiu Pew Research, mult mai mult decât videoclipul de pe locul doi al cercetărilor, mai exact Maroon 5’s “Girls Like You”.



Figura 1.1. Pagina de recomandari din aplicatia mobila YouTube Romania

Sudiul Pew Research, a dovedit faptul că unele videoclipuri au fost recomandate mai mult decât altele. Este o metodologie fascinantă, dacă nu complexă, pentru explorarea unuia dintre cele mai importante sisteme algoritmice din lume și una care a rămas în mare măsură neînțeleasă pentru cercetători, cu atât mai puțin utilizatorii săi.

Videoclipurile pentru copii au dominat cele mai recomandate 10 postări. De exemplu, melodiei „Bath Song” i s-a alăturat „Learn Colors with Spiderman 3D w Trucks Cars Surprise Toys Play Doh for Children”, „Wheels on the Bus | Nursery Rhymes & Kids Songs - Cocomelon (ABCkidTV) ”și, „Learn Shapes with Police Truck - Rectangle Tyres Assembly - Cartoon Animation for Children”. Toate sunt titluri de videoclip de o mixare de cuvinte cheie care, aparent, au sens pentru algoritmul de automatizare, dar nu și pentru utilizatori.

Dar dacă se privește pe scară largă nivelul la care a ajuns algoritmul YouTube în acest moment, conținutul considerat nerelevant pare a fi eliminat din pagina de recomandări. Printre primele 50 de recomandări, 43 dintre ele au fost videoclipuri muzicale [11], videoclipuri pentru copii sau concursuri TV [8].

YouTube dorește să recomande lucruri care sunt plăcute oamenilor, și cel mai clar semnal în acest sens este faptul că un anumit videoclip a fost plăcut și de alte persoane cărora le-a fost sugerat. Pew a constatat că 64% din recomandări au fost destinate videoclipurilor cu peste un milion de vizionări. Cele 50 de videoclipuri recomandate de YouTube cel mai des au fost vizionate în medie de 456 de milioane de ori fiecare. Popularitatea generează popularitate, cel puțin în cazul utilizatorilor (sau al roboților).

Pe de altă parte, dezvoltatorii YouTube au spus în lucrările anterioare[23] care descriau algoritmul de automatizare faptul că utilizatorilor le place conținutul recent publicat. Dar este nevoie de timp pentru că o postare să creeze un număr mare de vizualizări și să semnaleze algoritmului că merită promovat. Deci, provocarea devine cum să recomandați „videoclipuri noi pe care utilizatorii doresc să le vizioneze” atunci când aceste videoclipuri sunt noi pentru sistem și au un nivel redus de vizionări.

Cercetările lui Pew reflectă acest lucru: aproximativ 5% din recomandări au fost destinate videoclipurilor cu mai puțin de 50.000 de vizionări. Sistemul automat învață din performanța timpurie a unui videoclip și, dacă are succes, vizionările pot crește rapid. De exemplu, un videoclip extrem de recomandat pentru copii a trecut de la 34.000 de vizionări când algoritmul Pew l-a întâlnit prima dată în iulie, la 30 de milioane de vizualizări în august.

Pe măsură ce software-ul Pew a făcut alegeri, sistemul a selectat videoclipuri mai lungi. Este că și cum software-ul recunoaște că utilizatorul va fi în jur pentru o vreme și începe să ofere

videoclipuri mai lungi. Mai mult, algoritmul de automatizare Youtube a început să recomande videoclipuri mai populare, indiferent de cât de popular a fost videoclipul de pornire.

Aceste condiții, nu au fost aproape sigur codificate în procesul de luare a deciziilor algoritmice. La fel că majoritatea companiilor asemănătoare Google, YouTube folosește rețele neuronale de învățare profundă, un fel de software care își reglează rezultatele pe baza datelor introduse în sistem[15]. Nu un inginer angajat în cadrul YouTube a spus „Arată oamenilor videoclipuri pentru copii care sunt din ce în ce mai lungi și mai populare”, ci mai degrabă sistemul automat a dedus statistic că acest videoclip ar fi mult mai popular și dorit pe YouTube.

Pe baza cercetărilor, s-a constatat faptul că YouTube personalizează puternic recomandările bazate pe istoricul unui utilizator, ceea ce este imposibil de simulat peste tot. Ceea ce a testat algoritmul Pew sunt recomandările pe care YouTube le-ar oferi unui utilizator anonim. Totuși, majoritatea utilizatorilor YouTube sunt conectați și primesc recomandări pe baza istoricului de vizionare. Nick Seaver, un antropolog care studiază sistemele de recomandare la Universitatea Tufts, a spus că studiul algoritmului YouTube presupune faptul că un utilizator anonim generează un fel de „linie de bază” pe baza căreia conținutul va fi personalizat.

Peste 70% din videoclipurile recomandate de YouTube au apărut pe lista studiată de sistemul Pew o singură dată. Este imposibil de examinat modul în care sute de mii de videoclipuri se conectează la fiecare prim videoclip aleatoriu atunci când există date atât de limitate despre fiecare.

Deci, cercetarea Pew lasă câteva întrebări cruciale fără răspuns. Oamenii vor să știe dacă YouTube radicalizează în mod regulat oamenii cu recomandările sale, așa cum a sugerat savantul Zeynep Tufekci. Acest studiu sugerează că YouTube împinge un utilizator anonim către conținut mai popular, nu mai marginalizat. Dar acest lucru s-ar putea să nu fie valabil pentru un utilizator obișnuit de YouTube cu un istoric de vizionare real, a spus Seaver.

Natura personalizării YouTube face ca recomandările sale reale să fie extrem de dificil de urmărit cantitativ. Este cu siguranță posibil că utilizatorul să cadă în niște găuri de iepure ale YouTube, dar cât de des YouTube conduce către o astfel de gaură este încă o întrebare aprinsă.

În ceea ce privește contul YouTube destinat copiilor, întrebarea care se pune cel mai des este procentul în care copiilor li se recomandă videoclipuri care nu sunt destinate pentru media

lor de vârstă. Optzeci la sută dintre părinți au spus că, cel puțin ocazional, își lăsau copiii să urmărească YouTube, iar dintre ei, peste 60% au detaliat despre copilul lor faptul că „a întâlnit pe YouTube conținut pe care îl considerau inadecvat copiilor”

Răspunsul companiei la aceste tipuri de sugestii îngrijorătoare a fost că YouTube nu este pentru copii. Copiii ar trebui să folosească doar aplicația YouTube Kids, care a fost construită ca un spațiu sigur pentru ei, afirmând: „Protejarea copiilor și a familiilor a fost întotdeauna o prioritate absolută pentru noi. Deoarece YouTube nu este destinat copiilor, am investit semnificativ în crearea aplicației YouTube Kids pentru a oferi o alternativă special concepută pentru copii. ”[2]

1.3. Motivarea alegerii temei

Am ales această temă deoarece în ziua de astăzi rețelele de socializare cum ar fi: Instagram, Facebook, Twitter, Youtube sunt foarte utilizate de majoritatea persoanelor și este un subiect de interes pentru foarte mulți utilizatori.

Algoritmii de automatizare social media joacă un rol foarte important în capacitatea marketerilor și a proprietarilor de afaceri de a avea un impact online. Și, deși algoritmii sunt adesea dificil de înțeles, trebuie ca utilizatorii să învețe să lucreze cu ei și nu împotriva lor, dacă se dorește atingerea succesului. Din păcate, în zilele noastre, algoritmii de automatizare din spatele social media lucrează împotriva utilizatorilor, dacă putem spune așa deoarece tot mai multe persoane stau o perioadă îndelungată de timp pe rețelele de socializare și uită de atribuțiile pe care le au în viață de zi cu zi. De exemplu, la alegerile prezidențiale din S.U.A din 2016, s-a presupus faptul că Cambridge Analytica ar fi colectat datele utilizatorilor Facebook și le-a creat profile psihologice, pe baza cărora utilizând algoritmi de automatizare, le-au fost recomandate utilizatorilor anumite mesaje politice astfel încât să le influențeze votul.

Tema a fost aleasă tocmai pentru a înțelege ce se întâmplă de fapt în social media și în spatele a ceea ce noi vedem, pentru a ne arata că nu totul este atât de simplu precum credem și sunt ani de muncă pentru a dezvolta un algoritm de automatizare, care învață din click-urile pe care le dă utilizatorul.

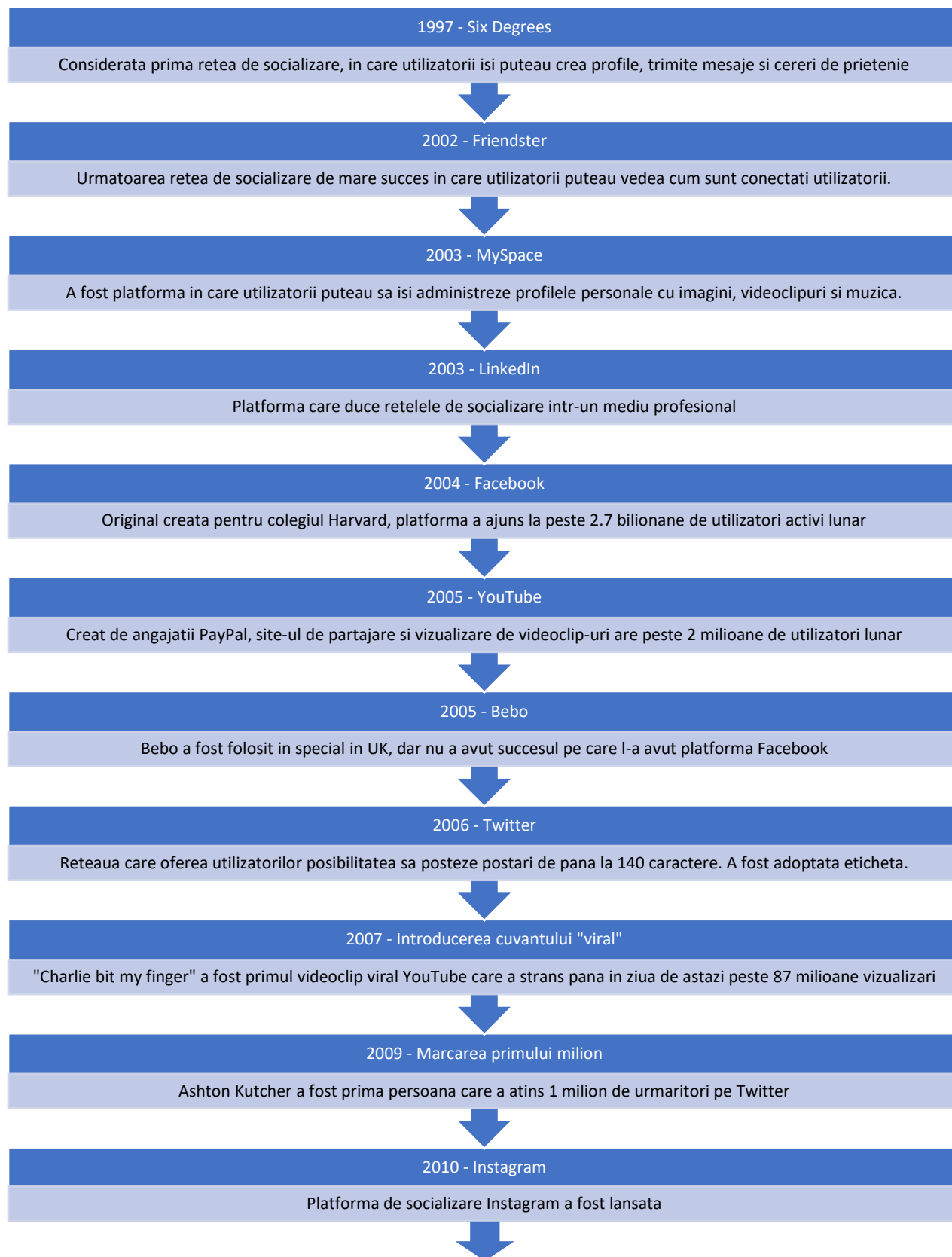
2. Cercetări privind algoritmi de automatizare social media

Acest capitol este organizat după cum urmează: O prezentare generală a algoritmilor de automatizare a rețelelor de socializare este descrisă în Secțiunea 2.1, în care se detaliază algoritmi Facebook, YouTube, Instagram. Studiul literaturii dedicate algoritmului de automatizare YouTube, regăsit în Secțiunea 2.2 cuprinde o prezentare generală a algoritmului de automatizare din spatele platformei YouTube, recomandarea videoclipurilor de pe platformă, un scurt exemplu, selectarea mai multor criterii care fac algoritmul să funcționeze, cum ajung videoclipurile să fie populare și reprezentarea unor caracteristici cheie de recomandare. Secțiunea 2.3 cuprinde cercetările teoretice originale ale autorului în ceea ce privește algoritmul YouTube și algoritmul Netflix, și o comparație între cei doi algoritmi de automatizare, împreună cu o scurtă descriere a algoritmului realizat în capitolul 3.

2.1. O privire de ansamblu asupra algoritmilor de automatizare asociați mediilor de socializare

Rețelele sociale joacă un rol important în viața oamenilor. Incepând din 2017 Facebook a fost al treilea cel mai vizualizat site de pe glob, după Google și YouTube. Cu cât exista mai mulți utilizatori pe platformele de socializare, cu atât este mai necesar să se creeze ordine și un algoritm care să creeze această ordine.

Rețelele sociale au fost utilizate pentru prima dată ca platformă pentru a păstra contactul cu prietenii și familia. Până în prezent, sunt încă folosite ca o modalitate de a rămâne în contact cu oamenii, dar este folosit și pentru partajarea fotografiilor, videoclipurilor și știrilor, pentru oportunități de carieră, pentru partajarea actualizărilor de stare, pentru participarea la discuții, scopuri de marketing și pentru a interacționa cu clienții, conform figurii 2.1.



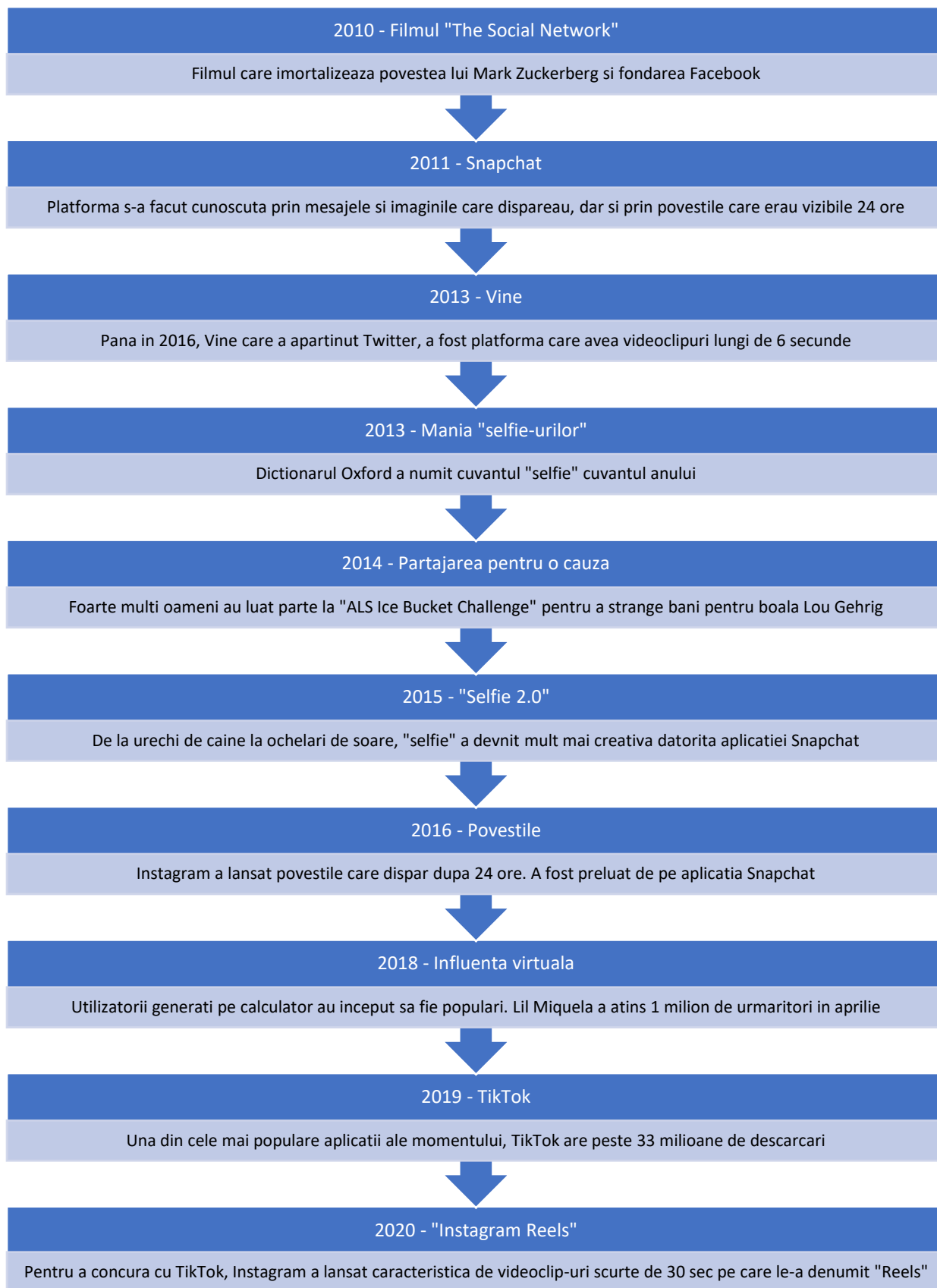


Figura 2.1 Evoluția rețelelor de socializare

Algoritmii de automatizare social media se schimbă întotdeauna, ceea ce face dificilă păstrarea pasului cu codurile care fac aceste platforme să existe. Nu există nicio modalitate de a cunoaște fiecare detaliu care se schimbă în algoritmi. În funcție de platforma care este utilizată, așteptările utilizatorului vis a vis de social media vor fi diferite. Algoritmul de pe fiecare rețea de socializare dictează locul în care un utilizator se clasează în anunțurile din rețelele sociale și în plasarea conținutului.

În continuare, vor fi prezentați diferiți algoritmi automatizare din spatele social media și cum ar trebui optimizat conținutul pentru fiecare rețea în parte.

2.1.1. Algoritmul de automatizare a deciziei asociat Facebook

Implicarea semnificativă a utilizatorilor este cheia pentru aspectul algoritmului Facebook. Algoritmul este creat pentru a intensifica importanța și vizualizarea postărilor locale, familiale și prietenoase, mai degrabă decât a postărilor de afaceri.

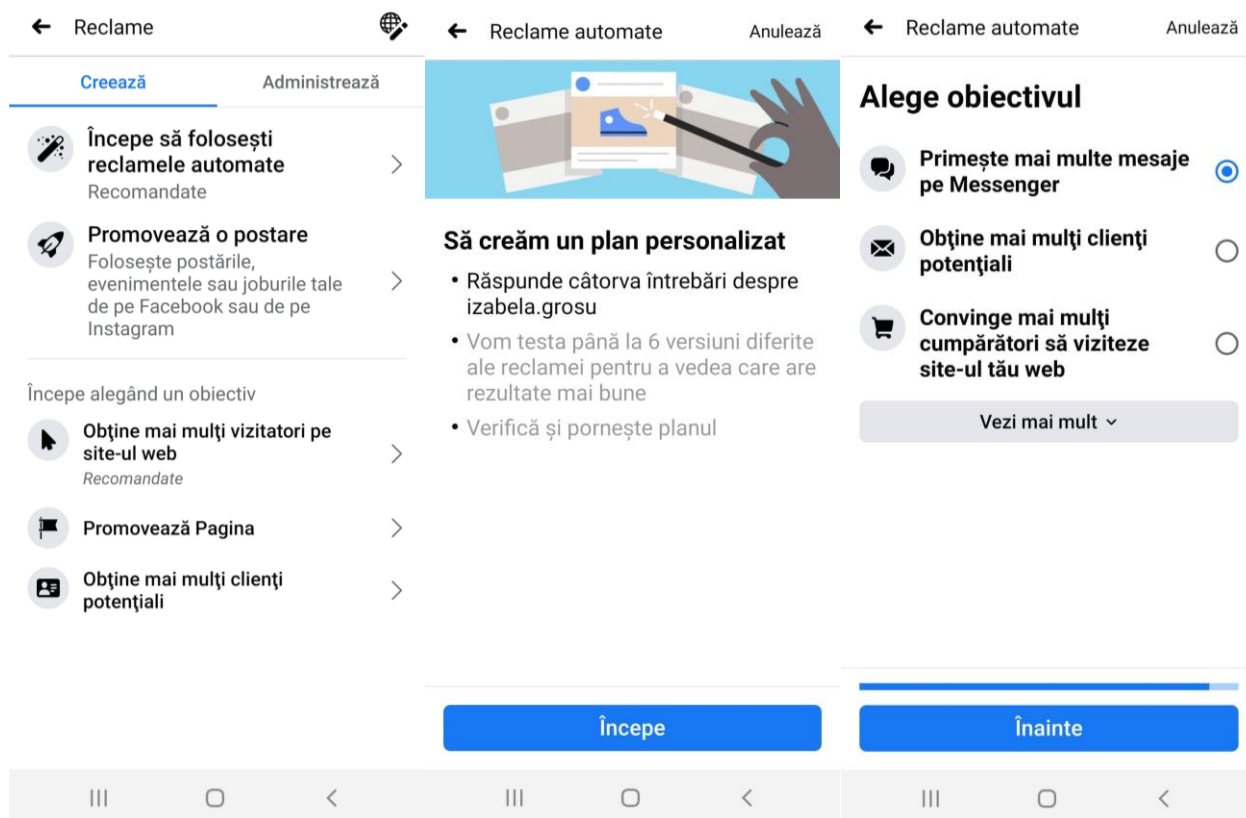


Figura 2.2. Automatizarea reclamelor platformei de socializare Facebook

Conținutul plătit este clasificat separat, dar reclamele sunt legate de implicarea utilizatorului, de ceea ce caută utilizatorul pe alte platforme, de exemplu Google, de interesele utilizatorului. În prezent, foarte multe firme apelează la automatizarea postărilor pe Facebook pentru a-și face reclama la propriile produse, pentru a avea mai multe aprecieri pe pagina, pentru a găsi cât mai mulți clienți care să le cumpere produsele. Pentru a se realiza automatizarea postărilor se poate folosi chiar instrumentul creat de Facebook denumit “Automated Ads”. Funcția oferă un plan de publicitate personalizat cu recomandări pas cu pas pentru reclamele potrivite pe care firma ar trebui să le difuzeze. De-a lungul timpului, anunțurile automate află ce funcționează cel mai bine și vor aduce îmbunătățiri și sugestii pentru campaniile firmei. Se pot crea până la 6 versiuni diferite ale unei singure postări, cu sugestii diferite pentru butoane, text și informații, după cum se poate vedea în figurile 2.2 și 2.3. După ce postarea a fost activă un timp, Facebook va începe automat să arate doar reclama cea mai populară. Câteva tipuri de automatizare în cadrul campaniilor publicitare Facebook sunt: alertele de performanță, care se trimit pe email cu

campania/postarea care are efect asupra utilizatorului, timpul la care este distribuita postarea, reguli de optimizare, cu ajutorul cărora, utilizatorul poate opri campaniile prea costisitoare.

Deși conținutul ușor de utilizat este esențial, o altă strategie pe care algoritmul Facebook a adoptat-o recent este gestionarea spamului. În primele trei luni ale anului 2018, Facebook a localizat și a șters peste 500 de milioane de conturi false

Facebook a fost original conceput pentru studenți de către Mark Zuckerberg în 2004, la Harvard University. Este o platformă de socializare care reușește să conecteze online, și să transmită postări membrilor familiei, prietenilor sau cunoscuților.

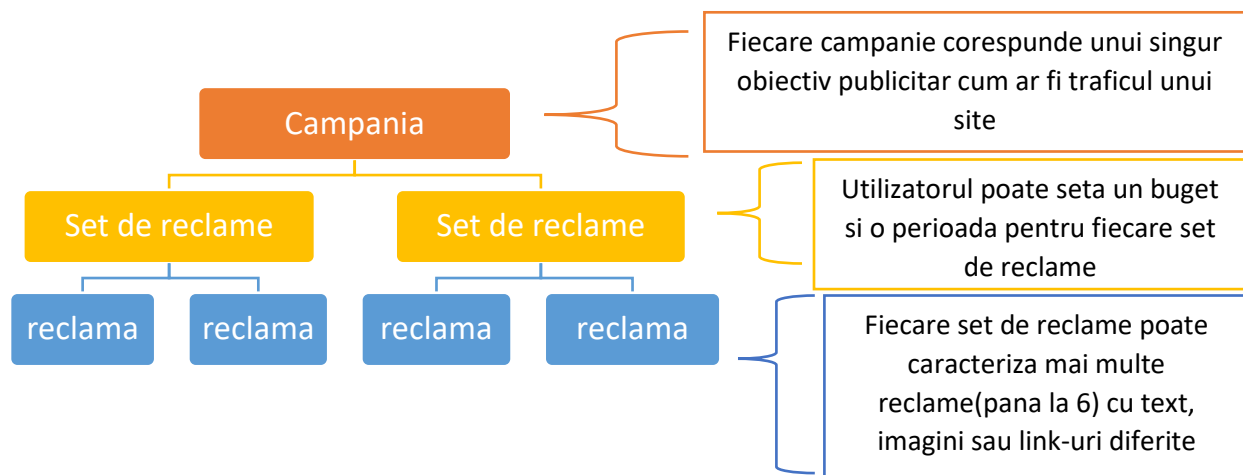


Figura 2.3. Structura campaniilor Facebook

În algoritmul Facebook există patru componente cheie care oferă utilizatoului o experiență mult mai bună pe platforma de socializare:

- Conținutul utilizatorului
- Semnale (ceea ce partajează utilizatorul sau postările la care reacționează)
- Predicțiile asupra semnalelor
- Scorul utilizatorului

Facebook, ca orice rețea de socializare dorește să păstreze utilizatorul logat pentru o perioadă mai lungă de timp. Platforma are în vedere mai multe aspecte pentru a decide ceea ce vede utilizatorul prima dată când se autentifică pe Facebook cum ar fi: timpul petrecut pe o anumită postare, timpul la care a fost distribuită postarea, dacă link-ul a fost trimis de către o altă persoană prin mesaj, numărul de comentarii și aprecieri, dacă postarea a fost distribuită de către o persoană cunoscută, timpul postării sau cât de informativă este postarea pentru utilizatorul respectiv.

Noul algoritm Facebook, după cum se poate vedea și în figura 2.4, este mult mai automatizat și focusat pe comentarii, distribuiri și reacții și încurajează utilizatorii să comunice prin reacții la postări. În funcție de postările la care reacționează utilizatorul, se va alege ce va apărea prima dată atunci când acesta se autentifică.



Figura 2.4. Momentele cheie ale algoritmului de *automatizare* Facebook

2.1.2. Algoritmul de automatizare a deciziei asociat YouTube

YouTube a fost fondat în 2005, și este unul din cele mai populare site-uri de vizionare de videoclipuri din zilele noastre. Milioane de videoclipuri au fost încărcate și distribuite pe YouTube de la trailere ale filmelor până la videoclipuri făcute de amatori.

O dezvoltare mai recentă și nu atât de cunoscută este automatizarea Youtube[24]. În loc de folosirea videoclipurilor ca un instrument de marketing, modelul automatizat de afaceri se focusează pe crearea de videoclipuri virale care generează un profit pentru canalul utilizatorului. De exemplu, reclamele. Un utilizator, în videoclipurile pe care le încarcă pe YouTube poate pune reclame de câteva secunde, iar dacă alt utilizator de click pe aceste reclame și cumpără, de exemplu produsul respectiv, deținătorul videoclipului încărcat va lua bani de pe urma “plasării de produs”, și aici se întâmplă automatizarea. O firmă, plătește un youtuber să îi facă reclamă, iar în urma numărului de produse comandate folosind un anumit cod, persoana care a încărcat videoclipul va primi la rândul ei bani în funcție de câte produse s-au vândut folosind acel link.

În continuare, se prezintă Ecranul inițial YouTube Kids după descărcarea aplicației, conform figurii 2.1[24].

Primul ecran este o pagină de bun venit care vorbește imediat cu „responsabilul”. Al doilea ecran solicită vârsta nașterii tatălui sau a mamei, pentru ca aplicația să se asigure că este adult. Al treilea ecran întreabă dacă adultul dorește să activeze sau nu căutarea în interiorul aplicației. Al patrulea ecran este o întrebare matematică, urmată de două ecrane de instrucțiuni despre siguranța platformei, sistemul de recomandare automatizată și orientări pentru blocare și reclamație, în cazul în care se găsește un conținut neautorizat. Ecranul final spune: „de acum, controlul este transferat către dvs.”.

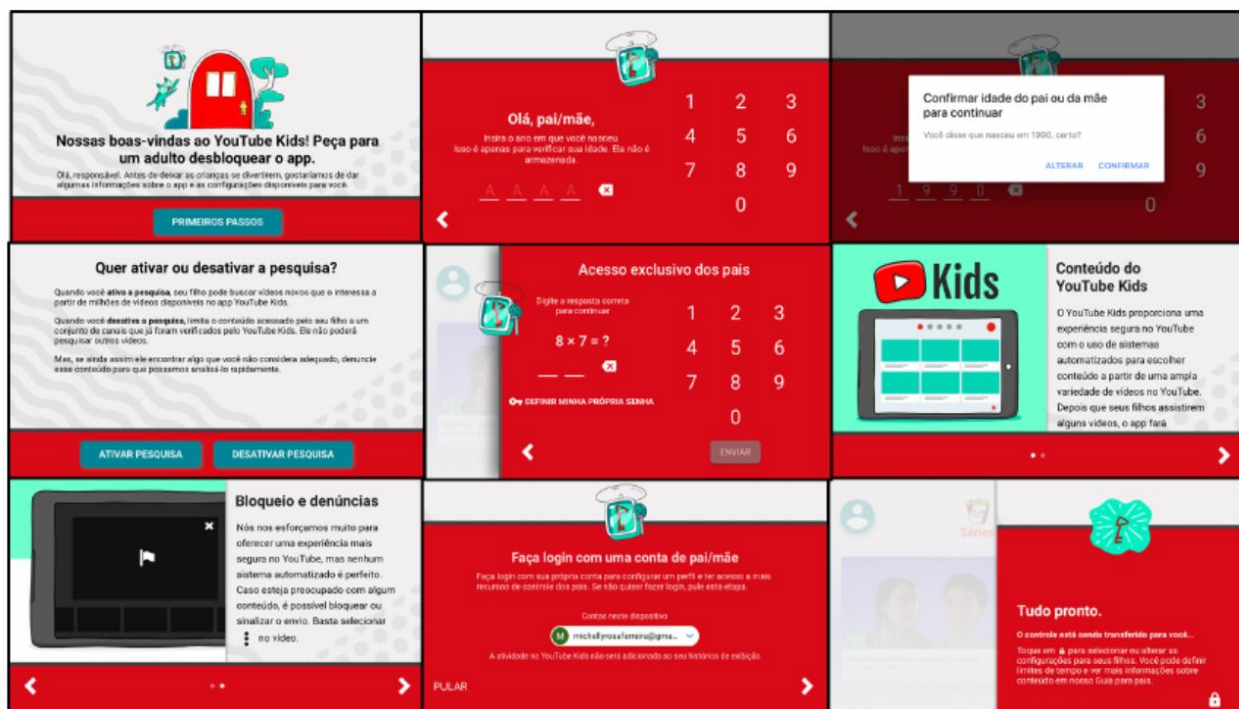


Fig. 2.1. Ecranul inițial YouTube Kids după descărcarea aplicației[24]

Algoritmul YouTube este un sistem automat care determină care videoclipuri sunt arătate atunci când utilizatorul revine pe site sau în aplicație. Ca și algoritmul Facebook prezentat anterior, algoritmul YouTube arată utilizatorului ceea ce dorește să vadă(sau ceea ce algoritmul crede că dorește să vadă).

Cu peste 400 de ore de videoclipuri încărcate în fiecare minut, YouTube are nevoie de o selecție automată pentru fiecare utilizator în parte. În algoritmul YouTube, există un sistem denumit ”caută și descoperă”, care nu doar decide care rezultat va fi văzut de utilizatori, dar de asemenea, determina care videoclip va fi arătat în secțiuni ale site-ului.

Există 6 chei principale ale algoritmului care influențează unde pot apărea videoclipurile utilizatorului: în rezultatul căutării, în pagina de videoclipuri recomandate, în pagina principală, în partea de videoclipuri populare ale momentului, în pagina de abonări ale utilizatorului, în bara de notificări.

Original, algoritmul YouTube era atent doar care videoclip poate avea un grad mai ridicat de vizionare, decât celelalte în ceea ce privește numărul de click-uri/videoclip. Dar vizionările afișau numai de câte ori a fost vizualizat un videoclip, acest fapt oferind satisfacție numai creatorilor de conținut ale căror videoclipuri aveau mai multe click-uri, netinându-se seama și de

câte minute a petrecut un utilizator pe respectivul videoclip. Astfel că, în 2012, YouTube a schimbat algoritmul în favoare creatorilor de conținut ale căror videoclipuri aveau mai multe minute de vizualizare, nemaitinându-se cont de numărul de click-uri.

În zilele noastre, algoritmul YouTube se concentrează pe cum interacționează audiența cu videoclipurile folosind inteligența artificială și automatizarea, care “învață din peste 80 de bilioane de biți de feedback de la utilizatori zilnic pentru a înțelege cum algoritmul să arate videoclipurile potrivite, la utilizatorul potrivit, la momentul potrivit”. Feedback-ul audienței include ce ar trebui și ce nu ar trebui utilizatorul să vadă, cât de mult timp petrece utilizatorul să vizioneze un videoclip(timpul de vizionare și timpul sesiunii pe fiecare vizita), aprecieri și displaceri.

2.1.3. Algoritmul de automatizare a deciziei asociat Instagram

Automatizarea platformei Instagram reprezintă utilizarea software-ului terților pentru a gestiona un cont, a îndeplini sarcini și / sau a interacționa cu utilizatorii fără un om prezent.

Instrumentele de automatizare Instagram pot fi de două tipuri:

- Servicii „care preiau în mod eficient” interacțiunile cu utilizatorii, precum aprecieri și comentarii
- Instrumente care *automatizează* Instagram în spatele scenei, având grijă de sarcini precum publicarea, raportarea și analiza

Sunt șanse ca un utilizator să fi întâlnit cel puțin o dată un „bot” Instagram. Dacă utilizatorul s-a confruntat vreodată cu spam de comentarii, vizualizări misterioase ale poveștilor sau urmăritori fantomă care dispar din senin, aceasta se poate considera ca activitatea unui „bot”.

Instagram se află de mulți ani în război împotriva „activității neautentice”. Au luat, de asemenea, măsuri pentru a închide instrumentele de automatizare Instagram utilizate pentru a atrage în mod artificial utilizatori prin aprecieri (cum ar fi închiderea Instagress în 2017) și „boții” care încalcă anumite reguli. Deși automatizarea Instagram atrage adesea o reputație proastă, nu toate instrumentele sunt spam.

Aceasta ne conduce la a doua categorie de instrumente de automatizare a rețelelor sociale, care se ocupă în mare parte de backend-ul conținutului utilizatorului.

Modul în care funcționează algoritmul de automatizare al platformei Instagram este descris în figura 2.5. Un utilizator realizează o postare. Acesta postare poate fi ștearsă de către utilizator în cazul în care s-a răzgândit și nu mai dorește ca postarea să mai fie publică, sau poate fi arătată la 10% din totalul urmăritorilor. Urmăritorii pot să comenteze și să aprecieze postarea ceea ce duce la arătarea postării la cât mai mulți utilizatori, sau pot să nu comenteze și nici să aprecieze postarea, ceea ce înseamnă faptul că utilizatorul care a realizat postarea nu va crește pe platforma Instagram. Dacă postarea este apreciată și este văzută de cât mai mulți utilizatori, aceștia pot să urmărească persoana care a realizat postarea respectivă sau pot alege să nu urmărească persoana respectivă, ceea ce duce la pierderea numărului de afișări ale postării. Utilizatorii care aleg să urmărească persoana care a realizat postarea, conduc la creșterea în topul de etichete și locații ale utilizatorului, care duce la creșterea acestuia, rezultând cât mai multe postări de la persoana respectivă.

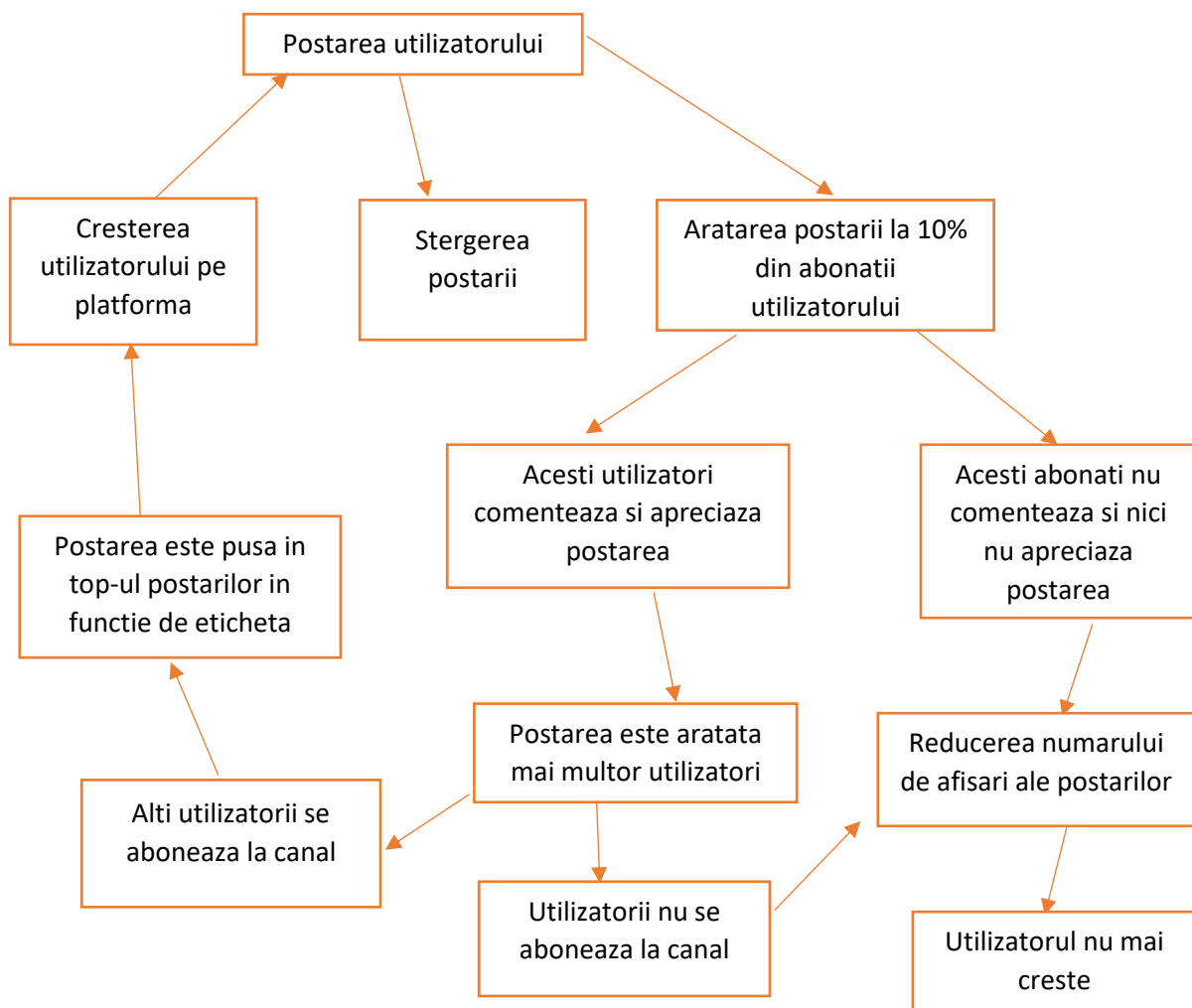


Figura 2.5. Modul de functionare a algoritmului de automatizare a platformei Instagram

De la automatizarea postărilor până la publicare, momentul în care urmăritorii sunt cei mai activi, până la descoperirea etichetelor pentru a primi cât mai multe aprecieri, se pot automatiza fragmente din contul Instagram fără ca utilizatorul să-și facă griji cu privire la spam sau la un avertisment de pe platformă. Acțiunile care se pot automatiza pe Instagram sunt prezentate în continuare:

- Aprecierile

Aprecierile sunt, fără îndoială, cea mai sigură interacțiune directă cu urmăritorii care poate să fie automatizată pe Instagram pentru ca toate aprecierile pe Instagram sunt forma universală de recunoaștere a platformei(acea inimioară care se dă la o postare).

Dacă un utilizator folosește un cont personal, este posibil să dea o apreciere la ceva fără să îi placă conținutul în sine (de exemplu: închiderea unei companii, un anunț serios etc.). Firmele de multe ori apreciază comentariile și întrebările clienților (indiferent dacă sunt pozitive sau nu).

Automatizarea aprecierilor va permite utilizatorului să își facă recunoscut contul mai multor urmăritori. Acestea fiind spuse, aprecierile reprezintă interacțiuni cu efort redus și, probabil, nu vor avea ca rezultat o mare aderență la contul unui utilizator prea mult.

- Comentarii

Automatizarea comentariilor pe Instagram nu este o idee foarte buna. De la comentarii potențial neadecvate (emoji-uri vesele pe o postare tragică) până la interacțiunea cu oameni care habar nu au cine sunt utilizatorii cărora le lasă mesaje, comentariile automate sunt chiar definiția spamului.

Acesta este exact motivul pentru care Instagram are la dispoziție procese pentru a bloca activitatea suspectă, cum ar fi comentariile automate. În teorie, interacțiunile mai frecvente, cum ar fi „Aprecieri” și comentarii, transmit vești bune algoritmului Instagram. În realitate totuși, automatizarea unor astfel de interacțiuni este un mod sigur de a fi raportat și blocat.

- Găsirea urmăritorilor

Poate că una dintre cele mai proactive utilizări ale instrumentelor de *automatizare* Instagram este descoperirea persoanelor pe care un utilizator dorește să le urmărească.

De exemplu, unele instrumente au capacitatea de a analiza conturi și etichete pentru a descoperi conturi relevante pentru publicul țintă. Fără a fi nevoie ca un utilizator să își facă griji cu privire la interacțiuni sau să se scufunde prin etichete, acest lucru poate fi un economisitor de timp sigur.

- A urmări sau a te dezabona de la anumite conturi

Urmează conceptul din spatele automatizării: cu cât un utilizator urmărește mai multe conturi, cu atât acel utilizator poate câștiga mai mulți urmăritori. Instrumentele de automatizare de multe ori au caracteristici care permit dezabonarea automată pentru a ajuta conturile să păstreze un raport mai bun între urmăritori.

La rândul său, răspândirea acestor tipuri de instrumente a dus, de fapt, la restricționarea urmăritorilor Instagram, astfel persoanele care se ocupa de gestionarea acestei platforme au impus câteva reguli:

- Pentru a ajuta la reducerea spamului, Instagram nu permite nimănui să urmărească mai mult de 7.500 de persoane.
- Oricine încearcă să urmărească peste 7.500 de persoane va vedea un mesaj de eroare, indiferent de câți urmăritori are. Oricine urmărește în prezent peste 7.500 de persoane a făcut acest lucru înainte ca modificarea numărului maxim de urmăritori să fie făcută publică.

- Mesaje directe (DM)

DM-urile automate sunt similare comentariilor *automatizate* din punct de vedere al spam-ului. Dacă un utilizator a avut vreodată de-a face cu DM-uri nesolicitate care ajung în „Solicitări de mesaje”, este probabil ca acestea să provină de la boți sau de la conturi automate. Aceste tipuri de mesaje automate sunt de obicei impersonale și probabil vor fi ignorate de cititori.

- Vizualizările poveștilor

Similar cu „Aprecierile”, vizionările poveștilor sunt o interacțiune cu efort redus, care nu fac mare lucru pentru a perturba alți utilizatori. Acestea fiind spuse, a avea conturi aparent aleatorii care se uită la Poveștile utilizatorului poate părea puțin ciudat. Din nou, acest concept care

păcăleşte cumva oamenii să se uite la un anumit cont este un lucru total greşit din punct de vedere etic şi nu este corect faţă de ceilalţi utilizatori.

- Programarea şi publicarea postărilor

Pentru a ţine pasul cu nevoia de a publica în mod constant pe Instagram, automatizarea poate însemna programarea conţinutului unui utilizator pe contul sau. Funcţii precum postarea încrucişată(partajarea conţinutului) din/în alte reţele de socializare şi timpul optimizat(ora postării) sunt deosebit de utile şi răspândite.

- Analiza contului

Raportul şi analiza contului pot fi automatizate pe Instagram. Se pot realiza rapoarte ale timpului petrecut pe instagram(conform figurii 2.6.a), a numărului de vizualizări ale canalului, sau chiar a link-urilor pe care un utilizator le-a deschis din Instagram(figura 2.6.b). Acest lucru poate ajuta companiile să vadă dacă au ales benefic un creator de conţinut şi suma pe care acest creator o vă primi în urma colaborării cu firma respectivă în funcţie de câte persoane accesează acel link.



a)

b)

Fig. 2.6. Analiza unui cont al platformei Instagram

Avantajele și dezavantajele automatizării Instagram

În continuare vom analiza cazurile contra și împotriva *automatizării* unui cont de Instagram.

Avantaje:

- economisirea timpului la realizarea sarcinilor mai mici și obositoare

Acesta este, cel mai mare avantaj. Creșterea prezenței Instagram în mod organic necesită cu siguranță o combinație de timp și strategii. A veni cu o strategie de conținut în sine este un angajament imens, înainte ca un utilizator să își facă griji cu privire la interacțiuni, cum ar fi comentarii și servicii pentru clienți.

Instrumentele de automatizare pot prelua o parte din sarcinile mai mici pe care utilizatorul le are de făcut. După cum s-a evidențiat mai sus, astfel de instrumente nu pot face totul, dar cu siguranță pot reduce timpul petrecut pentru sarcini obositoare.

- creșterea numărului de urmasitori mai repede

Automatizarea ajută fără îndoială utilizatorul să sporească numărul de urmăritori. Pur și simplu un creator de conținut trebuie să petreacă timp pe platformă și să interacționeze cu alte conturi în orice mod pentru a-și aduce la rândul lui urmăritori pe contul său. Desigur, aici este și un minus pentru utilizatorii cu care intractionati s-ar putea să nu fie interesați de alte conturi.

- descoperirea de informatii importante despre audiența contului

Nu există nicio îndoială că instrumentele de automatizare vor ajuta utilizatorul să afle mai multe despre publicul țintă. De la vizualizarea etichetelor pe care le folosesc oamenii până la sugerarea conturilor de urmat, roboții încurajează realizarea unei aprofundări în preferințele publicului canalului. Acesta este cu siguranță un lucru pozitiv.

Dezavantaje:

- Instagram nu este un fan al programului de *automatizare*

După cum am menționat mai înainte, Instagram are o istorie a instrumentelor pe care le consideră spam. Platforma își dorește ca utilizatorii să se angajeze în mod autentic în acțiunile pe care le realizează. Instagram are limitele și penalitățile sale pentru orice tactică pe care o consideră neoficială algoritmului lor.

- *automatizarea* slabă a luarii deciziilor arată performanțe îndoielnice ale sistemului

Utilizatorul trebuie să acorde credit publicului lui deoarece oamenii pot observa dintr-o privire un „bot”.

Automatizarea aprecierilor și comentariilor este o rețetă pentru o potențială criză de social media. Conturile de marcă ar trebui să sublinieze interacțiunile umane și să evite orice arată a spam.

Achiziționarea de aprecieri pe Instagram nu este un aspect bun pentru creatorii de conținut dar nici nu încearcă să promoveze un angajament neautentic.

Câteva exemple legitime de automatizare Instagram vor fi prezentate în continuare

Programarea postărilor bazată pe momentul în care oamenii sunt cei mai activi este unul din exemplele de automatizare acceptată de Instagram. Cu atât de mult conținut pe Instagram, orice porțiță pentru a ajunge la alți utilizatori atunci când sunt cei mai activi pe platforma este un plus.

Acesta este un exemplu foarte bun despre modul în care automatizarea și inteligența artificială[7] pot ajuta marketerii sociali să devină mai eficienți.

Un alt exemplu este analiza etichetelor pentru a descoperi noi urmăritori.

Săpatul prin etichetele Instagram pe cont propriu poate fi o pierdere de timp. Pe de altă parte, există site-uri care pot analiza automat contul și abonații pentru a descoperi etichete relevante pentru afacerea unei persoane. Aceste etichete se pot utiliza pentru a stimula conținutul sau pentru a descoperi potențialii clienți mai repede.

În concluzie, există multe instrumente de automatizare pe Instagram, dar condiția principală este că un utilizator să folosească aceste trucuri în mod corespunzător.

2.2. Studiul literaturii dedicate algoritmului YouTube

YouTube reprezintă unul dintre cele mai mari sisteme de recomandare industrială și cele mai sofisticate existente. YouTube este cea mai mare platformă din lume pentru crearea, partajarea și descoperirea de conținut video. Automatizarea algoritmului și recomandările YouTube sunt responsabile pentru a ajuta mai mult de un miliard de utilizatori să descopere conținut personalizat dintr-o bază de videoclipuri în continuă creștere. Acest capitol se va concentra asupra impactului imens pe care l-a avut recent învățarea profundă asupra sistemului de recomandări video YouTube. Figura 1.1. ilustrează recomandările pentru pagina de pornire a aplicației mobile YouTube. Recomandarea videoclipurilor YouTube este extrem de provocatoare din trei perspective majore:

- **Scala:** Mulți algoritmi de recomandare existenți dovediți că funcționează bine pe probleme mici nu reușesc să funcționeze la scara YouTube. Algoritmii de învățare distribuită foarte specializați și sistemele de servire eficienta sunt esențiale pentru gestionarea bazei masive de utilizatori YouTube.

- **Tendențele:** YouTube are o baza foarte dinamică cu multe ore de videoclipuri care sunt încărcate pe secundă. Sistemul de recomandare ar trebui să fie suficient de receptiv pentru a modela conținutul nou încărcat, precum și cele mai recente acțiuni întreprinse de utilizator. Echilibrarea conținutului nou cu videoclipuri bine stabilite poate fi înțeleasă din perspectivă explorării / exploatării.

- **Zgomot:** comportamentul istoric al utilizatorilor pe YouTube este inerent, dificil de prezis din cauza rarității și a unei varietăți de factori externi neobservabili. Rareori obținem adevărul de bază al satisfacției utilizatorilor și, în schimb, modelăm semnale de feedback implicite zgomotoase. Mai mult, metadatele asociate conținutului sunt slab structurate fără o ontologie bine definită. Algoritmii de automatizare YouTube trebuie să fie robuști față de aceste caracteristici particulare ale datelor de instruire. [3]

În conjugare cu alte domenii de produse de pe Google, YouTube a suferit o schimbare fundamentală de paradigmă către utilizarea învățării profunde ca soluție generală pentru aproape toate problemele de învățare. Sistemul este construit pe Google Brain [2], care a fost deschis recent sub denumirea de TensorFlow [1]. TensorFlow oferă un cadru flexibil pentru experimentarea

diferitelor arhitecturi de rețea neuronală profundă utilizând instruire distribuită pe scară largă. Modelele învață aproximativ un miliard de parametri și sunt instruiți pe sute de miliarde de exemple.

Spre deosebire de o mare cantitate de cercetări în metodele de factorizare a matricii [5], există relativ puține lucrări folosind rețele neuronale profunde pentru sistemele de recomandare. Rețelele neuronale sunt folosite pentru recomandarea de știri [3], citate [4] și evaluări de recenzii [6]. Filtrarea colaborativă este formulată ca o rețea neuronală profundă și autoencodere [8]. Elkahky a folosit învățarea profundă pentru modelarea pe mai multe domenii în ceea ce privește utilizatorii[9]. Într-un cadru bazat pe conținut, Burges a folosit rețele neuronale profunde pentru recomandarea muzicii [10].

2.2.1. Prezentare generală a algoritmului de automatizare din spatele YouTube

Structura generală a sistemului prezentat de recomandare este ilustrată în Figura 2.7. Sistemul automat este alcătuit din două rețele neuronale: una pentru generația de candidați și una pentru clasare. Rețeaua de generare candidat are ca intrare evenimente din istoricul activității YouTube a utilizatorului și preia un subset mic (sute) de videoclipuri dintr-un corpus mare. Acești candidați sunt setați să fie în general relevanți pentru utilizator și să aibă precizie ridicată. Rețeaua de generare candidat oferă o personalizare largă doar prin filtrare colaborativă. Similitudinea dintre utilizatori se exprimă în termeni de caracteristici grosiere, cum ar fi ID-urile de vizionări video, jetoane de interogare, de căutare și date demografice.

Prezentarea câtorva dintre „cele mai bune” recomandări într-o listă, necesită o reprezentare la nivel fin pentru a distinge importanța relativă dintre candidații cu vizualizare ridicată. Rețeaua de clasare folosește un algoritm de automatizare deoarece realizează această sarcină prin atribuirea unui scor fiecărui videoclip în raport cu o funcție obiectivă dorită, folosind un set bogat de caracteristici care descriu videoclipul și utilizatorul. Videoclipurile cu cele mai mari punctaje sunt prezentate utilizatorului, clasate după punctajul lor.

Abordarea în două etape a recomandării ne permite să facem recomandări dintr-un corpus foarte mare (milioane) de videoclipuri prin automatizare, fiind totuși siguri că numărul mic de videoclipuri care apar pe dispozitiv e personalizat și atractiv pentru utilizator.

În timpul dezvoltării, se utilizează pe scară largă valorile offline (precizie, rechemare, etc.) pentru a ghida îmbunătățirile iterative ale sistemului. Cu toate acestea, pentru determinarea finală a eficacității unui algoritm de automatizare sau model, la baza stă testarea A / B prin experimente live. Într-un experiment live, se pot măsura modificări subtile ale ratei de click, timpul de vizionare și multe alte valori care măsoară implicarea utilizatorilor. Acest lucru este important deoarece rezultatele live A / B nu sunt întotdeauna corelate cu experimentele offline.

După ce utilizatorul vizionează câteva din videoclipurile recomandate de YouTube, se realizează un feedback pe baza videoclipurilor vizionate, astfel încât se produce automatizarea procesului. Toate videoclipurile recomandate de YouTube și vizualizate de către utilizator, întorc un rezultat, un istoric al utilizatorului pentru ca algoritmul să învețe din click-urile pe care le da utilizatorul.

Burroughs definește publicitatea pe YouTube practic prin două strategii: prima este introducerea unor mici pauze de reclame în videoclipuri, plătiind creatorii de videoclipuri pentru fiecare vizualizare, iar a doua este un sistem de promovare a acestor videoclipuri, prin recomandare automată, utilizând algoritmi de comportament și rang [24]. Conform studiului său, algoritmi de automatizare sunt formule sau coduri social-tehnice în interiorul unei infrastructuri, care combină informațiile de date, procesează preferințele utilizatorului și în urma acestei procesări sunt redade rezultate/recomandări.

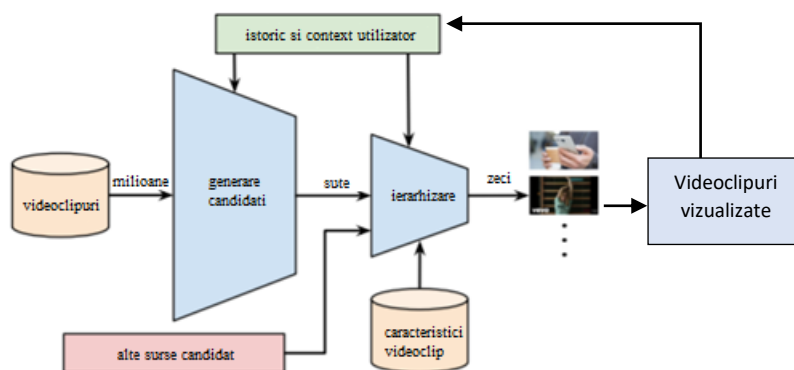


Figura 2.7. Arhitectura unui sistem automatizat de recomandări care demonstrează modul în care videoclipurile sunt primite și ierarhizate înainte de a prezenta doar o parte utilizatorului

2.2.2. Recomandarea videoclipurilor sub formă de clasificare

Deși există mecanisme explicite de feedback pe YouTube (degetul mare în sus / în jos, sondaje în cadrul produselor etc.), se folosește feedback-ul implicit [11] al vizualizărilor pentru a instrui modelul, unde un utilizator care duce la bun sfârșit un videoclip este un exemplu pozitiv. Această alegere se bazează pe ordinele de mărime disponibile, adică istoricul utilizatorilor, permițându-ne să producem recomandări adânc în coadă, unde feedback-ul explicit este extrem de rar. Intrarea sistemului în acest caz este un videoclip care e dus la bun sfârșit de utilizatorul. Acest videoclip poate fi notat cu valoarea 1 pentru ca algoritmul de automatizare să rețină alegerea utilizatorului, pentru ca pe viitor același utilizator să primească videoclipuri din aria lui de interes, care sunt, de asemenea, ieșirile sistemului.

O abordare populară este abordarea ierarhică [12]. În abordarea ierarhică, parcurgerea fiecărui nod din arbore implică discriminarea între seturi de clase care sunt adesea lipsite de legătură, ceea ce face ca problema clasificării să fie mult mai dificilă și performanțele să fie degradante.

Atunci când utilizatorul deschide platforma, trebuie calculate cele mai probabile N clase (videoclipuri) pentru a alege N -ul de top care trebuie prezentat utilizatorului. Notarea a milioane de articole sub o latență strictă de servire a videoclipurilor de zeci de milisecunde necesită o schemă de notare aproximativă, subliniată în numărul de clase. Sistemele de automatizare anterioare de pe YouTube se bazau pe criptare [13], iar clasificatorul descris aici folosește o abordare similară. Întrucât probabilitățile din stratul de ieșire nu sunt necesare la momentul difuzării, problema de notare se reduce la o căutare de vecinătate - cea mai apropiată în spațiul produsului pentru care pot fi utilizate biblioteci cu scop general [14]. A fost constatat faptul că rezultatele A / B (Testarea A / B (cunoscută și sub numele de bucket testing sau testarea split-run)) este o metodologie de cercetare a experienței utilizatorului. Testele A / B constau dintr-un experiment aleator cu două variante, A și B care include aplicarea testării de ipoteze statistice sau a „testării de ipoteze cu două eșantioane”, așa cum este utilizat în domeniul statisticii. Testarea A / B este o modalitate de a compara două versiuni ale unei singure variabile, de obicei testând răspunsul unui subiect la varianta A față de varianta B și determinând care dintre cele două variante

este mai eficientă.[16]) nu au fost deosebit de sensibile la alegerea celui mai apropiat algoritm de căutare a vecinilor.

Istoricul căutărilor unui utilizator este tratat în mod similar cu istoricul vizionărilor - fiecare interogare este simbolizată în unigrame și bigrame și fiecare simbol este încorporat. Odată mediate, interogările simbolizate și încorporate ale utilizatorului reprezintă un istoric de căutare dens rezumat. Această parte a algoritmului de automatizare este asemănătoare cu cea prezentată anterior, numai ca datele de intrare nu vor mai fi preluate din feedback-ul pe care îl da utilizatorul, ci vor fi aduse din istoricul căutărilor pe care un utilizator al platformei le va realiza.

Caracteristicile demografice sunt importante pentru furnizarea priorităților, astfel încât recomandările să se comporte în mod rezonabil pentru noii utilizatori. Regiunea geografică și dispozitivul utilizatorului sunt încorporate și concatenate. Funcțiile binare simple și continue, cum ar fi sexul utilizatorului, starea sau zona de conectare și vârsta sunt introduse direct în algoritm ca valori reale normalizate la $[0, 1]$ pentru ca automatizarea rezultatelor să se realizeze cât mai aproape de realitate.

2.2.3. Caracteristica „Exemplu de vârstă”

Videoclipuri care au multe ore sunt încărcate în fiecare secundă pe YouTube. Recomandarea acestui conținut recent încărcat este extrem de importantă pentru YouTube că produs.

Se observa în mod constant că utilizatorii preferă conținut recent, în detrimentul relevanței. În plus față de efectul de primă ordine de a recomanda pur și simplu videoclipuri noi pe care utilizatorii doresc să le vizioneze, există un fenomen secundar critic de bootstrapping și propagare a conținutului viral [20].

Sistemele de învățare automată, de automatizare prezintă adesea o părtinire implicită față de trecut, deoarece sunt instruite să prezică comportamentul viitor din exemple istorice. Distribuția popularității videoclipurilor este extrem de nestăționară, dar distribuția multinomială asupra corpusului produs de recomandator va reflecta probabilitatea medie a vizionării în fereastra de antrenament pe parcursul a câteva săptămâni. Pentru a corecta acest lucru, se alege vârsta

exemplului de antrenament că o caracteristică în timpul antrenamentului. La timpul de vizionare, această caracteristică este setată la zero (sau ușor negativă) pentru a reflecta faptul că modelul face predicții chiar la sfârșitul ferestrei de antrenament. [22]

Figura 2.4 demonstrează eficacitatea acestei abordări pe un videoclip ales în mod arbitrar [21].

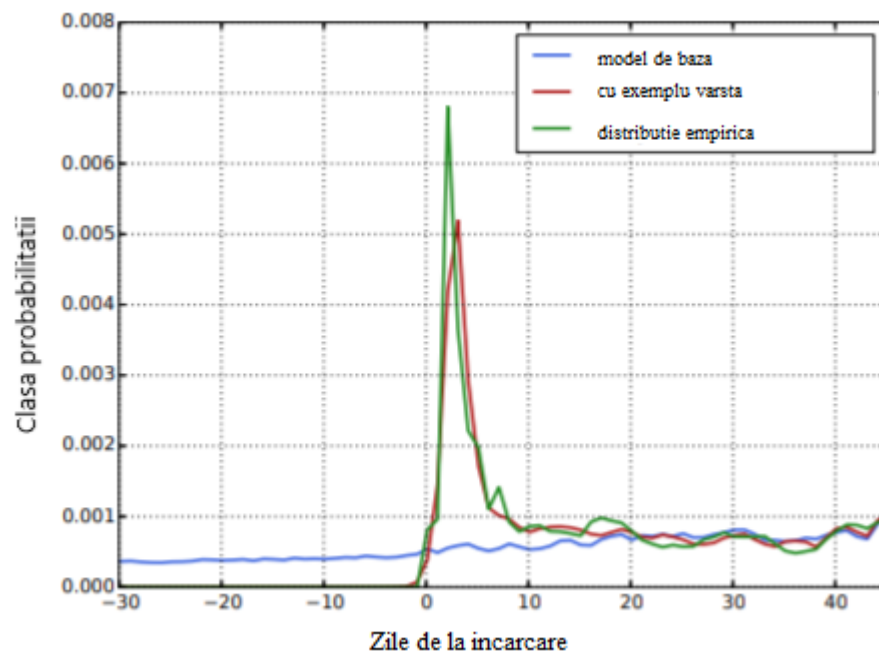


Figura 2.4. Pentru un anumit videoclip, modelul instruit cu exemplu de vârstă ca funcție este capabil să reprezinte cu exactitate timpul de încărcare și popularitatea dependentă de timp observate în date. Fără caracteristică, modelul ar prezice aproximativ probabilitatea medie peste fereastra de antrenament[3]

2.2.4. Selectarea etichetei și a contextului

Este important de subliniat faptul că automatizarea sistemului de recomandări de videoclipuri implică adesea rezolvarea unei probleme asemănătoare cu o problemă anterioară și transferarea rezultatului într-un anumit context. Un exemplu clasic este presupunerea faptului că prezicerea corectă a evaluărilor conduce la recomandări eficiente de filme. Se constată faptul că

alegerea acestei probleme de automatizare are o importanță foarte mare asupra performanței în testarea A / B, dar este foarte dificil de descoperit cu experimente offline.

Exemple de instruire ale algoritmului de automatizare sunt generate de toate vizualizările YouTube (chiar și cele încorporate pe alte site-uri) nu doar de vizualizarea recomandărilor care sunt pe platforma YouTube. În caz contrar, ar fi foarte dificil să apară un conținut nou. Dacă utilizatorii descoperă videoclipuri prin alte mijloace decât cele care se află direct pe platforma, se poate răspândi rapid această descoperire către alți utilizatori prin filtrarea colaborativă. O altă perspectivă cheie care a îmbunătățit valorile live a fost generarea unui număr fix de exemple de instruire pe utilizator, ponderând în mod eficient și egal utilizatorii în funcția de pierdere.

Oarecum contra-intuitiv, trebuie acordată o atenție deosebită reținerii informațiilor de la clasificator, pentru a împiedica modelul să exploateze structura site-ului. Să luăm ca exemplu un caz în care utilizatorul tocmai a emis o interogare de căutare pentru „taylor swift”. Întrucât problema prezentată se arată că o prezicere a următorului videoclip vizionat, un clasificator dat de aceste informații vă prezice faptul că cele mai probabile videoclipuri care vor fi vizionate sunt cele care apar pe pagina cu rezultatele căutării corespunzătoare pentru „taylor swift”. În mod surprinzător, reproducerea ultimei pagini de căutare a utilizatorului ca recomandări pentru pagina de pornire are o performanță foarte slabă.

Prin eliminarea informațiilor despre secvență și reprezentarea interogărilor de căutare cu un pachet de jetoane neordonate, clasificatorul nu mai este în mod direct conștient de originea etichetei.

Modelele naturale de consum/vizualizare ale videoclipurilor duc de obicei la probabilități foarte asimetrice de vizionare în comun. Seriile episodice sunt de obicei urmărite secvențial, iar utilizatorii descoperă deseori artiști într-un gen, începând cu cel mai popular, în general, înainte de a se concentra pe nișe mai mici. Prin urmare, s-a găsit o performanță mult mai bună care prezice următoarea vizualizare a utilizatorului, mai degrabă decât prezicerea unei vizualizări de durată aleatorie. Multe sisteme de filtrare colaborativă aleg implicit etichetele și contextul prin durata unui element aleatoriu și predicția acestuia din alte elemente din istoricul utilizatorului. Aceasta scurge informații viitoare și ignoră orice tipare de consum asimetrice. În schimb, „se derulează” istoricul unui utilizator alegând o vizualizare aleatorie și introducând doar acțiunile pe care le-a făcut utilizatorul înainte de vizualizarea etichetei de durată. Alegerea etichetelor și a contextului

de intrare pentru algoritmul de automatizare este dificil de evaluat offline, dar are un impact mare asupra performanței live. Aici, evenimentele solide sunt caracteristici de intrare în rețea, în timp ce evenimentele goale sunt excluse. Am constatat că prezicerea unui viitoare vizionări este mai bună în testarea A / B.

2.2.5. Clasamentul

Rolul principal al clasamentului este acela de a utiliza datele despre impresiile utilizatorului pentru a specializa și calibra predicțiile candidat pentru interfața respectivă cu utilizatorul. De exemplu, un utilizator poate viziona un anumit videoclip cu probabilitate mare de vizionare, dar este puțin probabil să facă click pe videoclipul respectiv din cauza alegerii greșite imaginii miniaturale. În timpul clasării videoclipurilor, un utilizator are acces la multe alte funcții care descriu videoclipul și relația utilizatorului cu videoclipul, deoarece în ierarhizare sunt înregistrate doar câteva sute de videoclipuri, față de milioanele înregistrate în generația candidată, conform figurii 2.3. Clasarea este, de asemenea, crucială pentru adunarea diferitelor surse candidat ale căror scoruri nu sunt direct comparabile.

Pentru a exemplifica acest lucru se poate folosi o rețea neuronală profundă cu arhitectură similară cu cea a generării candidat pentru a atribui un scor independent fiecărei impresii video. Lista videoclipurilor este apoi sortată după acest scor și returnată utilizatorului. Obiectivul final de clasare este în permanență reglat pe baza rezultatelor testării live A / B, dar este, în general, o funcție simplă a timpului de vizionare așteptat pe afișare. Clasarea în funcție de rata de click promovează adesea videoclipuri înșelătoare pe care utilizatorul nu le finalizează („clickbait”), în timp ce timpul de vizionare captează mai bine implicarea [17].

2.2.6. Reprezentarea caracteristicilor

În mod obișnuit, se folosesc sute de funcții în algoritmii de automatizare, repartizate aproximativ uniform între simple și continue. În ciuda promisiunii unei învățări profunde de a atenua sarcina inginerescă, natura datelor brute nu se pretează cu ușurință să fie introduse direct în rețelele neuronale avansate. Se consumă încă resurse ingineresti considerabile transformând

datele utilizatorilor și ale videoclipurilor în caracteristici utile. Principala provocare constă în reprezentarea unei secvențe temporale de acțiuni ale utilizatorilor și a modului în care aceste acțiuni arată un scor pe care utilizatorul îl da videoclipului în funcție de impresia lăsată. Se observă că cele mai importante semnale sunt cele care descriu interacțiunea anterioară a unui utilizator cu elementul în sine și alte elemente similare, potrivitându-se cu experiența altor utilizatori în clasarea videoclipurilor [4]. De exemplu, se ia în considerare istoricul trecut al utilizatorului cu canalul care a încărcat videoclipul înscris – numărul de videoclipuri urmărite de utilizator de pe acest canal, timpul când utilizatorul a vizionat un videoclip despre acest subiect. Aceste caracteristici continuate care descriu acțiunile anterioare ale utilizatorilor asupra articolelor conexe, alăturate sunt deosebit de puternice, deoarece se generalizează bine între elementele dispartate, lipsite de legătură. De asemenea, se consideră că este esențial să se propage informațiile de la generarea candidaților în clasament sub formă de caracteristici, de ex. sursele care au desemnat utilizatorul să vizioneze acel videoclip, scorurile atribuite de utilizator acelui videoclip, etc. Funcțiile care descriu frecvența afișărilor video din trecut sunt, de asemenea, esențiale pentru introducerea rezultatelor în recomandări. Dacă unui utilizator i s-a recomandat recent un videoclip, dar nu l-a vizionat, atunci algoritmul de automatizare va retrage în mod firesc această impresie la următoarea încărcare a paginii. Istoricul vizionărilor este o faptă inginerescă în sine care este vitală pentru a produce recomandări receptiv. Având în vedere exemplele de antrenament care sunt fie pozitive (s-a dat click pe videoclip), fie negative (nu s-a făcut click pe videoclip), se prezice timpul de vizionare așteptat până când un utilizator dă click pe un videoclip. Exemplele pozitive sunt adnotate cu timpul petrecut de utilizator vizionând videoclipul.

2.3. Cercetări practice privind algoritmi de automatizare prezentați

2.3.1. Algoritmul de automatizare YouTube

Acest subcapitol va începe cu modul în care se prezintă și modul în care funcționează recomandările cu YouTube.

YouTube este acum deținut de Google și, având în vedere obsesia Google pentru învățarea profundă, nu ar trebui să fie surprinzător faptul că YouTube a apelat la învățarea profundă și respectiv, algoritmi de automatizare pentru a-și susține recomandările. YouTube are câteva

provocări specifice, care trebuie reținute înainte că un utilizator să se decidă să aplice algoritmiilor de automatizare la propriile probleme. Dacă nu se întâmpină aceleași probleme, este posibil ca aceeași soluție să nu fie cea potrivită pentru respectiva persoană.

Prima problemă principală a platformei este scala, deoarece YouTube are peste un miliard de utilizatori, care urmăresc zilnic cinci miliarde de videoclipuri. Acest lucru este incredibil atunci când ne gândim pentru câți utilizatori și articole trebuie să calculeze recomandări algoritmiile existenți. Orice click realizat de utilizator trebuie să fie scalat de către algoritmul de automatizare și distribuit utilizatorilor extrem de eficient.

Următoarea problemă este noutatea videoclipurilor, deoarece peste 300 de ore de videoclipuri sunt încărcate pe YouTube în fiecare minut, iar algoritmiile de automatizare trebuie să țină seama atât de conținutul nou încărcat, cât și de datele acțiunilor individuale ale fiecărui utilizator care ar putea indica interesul lor despre un anumit lucru. Trebuie găsit un echilibru deliberat între recomandarea videoclipurilor și a conținutului nou, dar în mod clar totul se duce în favoarea conținutului nou.

Ultima mare problemă este zgomotul, majoritatea datelor de interes ale platformei constând din evaluări implicite, nu evaluări explicite. Așa cum am vorbit în subcapitolul anterior, datele de evaluare implicite sunt pline de pericole. Doar bazându-ne pe recomandările de videoclipuri pe care fac click utilizatorii se creează un sistem care este supus jocurilor și factorilor externi neobservabili. Cineva care face click pe un videoclip viral nu este neapărat o indicație a adevărului fundamental al intereselor acelui utilizator. Datele acestor utilizatori sunt, de asemenea, extrem de rare, iar atributele de conținut asociate videoclipurilor sunt structurate destul de slab fără o ontologie bine definită. Acesta este modul politicos al platformei YouTube de a spune că datele cu care lucrează sunt foarte greu de manevrat. Dar, pe baza rezultatelor întreprinse, se pare că algoritmiile de automatizare lucrează foarte bine în depășirea acestor provocări.

Structura generală a sistemului pentru generarea recomandărilor video de pe pagina de pornire mobilă YouTube a unui utilizator este ilustrată în Figura 2.8. Sistemul este alcătuit din trei faze: (1) generație candidată, în care elementele care apar în pagina de recomandări sunt selectate din milioane de videoclipuri, (2) clasarea, care comandă elementele care apar în pagina de recomandări și (3) politica, care impune nevoile de afaceri, cum ar fi cererea ca un anumit conținut să apară la o anumită poziție a paginii. Fazele (1) și (2) utilizează ambele rețele neuronale profunde.

Generarea de candidați este influențată în mod substanțial de comportamentul anterior al utilizatorului în sistem și calculează procentul în care se potrivesc videoclipurile cu preferințele utilizatorului. De exemplu, utilitatea comună este o măsură care se folosește: dacă un utilizator s-a bucurat de videoclipul A și mulți alți utilizatori care s-au bucurat de A s-au bucurat și de B, atunci B ar putea fi selectat în faza de generație candidată. Faza de clasare folosește, de asemenea, intens caracteristicile utilizatorilor, dar se bazează în plus pe caracteristici mai bogate ale articolelor (cum ar fi încorporările videoclipului într-un spațiu semantic). Faza de clasare tinde să ofere videoclipuri similare, predicții similare, ducând la fluxuri care au conținut repetitiv și, adesea, care rulează videoclipuri foarte similare. Pentru a atenua problema redundanței, la început, a fost introdusă o euristică în stratul nou, denumit „policy layer”, în care noi videoclipuri sunt adăugate printre videoclipurile pe care algoritmul de automatizare știe deja că vor fi plăcute de către utilizator, cum ar fi cerința că un utilizator individual să poată contribui cu mai mult de n articole în pagina de recomandări a oricărui utilizator. Deși această regulă este oarecum eficientă, aceasta interacționează destul de slab cu sistemul de recomandare de bază. Întrucât generația candidată și straturile de clasare nu sunt conștiente de această euristică, se fac predicții suboptimale irosind spațiu pe elementele care nu vor fi prezentate niciodată. Mai mult, pe măsură ce primele două straturi evoluează în timp, trebuie să se regăsească în mod repetat parametrii euristicii - o sarcină care este costisitoare și, prin urmare, în practică nu se realizează cu suficientă frecvență pentru a menține o mare parte din eficacitatea regulii. În cele din urmă, interacțiunile dintre mai multe tipuri de euristică produc, în practică, un algoritm de recomandare care este foarte greu de înțeles. Rezultatul este un sistem suboptim și dificil de evoluat.

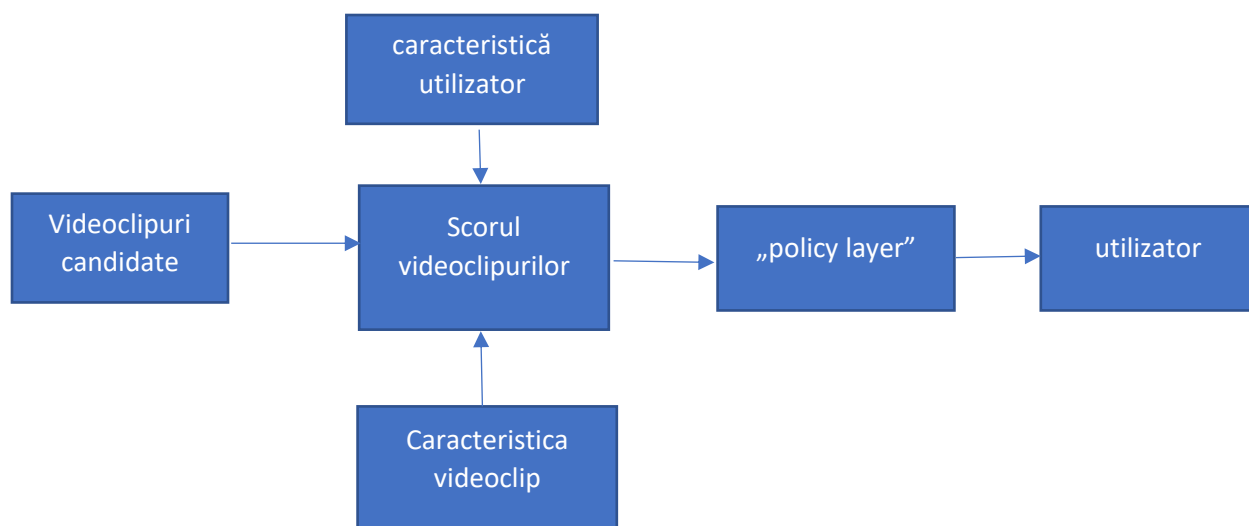


Figura 2.8. Structura generală a sistemului pentru generarea recomandărilor video de pe pagina de pornire mobilă YouTube a unui utilizator

În schimb, noua structura a sistemului pentru generarea recomandărilor video de pe pagina de pornire mobilă YouTube a unui utilizator are o abordare care utilizează procese de punct determinant (DPP). DPP este un model de recomandare setat care necesită doar două elemente: cât de bun este fiecare videoclip pentru utilizator și cât de asemănătoare sunt fiecare pereche de videoclipuri. Drept urmare, accentul se pune pe eliminarea redundanței. Se inserează DPP-urile înainte de stratul de politică, dar după stratul de scor al videoclipurilor (a se vedea Figura 2.9). Acest lucru permite să valorificarea investiției într-un marcator punctual extrem de sofisticat și, de asemenea, asigurăm faptul că politicile de afaceri sunt respectate.

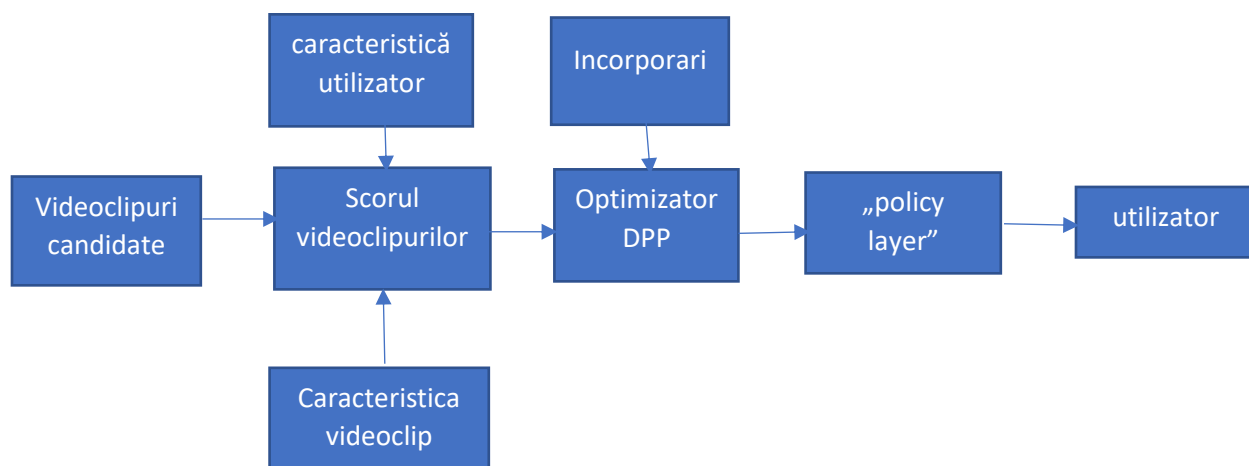


Figure 2.9. Structura noua a sistemului pentru generarea recomandărilor video de pe pagina de pornire mobilă YouTube a unui utilizator

Când un utilizator vizitează pagina de pornire YouTube, apar imediat recomandări care fac o treabă excelentă în captarea intereselelor actuale a utilizatorilor. Google a adoptat cu totul învățarea profundă, cu ideea de a o utiliza pentru aproape fiecare problemă de învățare automată pe care o întâmpină. Acest lucru permite concentrarea efortului ingineresc pe cadre de învățare profundă și folosirea lor ca instrument de scop general pentru a fi utilizat și de subcompaniile Google, cum ar fi YouTube. Astfel, în 2016 YouTube și-a mutat recomandările într-un cadru de învățare profundă, dezvoltat de Google Brain, care astăzi se numește TensorFlow. Anterior, abordarea lor se baza pe factorizarea matricială.

După cum am menționat anterior, se poate implementa factorizarea matricială cu o rețea neuronală, iar la început Google, a făcut exact acest lucru. Dar, în timp, abordarea lor a evoluat dincolo de factorizarea matricială.

Este interesant faptul că, deși YouTube are evaluări explicite cum ar fi degetul mare în sus / degetul mare în jos, algoritmi de automatizare nu le folosesc deloc pentru a genera recomandări, deoarece aceste date sunt prea rare. Nu sunt suficienți utilizatori care evaluează videoclipurile în mod explicit pentru că datele să fie utile. În schimb, algoritmi de automatizare YouTube, se bazează pe semnale implicite, cum ar fi videoclipurile pe care un utilizator le-a vizionat efectiv și căutat. Cu toate acestea, vizualizarea implicită și datele de căutare sunt în sine rare și, așa cum am spus în subcapitolul anterior, tratarea acestei rarități este o problemă imensă atunci când se încearcă să se aplice învățarea profundă la sistemele automate de recomandare.

Soluția găsită de dezvoltatori a fost împărțirea reprezentării rare a ID-urilor videoclipurilor și a jetoanelor de căutare pentru fiecare utilizator, într-o secvență de lungime variabilă de date rare, mapate pe un strat de o lungime fixă, adecvat pentru intrarea unei rețele neuronale într-un strat de încorporare. Pur și simplu, se mediază fiecare bucată din datele de intrare rare. Altfel spus, se împart datele de comportament rare ale fiecărui utilizator, în bucăți de lungime fixă și se ia media fiecărei bucăți, pentru a reduce acele date într-un strat de încorporare de lungime fixă, care poate fi folosit ca intrare în rețeaua neuronală, fapt ce restricționează această încorporare la cele mai populare videoclipuri sau termeni de căutare, pentru a o menține gestionabilă în funcție de amploarea situației.

Orice videoclip obscur pe care un utilizator îl vizionează ajunge să fie mapat la valoarea zero. Aceasta poate părea o alegere arbitrară, dar a fost experimentat pe alte modalități cum ar fi însumarea și luarea valorii maxime pentru fiecare componentă/videoclip. Sistemul YouTube de automatizare învață de fapt cel mai bun mod de a reduce dimensionalitatea datelor lor rare. În acest fel, s-a evitat problema instruirii sistemului doar pe videoclipuri sau jetoane de căutare care au fost invocate de fapt, de către orice utilizator individual.

În următorul pas, se combină acei vectori de vizionare video mediați și vectorii de căutare de jetoane, cu orice alte caracteristici de instruire care se doresc încorporați în model, cum ar fi locația geografică a utilizatorului, vârsta, sexul, etc. Toate aceste caracteristici concatenate sunt apoi alimentate ca intrare într-o rețea neuronală profundă, antrenată pe trei straturi. Primul start

are o lățime de 1.024 de unități, urmat de 512, iar apoi 256 în partea de sus. Acest lucru a oferit un echilibru bun de precizie, rămânând în același timp într-un buget mediu pentru partea hardware. Toate rezultatele sunt introduse într-o bază de date cu vecinii cei mai apropiați pentru fiecare videoclip, pentru a genera mai mulți candidați de recomandare, pe baza a ceea ce a găsit rețeaua neuronală profundă. Problema care se întâlnește este de a clasifica în mod optim toți acești candidați.

YouTube folosește algoritmi de automatizare pentru a clasifica candidații de recomandări în scopul a produce un set final de recomandări de top N, purtând denumirea de învățarea la rang. Algoritmii preiau întregul istoric al impresiilor pentru utilizator, adică fiecare videoclip care i-a fost prezentat, înainte de a viziona videoclipul care a generat această recomandare. Acestea sunt încorporate și mediatizate într-un mod similar cu modul în care datele de vizualizare rare au fost codificate pentru a genera candidați de recomandare. Limba utilizatorului și limba videoclipului sunt, de asemenea, încorporate într-un alt set de caracteristici. De asemenea, algoritmul de automatizare ia în calcul timpul scurs de la ultima vizionare de către privitorul unui videoclip pe acest subiect, dar și numărul de interacțiuni anterioare pe care le-a avut utilizatorul cu videoclipurile de acel gen. Algoritmul YouTube ia rădăcina pătrată a acestei valori, pătratul acesteia și valoarea în sine, toate preluate drept caracteristici de intrare separate. Acest lucru permite sistemului să descopere funcții superliniare și subliniare. Totul este normalizat înainte de a fi introdus în rețeaua neuronală, desigur, pentru a face ca toate aceste caracteristici să aibă o greutate inițială egală. Rezultatul tuturor este de fapt folosit pentru a prezice timpul de vizionare așteptat al fiecărui videoclip, de care depinde în cele din urmă clasamentul.

Dezvoltatorii YouTube nu vor să optimizeze prin algoritmi de automatizare doar predicția clickurilor pe un videoclip, deoarece acest lucru tinde să apară în videoclipuri clickbait de care oamenii nu sunt de fapt interesați. Cu toate acestea, dacă un utilizator urmărește un videoclip până la capăt, acesta este un indiciu mai puternic al faptului că acesta satisface interesele individului respectiv. YouTube este optimizat în funcție de minutele vizionate, nu de clickurile realizate. Și doar această decizie i-a motivat pe creatorii profesioniști de conținut să încarce videoclipuri din ce în ce mai lungi, într-un efort de a crește această valoare a timpului de vizionare. Astfel au introdus în mod accidental un nou mod de a fenta sistemul.

Cred că este interesant faptul că nevoia de a combate zgomotul datelor de click video cu datele minutelor vizionate a dus la crearea de către YouTube a unui sistem care încearcă să puna timpul de vizionare peste toate celelalte. Lupta împotriva acestui zgomot a dus la faptul că oamenii petrec tot mai mult timp pe YouTube, iar YouTube cheltuie din ce în ce mai multe resurse pentru a livra tot conținutul video. Este un exemplu de sistem de recomandare care are consecințe, probabil, neintenționate.

În continuare sunt prezentate câteva aspecte cheie importante ale algoritmului de automatizare. În primul rând, utilizatorul nu trebuie să se bazeze doar pe vizualizarea datelor atunci când instruieste un sistem de recomandare. Dezvoltatorii YouTube au venit cu cât mai multe semnale din comportamentul trecut al utilizatorului și din atributele utilizatorului și, unele caracteristici, au fost puse și sub formă de pătrate și rădăcini pătrate, în cazul în care ar putea fi găsite relații neliniare. Numai vizualizarea datelor în sine este fentată prea ușor și recompensează videoclipurile tip clickbait care nu reflectă de fapt interesele unui utilizator.

Este foarte frecvent că utilizatorii YouTube să vizioneze videoclipuri care fac parte dintr-o serie. O serie episodică de videoclipuri vă fi, în general, urmărită în ordine. Așadar, trebuie să încerce să prezică în mod specific următorul videoclip pe care îl va viziona un utilizator, și nu o vizualizare viitoare întocmită aleatoriu.

Un alt lucru important este că se poate aplica învățarea automată nu numai la generarea de candidați de recomandare, ci și la modul în care acești candidați sunt clasați.

Sistemele automate ale platformei produc un fel de scor final de recomandare pentru fiecare candidat care se poate sorta pentru a produce recomandări de top-N. Dar uneori există informații suplimentare care se pot încorpora în clasamentul respectiv. Și modul în care aceste informații sunt aplicate, pot fi învățate.

2.3.2. Algoritmul de automatizare Netflix

Netflix a publicat, de asemenea, modul în care algoritmi de automatizare produc recomandări, astfel încât, în continuare vom descrie acest lucru. Dezvoltatorii Netflix sunt la fel de deschiși precum Google când vine vorba despre modul de funcționare, astfel încât detaliile

despre modul în care funcționează algoritmi lor de automatizare nu sunt disponibile publicului. Dar, au împărtășit o mulțime de lecții pe care le-au învățat de-a lungul timpului și abordarea generală pe care o adoptă.

Spre deosebire de YouTube, Netflix nu are niciun mandat de a folosi învățarea profundă pentru fiecare problemă pe care o întâmpină. În schimb, se bazează pe o abordare hibridă care combină rezultatele din mulți algoritmi diferiți. Printre algoritmi de automatizare utilizați în trecut reamintim RBM și SVD ++. De asemenea, dezvoltatorii Netflix au prezentat faptul că algoritmul actual de automatizare combină abordările celui mai apropiat vecin, cu abordările de factorizare matricială pentru a încerca să obțină cele mai bune rezultate. Netflix prezintă o listă incompletă de metode pe care le utilizează precum regresie liniară, regresie logistică, SVD, factorizare matricială, RBM, reguli de asociere, mașini de factorizare, arbori aleatorii, arbori de decizie amplificați în gradient, mijloace k, propagare a afinității și procese Dirichlet . Acesta este modul în care Netflix spune că, în principiu, încearcă fiecare algoritm de automatizare și îi lasă pe toți să se lupte între ei când se confruntă cu problema generării de recomandări pentru un anumit utilizator. Probabil că este sigur să spunem că și acum se preocupă de găsirea celui mai bun algoritm de automatizare potrivit pentru platformă.

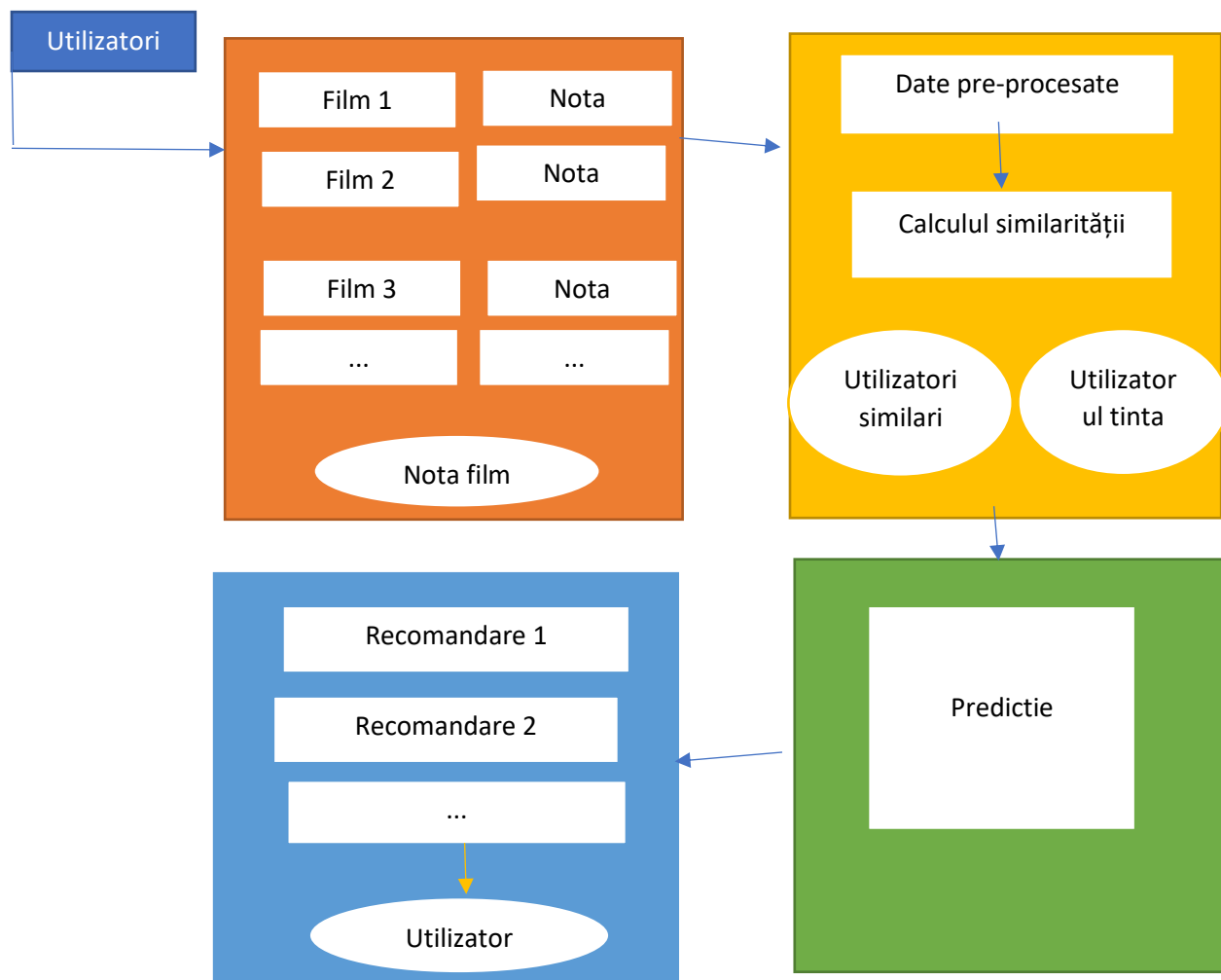


Figura 2.10. Algoritmul de automatizare Netflix

O expresie pe care Netflix o folosește în mod repetat în publicațiile lor este că totul este o recomandare.

Pagina principală Netflix este organizată în mai multe rânduri, fiecare rând conținând propria listă de recomandări pentru un anumit tip de recomandare. De exemplu, pagina de pornire include selecții de top, care probabil sunt rezultatele reale de top-N pentru fiecare utilizator în parte, după cum se poate vedea în fig. 2.11.

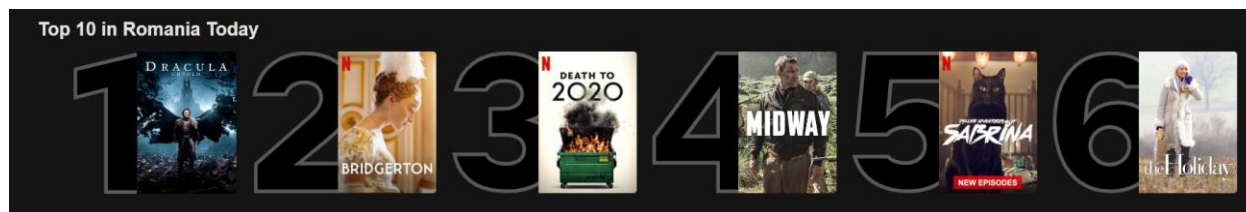


Fig. 2.11. Top N rezultate în România în ceea ce privește filmele/serialele

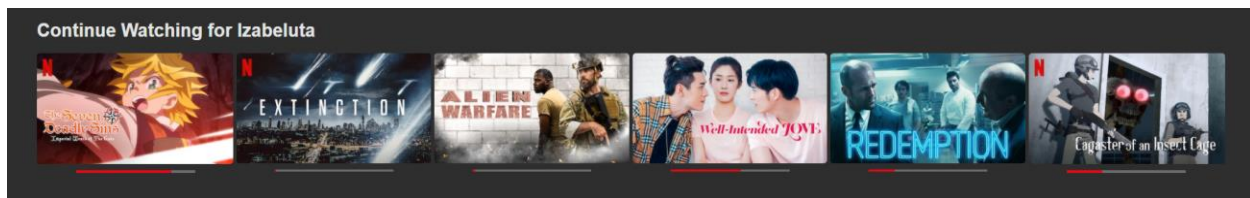


Fig. 2.12. Continuare vizionării unor filme/seriale ale unui utilizator Netflix

Continuarea vizionării(fig.2.12) este recomandarea bazată pe videoclipuri pe care utilizatorul respectiv le-a vizionat deja sau alte persoane din familia respectivă cu care respectivul își partajează contul, ca în figura 2.13.

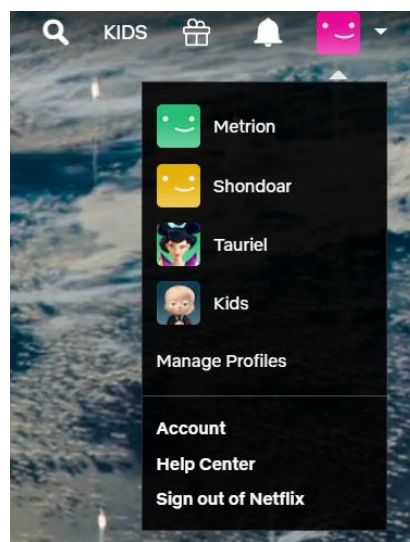


Figura 2.13. Exemplu de partajare a contului pe Netflix

De asemenea, există o listă de recomandări limitate la genul de care un utilizator este interesat cel mai mult, de exemplu SF-ul și fantezia, precum putem observa în figura 2.14.

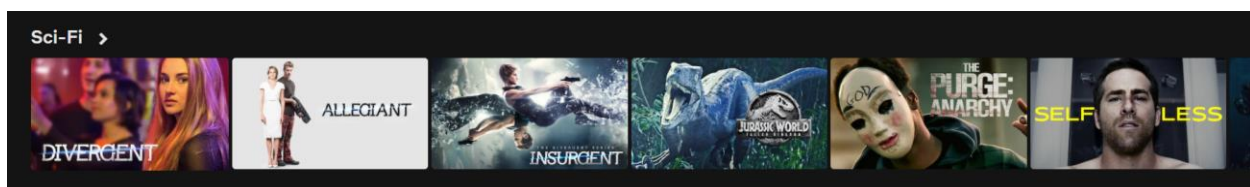


Fig. 2.14. Recomandări Netflix in funcție de gen

Dacă se va derula în jos, pe pagina principală, mai sunt nenumărate rânduri asociate cu recomandări pentru alte categorii care pot fi de interes pentru utilizator, recomandări pentru lansări

noi, pentru articole populare și pentru seriile originale Netflix, după cum se poate vedea în figura 2.15.

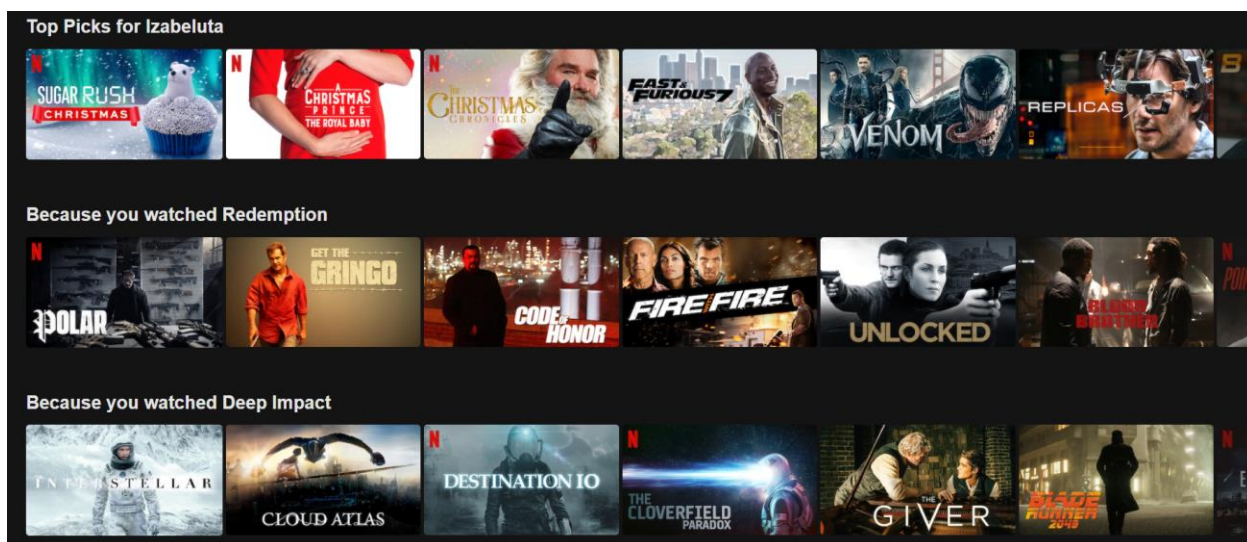


Figura 2.15. Recomandări Netflix și top N alegeri pentru utilizator în funcție de seriile vizualizate

Practic, tot ce se găsește ca o recomandare pe pagina lor de pornire este doar o serie de motoare de recomandare reglate pentru un anumit scop.

De asemenea, Netflix vine în ajutorul utilizatorului cu recomandări în procente pentru a vedea care este o posibilă următoare vizionare, conform figure 2.16.

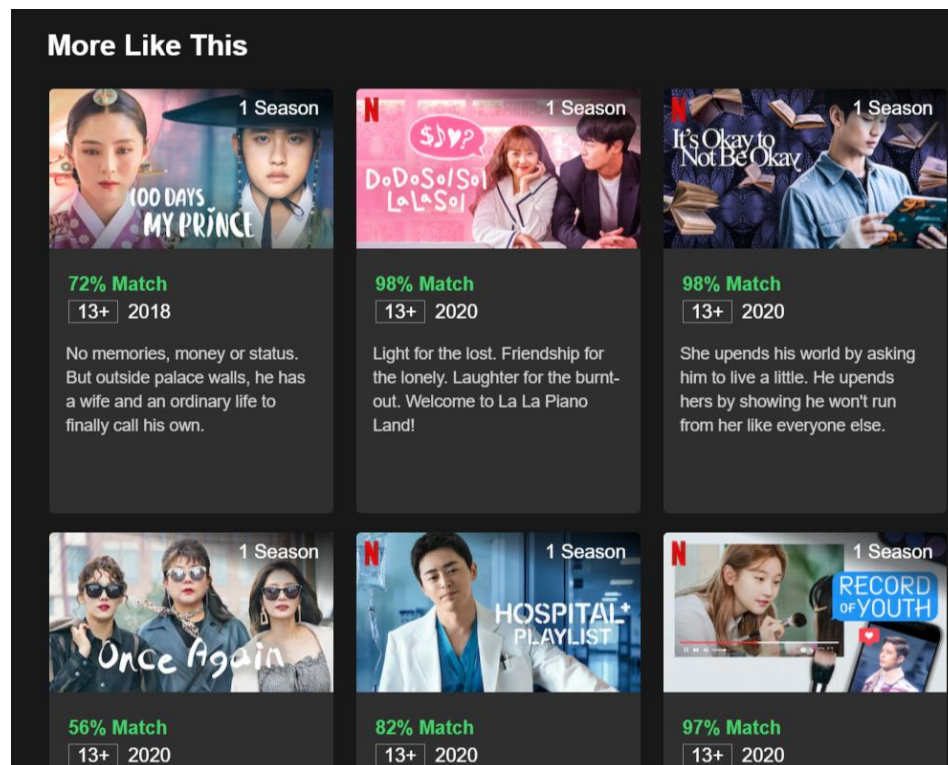


Figura 2.16. Recomandare in procente

Netflix a început să se bazeze pe motoarele de recomandare pentru a prezenta membrilor conținutul nou. A avea toate aceste rânduri de recomandare pe o singură pagină introduce propriul set de provocări. Aceasta înseamnă că Netflix nu numai că trebuie să personalizeze recomandările pentru filme și televiziuni, ci și să personalizeze ordinea în care prezintă utilizatorului aceste recomandări. Aceasta se numește optimizarea întregii pagini, iar utilizarea algoritmilor de automatizare pentru a optimiza selecția widgeturilor de evaluare individuală pe sloturile de pe o pagină este o tehnică care poate fi puternică pentru orice site web. Schema acestui algoritm poate fi văzută în figura 2.10.

Dacă se va deschide pagina de pornire Amazon.com, se va observa faptul că aceasta are o structură similară și se confruntă cu aceeași problemă de optimizare a paginii pentru a pune caracteristicile potrivite în sloturile potrivite pentru acea pagină.

Un alt lucru pe care Netflix îl spune în mod repetat este faptul că dezvoltatorii nu trebuie să se concentreze pe prezicerea corectă a valorilor de rating. Astfel, Netflix a eliminat în totalitate ratingurile de stele, astfel încât algoritmii de automatizare nu s-ar putea antrena pe această caracteristică chiar dacă ar dori. Vorbim aici despre modul în care au evoluat algoritmii Netflix,

concentrându-se inițial pe predicția ratingului, concentrându-se pe clasamente și, în cele din urmă, pe optimizarea întregii pagini. Deși nu anunță în mod specific acest lucru, nici YouTube nu încearcă să prezică ratingurile, ci optimizează minutele urmărite, iar funcția lor obiectivă este învățată în timp din testele A / B live, nu dintr-o măsură de precizie offline. Netflix se bazează, de asemenea, pe teste A / B live online pentru a-și regla sistemele automate.

Netflix observă, de asemenea, că datele de evaluare sunt inerent zgomotoase, ceea ce limitează utilitatea RMSE (Deviația rădăcină-medie-pătrat este o modalitate standard de măsurare a erorii unui model în prezicerea datelor cantitative), ca metrică pentru sistemul de recomandare. YouTube, de asemenea a abandonat în totalitate utilizarea evaluărilor explicite, așa că se pare că ambele platformele au ajuns la aceeași concluzie. De asemenea, similar cu YouTube, Netflix a separat problemele generării de candidați prin clasarea lor și a investit mult în propria abordare personalizată de învățare la rang, aceea de a genera o listă de top N ca o problemă de clasificare, în care fiecare slot din top N este o clasificare pe care se încearcă a se învăța dintr-un istoric trecut al utilizatorului.

Afișarea unor elemente populare este esențială pentru a produce încrederea utilizatorilor în sistemul de automatizare, pentru a-i determina să își facă un abonament.

Ca și în cazul YouTube, candidații de recomandări vin împreună cu scoruri sau predicții de evaluare care ar putea fi sortate pentru a obține o listă de top N, dar Netflix a dus această idee puțin mai departe și au aplicat algoritmi de automatizare pentru a găsi cea mai bună ordine pentru a prezenta acele rezultate finale încorporând mai multe informații, nu doar acele evaluări prezise.

Netflix, ca și YouTube, a descoperit faptul că, cu cât există mai multe funcții care instruiesc sistemul, cu atât mai bine. Nu este vorba doar despre videoclipurile care au fost vizionate în trecut, ci și despre modul în care au fost vizionate, chiar și circumstanțele în care se afla acum un utilizator ar putea afecta ce tipuri de conținut ar trebui să arate cel mai bine pe pagina principală.

Dacă se accesează Netflix de pe televizorul de acasă, poate că acesta este dispozitivul pe care abonatul îl utilizează personal și recomandarea unui conținut mai lung destinat adulților este potrivit în acel moment. În schimb, dacă copilul abonatului folosește contul Netflix pe o tabletă, formularea de recomandări bazate pe ceea ce s-a urmărit de pe tableta respectivă ar avea un conținut mai scurt, orientat spre copil, care are sens în acel context. Astfel, timpul ar putea face diferența.

S-ar putea ca utilizatorul să fie deschis la un conținut mai matur noaptea târziu, decât după-amiază când copiii sunt în jur. Chiar și ziua săptămânii ar putea conta. Poate un utilizator caută să piardă mai mult timp în weekend și, prin urmare, s-ar putea să prefere filme în weekend și emisiuni TV în timpul săptămânii. Acestea sunt toate exemple ale contextului în care se caută recomandări, iar un sistem bun automat de recomandare va utiliza toate aceste informații pentru a rafina rezultatele pe care le returnează.

Este, de asemenea, interesant faptul că, deși Netflix și YouTube sunt platforme de recomandare de videoclipuri, fac acest lucru în setări ușor diferite și cu cerințe ușor diferite, care au prezentat alegerile lor de algoritmi și tehnologie.

2.3.3. Algoritmul de automatizare realizat

Sistemul informatic pentru automatizarea deciziei prezentat are în vedere o pagina de creare cont, în care utilizatorul introduce câteva date personale cum ar fi adresa de email, nume utilizator și parola, date stocate în baza de date denumită “registration”, mai exact în tabelul “users”. Cu aceste date, utilizatorul se poate autentifica în pagina login.php pentru a deschide primul test, prin intermediul căruia se poate afla starea curentă a utilizatorului. După ce utilizatorul completează testul, poate afla rezultatul apăsând butonul “Afișează rezultatul”. Mărimile de intrare în cazul de față sunt întrebările testului. În continuarea testului de pe prima pagina se va afișa un mesaj, care indică starea utilizatorului din momentul realizării testului, însoțit de un videoclip și o nota pe care utilizatorul o poate da testului. În situația actuală, mărimile de ieșire din sistemul automat sunt: videoclipul, mesajul și ratingul. Ratingul dat de utilizator, împreună cu alegerile întrebărilor testului vor fi stocate în baza de date. Un astfel de test va fi regăsit și în pasul doi al algoritmului de automatizare pentru o acuratețe cât mai ridicată a recomandărilor. În a ultima pagina a site-ului se va face o medie aritmetică a celor două note date de utilizator, iar în funcție de aceasta medie, se vor afișa trei recomandări de videoclipuri, care reflectă cel mai bine starea utilizatorului. Scopul acestui algoritm de automatizare este reglarea stării utilizatorului cu ajutorul testelor de personalitate. Schema bloc a sistemului informatic pentru automatizarea deciziei este prezentată în figura 3.1.

3. Studiul și proiectarea sistemului informatic pentru automatizarea luării deciziei în social media

Am ales această temă deoarece platformele de tip social media sunt din ce în ce mai utilizate în ziua de astăzi. Scopul algoritmului de automatizare prezentat este reglarea stării psihice a utilizatorului cu ajutorul unor videoclipuri pe baza efectuării unor teste de personalitate. Mărimile de intrare în paginile testului de personalitate ale algoritmului de automatizare sunt imaginile prezentate sub forma de butoane radio. Utilizatorul poate alege imaginea care se potrivește cel mai bine stării lui curente. Mărimile de ieșire ale sistemului din cele doua pagini care cuprind testele, sunt mesajul cu starea care reiese din alegerile realizate în test, videoclipul care se potrivește cu starea curentă a utilizatorului și un rating. Utilizatorul folosind acest rating poate da o nota testului. Rezultatele testelor împreună cu nota sunt stocate în baza de date. În cea de-a treia pagină, pagina de recomandări, se preiau din baza de date cele doua note date de utilizator în urma prezentării celor doua videoclipuri, se face o medie aritmetică a acestora și se prezintă recomandări de videoclipuri cât mai apropiate de starea curentă a utilizatorului. Testele sistemului informatic pentru automatizarea deciziei au fost realizate pe familie, colegi de muncă, prieteni, pentru a afla rata de succes a testului. După realizarea testelor pe mai multe persoane, rata de succes a fost de 97%.

Social media sau media de socializare online sunt un grup de instrumente cum ar fi site-uri web și aplicații care funcționează cu ajutorul unui dispozitiv cum ar fi telefon mobil, tabletă, calculator, laptop conectat la internet. Aceste medii de socializare au fost realizate pentru a conferi comunicarea între utilizatori, crearea, distribuirea și schimbul de conținut (text, foto, video, audio, prezentări multimedia etc.) între utilizatori cum ar fi prieteni, colegi, familie, dar și între utilizatori necunoscuți.

3.1. Tipuri de medii de socializare

Kaplan și Haenlein propun o schemă care cuprinde șase tipuri de media de socializare:

1. proiecte la care colaborează mai mulți utilizatori (colaborative projects) prin crearea de conținut. Din această categorie fac parte site-urile de tip enciclopedie online (ex: [Wikipedia](#)) și site-urile ”semn de carte” (ex: [Del.icio.us](#), [Delicious](#))

2. bloguri – echivalentul unei pagini personale în media de socializare (ex: [Blogger](#), [Wordpress](#)). Kaplan și Haenlein includ în categoria bloguri și site-urile de tip viață în timp real – life streaming (ex: [Justin.tv](#))

3. comunități de conținut – scopul acestor comunități este distribuirea de conținut. Există comunități care sunt axate pe distribuirea de conținut: text (ex: [BookCrossing](#)), video (ex: [Youtube](#)), foto (ex: [Flickr](#)), audio (ex: [Audiobase](#)) și slide-uri (ex: [Slideshare](#))

4. rețele sociale – permit interacțiunea dintre utilizatori prin crearea de profile personale, prin invitațiile trimise prietenilor pentru a avea acces la aceste profile și prin trimiterea de mesaje și e-mailuri (ex: [Facebook](#))

5. jocuri online – lumi virtuale tridimensionale în cadrul cărora utilizatorii au posibilitatea de a-și alege un avatar și de a interacționa cu ceilalți oameni așa cum o fac și în viața reală (ex: [World of Warcraft](#))

6. lumi virtuale – utilizatorii folosesc un avatar cu ajutorul căruia intră într-o lume virtuală care copiază lumea reală (ex: [Second Life](#), [Habbo](#)) [18].

3.2. Resurse software

Programe și limbaje utilizate:

Geany - Geany este un editor de text GUI ușor care utilizează Scintilla și GTK, incluzând funcțiile de bază ale IDE. Este proiectat pentru a avea timpi de încărcare scurți, cu dependență limitată de pachete separate sau biblioteci externe de pe Linux.

HTML - HyperText Markup Language este un limbaj de marcare utilizat pentru crearea paginilor web ce pot fi afișate într-un browser. Scopul HTML este mai degrabă prezentarea informațiilor – paragrafe, fonturi, tabele ș.a.m.d. – decât descrierea semanticii documentului.[19]

CSS - CSS sau Cascading Style Sheets este un standard pentru formatarea elementelor unui document HTML. Stilurile se pot atașa elementelor HTML prin intermediul unor fișiere externe sau în cadrul documentului, prin elementul

3.3. Efecte pozitive și negative ale utilizării mediilor de socializare

3.3.1. Efecte pozitive

Câteva efecte pozitive pe care le are social media în rândul utilizatorilor sunt:

- Cunoașterea cât mai rapidă de oameni noi
- Comunicarea cât mai rapidă dintre persoane
- Comunicarea aproape gratis între persoanele care se află în străinătate, spre exemplu
- Promovarea cât mai rapidă și eficientă

3.3.2. Efecte negative

Câteva efecte negative pe care le are social media în rândul utilizatorilor sunt:

Reducerea stimei de sine

- Frica de a rămâne pe dinafară
- Perturbarea rutinei de somn
- Declanșarea sentimentelor de invidie
- Încurajarea comportamentelor nesănătoase

3.4. Proiectarea, realizarea și testarea aplicației

Schema bloc a sistemului informatic pentru automatizarea deciziei este prezentată în continuare:

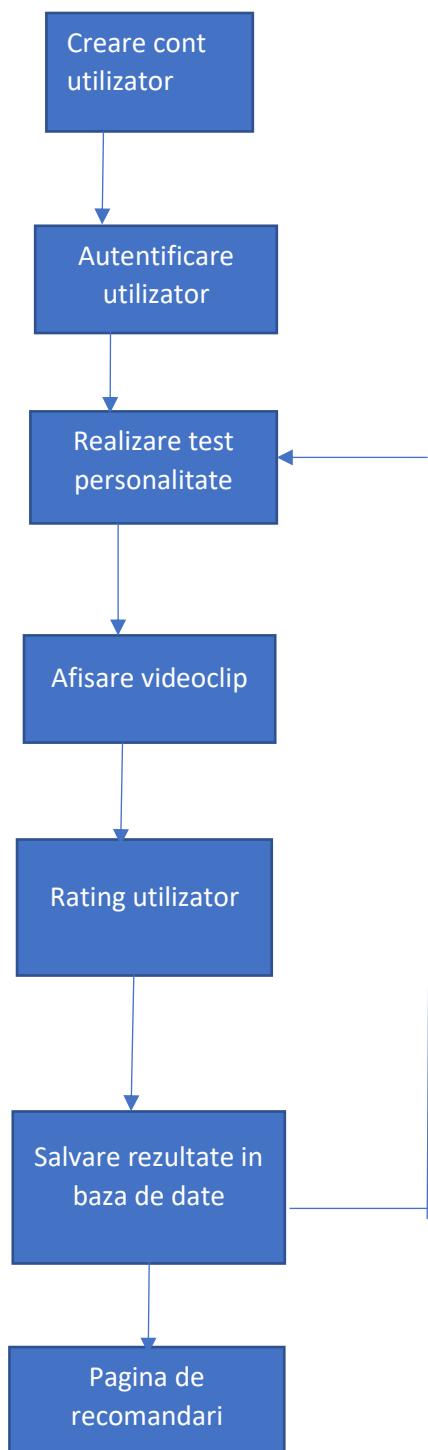


Figura 3.1. Schema bloc a sistemului informatic pentru automatizarea deciziei

Se deschide soft-ul Geany în care vom crea o filă nouă .css, .js, .php după cum se poate vedea în figura 3.2.

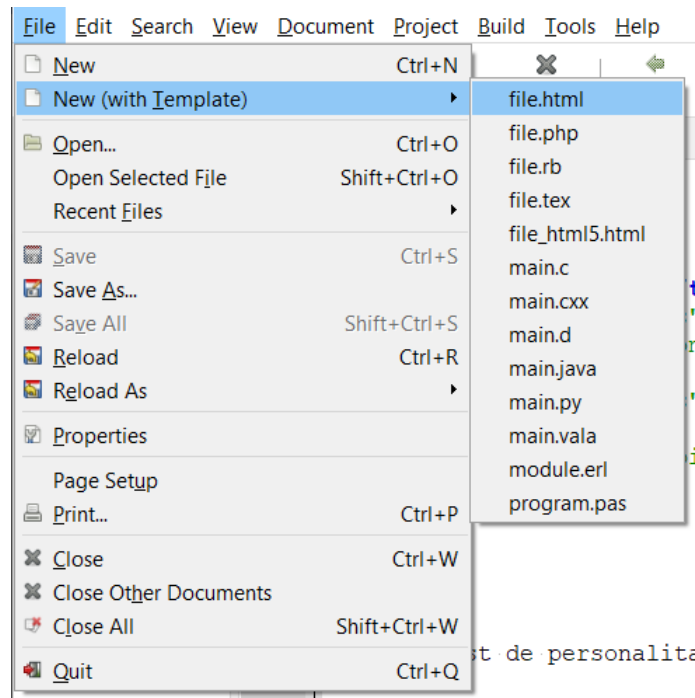


Fig. 3.2. Creare fila noua

Dupa ce deschidem o filă de tip PHP, vom începe să scriem codul. Predefinit, când se deschide fila .php, va apărea un template după cum se poate vedea în figura 3.3.

Toate fișierele trebuie să conțină declararea `<!DOCTYPE>` care nu este un tag html, ci este o informație către browser despre ce tip de document trebuie să deschidă. Elementul `<html>` este elementul rădăcină a unei pagini HTML. Elementul `<head>` conține informație meta despre pagina HTML. Elementul `<title>` specifică titlul pentru pagina HTML care este arătat în bara de titlu al browser-ului. Elementul `<body>` definește conținutul documentului și conține toate elementele vizibile cum ar fi paragrafe, imagini, tabele, butoane, etc..

```

1  <?php
2
3  ?>
4  <!DOCTYPE html PUBLIC "-//W3C//DTD XHTML 1.0 Strict//EN"
5  —>"http://www.w3.org/TR/xhtml1/DTD/xhtml1-strict.dtd">
6  <html xmlns="http://www.w3.org/1999/xhtml" xml:lang="en" lang="en">
7
8  <head>
9  —><title>untitled</title>
10 —><meta http-equiv="content-type" content="text/html; charset=utf-8" />
11 —><meta name="generator" content="Geany 1.37.1" />
12 </head>
13
14 <body>
15 —>
16 </body>
17
18 </html>
19

```

Fig. 3.3. Template PHP

În partea de `<head></head>` a paginii `index.html` am definit referința către stilul paginii – `style.css` și către pagina în care se regăsește javascript-ul - `main.js` după cum se poate vedea mai jos.

```

<head>
    <title>Delogare</title>
    <link rel="stylesheet" type="text/css" href="style.css">
</head>

```

În continuare voi prezenta câteva elemente principale care se regăsesc în pagina `test.php`. Între elementul `<body></body>` am definit conținutul documentului după cum urmează:

- `<h1></h1>` - tag utilizat pentru a defini text-ul „Bine ati venit!”
- `<form></form>` - o formă HTML este utilizată pentru a colecta intrările utilizatorului. Acest element este un conținut pentru diferite tipuri de elemente, în cazul nostru paragrafe, butoane radio, text, butoane tip submit.
 - `<p></p>` - tag utilizat pentru a defini paragraful „Va rog alegeti o imagine”, în care utilizatorul este rugat să facă o alegere.
 - `<input>` - utilizat în cazul de fata ca buton radio care selectează o variantă din doua posibile de răspuns la întrebare și un buton de tip submit
 - `<label></label>` - definește o etichetă pentru un element de tip `<input>`. În cazul de față, elementul oferă o imbunatatire a gradului de utilizare pentru utilizator. În

cazul în care utilizatorul face click pe imaginea din elementul , acesta comută controlul.

Atributul `for` al tag-ului `label` trebuie să fie aceeași cu atributul `id` al elementului `input`

- `` - tag utilizat pentru a introduce o imagine în pagina web
- `<video></video>` - tag utilizat pentru a arăta un video în pagina web

În fișierul `style.css` se regăsește stilul paginii, adică modul în care sunt afișate pe ecran elementele HTML. Selectorul arată care element urmează să fie stilizat. Blocul de declarație conține unul sau mai multe declarații separate de punct și virgulă. Fiecare declarație include o proprietate CSS și o valoare separate de două puncte. Mai multe declarații CSS sunt separate de punct și virgulă, și blocurile de declarație sunt înconjurate de acolade. [25] Codul specific stilului poate fi regăsit în Anexe.

```
h1 - selector
{
  text-align:center; - declarare
  ↑                ↑
  proprietate      valoare
}
```

Codul JavaScript din `main.js` începe cu o funcție care este definită pentru a realiza un anumit task. În cazul de față funcția `check()` verifică ce valoare a fost selectată prin butonul radio și afișează un videoclip în funcție de alegerile utilizatorului. Funcția `check()` începe prin declararea variabilelor. `document.quiz.question1.value`, unde `quiz` este numele formei și `question1` este numele input-ului, întoarce valoarea elementului `question1` din forma `quiz`. `Videos` și `messages` sunt doi vectori. În primul vector se regăsesc 4 videoclip-uri din care se va afișa un singur videoclip în funcție de alegerile pe care le face utilizatorul. În al doilea vector sunt mesaje cu stările pe care le are utilizatorul în momentul realizării testului. După ce utilizatorul da click pe butonul „Afișează rezultatul”, un videoclip și un mesaj din cei doi vectori vor apărea pe ecran. Codul JavaScript poate fi regăsit în Anexe.

Primul pas al algoritmului de automatizare, este crearea unui cont de către utilizatorul care dorește să completeze testele, pas regăsit în figura 3.4. Datele utilizatorului sunt stocate în baza de date „registration”, mai exact în tabelul „users”. Numele de utilizator și parola vor fi utilizate în pasul doi al sistemului, mai exact în pagina de autentificare utilizator, conform figurii 3.5

Bine ati venit pe pagina de inregistrare

Introduceti numele de utilizator

Introduceti adresa de email

Introduceti parola

Reintroduceti parola

Inregistrare

Aveti deja un cont? [Autentificati-va aici!](#)

Fig. 3.4. Pagina de inregistrare a utilizatorului

The image shows a user authentication page with a pink header and a white body. The header contains the text "Bine ati venit in pagina de autentificare!". The body contains two input fields: "Introduceti numele de utilizator" with the text "izabela" and "Introduceti parola" with masked characters ".....". Below the input fields is a black button labeled "Autentificare". At the bottom, there is a link that says "Nu aveti cont? [Faceti click aici pentru a va inregistra](#)".

Bine ati venit in pagina de autentificare!

Introduceti numele de utilizator

izabela

Introduceti parola

.....

Autentificare

Nu aveti cont? [Faceti click aici pentru a va inregistra](#)

Fig. 3.5. Pagina de autentificare a utilizatorului

Al treilea pas al sistemului de automatizare este completarea testului de personalitate. În cazul în care utilizatorul dă click pe o imagine, realizată sub forma unui buton radio, aceasta va rămâne selectată până când utilizatorul va face click și pe celelalte imagini, La final după ce utilizatorul va face toate alegerile, va apărea un videoclip în funcție de aceste alegeri, creatorul testului putând observa starea de spirit a utilizatorului, și astfel apărând un mesaj cu starea curentă a utilizatorului din momentul realizării testului, un videoclip și un rating, date stocate în baza de date după cum se poate observa în figura 3.6. Rating-ul este utilizat ulterior în pagina de recomandări.

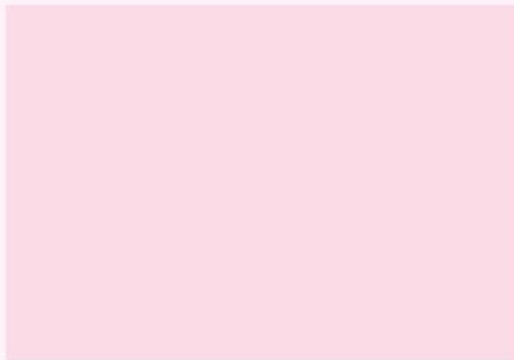
Bine ati venit!

In continuare va rugam sa completati urmatorul test pentru a afla preferintele dumneavoastra!

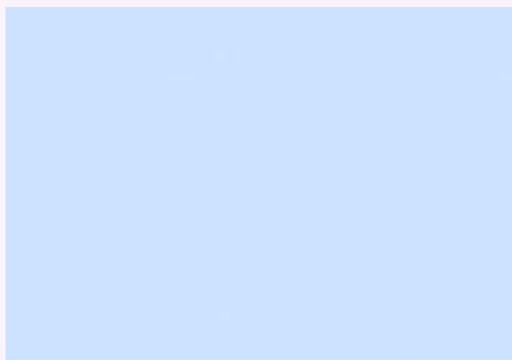
Va rog alegeti o imagine



Va rog alegeti o imagine



Va rog alegeti o imagine



Va rog alegeti o imagine



Va rog alegeti o imagine



Afiseaza rezultatul!

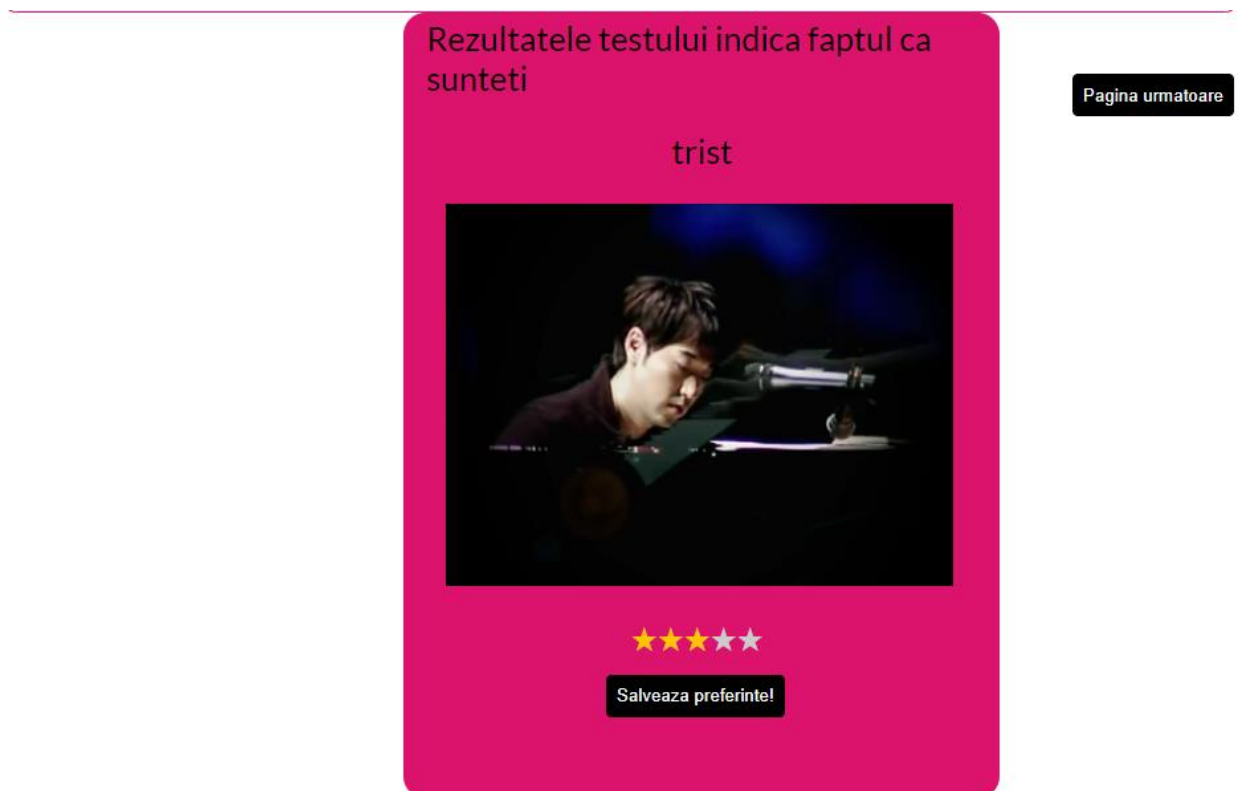


Fig. 3.6. Primul test de personalitate, insotit de afisarea rezultatului in urma testului realizat de utilizator

Același lucru este valabil și pentru pasul patru al sistemului, mai exact al doilea test pentru verificarea stării utilizatorului, conform figurii 3.7.

Pasul 2!

In continuare va rugam sa completati urmatorul test pentru a afla preferintele dumneavoastra!

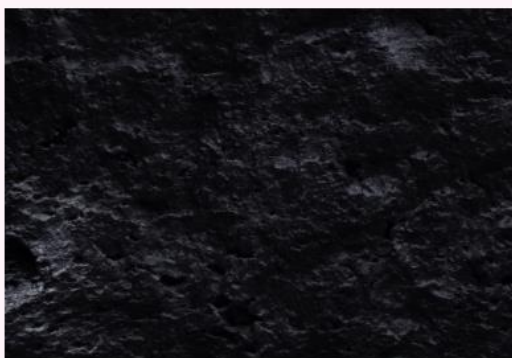
Va rog alegeti o imagine



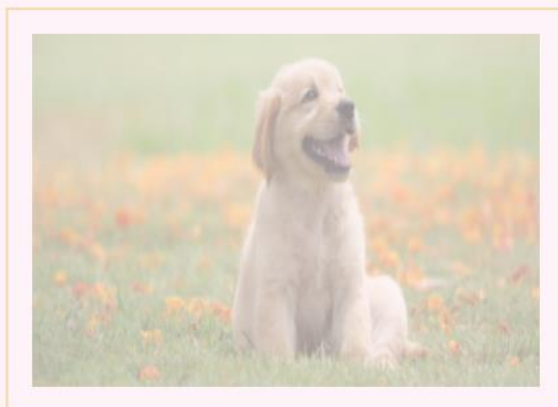
Va rog alegeti o imagine



Va rog alegeti o imagine



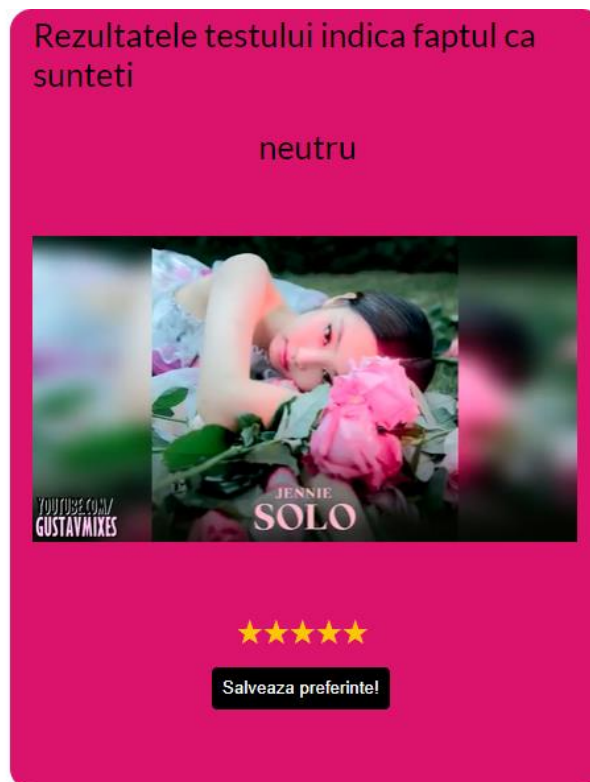
Va rog alegeti o imagine



Va rog alegeti o imagine



Afiseaza rezultatul!



Pagina urmatoare

Fig. 3.7. Al doilea test de personalitate, insotit de afisarea rezultatului in urma testului realizat de utilizator

În ultimul pas al sistemului informatic pentru automatizarea deciziei, se realizează media aritmetică a celor doua rating-uri date de utilizator, date preluate din baza de date, iar pe baza mediei aritmetice sunt afișate numai videoclipurile cât mai apropiate de starea curentă a utilizatorului, conform figuri 3.8.

Va multumim pentru timpul acordat! Nota pe care ati acordat-o recomandarilor noastre este 4

In continuare vor fi afisate mai multe videoclipuri in functie de starea dumneavoastra

Afiseaza recomandari!



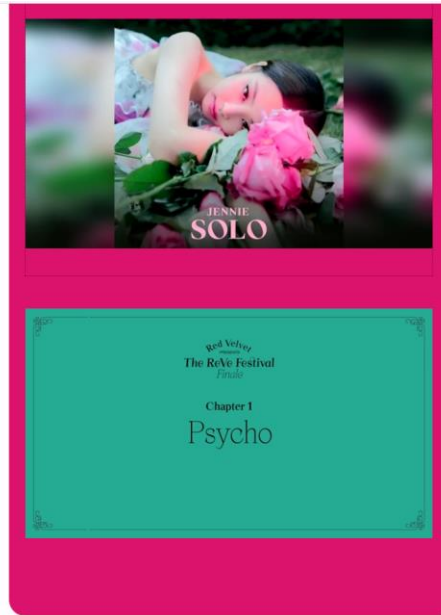


Fig.3.8. Pagina de recomandare

In figura 3.9, sunt afișate datele stocate din baza de date.

	id	username	email	password	alegere1	alegere2	alegere3	alegere4	alegere5	rating	alegere6
1	1	izabela	grosu.izabela2000@gmail.com	77e31e3a4b66c04c620f1d266b0a2c2	albastru	roz_deschis	albastru_inchis	br	pt	3	larna
6		alegere7	alegere8	alegere9	alegere10	rating2					

Fig. 3.9. Datele stocate in baza de date in urma testelor realizate si notelor acordate

4. Concluzii și direcții de cercetare ulterioare

Testarea aplicației a fost realizată chestionând mai multe persoane cunoscute, astfel încât, ponderea rezultată în urma testelor de personalitate și a notelor să fie cât mai apropiată de 100%. Prima etapă de testare a avut drept scop înțelegerea stării actuale a utilizatorului, și realizarea unui test care să fie cât mai apropiat de starea din momentul parcurgerii sistemului informatic pentru automatizarea deciziei. Cea de-a doua etapă a testării a avut drept scop alegerea videoclipurilor reprezentatoare pentru fiecare stare în parte.

Urmărind pașii de realizare ai algoritmului de automatizare parcurși în capitolul anterior, am observat faptul că un procent de 100% al satisfacției utilizatorului în urma testelor realizate nu poate exista.

Pe parcursul realizării algoritmului de automatizare, pornind de la realizarea chestionarului, până la pagina de recomandări, am întâmpinat o serie de probleme pe care a trebuit să le rezolv.

Din cauza faptului că testarea a fost realizată pe utilizatori cu personalități diferite, nu a fost ușor să aleg imaginile și videoclipurile care se întâlnesc în cele două pagini de chestionare. Una din problemele principale apărute a fost alegerea unei modalități cât mai ușoare de prezentare a întrebărilor chestionarelor, iar în urma testelor realizate am ajuns la concluzia că întrebările testelor să fie afișate sub forma de imagini, integrate în butoane tip radio, și stocarea rezultatelor în baza de date. Cea de-a doua problemă întâlnită a fost încărcarea stilului paginilor în timp real pe server, durata de așteptare pornind de la câteva minute, ajungându-se la câteva ore.

Din punct de vedere al costului, atât editorul de text Geany, cât și numele de host, localhost au fost gratuite.

Din punct de vedere al optimizării, inițial, am creat o pagină fără nicio legătură externă cu bază de date, care a constat într-un simplu chestionar, însă din cauza faptului că avea o pondere a notei utilizatorului foarte mică, în proporție de 25% din utilizatori acordau nota maximă, am decis să stochiez rezultatele testului împreună cu nota într-o bază de date, să schimb tipul întrebărilor, inițial fiind întrebări la care utilizatorul scria răspunsul într-o casetă, și să mai realizez și o a doua pagină de test, rezultatele și nota fiind de asemenea stocate în baza de date. În urma acestor

rezultate, a apărut și o pagina de recomandare în care videoclipurile sunt afișate pe baza mediei aritmetice a celor doua note date de utilizator.

Câteva direcții viitoare ar fi că utilizatorul să aibă libertatea să încarce propriile lui videoclipuri, care pot fi stocate într-o bază de date sau într-un folder cu id-ul utilizatorului, să existe o pagină de recomandare cu mai multe videoclipuri și teste de personalitate mult mai ample.

Bibliografie

- [1] M. Abadi, A. Agarwal, P. Barham, E. Brevdo, Z. Chen, C. Citro, G. S. Corrado, A. Davis, J. Dean, M. Devin, S. Ghemawat, I. Goodfellow, A. Harp, G. Irving, M. Isard, Y. Jia, R. Jozefowicz, L. Kaiser, M. Kudlur, J. Levenberg, D. Mané, R. Monga, S. Moore, D. Murray, C. Olah, M. Schuster, J. Shlens, B. Steiner, I. Sutskever, K. Talwar, P. Tucker, V. Vanhoucke, V. Vasudevan, F. Viégas, O. Vinyals, P. Warden, M. Wattenberg, M. Wicke, Y. Yu, and X. Zheng. TensorFlow: Large-scale machine learning on heterogeneous systems, 2015. Software available from tensorflow.org.
- [2] J. Dean, G. S. Corrado, R. Monga, K. Chen, M. Devin, Q. V. Le, M. Z. Mao, M. Ranzato, A. Senior, P. Tucker, K. Yang, and A. Y. Ng. Large scale distributed deep networks. In NIPS, 2012.
- [3] K. J. Oh, W. J. Lee, C. G. Lim, and H. J. Choi. Personalized news recommendation using classified keywords to capture user preference. In 16th International Conference on Advanced Communication Technology, pages 1283–1287, Feb 2014.
- [4] X. Glorot, A. Bordes, and Y. Bengio. Deep sparse rectifier neural networks. In G. J. Gordon and D. B. Dunson, editors, Proceedings of the Fourteenth International Conference on Artificial Intelligence and Statistics (AISTATS-11), volume 15, pages 315–323. Journal of Machine Learning Research - Workshop and Conference Proceedings, 2011
- [5] X. Su and T. M. Khoshgoftaar. A survey of collaborative filtering techniques. *Advances in artificial intelligence*, 2009:4, 2009.
- [6] D. Tang, B. Qin, T. Liu, and Y. Yang. User modeling with neural network for review rating prediction. In Proc. IJCAI, pages 1340–1346, 2015.
- [7] M. Oprea (2014-2015), Inteligență artificială, notițe de curs, UPG Ploiești
- [8] S. Sedhain, A. K. Menon, S. Sanner, and L. Xie. Autorec: Autoencoders meet collaborative filtering. In Proceedings of the 24th International Conference on World Wide Web, WWW '15 Companion, pages 111–112, New York, NY, USA, 2015. ACM.

- [9] A. M. Elkahky, Y. Song, and X. He. A multi-view deep learning approach for cross domain user modeling in recommendation systems. In Proceedings of the 24th International Conference on World Wide Web, WWW '15, pages 278–288, New York, NY, USA, 2015. ACM
- [10] A. van den Oord, S. Dieleman, and B. Schrauwen. Deep content-based music recommendation. In C. J. C. Burges, L. Bottou, M. Welling, Z. Ghahramani, and K. Q. Weinberger, editors, Advances in Neural Information Processing Systems 26, pages 2643–2651. Curran Associates, Inc., 2013.
- [11] D. Oard and J. Kim. Implicit feedback for recommender systems. In in Proceedings of the AAAI Workshop on Recommender Systems, pages 81–83, 1998.
- [12] F. Morin and Y. Bengio. Hierarchical probabilistic neural network language model. In AISTATS'05', pages 246–252, 2005.
- [13] J. Weston, A. Makadia, and H. Yee. Label partitioning for sublinear ranking. In S. Dasgupta and D. Mcallester, editors, Proceedings of the 30th International Conference on Machine Learning (ICML-13), volume 28, pages 181–189. JMLR Workshop and Conference Proceedings, May 2013.
- [14] T. Liu, A. W. Moore, A. Gray, and K. Yang. An investigation of practical approximate nearest neighbor algorithms. pages 825–832. MIT Press, 2004.
- [15] Moise A.(2021), Retele neuronale pentru recunoasterea formelor, notite de curs, UPG Ploiesti
- [25] Cărbureanu, M. - Programare Web. Ghid teoretic și practic, Editura Universității Petrol-Gaze din Ploiești, 2020.

Webografie

- [16] https://en.wikipedia.org/wiki/A/B_testing
- [17] <http://youtubecreator.blogspot.com/2012/08/youtube-now-why-we-focus-on-watch-time.html>. Accessed: 2016-04-20.
- [18] <https://www.nato.int/>
- [19] <https://www.w3schools.com/>
- [20] <https://www.javascript.com/>
- [21] <https://sproutsocial.com/insights/social-media-algorithms/>
- [22] <https://www.theatlantic.com/technology/archive/2018/11/how-youtubes-algorithm-really-works/575212/>
- [23] <https://static.googleusercontent.com/media/research.google.com/en//pubs/archive/45530.pdf>
- [24] https://www.researchgate.net/publication/341804453_The_Use_of_Algorithms_to_Target_Children_while_Advertising_on_Youtube_Kids_Platform_A_reflection_and_analysis_of_the_existing_regulation

Summary

In this paper, I aimed to represent the current state of a user. To accomplish this project, I used the Geany text editor and the hostname, localhost which were free.

The first chapter has presented the tools specific to social media automation algorithms, more precisely the general issues associated with social media, the current state of specialized research on social media automation algorithms, followed by the motivation for choosing the topic.

The next chapter has presented research on social media automation algorithms. In subchapter 2.1. it can be found an overview of the Facebook, YouTube and Instagram automation algorithms, followed by the subchapter dedicated to the study of the literature dedicated to the YouTube algorithm, and subchapter 2.3 contains practical research on the automation algorithms presented.

Chapter 3 contains the study and design of the computer system for automated decision making in social media.

After making this application, I deduced that it was very difficult to find out the status of a user and to create an automation algorithm to achieve this. Due to the very changing state of the users, it was difficult to perform a test that reflected their condition, but it is equally difficult to choose videos that universally fit these conditions.

This project aimed to implement an efficient way to deduce the status of a user based on this status to display music videos.