

Universitat de Lleida

PREMIS A TREBALLS DE RECERCA DE LA UdL
per a l'estudiantat de batxillerat i cicles formatius de grau superior

Motors de Recomanació, Una Arma de Doble Fil

Matías Mora Molina

Centre: Col·legi Claver de Lleida

Tutor: Josep Mallol Gurgui

Data: octubre, 2023



Motors de Recomanació, Una Arma de Doble Fil

Navegant en una societat on la llibertat
ens empresona

Col·legi Claver - Jesuïtes Lleida

2022-2023

2n de Batxillerat Tecnològic

Tutor: Josep Mallol



JESUÏTES Lleida
Collegi Claver - Raimat

Resumen

Este trabajo, como su nombre indica, pretende dar una doble visión a los motores de recomendación, enseñando que detrás de su gran labor en muchas de las tecnologías de hoy en día se esconde una potencial herramienta de manipulación. Además, aprovechando la doble faceta del protagonista de este trabajo, también intenta concienciar a la sociedad sobre la tendencia cultural a la que nos dirigimos.

Para demostrar dicha dualidad presenta un acercamiento en dos partes. En primer lugar, mediante un análisis teórico, donde estudia las características y funcionalidades tecnológicas que posee el algoritmo en cuestión y examina sus repercusiones y causas sociales. Posteriormente, complementa esa investigación con dos experimentos prácticos, también de ámbito tecnológico y social respectivamente. En el tecnológico, por medio de un programa informático que simula uno de estos sistemas de recomendación, procura demostrar la manipulación a la que estamos sometidos, mientras que en el social hace una reflexión sobre el propósito de la sociedad actual mediante un acercamiento a la población.

Los resultados adquiridos afirman la hipótesis, pues en el apartado teórico se corrobora el poder que estos algoritmos tienen sobre nosotros, mientras que en el práctico se prueba que este poder acaba afectando nuestro consumo y conducta, ya que es posible alterar aquello que se nos muestra por pantalla cuando navegamos.

Por último, concluye afirmando que los motores de recomendación contribuyen a la infinita rueda movida por el capital que se está creando en nuestro sistema, aunque esta no sea del todo responsabilidad suya. Si todo sigue igual, la insurgente inmediatez, la sobrecarga de estímulos y la infelicidad crónica seguirán reinando. ¿Es nuestra libertad realmente innegable?

Palabras Clave: Algoritmo, Recomendación, Manipulación, Sociedad, Impacto

Abstract

This work, as its name suggests, aims to give a double vision of recommendation engines, showing that behind their great work in many of today's technologies lies a potential tool for manipulation. Furthermore, taking advantage of the dual facet of the protagonist of the research project, this work also tries to raise awareness in society about the cultural trend we are heading towards.

To demonstrate this duality, this work presents a two-part approach. Firstly, through a theoretical analysis, where the technological characteristics and functionalities of the algorithm are studied and its social repercussions and causes examined. Subsequently, it complements this research with two practical experiments, also in the technological and social fields respectively. In the technological one, by means of a computer programme that simulates one of these recommendation systems, this work tries to demonstrate the manipulation to which we are subjected, while in the social one thinks about the purpose of today's society through an approach to the population.

The acquired results confirm the hypothesis, since in the theoretical section the power that these algorithms have over us is corroborated and in the practical one it is proven that this power ends up affecting our consumption and behavior, as it is possible to alter what is shown to us on the screen when we navigate through internet.

Finally, this work concludes by affirming that recommendation engines contribute to the infinite wheel moved by capital that is being created in our system, although this is not entirely their responsibility. If everything stays the same, insurgent immediacy, stimulus overload and chronic unhappiness will continue its kingdom. Is our freedom actually undeniable?

Key Words: Algorithm, Recommendation, Manipulation, Society, Impact

Agraïments

En primer lloc, m'agradaria agrair als dos tutors guia, Albert Felip i Josep Mallol, per totes les proposicions i suport que han estat necessaris per a realitzar el treball.

També sento un especial agraïment amb Josep Forné, qui m'ha ajudat tant en l'entrevista que vàrem dur a terme com en l'hora d'estructurar la part social, cosa essencial per a donar forma al producte final.

De la mateixa manera, vull agrair a la meva tutora, Marisa Samitier, per tota la constant ajuda que m'ha proporcionat incondicionalment, a més a més que la seva participació ha estat cabdal en l'organització i en el contingut del treball.

A més a més, hi ha hagut moltes altres persones com Eva Martín o Josep Marc Mingot, entre altres, que també han estat de gran ajuda, així que des d'aquí els ho agraeixo.

I per últim, és imprescindible agrair a la meva família, que des del primer moment m'han sabut donar l'acompanyament necessari per a no perdre el nord i fer d'aquest treball allò que finalment n'he fet.

Gràcies.

ÍNDEX

1. INTRODUCCIÓ.....	1
MARC TEÒRIC.....	4
2.1 Motors de Recomanació.....	4
2.1.1 Definició	4
2.1.2 Origen.....	7
2.1.4 Machine Learning	14
2.1.5 Aplicacions i Efectes.....	17
2.1.6 Funcionament	18
2.1.7 Fonts, Extracció i Tràfic de Dades.....	20
2.1.8 Tipus	26
2.1.8.1 Filtratge Basat en contingut	27
2.1.8.2 Filtratge Col·laboratiu.....	32
2.1.9 Avaluació de Recomanadors i Criteris a considerar.....	38
2.2 Motors de Recomanació a la societat.....	41
2.2.1 Neoliberalisme	41
2.2.2 Els Estímuls.....	45
2.2.3 Cultura de la Immediatesa	48
2.2.4 Echo Chamber i Filtratge Bombolla.....	52
MARC PRÀCTIC	56
3.1 Estudi de comparació de Motors de Recomanació.....	56
3.1.1 Elaboració Motor de recomanació basat en contingut	60
3.1.2 Resultats	70
3.2 Contextualització de l'entorn social.....	76
4. CONCLUSIONS.....	80
5. ANNEXOS	83

ANNEX 1 - Origen dels algorismes computacionals	83
ANNEX 2 – Origen dels Ordinadors	85
ANNEX 3 - Origen de la Intel·ligència Artificial	97
ANNEX 4 -Tipus de Machine Learning.....	104
ANNEX 5 - Entrevista Eva Martín.....	108
ANNEX 6 - Entrevista Josep Forné	108
ANNEX 7 - Codi del programa.....	108
6. REFERÈNCIES	109
6.1 WEBGRAFIA.....	109
6.2 BIBLIOGRAFIA.....	122
6.3 ÍNDEX D'IL·LUSTRACIONS	123

1. INTRODUCCIÓ

Fa uns mesos estava conversant amb un amic i em va dir que, just després de parlar amb el seu pare sobre unes vambes que li agradaven, li van aparèixer en un anunci. Això, inicialment, em va semblar irrellevant, ja que, encara que no ho pensem, és una realitat que tenim tan interioritzada que sovint obviem. Tanmateix, després de rumiar-ho una mica, em va despertar força interès. Volia saber què hi ha darrere d'aquests programes que tots usem però en el fons desconeixem i, també, donar resposta a aquest debat que tots hem tingut sobre si ens escolta el mòbil o allò que ens apareix per pantalla és simplement màgia.

Paral·lelament, ja fa anys que m'interessen molt els fonaments de la nostra societat i identitat, i entre llibre i llibre, cercant punts de vista filosòfics, econòmics i polítics sobre aquestes qüestions, sempre m'he quedat amb una pregunta incontestada: què ens caracteritza com a societat, tant individualment com col·lectiva?

Unint aquestes dues inquietuds vaig trobar un tema que, tot i ser fonamentalment de caire tecnològic, és també causa i/o conseqüència de les bases que componen aquesta era digital: els motors de recomanació. Aquestes eines, a grans trets, recomanen productes als usuaris segons el seu historial a la xarxa, de manera que, per pantalla, els surtin els elements més adients a les seves necessitats i gustos. Fan que, per exemple, la xarxa social “Tiktok” et mostri uns vídeos o uns altres.

D'ençà que la dependència del capital s'ha apoderat del sistema sociopolític en què vivim, estem cada vegada més exposats a estímuls imposats per agents que no controlen. Pensem que, amb la digitalització i el consegüent bombardeig d'informació constant, tenim un gran control sobre el nostre entorn, però això tan sols és un miratge que oculta la realitat actual.

Aquest control, encara que no ho creguem, comença amb nosaltres, amb les nostres dades, i aquí, el nostre protagonista té molt a dir-nos. És per això que, mitjançant l'estudi de com s'originen i es tracten les dades que s'usen per a nodrir els motors de recomanació, de com funcionen aquests internament i de l'impacte tan social com econòmic que tenen sobre nosaltres, s'acabarà de discernir la funció que desenvolupen a la nostra societat.

Per a centrar la recerca i fer més clar el propòsit de la mateixa, és important presentar uns objectius. Aquests, juntament amb la hipòtesi, formaran l'esquelet del treball, i determinaran també la metodologia emprada.

Endinsant-nos en els objectius, en primer lloc, el treball busca definir i exposar el procés tecnològic que hi ha darrere d'un motor de recomanació, fet cabdal per poder fer judici del seu ús. En segon lloc, busca demostrar que amb un petit canvi en el codi d'un motor de recomanació les seves recomanacions resultants varien completament. I per últim, busca relacionar les bases del sistema socioeconòmic actual amb els motors de recomanació, la qual cosa és la base del treball.

Mitjançant aquests objectius, i sobretot amb el segon, aquest treball pretén afirmar una única hipòtesi: que els motors de recomanació poden esdevenir una eina de manipulació ideològica i econòmica. Donem-li un cop d'ull.

Anteriorment, hem posat especial èmfasi a com el sistema socioeconòmic actual estableix un control sobre nosaltres. Nogensmenys, com podreu imaginar, això no ho identifiquem quotidianament amb els motors de recomanació, encara que una cosa és clara: aquests motors són molt presents dins la nostra vida. Netflix, Google, Facebook, Amazon o milers d'altres serveis els utilitzen, i allò que ens apareix per pantalla quan els emprem ve determinat per aquests algorismes recomanadors; no us fa pensar quelcom?

Tota aquesta informació que consumim i que depèn de factors que no controlem, en males mans, pot esdevenir un estri de manipulació, i hi accedim amb un simple clic. El que pretén la hipòtesi és comprovar que els motors de recomanació són manipulables per aquells que els creen, de manera que els resultats que se'ns mostren per pantalla també ho són.

Tot i això, com ja hem mencionat, per a relacionar un concepte de matís tan tecnològic com són els motors de recomanació i una realitat social com la que es pretén investigar, cal determinar bé el mode en què treballar. És per això que aquest treball consta, metodològicament, d'una divisió en 4 parts: la part teòrico-tecnològica, la part teòrico-social, la part pràctico-tecnològica i la part pràctico-social.

D'entrada, centrant-nos en el primer apartat teòric, el treball desenvolupa una recerca tecnològica que se centra en l'explicació formal i específica sobre què és un motor de recomanació. Aquí, s'estudien les maneres amb què les nostres dades s'estreuen de la xarxa, el seu ús i importància i, sobretot, s'analitzarà el funcionament, tipus, i altres consideracions a tenir en compte sobre aquests recomanadors.

En segon lloc, hi trobem una recerca social on s'investiguen algunes conseqüències de l'esclat postmodernista i l'estructura socioeconòmica neoliberalista, així com la cultura de la immediatesa, la sobrecàrrega d'estímuls o el nou “Echo Chamber”, transportant-ho al món dels motors de recomanació. D'aquesta manera, es tracta d'analitzar si aquests “simples” recomanadors de productes segons les necessitats de qui els usa tenen més a dir-nos del que creiem.

D'altra banda, el marc pràctic consisteix en, inicialment, l'elaboració d'un motor de recomanació on se simularà el procés que es fa en gairebé totes les webs en línia. Veurem com amb una petita modificació en el codi i dues execucions es comprova o desacredita la hipòtesi plantejada.

I per últim, tornant a l'anàlisi social, la part pràctica comprèn un capítol que se centra a fer un apropament a la realitat descrita en l'apartat teòric. Aquí, a partir d'enquestes realitzades a la població, es contesten diverses preguntes que deriven de la resposta a la hipòtesi, i es dona una visió més aplicada de tot l'explicat anteriorment.

Formalment, s'utilitzen eines de recerca com llibres i pàgines web, així com cursos per a poder realitzar la part pràctica tecnològica. Concretament s'han requerit coneixements del llenguatge de programació Python en la secció del Machine Learning aplicat. A més a més, s'han realitzat entrevistes a experts sobre el tema per a introduir altres punts de vista.

Els nostres temps van lligats a una suposada llibertat, però és aquesta certa?

Confrontar uns algorismes que poden adaptar-se perfectament a la realitat de cadascú amb una societat que anteposa la moneda abans que a la persona és un perill que no hem d'ignorar. Motors de recomanació, una arma de doble fil?

MARC TEÒRIC

2.1 Motors de Recomanació

2.1.1 Definició

No heu pensat mai per què quan dues persones obren qualsevol xarxa social al seu mòbil troben dues coses diferents per pantalla? Per què jo, per exemple, si obro Twitter, puc trobar-me una publicació de futbol mentre que el meu amic una campanya política?

Imagineu, prenent l'exemple de Twitter, que enlloc que es mostressin en pantalla publicacions concordes al que us interessa, a la gent que coneixeu, entre altres criteris, sortís a tothom el mateix sense cap mena de distinció. Oi que no tindria sentit?

Com bé pensareu, això és perquè internet, i sobretot les xarxes socials (dirigides a l'usuari), necessiten un ordre. Doncs aquest és el paper dels motors de recomanació. Fer que aquesta llista infinita d'informació sense sentit apparent prengui un ordre segons qui l'ha de llegir. Fan que internet no sigui un abocador d'informació sense solta ni volta, ordenant-la segons l'usuari que la cerca.

Posant-nos més tècnics, per definició, un motor de recomanació és una subclasse de sistema de filtratge d'informació dins del "Machine Learning" que ofereix suggeriments basats en elements que són rellevants per a un usuari concret.¹ Resumidament, els motors de recomanació són un conjunt d'algoritmes que reconeixen patrons de consum entre els usuaris i recomanen productes basant-se en això, encara que, com ja hem introduït, el seu funcionament i sobretot funció a la societat no és tan fàcil.

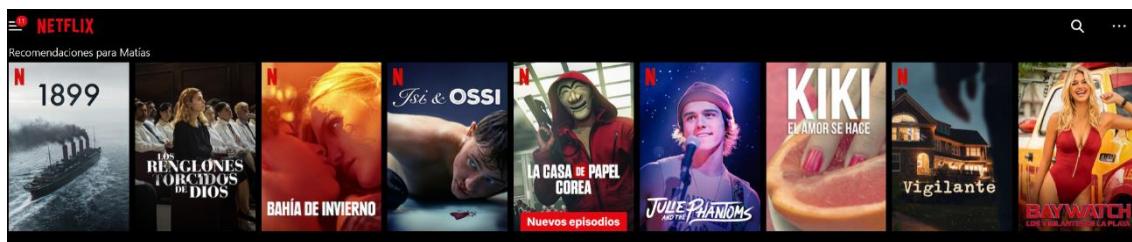
És important no confondre'l s amb recomanadors de valors arbitraris, ja que llavors estaríem parlant d'un altre tipus de "Machine Learning" que no ordena segons usuari. El que fa un motor de recomanació és predi la valoració que un usuari li donaria a un ítem mitjançant dades passades que l'usuari ha donat. Amb aquesta predicció, per exemple, pot

¹ (Recommender System - Wikipedia, n.d.)

crear una llista d'ítems que més s'apropen al comportament passat de l'usuari i recomanarlos.

Encara que el nom no inspiri a res en concret, és una tecnologia que trobem arreu en el nostre dia a dia. Totes les cerques de Google, posts a xarxes socials o fins i tot articles de botigues “online” ens apareixen davant la pantalla per alguna raó, i aquesta és els motors de recomanació. De fet, en el món “online”, és realment rar trobar quelcom que no estigui on està per cap raó, ja que molts dels milions d'articles que se'ns apareixen davant de la pantalla que creiem aleatoris són realment prediccions del nostre comportament fetes per fer-nos continuar consumint.

Un exemple de funció que utilitzem diàriament és la d’“Et Recomanem”, de Netflix, on segons les pel·lícules que hem vist anteriorment i la valoració que li hem posat, se'ns mostren certes recomanacions.



Il·lustració 1: Funció "Et Recomanem", de Netflix

Font: Netflix

Cal dir, de la mateixa manera, que encara que en termes específics semblin un estri fet només pel benefici de les empreses, aquests motors són tecnologies vertaderament útils, i que, a termes pràctics i superficials, ens ajuden molt a tots. Quan pretenem cercar a internet quelcom o entrem a una pàgina web com “Amazon”, és de gran ajuda que una base de dades ens ajudi a trobar allò que volem fàcilment mitjançant el nostre comportament a la xarxa. Sense ells, hauríem de navegar per tota la xarxa fins a trobar allò que requerim, i això faria que internet, encara que menys additiu, fos una enciclopèdia infinita de termes desglossats. Podríem dir que els motors de recomanació ens ordenen personalitzadament els resultats de qualsevol cerca.

Per exemplificar-ho, aquesta fàcil comparativa entre dos usuaris diferents de la xarxa social “Twitter” i la mateixa cerca “Juego” mostra, a grans trets, el funcionament d’aquests algorismes.



The image shows two side-by-side screenshots of a Twitter search interface. Both screens have a search bar at the top with the query "juego".

Left Screen (User 1):

- Top Result:** User **micol** (@tizianamicols) posted 19 hours ago: "Thiago le dice a Daniela que hablo con la Tora y le dijo que salve a quien ella quiera, y que deje de pensar en sacar a Agustín porque él no se va a ir, si lo salva a Juan se va Nacho." Below the tweet: "No puedo creer por fin alguien entendió el **juego** y fue Thiago, el verdadero plot twist #Gh2022".
- Second Result:** User **;indigo** (@myglosshy) posted Nov 10: "El álbum de NamJoon se llama Índigo que en coreano su escritura sería: "남색/namsaek". Al mismo tiempo, esto sería un **juego** de palabras de Nam (NamJoon) + Saek (Colores), de forma abreviada esta palabra significaría "Los colores de NamJoon" Oh, WOW @BTS_twt".

Right Screen (User 2):

- Top Result:** User **VarskySports** (@VarskySports) posted 19 hours ago: "Cesar Luis Menotti. Nadie lo explica este **juego** mejor que él." Below the tweet is a video thumbnail of Cesar Luis Menotti speaking.
- Second Result:** User **Veronica Brunati** (@Veronica_Brunati) posted 3 hours ago: "Valorant ESP 🇪🇸 @Valorant_ES - 3h Meme 😂: "Juego en Equipo" 😂 | #VALORANT". Below the tweet is a blacked-out video thumbnail.

Il·lustració 3: Cerca Usuari 1: "Juego"
Font: Twitter

Il·lustració 2: Cerca Usuari 2: "Juego"
Font: Twitter

Independentment del contingut, es pot observar com, segons l’usuari que faci la cerca, el resultat obtingut varia. Aquesta propietat fa que, com hem dit anteriorment, les cerques s’adeqüin a cadascú gràcies al motor de recomanació pel que les seves dades són processades. Factors com la localització, la gent que segueixes, o fins i tot persones qualsssevols amb un historial de cerca similar al teu són factors cabdals per aquesta recomanació.

2.1.2 Origen

La cerca de l'òptim ha estat sempre una gran prioritat i necessitat humana, ara a més a més potenciada pel capitalisme. Tanmateix, no sempre hi ha hagut intel·ligències artificials que compleixin aquesta necessitat i que ens ofereixin allò que el nostre cervell vol rebre com fan els motors de recomanació, sinó que hem utilitzat altres tècniques. Moltes de les tecnologies que per a molta gent es consideren noves i futuristes, són només una versió informatitzada d'un aspecte que defineix la nostra espècie.

Els sistemes de recomanació poden ser un concepte nou en l'àmbit del màrqueting i de la investigació, però han estat dominant la societat des de fa temps. De fet, l'acte de recomanar no només és present en l'ésser humà, sinó que també el podem trobar en el funcionament psicomotor d'altres espècies com, per exemple, les formigues. És molt comú el fet de veure formigues corrent per casa nostra caminant en fila darrere d'altres que han trobat menjar. Això es deu al fet que les formigues han evolucionat genèticament per deixar marcadors que actuïn com a recomanació per a altres formigues, mostrant-los el camí cap al menjar.

Apropant-nos més al que ara entenem com una recomanació tecnològica, trobem també el comportament humà, en aquest cas explícit mitjançant certs costums o formes de relacionar-se en societats passades, a les civilitzacions antigues. Aquestes, com apunt, van formar-se entre el període 4000 a 1200 aC. Si parlem de recomanacions en aquella època, possiblement aquestes podrien anar des de quin cultiu conrear i durant quin temps, fins a temes molt més decisius en la presa de decisions de l'individu com per exemple quina religió seguir. Més tard, durant el període de colonització entre els segles XI-XVIII, les recomanacions també van ser presents a l'hora de decidir quin territori colonitzar d'acord amb els aspectes fructífers com la fertilitat de la terra, la mà d'obra (per a l'esclavitud) i molts altres recursos. Per exemple, quan hi havia regnes arreu del món, el rei solia rebre suggeriments dels seus ministres o consellers sobre gairebé tots els assumptes que afectaven el regne. De fet, hi havia un grup de ministres que exclusivament s'encarregaven d'exposar el seu punt de vista davant del rei per ajudar-lo a prendre la decisió que ell acabés prenen com a bona amb major facilitat. Suggerir quin regne conquerir davant d'un altre, quines polítiques s'han d'executar pel benefici de les masses,

entre altres innumerables qüestions són exemples del gran pes que tenen i han tingut les recomanacions en la història de l'espècie. Els exemples anomenats són decisions reials amb conseqüències a gran escala, però a un nivell inferior i quotidià també trobem aquestes recomanacions com un factor molt important. Sense anar gaire lluny respecte a l'anterior apartat, quan antigament les famílies solien concertar matrimonis, normalment era algú de la mateixa família el que troava un candidat per concertar. Aquestes eren, de nou, recomanacions.

Apropant-nos més encara en l'actualitat, les persones que demanen opinions d'altres sobre on comprar els millors productes, quin destí és bo per a les vacances, o el camí més curt per arribar a alguna destinació, entre altres, són recomanacions presents en la societat des de fa molt de temps. No hi havia ordinadors al vell món, però tot i això la gent solia rebre les recomanacions dels seus companys. Ara tenim aplicacions que ens diuen on comprar els millors productes segons el que busques, on anar de vacances segons el preu i gustos i la millor ruta per anar a qualsevol lloc segons el tràfic. Però el fet que no sempre hagi estat així no exclou que les recomanacions hagin estat sempre presents.

El punt d'inflexió que ens ha portat a la concepció de recomanació que tenim ara va ser la revolució industrial i la seva conseqüent globalització. Amb aquesta, no només les opcions en molts àmbits van augmentar, sinó que l'era de l'ordinador i de la computació informàtica es va concebre i va revolucionar el mercat global. La gent tenia tantes opcions davant seu que això portava a una certa confusió sobre quin producte o servei realment complia amb les seves necessitats. Va ser llavors que es va entendre la necessitat de tenir un sistema que facilités aquesta selecció i que erradiqués els dilemes de les masses, creant la idea del sistema de recomanació actual.

Els humans confiem gairebé a cegues en recomanacions de fonts externes a nosaltres per gran part de les nostres decisions quotidianes, ja que les opcions entre les quals ens veiem obligats a triar i la complexitat d'aquestes, augmenten dia rere dia. Per exemple, a l'hora de comprar un mòbil nou, visitem un gran nombre de pàgines web on ens guiem d'opinions basades en l'experiència de tercers, fonts alienes a un mateix. El mateix exemple s'aplica a la compra d'un vehicle, a veure una pel·lícula, a llegir un llibre, entre moltíssimes altres. De fet, fins i tot es pot aplicar a l'hora de trobar parella, acció que ara les pàgines web de cites "online" ens faciliten, fet comparable a què feien abans les

famílies amb el matrimoni concertat esmentat abans, però només amb uns anys de diferència.²

Els sistemes de recomanació, però, van evolucionar com a àrea de recerca independent a mitjans dels anys setanta a la Universitat de Duke, als Estats Units. Des d'aquell moment, aquest àmbit de recerca ha anat desenvolupant-se fins a constituir la base de les grans pàgines web comercials, sistemes de sondejos polítics, entre altres. El primer sistema de recomanació concretament va ser Tapestry1, el qual va ser dut a terme al Centre de Recerca de Xerox Palo Alto, a principis del 1990. Tot i que aquest va ser el primer amb finalitat comercial, es diu que anteriorment hi varen haver altres amb un objectiu de recerca, com bé va ser Grundy, un motor de recomanació que l'any 1979 funcionalment recomanava llibres. La motivació que va portar al desenvolupament de Tapestry1 va ser l'abundant nombre de correus electrònics, la majoria innecessaris, que hi havia a les bústies de molts usuaris, cosa que feia molt difícil accedir al contingut que aquest volia. El que es va fer per superar aquest problema fou fer que els mateixos usuaris creessin una llista de contactes finita, de manera que els correus provinents d'una font aliena a aquella llista simplement es posaven al correu brossa, similar a com succeeix actualment.

Aviat els sistemes de recomanació es van fer més populars i ara juguen un paper important en llocs d'Internet com Amazon, Netflix, YouTube, Google, Tripadvisor, entre milions d'altres. De fet, es pren en consideració que en molts casos és precisament la qualitat del sistema de recomanació el que fa un servei o pàgina web sobresortir davant d'una altra.

Òbviament, hi ha hagut un gran salt des del primer sistema fins a l'actualitat, amb ajuda de tecnologies com la intel·ligència artificial, la interacció home-ordinador, la recuperació d'informació i moltes altres que després explicarem; però la base és clara: la mateixa naturalesa humana.

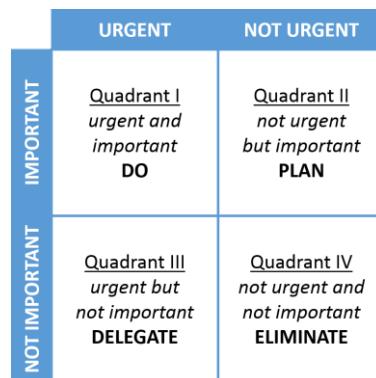
² (*Evolution of Recommender Systems from ancient times to modern era: A survey.*
(n.d.). Retrieved September 29, 2022)

2.1.3 Algorismes

Els motors de recomanació, endinsant-nos en el seu concepte, estan conformats per certs algorismes i es classifiquen en molts tipus, però per entendre això s'ha de tenir abans clar què és essencialment un algorisme.

I és que un algorisme no és més que un conjunt finit d'instruccions o passos que serveixen per a executar una tasca o resoldre un problema, i per molt que últimament aquest mot hagi pres el complex sentit de ser purament informàtic, els algorismes han estat sempre presents en les nostres vides.³

Etimològicament, la paraula és d'origen àrab, i s'estima que es va crear al segle IX. Això es deu al fet que va ser el matemàtic Mohammed Ibn Musa-al-Khwarizmi el que va establir el nom després de desenvolupar maneres de possibilitar que amb un nombre finit de processos fos possible resoldre equacions de primer i segon grau. Ençà, la paraula es limitava a les normes de l'aritmètica aràbiga, però al segle XVIII es va expandir el seu significat pel qual coneixem ara.



Il·lustració 4: Matriu de Covey: exemple d'algorisme

Font: Wikipedia

Bàsicament, els algorismes són una sèrie d'instruccions seqüencials emmagatzemades en una memòria, que una darrere l'altra, ens faciliten la realització d'accions i programes, és

³ (*Algorisme* - Viquipèdia, l'enciclopèdia Lliure, n.d.)

a dir, una sèrie de càlculs. Adaptat a la nostra vida quotidiana, un algorisme pot ser una simple recepta de cuina, el procés generalitzat d'un exercici de matemàtiques, o, fins i tot, l'anàlisi mètrica d'una poesia. De fet, inclús el nostre cervell fa càlculs complexos algorítmics a l'hora de prendre decisions, així com la gestió de prioritats, que podem correspondre conscientment a estudis com la matriu de Covey o altres processos lògics. Seguint amb aquest tema, el guanyador del Premi Nobel de Medicina al 2014, Edvard Moser, precisament va investigar sobre els algorismes paral·lels ⁴interactius amb què funciona el nostre cervell; en altres paraules, com computen les nostres neurones.

Sembla que, en traslladar els algorismes al territori quotidià, el seu funcionament s'hagi simplificat; però la realitat és que aquesta agrupació de dades i seqüència de passos pot prendre moltes formes diferents (com ja hem aproximat amb l'estudi Nobel esmentat anteriorment). Podem trobar algorismes computacionals i no computacionals, i dins d'aquests últims qualitatius i quantitatius, encara que es divideixen en moltes altres seccions segons la complexitat de la seqüència.⁵

A partir d'aquest punt, coneixerem les regles bàsiques de l'estructura tripartida que ha de conformar un algorisme per a ser considerat com a tal. Bàsicament, aquests han d'estar composts per tres parts principals, que són: entrada, procés i sortida, independentment de tractar-se d'algorismes computacionals, algorismes no computacionals, algorismes qualitatius o algorismes quantitatius.

La fase d'entrada de l'algorisme es correspon amb la dada d'entrada amb què cal alimentar-lo. Bàsicament, es tracta de la informació que es lliura a l'algorisme, és a dir, els valors de dades amb què haurà d'ofrir un determinat resultat.

La segona fase d'un algorisme, el procés, ocupa els càlculs necessaris per processar les dades amb què hem alimentat l'algorisme. Per tant, en aquest punt es processarà la

⁴ Un algorisme paral·lel és, en oposició als algorismes clàssics o algorismes seqüencials, un algorisme que pot ser executat per parts en el mateix instant de temps per diverses unitats de processament, per finalment unir totes les parts i obtenir-ne el resultat correcte.

Font: Wikipedia

⁵ (*> Algoritmos Cotidianos » Qué Son, Utilidad, Tipos y Ejemplos*, n.d.)

informació lliurada a l'entrada de l'algorisme i amb la qual s'ha d'arribar al resultat esperat. La tercera i última fase d'un algorisme, la sortida, és on s'obté el resultat de l'operació, és a dir, la transformació de les dades que van ser proporcionades a la fase d'entrada i desenvolupades en la fase de procés.⁶

Un exemple d'algorisme no computacional qualitatiu, com pot ser fregir un ou, prendria aquesta forma:

1. Obrir la Nevera
2. Agafar un ou cru
3. Agafar la paella
4. Posar oli a la paella
5. Obrir el foc
6. Posar la paella sobre el foc
7. Esperar X segons fins que s'escalfi l'oli
8. Trencar la closca de l'ou
9. Posar la clara i el rovell a la paella
10. Esperar X segons fins que l'ou es fregeixi
11. Treure l'ou de la paella
12. Servir l'ou al plat

De la mateixa manera, aquest procés també pot prendre complexitat, per exemple, afegint operacions condicionals, com bé exemplifica aquest algorisme per buscar un parell de sabates per una festa.

1. Entrar a una botiga de sabates
2. Agafar un parell de sabates
3. Són sabates de festa?
SI: (anar al pas 4) – NO: (tornar al pas 2)
4. Hi ha la talla pertinent?
SI: (anar al pas 5) – NO: (tornar al pas 2)
5. El preu és viable?
SI: (anar al pas 6) – NO: (tornar al pas 2)
6. Comprar el parell de sabates elegit

⁶ (*¿Qué Es Un Algoritmo? ¿De Donde Surge?* – Programa En Línea, n.d.)

Aquest tipus de procés, però, també pot agafar forma d'un codi de programació, i conformar la base d'aplicacions com per exemple Google Maps, intel·ligències artificials, entre altres. De la mateixa manera, també pot ser el mecanisme d'una màquina a vapor, o la base de termes matemàtics com l'algorisme d'Euclides⁷, un mètode eficaç per calcular el màxim comú divisor.

Bàsicament els algorismes, així com les recomanacions, són un concepte molt important dins del nostre dia a dia, fins i tot formen part de la naturalesa humana, i a diferència de la creença popular, són presents d'ençà que tenim ús de consciència i memòria com a espècie. Els algorismes informàtics, però, no van ser una tecnologia que va sortir d'un dia a l'altre, sinó que estan envoltats d'una història que comença molts anys enrere.

Com hem dit, els algorismes com a tal porten existint d'ençà que el primer tipus de memòria (ésser viu) va realitzar una sèrie de passos finits i definits. No obstant això, la seva traducció a un llenguatge computacional inanimat, va tardar una mica més. Per a saber-ne més, la història dels algorismes informàtics es troba a:

→ANNEX 1 - Origen dels algorismes computacionals

Actualment, els algorismes informàtics són una part gairebé essencial del nostre dia a dia, ja que es troben en tots els processos que passen en una màquina computacional (ordinador, mòbil...). Estem envoltats de processos mil·limetrats i desconeguts per la gran majoria, que componen les tecnologies que dia rere dia, cegament, utilitzem. És veritat que els algorismes informàtics han evolucionat en estructura i aplicació, però la seva essència es troba dins de tots nosaltres. Actualment, en un món computacional, som només una petita variable.

⁷ Euclides, també conegut com a Euclides d'Alexandria, (va viure cap al 300 aC), fou un matemàtic grec, conegut avui dia com a pare de la geometria. El seu treball més famós fou els Elements, considerat sovint el llibre de text de més èxit de la història de les matemàtiques. S'hi dedueixen les propietats dels objectes geomètrics i dels nombres naturals a partir d'un petit conjunt d'axiomes.

Font: Viquipèdia

2.1.4 Machine Learning

Com hem mencionat anteriorment a l'apartat d'origen, una de les tecnologies més importants que compon l'estructura d'un motor de recomanació és el “Machine Learning”.

El “Machine Learning”, com el seu nom indica, és una disciplina que, a través d'algorismes, dona als ordinadors la capacitat d'identificar patrons de dades massius i elaborar prediccions. És un subcamp de les ciències de la computació i una branca de la intel·ligència artificial. El seu objectiu és trobar tècniques per a tornar un procés computacional “intel·ligent”, és a dir, que pugui prendre decisions segons el context.⁸

Ara mateix la seva aplicació és cabdal en àmbits com els assistents虚拟 o els vehicles autònoms, però la seva història va una mica més enrere. Com diu el científic de dades José Luis Espinoza, “En definitiva, el Machine Learning és un mestre del reconeixement de patrons, i és capaç de convertir una mostra de dades en un programa informàtic capaç d'extreure inferències de nous conjunts de dades per als quals no ha estat entrenat prèviament.”

L'estadística és la base fonamental d'aquest aprenentatge automàtic, ja que, essencialment, aquest consisteix en una sèrie d'algorismes capaços d'analitzar gran quantitat de dades per decidir quin és el resultat òptim per un determinat problema.

De tota manera, creiem que és pertinent conèixer abans l'inici d'allò que entenem com a ordinador i l'origen de les idees primitives que estableixen la idea del “Machine Learning” per a entendre correctament com s'han pogut ajuntar aquests dos elements i crear aquesta tecnologia tan innovadora.

En primer lloc, la història del que coneixem com “ordinador”, es troba aquí.

→ANNEX 2 – Origen dels Ordinadors

⁸ (Alpaydin, 2020; *How to Explain Machine Learning in Plain English / The Enterprisers Project*, n.d.; *What Is Machine Learning? / IBM*, n.d.; Nilsson, 1998)

Tot seguit, la història del “Machine Learning” i la intel·ligència artificial es troba aquí:

→ANNEX 3 - Origen de la Intel·ligència Artificial

Aquesta aplicació i branca de la intel·ligència artificial, la qual és la base dels algorismes que aquest treball tracta, dona possibilitats quasi infinites a l'ésser humà pel que fa a control de dades i sobretot en quant a l'aplicació d'aquestes. Un sistema en el qual el seu patró decisiu està basat en milions de reproduccions, pot arribar a conclusions i resultats gairebé perfectes, i és aplicable a gairebé totes les disciplines del món en què vivim.

La “màgia” del “Machine Learning” és la seva capacitat de construir i adaptar un arbre de decisions d'acord amb les dades conegeudes, i aplicar a més fórmules heurístiques a cada node de l'arbre, constraint un sistema d'inferències⁹. A més a més, ho fa de forma recursiva, és a dir, construeix un petit arbre de decisió per cada branca que va creant, tornant fins i tot a tenir en compte de nou dades que ja han servit per arribar o descartar la branca actual.¹⁰

Dins del “Machine Learning”, com en totes les disciplines, també hi ha una certa classificació basada en el procés d'aprenentatge: l'aprenentatge supervisat, no supervisat i reforçat. El que utilitzen la majoria de motors de recomanació és el segon.

Per saber-ne més, els tipus de “Machine Learning” es troben aquí

→ANNEX 4: Tipus de Machine Learning

De tota manera, el més important d'aquest apartat és entendre el seu paper dins dels motors de recomanació. Sovint malinterpretarem aquest concepte, i pensem que el

⁹ En l'àmbit de la intel·ligència artificial, un motor d'inferència és un component del sistema que aplica regles lògiques a la base de coneixement per deduir nova informació.

Font: Wikipedia

¹⁰ (*La Historia Del Machine Learning - ELearning Actual*, n.d.-a; *Machine Learning. Qué, Cómo, Cuándo, Dónde*, n.d.)

“Machine Learning” és algun tipus de màgia o de principi que introduïm a un codi i fa que el nostre algorisme es torni intel·ligent, i no és ben bé així.

Realment el “Machine Learning” és una àrea computacional que interpreta patrons i estructura dades per a poder prendre decisions, sempre establertes en un codi que el delimiti. D'aquesta manera el “Machine Learning” permet crear algorismes intel·ligents que tinguin com a exemple a seguir grans quantitats de dades.¹¹

Precisament això és el que fa amb els motors de recomanació. Per ser més exactes, els motors de recomanació són una branca del “Machine Learning”, ja que per definició són processos computacionals que determinen una recomanació guiats per quantitats molt grans de dades.

El Machine Learning, dins d'aquest àmbit, no és cap ànima lliure que pren decisions per si sol, cosa que sovint pensem, sinó que ens ajuda a relacionar productes molt eficientment. L'essencial que hem de saber és que això vol dir que mitjançant moltíssims exemples (cal pensar que un ordinador arriba a fer 100 milions de càlculs per segon), la "màquina", és a dir, el recomanador, estableix patrons entre productes, i d'aquesta manera recomana segons els exemples que ha rebut al passat.

És gràcies a això que els recomanadors poden arribar a fer recomanacions que primerament semblen impossibles de fer amb l'historial de dades i valoracions de l'usuari, però mitjançant milions de càlculs i establint milers de patrons de consum entre diferents productes, ho recomana.

Per exemple, tot i que un usuari mai hagi valorat positivament una pel·lícula de terror, un recomanador pot recomanar-li'n pel fet que les persones que han valorat positivament les mateixes pel·lícules que l'usuari també ha valorat, n'han valorat positivament de terror.

¹¹ (*What Is Machine Learning - ML - and Why Is It Important? / NetApp, n.d.*)

Aquest és l'aport del Machine Learning, una intel·ligència artificial que sembla que pren decisions tan humanes com és recomanar algú mitjançant milions i milions de dades. Per això les dades són tan importants, així com la nostra petjada a la xarxa; D'elles es determina tot.

2.1.5 Aplicacions i Efectes

Després d'aquesta extensa contextualització sobre el món que envolta els motors de recomanació, cal endinsar-se en l'ampli món que aquests sistemes intel·ligents anomenats motors de recomanació ens proporcionen, repassant el seu concepte.

Els motors de recomanació són sistemes avançats de filtratge de dades que prediuen quin contingut, producte o servei és probable que consumeixi o interacció un client. No cal mirar lluny per veure'n en acció. Cada vegada que algú tria un programa de televisió amb la funció "També et pot agradar..." de Netflix o compra un producte que recomana Amazon, està utilitzant potents motors de recomanació. Aparentment, els motors de recomanació són funcions de "win-win" tant per als clients com per a les empreses que els implementen, ja que el client obté allò que segons el seu perfil requereix (o creu que requereix) i l'empresa té un benefici econòmic.

Cal dir que els motors de recomanació només processen dades, és a dir, no són els complets responsables de què cliquem un ítem o no. M'explico, si a una pàgina web hi ha un motor de recomanació molt ben entrenat que et mostra allò que més s'adqua als teus desitjos consumistes, però no hi ha una bona interfície visual, no servirà de res. Podríem dir que els motors de recomanació són la intel·ligència genotípica, la que presenta aquells elements que més s'adequen a cada usuari, però després hi ha altres branques com el neuromàrqueting que s'encarreguen de l'apartat més "visual".¹²

¹² (*Recommendation Systems: Applications and Examples in 2022*, n.d.)

Els motors de recomanació milloren àmbits de la informàtica aplicats al públic. Entre altres aplicacions , trobem:

- Pàgines web destinades a la venda (Netflix, Amazon)
- Jocs d'ordinador
- Bases de dades
- Filtratge d'informació (Separadors automàtics d'informació)
- Sondejos Polítics (Bases de dades amb activitat política)
- Bancs (Bases de dades amb activitats monetàries que recomanen què fer)

Generalment, un motor de recomanació és molt útil i beneficiós per a qualsevol empresa que s'encari a algun d'aquests àmbits. Primerament, la satisfacció de l'usuari que utilitza el servei augmenta, ja que si les recomanacions són adients, trobarà allò que (suposadament) necessita o cerca abans d'haver-ho de buscar. Això, obrint pas al segon punt, conseqüentment donarà més fiabilitat al servei web, i de la mateixa manera els ingressos seran majors.¹³ Com hem vist, dins del model de vida en què vivim, on el consum i la rapidesa són prioritaris, els motors de recomanació són realment útils. Ara bé, és això bo o dolent? Això, com veurem a l'apartat social, entra dins del criteri de cadascú, i del seu judici dins del sistema socioeconòmic.

2.1.6 Funcionament

Fins ara hem estat veient molts aspectes tècnics que potser ens allunyen del concepte principal dels motors de recomanació. Com veurem posteriorment, aquests prenen diverses formes segons el tipus, però que és necessari fer un resum de com, generalment, funciona un motor de recomanació. Ho farem mitjançant passos:

1. Agafa l'historial de dades d'un usuari.
2. Agafa un ítem aleatori (dins de tots els que té a la base de dades) que pot recomanar.

¹³ (*Real World Applications of Recommendation Engine - Muvi One*, n.d.)

3. Compara aquest ítem amb tots els ítems que ha valorat l'usuari, guiant-se per la semblança entre els ítems, que determina gràcies al Machine Learning que prèviament hem entrenat.
4. Estableix una predicció de la valoració que li posarà l'usuari a l'ítem triat
5. Repeteix això amb tots els ítems
6. Elabora una llista dels ítems amb màxima predicció de valoració
7. Recomana la llista

Posem-ho amb un exemple, en aquest cas amb pel·lícules (es fàcil de trobar semblances):

1. El motor rep un historial de l'usuari on sap que ha valorat positivament “Avatar” i negativament “Fast and Furious”.
2. Dins de totes les pel·lícules que té a la base de dades, tria la Pel·lícula “Pocahontas”.
3. Veu que (gràcies al Machine Learning) “Pocahontas” és semblant a “Avatar”, però no a “Fast and Furious” (cosa que és positiva).
4. Prediu que l'usuari li posaria un 4.5/5 a la pel·lícula “Pocahontas”.
5. Repeteix això amb altres pel·lícules.
6. Veu que la pel·lícula que major predicció de valoració té de totes les que ha fet és “Tarzán”.
7. La recomana.

Aquest procés el podríem trobar (encara que més complexe) a la funció anteriorment esmentada de “Et Recomanem”, de Netflix. Cal dir que no sempre pren aquesta forma, ja que en alguns casos concrets com quan hi ha milions de ítems això no és eficient, però l'esquema general és aquest.

2.1.7 Fonts, Extracció i Tràfic de Dades

Quan hem definit motor de recomanació hem esmentat reiterades vegades que hem de partir d'una base de dades molt gran. Tot i això, no hem acabat de definir quines dades extraiem ni per a què serveixen.

Primerament, definirem les dues naturaleses de dades que podem obtenir d'un usuari per a després fer recomanacions segons aquestes. En primer lloc, estan les dades explícites, que són les qualificacions que l'usuari explícitament posa als articles. Si ho traduïm a l'àmbit de les xarxes socials, serien els “likes”. Aquest tipus de dades també poden prendre forma d'una valoració numèrica, com per exemple seria valorar de l'1 al 10 una pel·lícula. Per tant, si haguéssim de definir aquest terme, podríem dir que són les dades que el motor de recomanació rep d'un usuari directament comunicades des d'ell, atès a la seva implicació en la valoració d'un ítem.

En canvi, les implícites són aquelles dades que no estan activament definides per l'usuari, com per exemple, el temps de visualització d'una pel·lícula. Encara que l'usuari no hagi donat una valoració a l'ítem en qüestió podem suposar que, a partir de cert nombre de minuts, l'ítem ha agradat a l'usuari. Aquestes dades poden prendre moltes formes. Per exemple, historials de cerca, historials de compra (que són molt fiables a l'hora de classificar dades), comentaris, hores de visualització, productes més vists, entre altres. Per exemplificar-ho un altre cop amb les xarxes socials, podríem dir que aquestes dades són les que l'usuari dona, per exemple, quan entra a un mateix perfil repetits cops, indici de què aquell perfil li agrada. Per raons òbvies, això farà que la fiabilitat de les dades implícites sigui menor, ja que, per exemple, algú es pot deixar el mòbil obert sense voler i visualitzar un vídeo molts cops.

Per a fer-nos una idea, només l'1% dels ítems visitats pels usuaris són valorats, així que encara que les dades explícites tinguin més rellevància atès a la seva fiabilitat, les dades implícites són les que acaben determinant una recomanació. És curiós que, per exemple, Netflix no destaquí pel nombre de persones que valoren un article, però el seu sistema de recomanació sigui molt bo. Això és tot gràcies a les dades implícites.

A més a més, les dades explícites també tenen un problema. Els criteris “d’avaluació” són un concepte molt relativista, i depèn bastant de la cultura i del país. Potser a un lloc del món és molt normal valorar les coses amb la màxima puntuació, però a un altre país no ho és tant. En aquest cas, encara que la satisfacció de dues persones sigui la mateixa, les dades rebudes no serien iguals.¹⁴

Un dels tòpics més controvertits dels que es parla dins d'aquest món és si realment existeix el tràfic de dades “online” i si hi ha llocs que no sabem d'on ens extreuen dades.

Un exemple de tot això són les milers de vegades esmentades “cookies”. Aquesta desagradable alerta que rebem en entrar a gairebé qualsevol lloc web i que sense parar-hi gaire compte acceptem té molta relació amb tot això. El que fem en acceptar-ho és donar permís a la web a que prengui dades com l'idioma del nostre navegador, interessos en funció de la ubicació, que es guardi o no el nostre usuari, entre altres. Les “cookies” són la manera de “legalitzar” aquesta recollida de dades implícita, que després pot ser utilitzada per recomanar un tipus de publicitat, servei, o article determinat.¹⁵

Tot i això, les “cookies” que verdaderament ens importen son les de tercers, que són aquelles que no provenen de la pàgina a què accedeixes, sinó d'altres empreses. Aquestes són generades per proveïdors externs a través d'anuncis, píxels de seguiment o similars. Recullen informació rellevant com l'edat, la ubicació, el sexe i el comportament dels usuaris, cosa que les converteix en potents eines per al màrqueting “online”, sobretot en l'àmbit de la publicitat personalitzada, i, per tant, dels motors de recomanació.¹⁶

Concretament, atès que als analistes web els hi interessa sobretot el comportament de navegació dels usuaris, les “cookies” de tercers documenten principalment l'historial de subpàgines visitades d'una web. D'aquesta manera, la informació veritablement rellevant l'obtenen aquestes cookies de seguiment quan “reconeixen l'usuari” en una altra web.

¹⁴ (*El Filtrado Colaborativo y Los Motores de Recomendación – Juan Barrios*, n.d.-b; *The Difference Between Implicit and Explicit Data for Business / by Tim Clayton / Blog / Mirumee*, n.d.; *What Is Implicit Data? - Definition from WhatIs.Com*, n.d.)

¹⁵ (*Qué Significa Aceptar Las Cookies al Entrar En Una Página Web*, n.d.)

¹⁶ (*What Is a Third-Party Cookie?*, n.d.)

Com que el navegador es torna a comunicar amb el mateix servidor d'anuncis, aquest pot rastrejar els moviments de l'usuari a Internet. És el seu comportament a la web el que revela informació central sobre els seus interessos i el seu patró de consum. D'aquesta manera es crea el perfil d'usuari, que permet la publicitat personalitzada que tant coneixem.

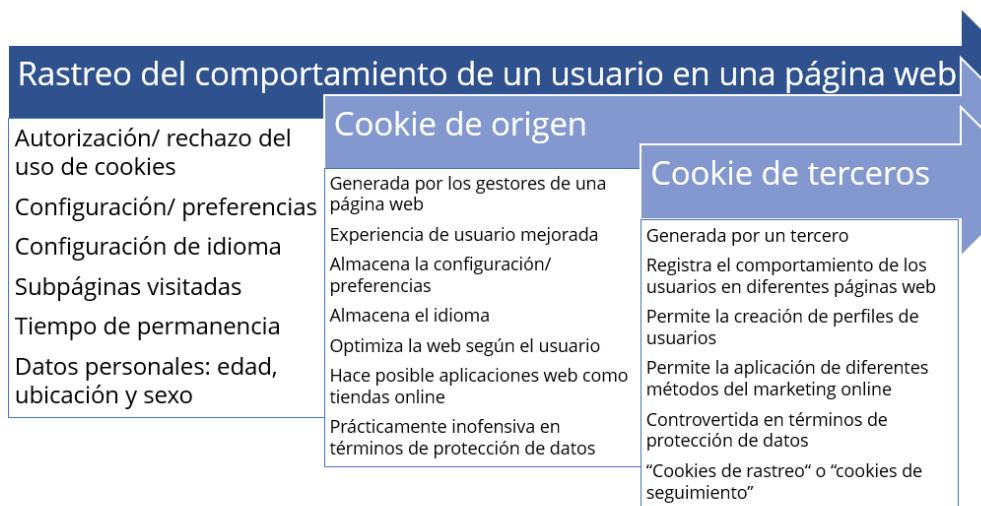
Aquest és un exemple del seu funcionament:

1. En primer lloc, la pàgina informa sobre l'ús de “cookies” amb un missatge aclaridor sobre l'ús de cookies. Si l'usuari accepta, el navegador generarà dues “cookies”: una “cookie” d'origen del gestor de la pàgina i una altra de tercera de l'Ad Server¹⁷. A partir d'aquell moment les dues “cookies” recolliran informació sobre el seu comportament a la pàgina web.
2. La pàgina web mostra una botiga en línia de moda a la qual l'usuari li vol donar una ullada. Com que està buscant una bossa marró, busca a totes les subpàgines que ofereixen aquest tipus de producte. Hi ha dues bosses que li agraden i, per no perdre'ls de vista, les afegeix a la cistella de compra. Una “cookie” d'origen anota aquesta informació perquè no es perdi, fins i tot si en aquell moment l'usuari no està registrat amb el compte. La “cookie” de tercera també ho anota, encara que els seus motius són altres: vol saber els tipus de productes en què l'usuari està interessat.
3. L'usuari no presta molta atenció a l'anunci de l'agència de viatges, ja que aquest any ja ha planejat les vacances, així que no hi fa clic.
4. Com que les bosses que ha trobat no el convencen del tot, obre una altra finestra al navegador i accedeix a una altra botiga “online”. Almenys així pot comparar preus.

¹⁷ Un Ad Server o servidor d'anuncis és un conjunt de maquinari i programari que permet emetre o servir, un conjunt d'anuncis dins dels espais publicitaris dels diferents llocs web.

Font: Wikipedia

5. També aquesta botiga “online” conté un avís sobre l’ús de les “cookies” i crea una “cookie” d’origen. A més, apareix publicitat del mateix anunciant que abans, el servidor del qual torna a reconèixer l’usuari, ja que la “cookie” de tercets d’aquest proveïdor ja existeix. Aquí els anuncis segueixen mostrant la publicitat d’una agència de viatges.
6. Encara que al principi l’usuari va entrar per mirar bosses de color marró, s’adona que també necessita un abric, així que per què no visitar també subpàgines amb aquest tipus de producte. Ambdues “cookies” registren aquesta acció.
7. Finalment, l’usuari es decideix: comprrà l’abric, la bossa la deixa per a un altre moment. La “cookie” de tercets així ho anota, i l’Ad Server interpreta que l’usuari continua interessat a trobar una bossa marró. Tanca la finestra del navegador de les dues botigues. La sessió s’ha acabat i les galetes estan desactivades, però no s’esborren del disc dur. (*Cookies de Terceros: El Rastro Digital Bajo La Lupa*, n.d.-a)



Il·lustració 4: Funcionament i Tipus de Cookies

Font: Ionos

És precisament aquí on entra tot aquest mercat de dades de què tots hem sentit parlar, i tot aquest tema frega uns límits ètics que no s’acaben de delimitar, encara que la política d’extracció de dades sigui realment forta a l’UE. Sobre aquest dilema ètic es va també

parlar a l'entrevista amb Josep Forné¹⁸, i entre altres coses es va acabar concloent amb l'actual relativisme ètic i la gairebé nul·la importància que aquesta branca del pensament humà està rebent atès al neoliberalisme. A molt poca gent li importa si una cosa és ètica o ho deixa de ser si això provoca un creixement econòmic. Això ens passa amb les nostres dades, puix que o hi ha una gran desinformació sobre el tema o simplement hi ha un general conformisme, com parlarem posteriorment a la part pràctica social.

Les nostres dades tenen un valor cabdal. Cal entendre això perquè cada petjada que es deixa a internet pot ser utilitzar per serveis web de moltes maneres. Això dona explicació a què quan busquem reiterades vegades quelcom a internet ens acaben apareixent anuncis d'aquell article o servei a tot arreu.

Primerament, sembla molt exagerat dir que s'ha desvirtuat l'ètica i que la privacitat cada dia pren una menor importància, però tenint en compte que Google té una pàgina web on et diu les seves prediccions sobre la teva edat, estat civil, sexe, si vius de lloguer o no, si tens fills i quants, la teva ocupació, estudis, entre altres, i tot això només amb el teu historial web, potser no ho sembla tant. Aquesta web es diu MyAnalytics, i qualsevol persona que tingui un compte de Google hi està registrat.

Amb això venim a dir que tot el que fem a la xarxa s'utilitza d'alguna forma. Nosaltres pensem que acceptant la típica alerta de "cookies", o valorant una foto a Instagram res canvia, però ens oblidem que darrere de tot això hi ha màquines que fan 100 milions de càlculs per segon que juguen amb moltíssimes variables.

Això ho sumem a que els motors de recomanació utilitzen Machine Learning, és a dir, una intel·ligència artificial que aprèn dels milions d'exemples que rep dia a dia a internet i els reflecteix en una recomanació i la conclusió és clara. Pel simple fet d'entrar dins d'una pàgina web és possible que un motor de recomanació detecti patrons que alterin tot el que veiem a internet, és a dir, tot està fet al mil·límetre i, sovint, no en som conscients.

¹⁸ accessible a → ANNEX 6 - Entrevista Josep Forné

La realitat és que, tot i això, hi ha molts mites sobre l'extracció de dades. I és que no, el nostre mòbil no ens “escolta” ni “llegeix” les nostres converses de “Whatsapp”, ja que aquestes dades si que estan restringides. El que realment passa és que quan parlem alguna cosa amb algú, hem generat un interès que tard o d'hora acabem introduint en algun portal web que deriva en aquesta publicitat. Recordem que els motors de recomanació són experts en predir, i amb una gran quantitat de dades d'un usuari i de molts semblants a ell (com explicarem a tipus) pot arribar a predir un comportament web que ni el mateix usuari sap.¹⁹ A més a més, com podem corroborar amb els ignorats termes de privacitat d'aplicacions com Instagram o Tiktok, ara aquestes tecnologies tenen accés a, per exemple, l'historial de reconeixement d'allò que introduïm al teclat, és a dir, escrivim. Probablement això és una de les explicacions d'aquest fenomen que tant ens sobta.

De fet, Google li ha donat una explicació detallada a això, que es pot resumir en els següents punts:

- **La manera com interactues amb altres elements publicitaris.** Si han aparegut x anuncis i t'hi has mostrat interessat, n'apareixeran molts més en relació amb un mateix producte o servei.
- **La cerca orgànica des del navegador.** Si, per exemple, feu una cerca des del vostre dispositiu mòbil o ordinador “Sabatilles per fer esport d'home” heu d'estar preparats per a l'allau d'anuncis publicitaris que rebreu sobre això. Sobretot si les pàgines a les quals vau accedir estan registrades a Google Adsense.
- **La teva interacció a les xarxes socials.** A les xarxes socials a l'hora d'interactuar amb els teus amics, familiars o influències estàs immers en un target. D'aquesta manera, apareixeràs com a client potencial de moltes empreses i interactuaràs amb anuncis que les persones d'aquest mateix target (amics, influències) han valorat positivament.

¹⁹ (*Publicidad Misteriosa: ¿nuestros Móviles Nos Escuchan?*, n.d.)

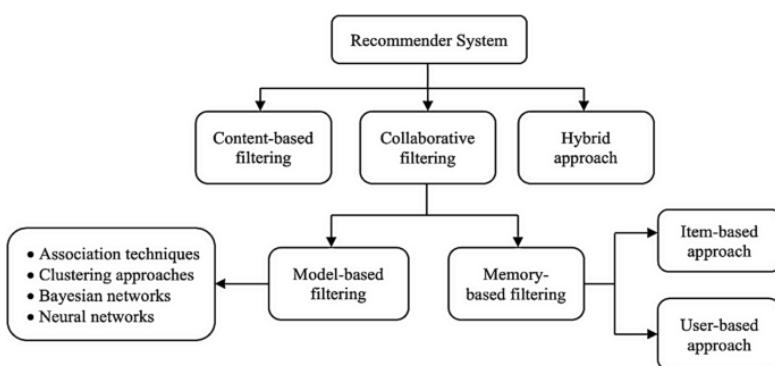
- **Gustos, preferències, edat, sexe i regió.** Com expliquem al punt anterior, formeu part d'un target potencial i si entreu en determinades característiques també en rebreu anuncis.²⁰

2.1.8 Tipus

Aquests motors funcionen amb el principi de trobar patrons en les dades de comportament del consumidor, encara que la manera d'agrupar les dades no sempre és la mateixa.

Hi ha tres diferenciacions clares dins dels motors de recomanació, basades en la manera en què el sistema processa les dades, el “filtratge” d'aquestes dades. Aquests són el filtratge col·laboratiu, el filtratge basat en contingut, i un híbrid d'aquests dos. També podríem agrupar aquests tres en un gran grup anomenat Sistemes de Recomanació Personalitzats (els que utilitzen Machine learning) però com no parlarem de l'altre tipus (els que no n'utilitzen) perquè no són presents estadísticament als sistemes moderns, creiem que no és adient fer-ho.

Cal també esmentar que amb l'auge dels motors de recomanació no és una exageració dir que gairebé cada setmana trobem variacions dins de cada tipus de motor de recomanació, ja que les tecnologies que empren les grans multinacionals com Netflix requereixen constant evolució, però ens centrarem en les bases.



Il·lustració 5: Agrupació de tipus de motors de recomanació.

²⁰ (*¿Por Qué Me Aparece Publicidad de Lo Que Hablo? Esta Es La Respuesta*, n.d.)

Font: *Journal of Big Data*

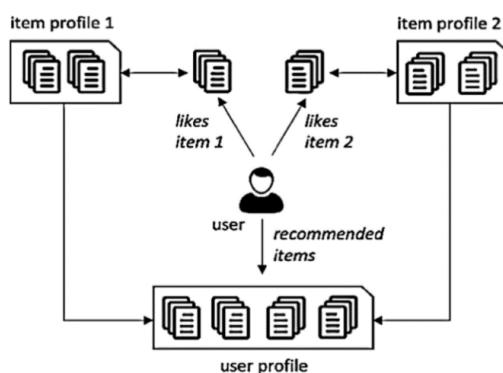
2.1.8.1 Filtratge Basat en contingut

En primer lloc, trobem el menys informàticament complex de tots, el filtratge basat en contingut. Aquest tipus de motor de recomanació recomana ítems només basant-se en els atributs d'aquests ítems, sense tenir en compte les dades de comportament de l'usuari, com si farien els de filtratge col·laboratiu.

Per exemplificar-ho, en el cas dels llibres podríem dir que el llibre *1984* és un ítem, i els seus atributs serien “George Orwell” (autor), ciència-ficció, entre altres aspectes que ens podem inventar.

Aquest tipus de motor de recomanació és efectiu si, seguint amb l'exemple, només ens volem centrar en l'autor i recomanar segons això, ja que recordem que només treballa amb atributs.

Quan un usuari dona una valoració positiva a un ítem (un llibre), aleshores, tots els atributs presents en aquest ítem creen un perfil d'usuari que combina tots els atributs dels ítems valorats positivament per l'usuari. Els elements presents en aquest perfil d'usuari es recomanen a l'usuari seguint l'esquema següent:



Il·lustració 6: Filtratge basat en contingut

Font: *Journal of Big Data*

Lògicament, els ítems que tenen una major correlació amb els elements que agraden a l'usuari són els que volem recomanar. Per fer això, però, haurem de determinar una manera de relacionar els ítems valorats per l'usuari amb l'altre que potencialment volem recomanar, però això és un apartat que, tot i que aproximarem ara, és millor entendre funcionalment a la part pràctica.

	Feature 1	Feature 2	Feature 3	Feature 4
Product 1	1		1	2
Product 2		1	4	
Product 3	3			1
User Data	2		1	1

Il·lustració 7: Matriu de 2 dimensions entre ítems i atributs

Font: Upwork

Per fer una petita aproximació teòrica a com és aquesta relació entre nombres, utilitzarem aquesta simple taula de valors.

En primer lloc, tenim tres ítems (“Product”), els quals tenen, de l’1 al 4, un índex de rellevància per cada atribut (“Feature”). Per exemple, l’atribut 1 pren gran rellevància al producte 3. Després, a la part inferior tenim un perfil d’usuari amb també una certa rellevància segons atribut.

Un cop s’ha assignat un valor numèric als atributs i als interessos de l’usuari cal identificar un mètode per identificar les similituds entre aquests ítems i interessos. Una fórmula molt bàsica seria el producte escalar, encara que n’hi ha moltíssimes. Per calcular el producte escalar s’ha d’utilitzar la fórmula següent, $\sum i = piui$. (on pi és el valor del producte i ui és el valor d’interès per cada columna i .

A la taula anterior, el nivell d’interès de l’usuari amb el “Product 1” es pot estimar en $2 * 1 + 1 * 1 + 1 * 2$, que és igual a 5. De la mateixa manera, l’interès pel “Product 2” serà $1 * 4 = 4$ i $2 * 3 + 1 * 1 = 7$ al “Product 3”. Per tant, el producte 3 serà la recomanació principal de l’algorisme per a l’usuari.

Altres mètodes de relació podrien ser la correlació de Pearson, el coeficient de Jaccard i la similitud del cosinus, de la que en parlarem a la part pràctica. Aquestes, però, demanen una agrupació de les dades que no teníem en aquest exemple.

Com ja heu vist, el filtratge basat en contingut es basa en l'assignació d'atributs als ítems de la base de dades, de manera que l'algorisme sàpiga alguna cosa sobre cada objecte. Per tant, és essencial que aquesta tria d'atributs rellevants es faci correctament, i d'això pot anar l'èxit d'una empresa. Per assolir això, moltes d'aquestes recorren a utilitzar equips d'experts en la matèria per assignar atributs a cada ítem manualment. Per exemple, Netflix, ha contractat guionistes per valorar els programes en aspectes que van des dels llocs de rodatge i els actors fins a les trames, el to i els efectes emocionals. Els atributs resultants, usats pel recomanador, es combinen algorítmicament per agrupar pel·lícules que comparteixen aspectes similars. Com amb tot, cada cosa té la seva escala. Com major sigui el nombre d'atributs rellevants amb els quals treballem, resultats menys distants a la realitat obtindrem.²¹

Dins del perfil d'usuari, els atributs que apareixen en diversos ítems tenen més ponderació que els que apareixen amb menys freqüència. Això ajuda a establir un grau d'importància perquè no tots els atributs d'un ítem són igual d'importants per a l'usuari. A partir de les ponderacions i els historials d'atributs, el sistema de recomanació produceix un model únic de les preferències de cada usuari.

El model consta d'atributs que a l'usuari pot agradar o no en funció d'activitats passades, també ponderades per importància. Els models d'usuari es comparen amb tots els ítems de la base de dades, als quals se'ls assigna puntuacions en funció de la seva similitud amb el perfil d'usuari.

Aquí teniu un exemple: suposem que has escoltat "The Show Must go On" de Queen, "Every Day is Like Sunday" de Morrissey i "Like a Rolling Stone" de Lou Reed. Un bon sistema de recomanació amb els atributs adients pot reconèixer que t'agraden els artistes de rock i les cançons líricament tristes. Per tant, és normal que per aquesta raó rebis

²¹ (*What Is a Content-Based Recommendation System in Machine Learning?/ Analytics Steps*, n.d.)

recomanacions per a més cançons tristes d'aquests i d'altres artistes de rock, com ara podria ser "Creep" de Radiohead. El sistema de recomanacions també pot suggerir diferents tipus de cançons de Queen, encara que no siguin tristes, pel simple fet que l'artista t'agrada; és a dir, que les recomanacions no s'han de basar en una mescla de tots els atributs que t'agraden, sinó pot recomanar certes cançons mitjançant atributs singulars. Tot i així, com que aquesta recomanació no compleix tots els atributs d'ítem que s'han assignat al teu perfil, aquestes seleccions rebran una puntuació assignada més baixa, és a dir, apareixeran menys. Això és bo, ja que atenua una de les pitjors repercussions que tenen aquests sistemes de recomanació personalitzats, l'"Echo Chamber", del que més endavant parlarem.²²Aquí tenim certs aspectes positius d'aquest tipus de procés de dades, que per fer més entenedora la comprensió, compararem amb el seu altre competidor, el filtratge col·laboratiu.

En primer lloc, no calen dades d'altres usuaris per començar a fer recomanacions. A diferència del filtratge col·laboratiu, el filtratge basat en contingut no necessita dades d'altres usuaris per crear recomanacions. Un cop un usuari ha cercat i explorat alguns articles i/o ha completat algunes compres, un sistema de filtratge basat en contingut pot començar a fer recomanacions rellevants. Això el fa ideal per a empreses que no tenen un grup enorme d'usuaris per provar. De tota manera, també funciona bé per als venedors que tenen molts usuaris però un nombre reduït d'interaccions d'usuari en categories o nínxols específics.

A més a més, atès que basa les seves recomanacions solament en l'usuari a recomanar, les recomanacions són molt rellevants i concretes per a l'usuari. Els sistemes de recomanació basats en contingut es poden adaptar molt als interessos de l'usuari, incloses recomanacions per a productes nínxol²³, ja que el mètode es basa a fer coincidir les característiques o els atributs d'un ítem de base de dades amb el perfil de l'usuari. Per exemple, el filtratge basat en contingut reconeixerà les preferències i els gustos d'un usuari específic, com ara salses picants fets a Vic amb pebrots escocesos ecològics, i

²² (*Sistemas de Recomendación de Contenido Con Machine Learning – Cleverdata*, n.d.)

²³ Els productes de nínxol són articles que pertanyen a categories de productes més definides que atrauen i s'orienten a grups de clients més petits i específics amb necessitats particulars.

recomanarà productes amb els mateixos atributs. En un col·laboratiu, seria molt difícil que es trobés una coincidència d'atributs tan gran entre dues persones (ja que és molt concret) i, per tant, la recomanació seria menys precisa. A més a més, aquest procés també és valuós per a les empreses amb una àmplia gamma d'un sol tipus de producte, com ara els telèfons intel·ligents, on les recomanacions s'han de basar en moltes funcions discretes.

Addicionalment, les recomanacions són transparents per a l'usuari. Les recomanacions molt rellevants projecten una sensació d'obertura cap a l'usuari, reforçant el seu nivell de confiança en les recomanacions ofertes. En comparació, amb el filtratge col·laboratiu, és més probable que es produixin casos en què els usuaris no entenen per què veuen recomanacions específiques. Per exemple, suposem que un grup d'usuaris que han comprat un paraigua també s'han comprat abrics de pluja. Un sistema col·laboratiu pot recomanar abrics de pluja a altres usuaris que han comprat paraigües però que no els interessin i que mai no han navegat ni comprat aquest producte.

Un aspecte també molt important és que evita el problema de “l'inici lent”. El filtratge col·laboratiu crea un escenari potencial d'inici en fred quan un lloc web o una comunitat nova té pocs usuaris nous i no té connexions d'usuari. Tot i que el filtratge basat en contingut necessita algunes aportacions inicials dels usuaris per començar a fer recomanacions, la qualitat de les primeres recomanacions generalment és millor que un sistema col·laboratiu que requereix l'addició i la correlació de milions de punts de dades abans d'optimitzar-se.

I per últim, els sistemes de filtratge basats en contingut són generalment més fàcils de crear. La ciència de dades darrere d'un sistema de filtratge basat en contingut és relativament senzilla en comparació amb els sistemes de filtratge col·laboratiu destinats

a imitar les recomanacions d'usuari a usuari. El treball real en el filtratge basat en contingut és assignar els atributs.

Així i tot, aquest sistema també té les seves febleses, i les classificarem de la mateixa manera.

Hi ha una manca de novetat i diversitat. Hi ha més recomanacions que rellevància. Suposem que t'ha agrat la cançó “Thriller” de Michael Jackson. És probable que també t'agradi “Black or White”, però hi ha una gran probabilitat que no necessitis un sistema de recomanació per dir-te això. Per tant, per ser útils, els motors de recomanació han de tenir resultats diversos i inesperats, que normalment són més difícils de trobar.

Altrament, l'escalabilitat és un repte. Cada vegada que s'afegeix un nou producte o servei o contingut nou, s'han de definir i etiquetar els seus atributs. La naturalesa àrdua i interminable de les assignacions d'atributs pot fer que l'escalabilitat sigui difícil i requereixi molt de temps.

I per finalitzar, els atributs poden ser incorrectes o inconsistents. Les recomanacions basades en contingut només són tan bones com els experts en la matèria que etiqueten elements. Potencialment, milions d'elements necessiten atributs assignats, i com que els atributs poden ser subjectius, molts poden estar etiquetats incorrectament. És primordial un procés que garanteixi que els atributs s'apliquen de manera coherent i precisa. En cas contrari, un sistema de recomanació basat en contingut no funcionarà com s'ha previst.

2.1.8.2 Filtratge Col·laboratiu

Després d'aquesta extensa explicació sobre els sistemes de recomanació basats en contingut, cal parlar dels sistemes de recomanació amb filtratge col·laboratiu, que a trets generals venen a ser conceptualment contraris.

El filtratge col·laboratiu, a grans trets, filtra la informació utilitzant les interaccions i les dades recopilades pel sistema d'altres usuaris. Es basa en la idea que aquelles persones

que van estar d'acord en la seva avaliació de determinats ítems probablement tornaran a estar d'acord en el futur.

El concepte és senzill: quan per exemple volem veure una pel·lícula que no hem vist mai, sovint demanem recomanacions als nostres amics. Naturalment, confiem més en les recomanacions d'amics que comparteixen gustos semblants als nostres, i aquest tipus de procés de dades fa més o menys el mateix. La majoria dels sistemes de filtratge col·laboratiu apliquen l'anomenada tècnica basada en índexs de semblança, encara que ara mateix està més esbiaixat. Amb aquesta, es selecciona un nombre d'usuaris en funció de la seva similitud amb l'usuari actiu (després desglossarem aquesta similitud en els tipus). Amb aquesta informació, es recomana aquells articles comprats per un usuari decretat semblant a l'actiu per tal de trobar un patró en la compra de l'usuari actiu, i així incitar-lo a consumir el recomanat. La inferència per a l'usuari actiu es fa calculant una mitjana ponderada de les valoracions dels usuaris seleccionats segons els ítems consumits per aquests.

Bàsicament, els sistemes de filtratge col·laboratiu es centren en la relació que hi ha entre els usuaris i els ítems que aquests han puntuat, relacionant dos o més usuaris segons la puntuació que aquests han establert sobre un ítem. D'aquesta manera, es recomanen aquells ítems (no consumits ja per l'usuari a recomanar) que un altre usuari establert com a semblant ha puntuat positivament.

Aquest tipus de filtratge de dades es divideix en dos subtipus, els anomenats filtratges de dades col·laboratius de model i de memòria.

Encara que els filtratges col·laboratius de model són considerats un subtipus d'aquest tipus de motor de recomanació, componen una branca immensa del Machine Learning. A més a més, els algorismes de “Machine Learning” que comprenen aquesta branca no tenen gaire a veure un amb l'altre i aprofundir en cada un d'ells seria inviable en aquest TR, així que ens limitarem a esmentar els més famosos.

Dintre dels filtratges col·laboratius de model trobem les xarxes neuronals, les xarxes Bayesianes, i la descomposició en valors singulars. Tots aquests models usen matrius complexes pels seus càlculs.

En canvi, ens centrarem més en l'explicació de l'altre subtipus de filtratge col·laboratiu, que és el de memòria, que representa la idea general del que és un motor de recomanació amb filtratge col·laboratiu. Les tècniques de recomanació basades en memòria s'apliquen a dades en brut sense preprocessament. Això fa que siguin fàcils d'implementar i les recomanacions resultants siguin generalment fàcils d'explicar.

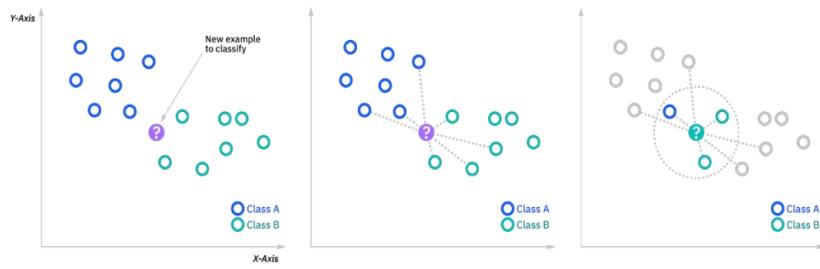
El principal inconvenient dels filtratges basats en memòria és el requisit de carregar una gran quantitat de memòria en línia. El problema és greu quan la matriu de qualificació es fa tan gran en una situació que hi ha moltíssimes persones que utilitzen el sistema. Els recursos computacionals es consumeixen molt i el rendiment del sistema baixa, de manera que el sistema no pot respondre la sol·licitud de l'usuari immediatament. L'enfocament basat en models pretén resoldre aquests afers d'una manera bastant eficient. Això ho fa, principalment, eliminant dades innecessàries per la recomanació a establir al moment, cosa que fa el sistema més ràpid, escalable i dimensionalment menor.²⁴

Aquesta branca també es divideix en dues seccions molt importants, les quals es diferencien per un tret molt important. Els basats en ítem realitzen aquesta “col·laboració” entre ítems, és a dir, busquen l'ítem similar al valorat positivament per l'usuari, mentre que els basats en usuari realitzen la “col·laboració” entre usuaris, és a dir, recomanen allò que un usuari similar a l'usuari que volem recomanar ha valorat positivament.

D'una banda, els sistemes de filtratge col·laboratiu de memòria basats en l'usuari es centren en l'enfocament del "veí més proper"²⁵ per a les recomanacions. Aquest mira els patrons de valoració dels usuaris del sistema i troba els “veïns” més similars en termes de consum, és a dir, els usuaris amb valoracions més semblants. Cal dir, a més a més, que van ser els primers que es van idear dins d'aquesta diferenciació.

²⁴ (*Model-Based Recommendation System / Towards Data Science*, n.d.)

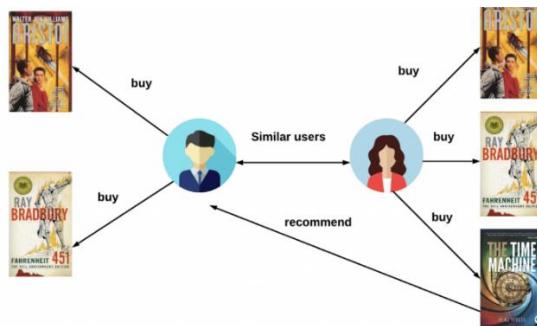
²⁵ L'algorisme de k veïns més propers, també conegut com a KNN o k-NN, és un classificador d'aprenentatge supervisat no paramètric, que utilitza la proximitat per fer classificacions o prediccions sobre l'agrupació d'un punt de dades individual.



Il·lustració 8: Diagrama KNN

Font: IBM

Aquest enfocament agrupa els usuaris (mitjançant un recull de dades explícit o implícit) que tenen comportaments similars, contactes comuns o dades demogràfiques properes, i obviament comportaments com a consumidors similars. Els llocs de xarxes socials incorporen aquest enfocament per recomanar usuaris a altres usuaris en funció dels seus patrons de comportament, com exemplificarem ara. A més, aquest enfocament també s'utilitza per a anuncis orientats i elements suggerits basats en altres usuaris que tenen opcions i preferències similars. Un exemple típic d'aquest enfocament és la categoria "amics suggerits" que es mostra a Instagram o qualsevol altra xarxa social, que recomana persones que els usuaris poden conèixer en funció dels seus contactes virtuals i preferències similars.



Il·lustració 9: Funcionament d'un sistema de recomanació amb filtratge col·laboratiu basat en l'usuari

Font: Things Solver

D'altra banda, els filtratges basats en ítem exploren la relació entre una parella d'ítems que es correlacionen en diferents compres de diferents usuaris (l'usuari que va comprar Y, també va comprar Z).

En mesurar la similitud entre productes i inferir les valoracions respectives, es recomanen ítems als usuaris en funció de les seves dades històriques i de l'historial interactiu. Encara avui llocs de comerç electrònic com Amazon i Flipkart fan servir sistemes de recomanació basats en ítems per recomanar productes similars als usuaris filtrant-los amb l'ajuda del passat d'un usuari, és a dir, relacionant la compra d'un ítem amb una altra.²⁶ Un clar exemple seria la característica d'Amazon que ens recomana, just abans de pagar, productes complementaris a aquells que estem a punt de comprar. Segons moltes estadístiques, això incrementa severament les vendes de l'empresa, ja que, com parlarem posteriorment, ens genera una necessitat perfectament col·locada allí perquè (estadísticament) la comprem.



Il·lustració 10: Funció Amazon (Compra-ho amb); si primerament elegeix comprar uncomandament per a videojocs.

Font: Amazon



Il·lustració 11: Funcionament Sistema de recomanació amb filtratge col·laboratiu basat en l'ítem

Font: Things Solver

²⁶ ¿Qué Es El Filtrado Colaborativo?, n.d.; Various Implementations of Collaborative Filtering / by Prince Grover

En resum, la diferència entre aquests dos tipus és que, en l'esmentat en primer lloc, es relaciona dos usuaris segons la similitud de les seves valoracions, i en l'esmentat posteriorment es busca un patró entre dos o més ítems comprats pels mateixos usuaris. Imaginem-nos que hi ha quatre persones que les seves valoracions positives (o reproduccions continuades) en una pàgina web es mostren així:

P1 = Blues, Pop, Funk, Jazz, Rock; P2 = Rock, Funk, Jazz, Country;

P3 = Jazz, Blues, Funk; P4 = Funk, Rock, Blues.

En aquest cas, si ens centrem en les recomanacions de P3, sempre basant-nos en el concepte dels tipus i no aplicat exactament a un motor real (ja que necessitaríem més dades) el primer tipus li recomanaria el Pop, ja que el sistema observa una estructura similar a P1 i a aquest li agrada el Pop. En canvi, amb la segona, el sistema agruparia el Funk i el Rock, car és un patró repetit a P1, P2, i P4, i recomanaria el Rock.

Tot i així, la similitud entre els ítems és sempre més estable que la semblança entre els usuaris, puix que un llibre de matemàtiques sempre serà un llibre de matemàtiques, però l'usuari pot canviar d'opinió (alguna cosa que li va agradar la setmana passada potser no li agrada la setmana que ve). Un altre avantatge és que hi ha menys productes que usuaris, el que ens porta a la conclusió que una relació ítem-ítem amb puntuacions de semblança serà més petita que una relació usuari-usuari. També l'enfocament basat en ítem és un millor enfocament si un nou usuari visita la web, mentre que l'enfocament basat en l'usuari és problemàtic en aquest cas, ja que no s'ha rebut encara suficients dades sobre l'usuari.²⁷

Per finalitzar, aquest esquema és útil per entendre les principals diferències entre els dos grans tipus de motors de recomanació esmentats en les anteriors pàgines.

²⁷ (Aditya et al., n.d.; Bokde et al., 2015; Boström et al., 2017; (*PDF Model-Based Approach for Collaborative Filtering*, n.d.; *Sistemas de Recomendación*)



Il·lustració 12: Comparació entre els dos grans tipus de motors de recomanació

Font: *StatDeveloper*

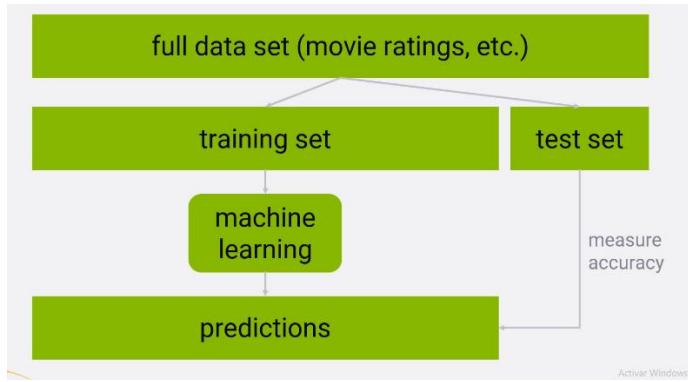
2.1.9 Avaluació de Recomanadors i Criteris a considerar

Una part molt important del procés de creació d'un motor de recomanació, com veurem a la part pràctica posteriorment, és calcular la seva qualitat. Aquest concepte és bastant subjectiu, però hi ha una sèrie de mètriques que s'utilitzen per a establir-lo.

Cal mencionar que tot aquest procés no és rellevant pel consumidor, ja que tot forma part d'un sistema d'avaluació intern.

Per a comprovar la qualitat d'un motor de recomanació es pot usar un mètode anomenat “train/test split”. Aquest procés el que fa és dividir les dades de valoracions passades en disposició del motor en dos grups, un que es farà servir per entrenar (“train”) i un altre per practicar (“test”).²⁸

²⁸ (*Train-Test Split for Evaluating Machine Learning Algorithms*, n.d.)



Il·lustració 13: Funcionament del mètode "Train/Test split"

Font: Sundog Education

Amb les dades que empri per entrenar aprendrà les relacions entre ítems i usuaris que després utilitzarà per recomanar. Subseqüentment, quan hagi acabat de processar les dades del set d'entrenament, amb les dades que quedin per practicar (que no ha vist el motor) es posarà a prova.

Per exemple, si un usuari ha valorat amb un 5/5 la pel·lícula “Jumanji” el motor de recomanació usará el que ha “aprèss” entrenant amb l’altre set de dades per endevinar-ho.

Després d’haver reiterat aquest procés diverses vegades, es podrà calcular la qualitat del motor de recomanació segons el que s’hagi allunyat la recomanació estimada del valor real establert per l’usuari.

Les dues mètriques de precisió més famoses són, en primer lloc, l’error Absolut Mitjà (MAE). Bàsicament, és un càlcul molt simple, ja que només s’ha de restar el valor estimat de recomanació al real, aplicar-li el valor absolut i fer la mitjana del sumatori de totes les reiteracions.

$$\frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i|$$

Il·lustració 14: Fórmula de l’Error Absolut Mitjà

Font: Codingprof

En segon lloc, tenim la més famosa, que és l'Arrel de l'Error Quadràtic Mitjà (RMSE). Aquest procés, en canvi, eleva al quadrat la diferència entre el valor estimat i el real, fa la mitjana i acaba fent l'arrel quadrada. La diferència amb el MAE és que aquest sistema penalitza més als valors molt distants.

$$RMSE = \sqrt{\sum_{i=1}^n \frac{(\hat{y}_i - y_i)^2}{n}}$$

Il·lustració 15: Fórmula de l'arrel de l'error quadràtic mitjà (RMSE)

Font: GeeksforGeeks

També hi ha altres maneres d'avaluar el sistema, com són el “hit-rate”. Aquest mètode el que fa és, amb les dades d'un sol usuari, crear una llista de n millors ítems. Llavors, el motor de recomanació ha d'esbrinar quins ítems hi ha en aquella llista, quedant com a resultat els encerts dividits entre el nombre d'ítems de la llista (n).

A més a més, hi ha tècniques com el “Leave-One-Out Validation”. El que fa és, igual que l'anterior, crear una llista d'n millors ítems sobre un usuari. Tot seguit n'esborra un, i el motor de recomanació ha d'esbrinar, mitjançant les dades que ha processat al set d'entrenament, l'ítem que falta.

Encara això, la precisió no és l'únic valor que s'ha de tenir en compte a l'hora de crear un sistema de recomanació.

Un valor que s'ha de tenir en compte és la cobertura, o “coverage” en anglès, que és el percentatge de recomanacions possibles que pot proveir un sistema de recomanació. Més tècnicament, vindria a ser el percentatge de parelles (usuari, ítem) que poden ser predites pel motor de recomanació. Bàsicament mesura si, de tots els teus ítems, els mesures tots.

Un altre valor bastant important és la diversitat, que calcula la diferència entre atributs dels ítems que el motor de recomanació recomana. Exemplificant-ho, si a una persona que li agrada Star Wars I només li recomanes Star Wars II i Star Wars III, la diversitat serà baixa. Sembla bastant subjectiu, però com observarem després al codi de la part pràctica, és measurable.

Per últim, trobem la popularitat. Com ja diu el nom, això mesura la popularitat dels elements que recomanes. Per exemple, si a una persona que li agrada una pel·lícula aleatòria de ciència-ficció li recomanes Star Trek, la popularitat serà alta.

Aquests valors, encara que són computables i numerables, no tenen cap xifra que sigui correcta. En altres paraules, no és fàcil determinar quin índex de popularitat, diversitat o cobertura es vol, ja que això depèn de cada motor de recomanació. Equilibrar la precisió amb aquests valors de vegades també és el que acaba diferenciant un motor de recomanació d'un altre. Per exemple, podem pensar que sempre és bo tenir la popularitat alta, però si una persona veu que tots els llibres que se li recomanen són best-sellers, deixarà de tenir interès en el motor de recomanació. Trobar un equilibri entre aquestes magnituds i exercicis d'entrenament és realment complicat.²⁹

2.2 Motors de Recomanació a la societat

2.2.1 Neoliberalisme

Vivim en un món on el capital s'ha apoderat de les nostres vides. Tant és el fet que la nostra manera de pensar està condicionada per un sistema socioeconòmic que ni tan sols hem escollit. El fet de materialitzar-ho tot o de moure'ns pel més "eficient" ens allunya del que per molta gent és la naturalesa humana. Aquest tipus d'interacció entre béns i serveis s'anomena neoliberalisme, i encara que el seu fonament ideològic no comporti explícitament la declaració que acabo de realitzar, les seves conseqüències pràctiques ens han portat al fonament de la societat actual. No és casualitat que una màquina que prediu allò que volem (motors de recomanació) tingui tant d'èxit.

Tècnicament, el terme neoliberalisme es refereix a la filosofia politicoeconòmica que propugna la mínima intervenció d'organismes estatals sobre l'economia i la política, fent que el paper de l'estat es redueixi a garantir les llibertats individuals.³⁰

²⁹ (Kaminskas & Bridge, 2016)

³⁰ (*Per Una Dissecció de La Postmodernitat*, n.d.)

Generalment se'l comprèn com un ressorgiment dels preceptes del liberalisme sorgit entre els segles XVII i XVIII, on els burgesos de l'època desitjaven deslliurar-se de l'absolutisme monàrquic i viure en una societat amb més llibertats econòmiques i individuals. Encara que la idea va ser concebuda el 1938 i és hereva de les teories neoclàssiques de finals del segle XIX, va entrar en aplicació cap a la dècada del 1970, tenint com a models el govern de Ronald Reagan i Margaret Thatcher. Es diu que és una conseqüència del postmodernisme, que com bé diu el filòsof Jean-François Lyotard, té un impacte en la condició humana que ens fa viure en una llibertat falsa.³¹

En termes generals aquest sistema és una radicalització del capitalisme, i no és casualitat que justament sorgís quan el capitalisme social potenciat pel comunisme va deixar de tenir força en els sectors productius de la societat occidental. Concretament, té com a ideals “filosòfics” que el mercat ho sosté tot, incloent-hi la igualtat social i l’economia, que la privatització és la manera més eficient de garantir una estabilitat i sobretot sosté que la globalització i el sistema capitalista és origen i final de la intervenció humana.³²

Això, en termes econòmics, acaba decretant una nul·la intervenció de l'estat en la regulació del comerç, en la iniciativa privada, i en el dèficit públic, argumentant que les polítiques de protecció social com l'estat del benestar són perjudicials per a l'economia, i que l'equilibri salarial ja es produeix amb l'oferta i demanda.³³

En primera instància aquests preceptes semblen inofensius, però la seva pràctica ens ha portat per un altre camí. Se'ns presenten dos problemes: el primer, com passava amb el capitalisme, és que el sistema fomenta una gran desigualtat entre classes, ja que la poca intervenció de l'estat ajuda al ric a fer-se més ric i al pobre no se l'ajuda a sortir de la pobresa. El segon, és la seva conseqüència, i és que el sistema pretén ocultar aquesta desigualtat amb una trista llibertat convertida en consum i desitjos, i això fa que, si ho sumem al conformisme que defineix l'època, no sigui fàcil un canvi de sistema com va

³¹ (*La Condició postmoderna, per Jean François Lyotard*).

³² (*¿Por Qué El Neoliberalismo Provoca Desigualdad?*, n.d.)

³³ (*13 Características Del Neoliberalismo - Significados*, n.d.; *Neoliberalism - Overview, History, Characteristics, and Schools of Thought*, n.d.; *Neoliberalism: What It Is, With Examples and Pros and Cons*, n.d.; *Neoliberalismo: Definición, Ventajas y Características*, n.d)

passar a zones com Rússia a principis del segle XX. És almenys sorprendent que en una societat on, segons el laboratori de les desigualtats mundials, el 10% de les persones acumulen el 76% de la riquesa global no hi hagi cap “odi” popular cap al sistema capitalista.³⁴



Il·lustració 16: Dibuix referent al Neoliberalisme

Font: El País

Pragmàticament, és molt fàcil de trobar exemples, ja que el conegut somni americà, o en alguns llocs com l'article de Gerardo Sánchez Guadarrama, ja reconegut com a esclavització postmoderna, idealment, ho ratifica.³⁵

Bàsicament, el que ve a dir el somni americà és que qualsevol pot arribar a l'èxit econòmic independentment de la seva classe social, situació econòmica o circumstància de vida. Clar és l'exemple d'una entrevista que se li va fer a una persona pobra dels EUA. Se li preguntava per què no preferia una major intervenció de l'estat per a millorar la seva condició en lloc de la privatització massiva. La persona defensava que el seu somni sempre havia estat ser advocat, i creia que quan ho aconseguís, una intervenció de l'estat en forma d'impostos destinats a ajudes li dificultaria el camí, com també li passaria al seu fill quan arribés a l'NBA.³⁶

De fet, que l'esperança d'algú pugui fonamentar-se en una probabilitat tan mínima recorda al conegut “opi del poble” que esmentava Karl Marx quan parlava de la religió.

³⁴ *Neoliberalismo: Qué Es, Características y Representantes Principales - Toda Materia*, n.d.

³⁵ (*El Sueño Americano o La Nueva Forma de Esclavitud Moderna*, n.d.)

³⁶ (“*El Sueño Americano Es Una Farsa*”: *Qué Es La “Trampa de La Meritocracia” (y Cómo Afecta Tanto a Pobres Como a Ricos En Estados Unidos)* - *BBC News Mundo*, n.d.; *Historia Del Sueño Americano y Sus Consecuencias - EEUU - Internacional - ELTIEMPO.COM*, n.d.)

El neoliberalisme sembla una política fàcil, però és una manera de ficar-se dins del nostre cap, de canviar tota aspiració humana per un bé inexistent. Creiem que com que l'estat és l'òrgan regulador de la societat, hi recau tot el poder, però realment el poder el tenen els diners; ara, objectiu de tothom. Tendim a l'autodestrucció de la condició humana.³⁷

El neoliberalisme delimita la grandesa de l'ésser humà a la capacitat de generar ingressos monetaris i exacerba l'individualisme i la carrera per guanyar i posseir béns. Desliga la cobdícia, la corrupció i la violència i, en generalitzar-se en els grups socials, destrueix socialment la comunitat. S'imposa així un ordre de valors on preva la llibertat individual per accedir al consum de satisfaccions i plaers. ³⁸

Podríem dir que, pel sistema, més que humans som productes o consumidors de productes, ja que el motor de tot és simplement el capital. Ens trobem en deshumanització social que, tot i que està encarada de diferent manera, ens presenta idealment una societat distribuïda com va fer Aldous Huxley al seu llibre *Un Món Feliç*.

Allí veiem desvirtuat el concepte de felicitat, veiem la pèrdua del valor de l'individu, veiem la naturalesa prohibida per la seva nul·la aportació al consumisme, entre altres molts factors que ens fan pensar que un estat així seria insostenible. Allò que fa prosperar la societat d'Aldous Huxley és una droga que es subministra a tots els ciutadans que els fa sentir una “felicitat artificial” que reemplaça tots els valors eliminats que fonamentalment ens defineixen com a humans. L'empatia, la família, l'estima, tot això s'elimina amb una simple droga.

Bé, doncs precisament això és el que fa l'ànsia del consum a la nostra societat. Com bé explicarem posteriorment, els humans d'ara tenim una gran addicció al consum, el qual ens genera petites dosis de dopamina que ara denominem com a “felicitat”. Aquesta necessitat només fa que retroalimentar al sistema, i conseqüentment, apropar-nos cada vegada més a la dictadura del capital.

³⁷ (*La societat autófaga*, Anselm Jappe).

³⁸ (*Ensayos sobre el posmodernismo*, Fredric Jameson)

2.2.2 Els Estímuls

Què ens fa qui som? Aquesta pregunta ha estat sempre un motiu de debat entre els pensadors, però fora del camp filosòfic s'ha determinat la resposta. Òbviament, la genètica té un paper molt important a l'hora de fonamentar el nostre caràcter i trets físics, però l'entorn és el factor que més pes té en la nostra manera d'actuar i de prendre decisions. New York Times, mitjançant la història d'una mare i els seus dos fills ens demostra que la diferència és realment aclaparadora. La història ens presenta com dos bessons genèticament iguals, només al néixer, ja presentaven diferències. Mentre que el primer s'havia vist pressionat durant diverses hores per una pelvis inflexible i va passar tot el seu primer dia alerta i mirant desconfiat el seu voltant, l'altre es va deslliurar de l'estrés perinatal i va dormir plàcidament durant tot el dia. Aquestes diferències, si les sumem amb la different rutina del dia a dia, es multipliquen. Per això es troben més indicis de delinqüència en sectors desfavorits; no perquè genèticament siguin inferiors, sinó perquè el seu entorn ha fet construir la seva conducta d'una altra manera. A més, ajuntant els dos termes, el mateix entorn pot fer expressar uns gens o uns altres, el que es diu epigenètica. Com diu el periodista Erik Engquist, “els gens determinen el teu potencial, però el teu entorn determina qui seràs.”³⁹

Per aquesta raó, és cabdal centrar-se en la formació dels nens a partir d'un sistema que inculqui valors profitosos pel bé comú (desatenent-nos de relativismes), ja que és d'allí on més estímuls de l'entorn podrien rebre. Encara que de la mateixa manera, que l'entorn sigui predominant en la nostra conducta, fa realment fàcil manipular persones.

Com ens diu Mark Fisher al seu llibre *Realisme Capitalista*, la quantitat d'estímuls que ens predisposen al consumisme és insostenible. Mentre que els nostres gens i el nostre cervell han evolucionat molt poc des de l'inici de la nostra espècie, la quantitat d'informació que rebem de tots els llocs no para d'augmentar.⁴⁰

³⁹ (*Ensayos sobre el posmodernismo*, Fredric Jameson)

⁴⁰ (*Bermúdez de Castro: “Nuestro Cerebro Ha Cambiado Muy Poco Desde Hace 200.000 Años”* - Valencia Plaza, n.d.)

Constantment veiem una publicitat al carrer, un segon després rebem una notificació al mòbil, un segon després una trucada, i així fins a modelar-nos en un consumisme ara ja intrínsec des de que naixem.

Tant és així que, com diu Alexander Lowen a *El narcicismo, la enfermedad de nuestro tiempo*, hi ha moltes persones que directament no imaginen una societat fora dels límits del capitalisme, obviant els moltíssims anys d'història on el comunitarisme va ser la forma d'agrupar-se en societat predilecta pels humans.

Recollint el que diu Gilles Lipovetsky a *La felicitat paradójica, ensayos sobre la Sociedad del hiperconsumo* i Jean Baudrillard a *La ilusión del fin, la huelga de los acontecimientos* i a la tesi doctoral de Sergio López, *Problemas del primer mundo*, podem afirmar que no tenim capacitat per processar tots aquests estímuls, persones i informació diàriament, i això ens fa inoperants a l'hora de transformar el nostre entorn. No som capaços de deixar marca enllloc ni que res ens deixi marca, ja que el món va més de pressa que el nostre cervell processant-lo. Ens sentim flotant en un buit immens que la falsa felicitat del consumisme intenta tapar. Això es veu potenciat amb el fet que depenem de tercers per tot, ja que no som capaços de fer res sense que una transacció capitalista hagi de produir-se. Si volem fer una casa hem de comprar terreny, eines, materials, entre altres. El nostre cervell no pot gestionar tants processos complexos. Els ritmes productivistes del capitalisme acaben amb els cicles biològics de la natura, i això ens sotmet a un estrès mental crònic.

Aquí entren els esmentats motors de recomanació, que són un mirall d'aquesta realitat.



Il·lustració 16: Scroll Infinit
Font: Clickio

Els motors de recomanació són una gran manera de distribuir aquestes distraccions o articles consumistes que contribueixen a la “felicitat”, ja que com hem dit, mostren a l’usuari l’ítem o article que més s’adequa al seu perfil comercial. Seria molt més difícil tenir enganxada a la població a, per exemple, Instagram, si els posts recomanats a cada persona no tinguessin res a veure amb el que han vist anteriorment. És per això que quan

per exemple busquem per internet sabates Nike ens acaben sortint sabates Nike a gairebé tots els anuncis.

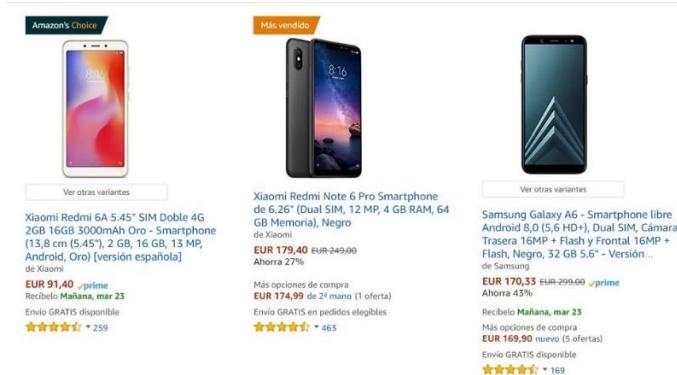
Això, sumat a tecnologies com l'Scroll infinit, que permeten a l'usuari accedir a contingut infinitament sense saber quin article visualitzarà després de l'actual, fan que cada vegada que es generi un contingut virtualment, el nostre cervell rebi un cop de dopamina que ens incita a seguir consumint.

Això ho fa mitjançant un “plaer” que sovint es confon amb la felicitat, encara que no és ben bé així. Neurobiològicament, la dopamina és l'hormona del “més”, fent que quan la segreguem instantàniament vulguem consumir més d'allò que ens la fa produir. Això es pot veure fàcilment a les xarxes socials, i més amb els motors de recomanació, que mostren allò que més probablement produirà un pic de dopamina a l'usuari. Aquesta cerca de plaer, però, no acaba tractant-se del plaer en sí, sinó d'alleujar el malestar que ens produeix l'absència d'aquest plaer instantani. Això fa que aquestes tecnologies siguin molt poderoses, ja que la seva immensa presència d'estímuls crea un cicle de dopamina que ens lliga encara més a elles.

Per tant, si es controla el que mostra per pantalla un motor de recomanació pot esdevenir una gran eina de manipulació econòmica i ideològica. Moltes vegades acabem consumint allò que ens apareix primer per pur conveni i mandra, i per aquesta mateixa raó els sistemes de “et pot interessar” de webs com Amazon funcionen tan bé. Aquesta potencial manipulació de les dades no es troba només en una suposició teòrica, sinó que convivim amb ella en molts serveis que freqüentem cada dia. Un àmbit que ho corrobora, com ens diu Eva Martín (entrevista annex 5), que va fer la tesi doctoral sobre recomanació d'hotels, és el d'hotels, amb l'exemple de Booking. Aquesta empresa, a grans trets, recomana hotels a qualsevol persona que vulgui fer un viatge segons la zona, preu, i altres característiques que sol·liciti l'usuari.

El problema és que amb un petit canvi en el seu sistema de recomanació, la pàgina web va deixar de recomanar hotels petits com a primeres opcions a canvi de recomanar hotels grans, cosa que va provocar moltíssimes pèrdues econòmiques als hotels petits.

Bàsicament, es va afegir l'opció de “Booking recomana”, la qual era una etiqueta que s'afegia a certs hotels i feia que sortissin els primers a l'hora de fer una cerca, més o menys com l’”Amazon’s Choice” que presenta Amazon. El problema fou que el criteri no era una selecció objectiva dels millors hotels, sinó que s'havia de pagar una quota per tenir aquella etiqueta, fent que els hotels més grans s’ho permetessin i els més petits no. Això demostra no només que normalment l'opció més triada és la primera que veiem, sinó que a més a més això és fàcilment manipulable.



Il·lustració 17: Funció Amazon's Choice

Font: Amazon

Això també es pot traslladar al camp ideològic, on si només veiem un tipus d’opinió tendim a creure’ns-la més. Aquest fenomen, però, l’explicarem millor a l’apartat de l’Echo Chamber.

2.2.3 Cultura de la Immediatesa

Un dels aspectes culturals més remarcats de la modernitat tardana actual és la immediatesa, aquest desig i necessitat de tenir tot el que volem al moment, en un sol clic. Això, però, no és natural ni beneficis a llarg termini, ja que ens converteix en addictes a la gratificació instantània i incapços de lluitar per un objectiu no assolible en tan sols uns segons.

Aquesta cultura s’ha format a partir de la revolució digital, un fenomen que ha eliminat el component manual de gran part de les professions i oficis, ha facilitat la cerca d’informació mitjançant internet (encara que no ha desenvolupat un major desig de

coneixement) i ha facilitat processos quotidians com el transport, l'anàlisi de dades, entre altres.

Atès a aquest progrés tecnològic, gràcies a la Llei de Moore sabem que la capacitat informàtica es duplica cada dos anys, alhora que està a disposició de gairebé tothom.

Aquest fet produeix que tot el que fem o cerquem tingui una resposta immediata. Fa uns anys, per exemple, era molt més difícil trobar certa informació molt específica, ja que s'havia de recórrer a l'ús d'enciclopèdies o d'experts. També és molt més fàcil la comunicació, ja que podem parlar amb una persona de Xina o Canadà en qüestió de segons, mentre que abans era un procés gairebé inviable. Fins i tot, ara podem realitzar compres amb el mòbil per optimitzar el temps que suposava treure la targeta de crèdit. El problema d'aquest sistema és que com més facilitats tenim, més depenents ens tornem d'aquestes. De fet, hi ha molts estudis que determinen que aquests mateixos avenços tecnològics ens han fet canviar la nostra percepció de la realitat, tornant-nos en individus impacients que inconscientment pretenen obtenir allò que volen en qüestió de segons, minuts, o com a molt hores.⁴¹

Com diu la doctora Laia Molina presentant el moviment “Slow”, “amb l'aparició de les noves tecnologies estalviem temps, perquè fem més activitats amb menys temps i més còmodament”. Adquirir una entrada per un concert, imprimir el bitllet de tren o bé comprar per internet són activitats que ens permeten guanyar temps. Amb aquest temps, però, encara fem més activitats, de tal manera que rarament és emprat per a la distensió, la contemplació o, senzillament, el lleure.

I, al final, no tenim temps per a res. Coses importants com la família, els amics i altres dèries sempre queden en un segon pla. També el cultiu del pensament, que suposadament és una de les coses que identifica més i millor la nostra espècie, resta a l'espera. La pressa ens aclapara. En aquest context precipitat necessitem temps per descobrir el temps.

⁴¹ (*Advantages and Disadvantages of Technology Advances by Lynda Moultry Belcher, Demand Media, n.d.; The Psychology of Instant Gratification and How It Will Revolutionize Your Marketing Approach, n.d.*)

Això ens porta a efectes secundaris com bé és la característica “pressa”, que defineix el nostre temps o la ja esmentada impaciència.

La pressa regeix la nostra vida amb mà de ferro. Ens imposa la pressió que tot s'ha de fer d'una manera determinada i en un marc de temps limitat. Sota el seu efecte, les hores del dia semblen no ser suficients per complir les nostres obligacions laborals i les nostres responsabilitats familiars. El seu únic objectiu és transformar la nostra existència en una cursa de màxima velocitat sense treva ni fi. Igual que un virus, resulta extremadament contagiosa i molt perjudicial.

Per si no n'hi hagués prou, sol anar de la mà de la tòxica impaciència. Un verí que resulta letal per al nostre benestar, la nostra capacitat de relativitzar i el nostre bon humor. Sota la seva influència, perseguim la gratificació instantània sense tenir en compte les repercussions que això pot tenir sobre els altres i sobre la nostra vida. Quan cedim als seus dictats busquem resultats immediats, i perdem interès a esforçar-nos per a aconseguir metes a llarg termini.

Totes dues limiten en gran manera la nostra tolerància a la frustració, i ens porten a la reactivitat i a la precipitació, fent-nos esclaus de la insatisfacció, l'estrés i l'ansietat. La pressió i la sensació d'angoixa ens fan prendre decisions amb el pilot automàtic, sense temps per planificar o preveure. Com a conseqüència d'aquesta cultura de la immediatesa, abocada al fer i al tenir, a penes ens queda temps per ser. I aquesta realitat encara es veu més accentuada a causa de l'avenç imparable de les noves tecnologies i la presència de les pantalles com un nou element indispensable a la nostra vida. Cada vegada estem més endollats a una realitat virtual en què els nostres desitjos immediatament es multipliquen. I el cert és que ja ningú no posa en dubte que aquesta activitat frenètica minva la nostra salut física i emocional. Potser és el moment de qüestionar la premissa generalitzada que la pressa i la impaciència ajuden d'alguna manera a obtenir millors resultats en menys temps.⁴²

⁴² (*La Inmediatez En La Sociedad Actual / Blog / Centro Cristina Andrades, n.d.; SOBRE EL PROBLEMA DE LA INMEDIATEZ O LA NECESIDAD DE GRATIFICACIÓN INMEDIATA: “Su Relación Con La Autodestrucción” – Centro Psicológico SELF, n.d.*)

Resumint-ho, com diu Eckart Tolle, “El que causa malestar és estar al present desitjant estar en un futur”. El concepte “esperar” ara connota quelcom negatiu, i la vinculem amb moments irritants de desesperació, per això fugim del present. Si, com presenta *Juan Salvador Gaviota*, de Richard Bach, aprenguéssim l’art de la paciència, aprendríem a extreure molt més de l’”ara”. De vegades, és tan fàcil com preguntar-se , De què serveix anar més ràpid si estem a la carretera equivocada?

Aprendríem que, per molt que les tecnologies ens facin anar més ràpid, hi ha aspectes de la vida humana que no es poden adaptar al guió marcat per la cultura de la immediatesa.

⁴³

Aquí és, de la mateixa manera que amb l’apartat anterior, on entren els motors de recomanació. Actualment, gran part de les nostres “necessitats” són simplement desitjos per complir una gratificació instantània que anomenem “felicitat” en el món capitalista. Vivim en un món de distraccions, distraccions i més distraccions, com bé diu el filòsof Josep Maria Forné a la seva entrevista, i recordem que això simplement intenta omplir un buit que mai acaba d’omplir, creant-nos insatisfacció crònica (entrevista a Annex 6).

Aquests sistemes de recomanació, tot i la seva important labor, fan un gran treball preservant aquest sistema de consum immediat. Si, per exemple, per trobar aquell vídeo de la teva xarxa social preferida que et genera aquest pic de dopamina l’haguessis de buscar manualment, és a dir, si haguessis de renunciar a la immediatesa que els motors de recomanació ofereixen, deixaries de generar aquesta necessitat, i el triomf del tecnològic sobre l’analògic no seria tant clar (analogia a *Elogio del Papel*, Roberto Casati). ⁴⁴Per tant, la cultura de la immediatesa en la societat postmoderna és un factor que contribueix a l’èxit dels sistemes de recomanació, a la vegada que els sistemes de recomanació contribueixen al nostre desig d’immediatesa, produint-se així la retroalimentació esmentada a l’apartat de Neoliberalisme.

⁴³ (*La Cultura de La Inmediatez - MetamorfosisMetamorfosis*, n.d.)

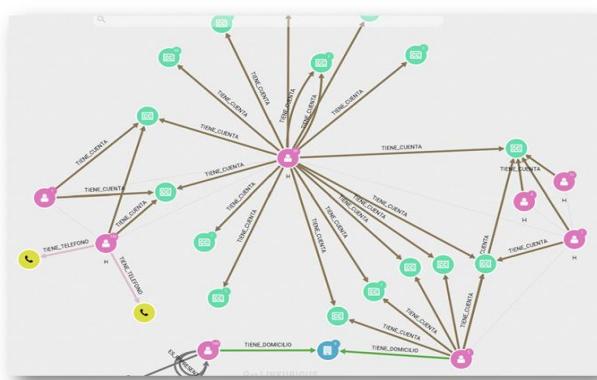
⁴⁴ (*The culture of Speed: The coming of Immediacy*, per Jon Tomlison).

Com també s'esmenta a l'entrevista amb Josep Maria Forné sobre el llibre *Elogio del Papel*, per Roberto Casati, aquesta lluita entre l'analogisme i la tecnologia hauria d'esdevenir un equilibri. Actualment, segons Casati, atès a l'extensió dels sistemes tecnològics en la nostra rutina estem vivint un colonialisme digital que treu molts valors analògics essencials pel coneixement i vida humana. Per exemple, ens està convertint en éssers superficials que ens preocupem més per com ens veiem a les xarxes socials que per la natura, pel món. Processem molta informació de manera constant, urgent i immediata, fent que la nostra lectura deixi de ser crítica i profunda, cosa que no passa amb un llibre. Com es diu en el mateix llibre, “sabem de moltes coses diferents però molt poc de cadascuna”.⁴⁵

2.2.4 Echo Chamber i Filtratge Bombolla

En aquest apartat ens agradaria parlar d'un fenomen produït principalment pels motors de recomanació que esdevé un problema en la polarització de la ideologia de les masses així com en l'aïllament no només ideològic sinó social.

Com a introducció molt lligada a l'apartat d'estímuls, voldríem citar la conferència “Parlem sobre Matemàtiques i Salut” duta a terme el 2 de febrer a Lleida per la matemàtica Clara Grima. A la dita conferència es va parlar sobre com, mitjançant un sistema de nodes (teoria de grafs), es pot representar la xarxa de connexions d'una persona, així com els estímuls que rep.



⁴⁵ (*Elogio del papel*, Roberto Casati).

Il·lustració 18: Sistema de Grafs i nodes

Font: Grapheverywhere

Es van originar debats molt interessants a partir d'aquest concepte, com que el que fan les sectes per reclutar a algú és tallar totes les seves connexions externes a la secta per tal que al seu “sistema de nodes” només hi aparegui la secta en qüestió. D'aquesta manera, tots els estímuls que rep el predisposen a unir-se a aquella secta i ho acaba fent per falta de context real. La realitat d'aquella persona, atès als estímuls que rebia en aquell moment, es relativitzava a simplement la secta.

Amb aquest exemple tan fàcil es demostra la importància dels estímuls, que en un sistema neoliberalista, com ja hem esmentat anteriorment, predisposen a un cicle consumista sense fi.

Doncs més o menys una cosa així passa amb l'Echo Chamber i el filtre bombolla. Quan una persona només recerca sobre una opinió a, per exemple, una xarxa social, el motor de recomanació d'aquesta xarxa social recomanarà ítems semblants al d'aquesta opinió, és a dir, ítems que la reforcen. Això fa que la persona només rebi informació sobre una opinió, i que, per tant, només vegi una part de la realitat. Quan un algorisme de recomanació fa això, diem que ha fet “Echo Chamber”.

En termes més tècnics, el filtre bombolla i el seu derivat “Echo Chamber” són el resultat d'una cerca personalitzada en la qual l'algorisme recomanador d'una pàgina web selecciona, a través de prediccions, la informació que a l'usuari li agradaria veure, basant-se en informació sobre ell mateix i en la seva interacció a la xarxa (dades implícites/explícites). El fenomen esdevé quan tota la informació que l'algorisme rep d'un usuari pertany a un grup d'informació molt concret. Com a resultat, aparta als usuaris de la informació no-aliniada amb els seus punts de vista i els aïlla en bombolles ideològiques i culturals, per tal que aquests es trobin còmodes a la xarxa i la seva activitat “online” es perllongui.⁴⁶

⁴⁶ (Eli Pariser: *Eli Pariser: Cuidado Con La “Burbuja de Filtros” En La Red* | TED Talk, n.d.; *Filter Bubbles in Internet Search Engines* - BBC News, n.d.)

Analògicament, a l'hora de nodrir-te d'informació per tenir una opinió, recerques sobre aquest tema i fent-ho et trobes tota mena d'opinions. A partir de totes aquestes, formes la teva, i d'aquesta manera, amb un procés de reiteració pots contrastar arguments a favor o en contra del que anteriorment pensaves i formar una nova opinió. Això, estadísticament, fa que les opinions es centralitzin, però els motors de recomanació fan el contrari. Quan una persona visualitza un contingut durant x temps sobre qualsevol cosa, el sistema detecta que l'hi ha agradat, i, per tant, continua recomanant el mateix. Això fa, per exemple, que quan comprem vambes només ens surtin anuncis de vambes, però és clar, quan es transporta a l'àmbit ideològic la cosa és més retorçuda. El sistema de recomanació, automàticament, fa que només accedim a veure l'opinió que inicialment hem valorat positivament, i fa que la reforcem i retroalimentem sense tenir en compte les altres. Imagino que serà familiar la següent situació. Prens el teu telèfon mòbil, obres la teva xarxa social preferida i comences a revisar el “timeline”. Apareix una informació, una notícia que et crida l'atenció, a continuació veus com un dels teus contactes l'ha compartit i comentat, i si continues desplaçant la pantalla del mòbil, aquesta mateixa notícia o notícies del mateix caire comencen a aparèixer una vegada i una altra.

Per tant, aquest succès fomenta la tendència humana de l'homofilia, és a dir, a relacionarnos amb aquells que més ens hi assemblem, i també al biaix d'informació, que és el fet de tenir en compte esdeveniments de manera que confirmin les nostres preconcepcions o hipòtesis.⁴⁷

Per tant, el fet de només sentir una opinió, com ja hem esmentat, condueix a una radicalització d'aquesta, ja que només es troba aprovació i més aprovació d'allò que es creu. Un exemple que ho prova és una història viscuda el juny de 1954 a Oklahoma, ara ja estudiada com a fonament de la psicologia social. Allí, dos grups de persones acabades de conèixer, tots blancs i protestants, es van dividir en dues cabines dels boys scouts. Uns es van decidir anomenar “Eagles” mentre que uns altres “Rattlers”. Durant la primera setmana, els dos grups es van conèixer separadament de l'altre grup duent a terme activitats vàries, fent que cada grup tingués una pròpia jerarquia social.

⁴⁷ (*Confirmation Bias - The Skeptic's Dictionary - Skepdic.Com*, n.d.; McPherson et al., 2003)

Al final de la setmana, però, aquests dos grups es van ajuntar, i van decidir fer un torneig. Això va provocar moltes discussions, que sempre es resumien en el mateix. El primer dia, els “Rattlers” van guanyar a totes les competicions i van alçar una bandera que els “Eagles” van cremar immediatament, mentre justificaven la seva derrota per una injustícia. El segon dia va passar exactament el contrari, els “Eagles” ho van guanyar tot, fent que els “Rattlers” ataquessin la seva cabina a la nit i que els “Eagles” contraataquessin al matí. Cada vegada els grups es tornaven més agressius envers l’altre grup i més ferms envers el seu. Uns justificaven que els bats de l’altre equip eren millors, es robaven entre grups, entre altres moltes coses.

Bàsicament, el que venim a dir és que els participants eren molt semblants. L’únic que va fer falta fou tenir un nom diferent i un torneig per fer odiar el grup contrari. El mateix passa, com diu Lillian Mason, amb la política demòcrata i republicana estatunidenca. El gener de 2018, Facebook va anunciar que estava canviant l’algorisme que utilitzava per determinar quines publicacions veuen els usuaris al seu canal de notícies. Aparentment, el canvi es va dissenyar per promoure "interaccions significatives entre les persones". Se suposava que el nou algorisme havia de fomentar la "discussió d'anada i tornada" augmentant el contingut que provocava reaccions emocionalment fortes. El nou sistema, segons la majoria de comptes, va resultar encara pitjor que l'antic. Com potser s'hauria d'haver previst, els missatges que van tendir a provocar més reacció van ser els més provocadors políticament. Així, el nou algorisme va produir una mena de cicle vicios: com més indignació inspirava una publicació, més es promocionava. La gent, per tant, va començar a veure publicacions en contra dels seus ideals que anteriorment no veia, i per això el descontentament va ser tan gran.



Il·lustració 19: Exemple gràfic del funcionament de l'Echo Chamber
Font: The New York Times



Il·lustració 20: Exemple gràfic del funcionament de l'Echo Chamber
Font: The Conversation

Aquí entra l'”Echo Chamber” produït pels motors de recomanació. A Facebook, la gent es fa "amic" de persones amb opinions similars, ja siguin els seus amics genuïns o celebritats i altres personatges públics que admiren. Exemplificant-ho amb l'homofonia, els partidaris de Trump acostumen a escoltar d'altres partidaris de Trump, i els que odien Trump d'altres que odien Trump. Un estudi d'investigadors de Facebook va demostrar que només aproximadament una quarta part del contingut de notícies que els demòcrates publiquen a la plataforma és vist pels republicans, i viceversa. Un estudi sobre l'ús de Twitter va trobar patrons similars. Mentrestant, una infinitat d'estudis, molts que es remunten a abans que Internet es somiés mai, han demostrat que, quan la gent confereix amb altres que estan d'accord amb ells, les seves opinions es tornen més extremes. Els científics socials han batejat aquest efecte com a "polarització grupal" i molts es preocupen que la web s'hagi convertit en un pou d'ideologies on cadascú retroalimenta la seva.⁴⁸

Vivim en una societat on es normalitza el fet que tothom té una opinió vàlida i justificada, cosa que sumada amb l'”Echo Chamber” fa que les opinions externes a les nostres no afectin a la nostra visió de les coses. De fet, és precisament el contrari, ja que crea un conflicte del que ningú treu res, ja que ningú està disposat a trencar esquemes per escoltar altres opinions, sumat al fet que, probablement per una causa comú, cada vegada és menys corrent relacionar-se amb persones alienes a la nostra forma de veure les coses.⁴⁹

MARC PRÀCTIC

3.1 Estudi de comparació de Motors de Recomanació

Fins ara, mitjançant la part teòrica, hem observat com un motor de recomanació no només pot prendre moltes formes i complexitats sinó que té un gran impacte en les nostres vides.

⁴⁸ (*How Politics Got So Polarized / The New Yorker*, n.d.)

⁴⁹ (Adams et al., 2012; Barberá et al., 2015; Cinelli et al., 2021; *La Cámara de Eco o Cómo La Red Te Muestra Sólo Lo Que Quieres Ver - Beers&Politics*, n.d.)

Per tant, no és rar que diguem que el procés de creació d'aquests extensos algorismes és realment difícil, puix que no oblidem que pretén predir el comportament d'un organisme complex com som els humans en un nombre finit d'instruccions iterables.

És per això que molts programadors troben difícil definir el concepte, i més aviat pensen que la seva creació és més aviat un art que una ciència, ja que intentar entrar dins del cap dels consumidors i construir models de les seves preferències en un codi no té cap fórmula general.

Bé, pel que hem definit i donat per entès, un motor de recomanació recomana aquells articles que l'usuari té més probabilitats de consumir, el que és més adient per l'usuari. És a dir, hi ha un codi que fa que, per exemple, Netflix recomani a un usuari qualsevol pel·lícula atès als seus gustos. Seguint aquesta lògica, la recomanació no hauria d'anar basada en funció del codi, sinó de la valoració de l'usuari.

És a dir, si un motor ens mostra allò que més desitgem, podem pensar que allò que es mostra només va en funció de nosaltres mateixos.

Doncs bé, el que volem provar amb aquesta part pràctica és que, al contrari, això és al revés, i el que se'n mostra per pantalla és susceptible a canvis que es fan al codi, que entre altres raons poden establir-se segons interessos de les empreses. Resumidament, provant això es demostrarria que els motors poden esdevenir potencialment una eina de manipulació, que vet aquí vindria a ser la nostra hipòtesi.

Anem a profunditzar-ho; si les cerques són fruit de l'aprenentatge per reiteració d'una intel·ligència artificial, podríem suposar que amb els suficients valors per fer proves tots els motors de recomanació arribarien a una mateixa llista d'articles final que recomanarien a l'usuari. És a dir, partint d'un codi o un altre, la recomanació resultant de cada motor de recomanació seria la mateixa, puix que les recomanacions que rebem són efectives. Això, dit d'aquesta manera sembla absurd, però si pensem en els motors que ens envoltem cada dia no ho és tant. Quan busquem un hotel, estem posant tota la nostra confiança i també cartera en un sistema de recomanació. Probablement, de fet, el nostre viatge anirà depenent del que un algorisme pensa que és el millor per tu. Això és un procés bastant rutinari, ja que recordem que els motors de recomanació també s'utilitzen en xarxes socials, botigues “online”, entre altres..

De fet, com ens explica Eva Martin amb el cas anteriorment esmentat de Booking, tendim a confiar molt en les recomanacions. De fet, hi confiem tant (sigui per mandra o pel que sigui) que si un hotel en concret deixa de ser recomanat a la primera pàgina de cerca perd moltíssimes vendes.

Per tant, seguint aquest procés lògic podem afirmar que els motors de recomanació, mitjançant recomanacions adients, afecten directament el nostre consum o conducta. El nostre consum pot ser afectat per motors de recomanació encarats al servei de compra “online”, mentre que la nostra conducta la poden modificar els resultats que obtenim, per exemple, a una xarxa social, com hem detallat a l’apartat teòric social anteriorment.

El que prenem com a hipòtesi, però, és que un petit canvi en una línia de codi pot acabar afectant a gran escala els resultats assolits en una cerca, i això ho posarem a prova de l’única manera possible, programant-ho.

A través d’un codi fet amb Python, el llenguatge de programació més adient per aquesta branca de la informàtica, elaborarem un motor de recomanació que ho provarà. Aquest prendrà forma d’un motor de recomanació basat en contingut. Per provar la hipòtesi primer sotmetrem les dades d’un usuari en concret al motor de recomanació inicial. Tot seguit, sotmetrem les mateixes dades a aquest motor de recomanació amb una petita diferència en el codi. Canviarem aspectes molt puntuals d’aquest segon motor de recomanació, així com un simple nombre dins de les incomptables línies de codi que té, i comprovarem els resultats.

Si la nostra hipòtesi és certa, sotmetent aquests motors de recomanació a les mateixes dades i al mateix sistema d’avaluació de resultats, en executar-los rebrem recomanacions diferents, provant que les recomanacions que rebem dia a dia estan més controlades del que pensem.

Aquesta prova pot ser realitzada amb molts tipus de dades, però en el nostre cas utilitzarem pel·lícules, ja que són un tipus de dada fàcilment classificable en atributs i els resultats que obtenim són fàcils de diferenciar.

El funcionament general dels dos motors de recomanació serà fàcil. Bàsicament, mitjançant moltes dades introduïdes al sistema, demanarem al motor de recomanació quin

rànquing de pel·lícules recomana a una persona en concret. El motor farà càlculs estadístics i processos d'elecció, i presentarà uns resultats. Paral·lelament, es determinarà l'eficàcia del motor de recomanació comparant els resultats obtinguts i la seva semblança amb les pel·lícules que ja havia valorat l'usuari en qüestió. Per tant, com a resultat final aconseguirem dos llistes de les 10 pel·lícules que cada motor de recomanació recomana a un usuari de la base de dades en concret, és a dir, les que ha predit que tindrien major valoració. També prepararem una llista de recomanacions aleatòries per a provar que els sistemes de recomanació funcionen, i que el seu error és menor al d'aquestes recomanacions aleatòries.

Com és obvi, el nombre de variables que es tindran en compte no són comparables al d'un motor de recomanació fet per una gran empresa, ja que les dades de què disposem no ho permeten, atès que no tenim l'historial implícit de cada persona, i és inviable programar un codi sol on hi hagi tants processos a calcular.

Centrant-nos en els motors, primer de tot necessitarem unes dades amb les quals partir i entrenar la nostra intel·ligència artificial, i no precisament poques, ja que per fer una predicción d'aquest caire fan falta molts exemples i atributs a tenir en compte.

Tenint en compte això, hem d'entendre que les dades que emprarem són la major limitació del nostre sistema. Factors com el nombre d'atributs o la fiabilitat de les dades poden limitar la qualitat de les nostres recomanacions o la quantitat de funcions que podem afegir al sistema.

En el nostre cas utilitzarem un paquet de dades anomenat MovieLens, el qual ens proporciona més de 9000 pel·lícules, classificades per atributs com any, director, o gènere valorades per gairebé 1000 usuaris. Aquesta fiabilitat no és completa, però per raonar la hipòtesi que plantegem és suficient. A més a més, com que aquest projecte no és una pàgina web interactiva, no compta amb tipus de dades implícits, sinó explícits, ja que només disposa de les dades donades per MovieLens.

Cal també anomenar una llibreria de Python que actualment empren gairebé tots els motors de recomanació anomenada “Surprise”. Gràcies a aquesta podem entrenar motors de recomanació d'una manera molt més eficient, puix que aquesta fase la seva creació és

comuna a gairebé tots. Calcula funcions com el RMSE, MAE, que ens ajuden a veure l'error del nostre sistema de recomanació i també generen els càlculs com l'índex de similitud entre dos ítems, cosa que ja parlarem després.

Per conveni, el codi està escrit en anglès, ja que totes les llibreries i funcions a Python utilitzen aquest idioma. Per facilitar la lectura, però, hem comentat algunes línies en Català.

El codi al que farem referència durant tota l'explicació està complet a :

➔ **ANNEX 7 - Codi del programa**

3.1.1 Elaboració Motor de recomanació basat en contingut

En aquest apartat programarem el motor de recomanació base, del qual després modificarem valors per provar la hipòtesi. Hem elegit aquest tipus de motor i no un altre perquè l'abast de les nostres dades està en un llindar molt petit, i fer un motor de recomanació amb filtratge col·laboratiu no tindria gaire sentit.

Abans, però, creiem necessari fer un petit resum de tot el que farà el motor de recomanació, per fer més clar el procés extens que el segueix:

1. Primer agafa dades d'una base de dades que he descarregat d'internet (amb 100.000+ valoracions de pel·lícules).
2. Després entra l'algoritme de manera que el motor de decisió de la intel·ligència artificial sigui més eficaç, i les relacions entre pel·lícules les faci òptimament.
3. Posteriorment, agafa un usuari que nosaltres escollim i determina les recomanacions per aquell determinat usuari.
4. Per acabar, el codi evalua aquestes recomanacions mitjançant diferents mètodes. Com el RMSE o el MAE.

Ara sí, endinsem-nos dins de l'elaboració del motor de recomanació.

Aquest motor de recomanació, si bé recordem l'esmentat a la part teòrica, només utilitza els atributs de cada element per recomanar-ne un altre. La base de dades que usem, MovieLens, no ens dona gaires atributs amb els quals treballar, així que tindrem en compte el gènere de la pel·lícula i l'any que va ser estrenada.

Hi ha 18 possibles gèneres per cada pel·lícula, així que necessitem una manera de mesurar la relació entre dues pel·lícules mitjançant els gèneres de cada una. Per exemple, una pel·lícula amb els gèneres misteri, terror i aventura tindrà més relació amb una altra pel·lícula amb gènere aventura que una amb gènere ciència-ficció. Per fer aquest càlcul farem servir el mètode de la similitud del cosinus⁵⁰. El funcionament és senzill.

Imaginem un eix de coordenades, on l'eix X és, per exemple, el gènere de comèdia i l'eix y és el gènere d'aventura. Si una pel·lícula, com és Toy Story, té l'atribut comèdia i aventura, es situarà al punt [1,1] del pla, mentre que si una pel·lícula només té d'atribut el gènere comèdia es situarà al punt [1,0].



Il·lustració 21: Funcionament similitud cosinus

Font: Sundog Education

⁵⁰ La similitud del cosinus és una mesura de similitud que es calcula entre dos vectors diferents de zero dins de l'espai intern del producte que mesura el cosinus de l'angle entre ells. El cosinus de 0° és 1, i és menor que un per a qualsevol angle que es trobe a l'interval $(0, \pi)$ radians, així que aquest cas es tracta d'un càlcul que dona origen a un judici d'orientació i no de magnitud .

Font: Grapheverywhere

En aquest cas, l'angle que es forma entre els vectors d'aquests dos punts és 45 graus. Com major sigui l'angle menys semblants seran les pel·lícules, així que la raó trigonomètrica a emprar és el cosinus. En aquest cas, el cosinus valdria 0,7. Si dues pel·lícules estiguessin en el mateix punt l'angle seria 0, i, per tant, el cosinus 1. Per realitzar aquest càlcul també podríem haver emprat la distància euclidiana, el coeficient de correlació de Pearson o molts altres mètodes, però aquest és el més fàcil d'entendre.

A la pràctica, però, representar gràficament la relació entre gèneres no serà així de fàcil, ja que hauríem d'escalar aquest exemple amb 2 gèneres a 18. Bàsicament, hauríem d'imaginar cada gènere com una dimensió espacial, i no és viable.

Per tant, en lloc d'això imaginarem que cada pel·lícula té un “set” de 18 coordenades diferents, que oscil·len entre 0 i 1, dependent de si la pel·lícula té un gènere o no. Així quedaria:

Movie	action	adventure	animation	children's	comedy	crime	documentary	drama	fantasy	film-noir	horror	musical	western	mystery	romance	sci-fi	thriller	war	western2
Toy Story	0	1	1	1	1	0	0	0	1	0	0	0	0	0	0	0	0	0	0
Jumanji	0	1	0	1	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0
Grumpier Old Men	0	0	0	0	1	0	0	0	0	0	0	0	0	0	1	0	0	0	0
Waiting to Exhale	0	0	0	0	1	0	0	1	0	0	0	0	0	0	0	1	0	0	0
Father of the Bride	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0

Il·lustració 22: Representació de pel·lícules per gènere

Font: Sundog Education

Traduïda en codi, aquesta funció serà molt important pel motor de recomanació. Matemàticament, la fórmula que determina la similitud de cosinus per i nombre de dimensions és la següent:

$$\text{CosSim}(x, y) = \frac{\sum_i x_i y_i}{\sqrt{\sum_i x_i^2} \sqrt{\sum_i y_i^2}}$$

Il·lustració 23: Fórmula de la similitud del cosinus

Font: Researchgate

On les pel·lícules en qüestió son x i y , la similitud de cosinus és la relació entre aquests termes. Al numerador podem observar el sumatori de productes de x i y a cada dimensió, és a dir, a cada gènere, i al denominador, en canvi, el producte de les arrels quadrades del sumatori de tots els quadrats de cada gènere.

Explicat d'aquesta manera apareix un procés difícil, però realment és senzill posat en un codi. Aquest codi, després d'executar la fórmula esmentada ens retornarà un valor de similitud entre dos pel·lícules basant-nos només en el gènere, cosa que després utilitzarem.

```
def SimilitudGenere(self, peli1, peli2, genere):
    genere1 = genere[peli1]
    genere2 = genere[peli2]
    sumaxx, sumaxy, sumayy = 0, 0, 0
    for i in range(len(genere1)):
        x = genere1[i]
        y = genere2[i]
        sumaxx += x * x
        sumayy += y * y
        sumaxy += x * y
```

Il·lustració 24: Funció similitud cosinus explicada en codi.

Font: Imatge i Codi de l'autor

L'altre factor que volem considerar és l'any d'estrena de cada pel·lícula. Hi ha moltes maneres d'establir un valor de similitud entre dos anys diferents, i per això cal centrar-se en la naturalesa de les dades que s'empra.

Per això hem decidit utilitzar una funció molt senzilla que segueix aquest esquema:

1. Guarda en una variable el valor absolut de la resta dels anys entre una pel·lícula i una altra
2. Guarda una variable el resultat d'elevat el nombre e a la variable anterior dividida entre 10
3. Retorna aquesta variable

```
def computeanyssimilitud(self, peli1, peli2, anys):
    diff = abs(anys[peli1] - anys[peli2])
    sim = math.exp(-diff / 10.0)
    return sim
```

Il·lustració 25: Funció similitud entre anys codificada

Font: Imatge i Codi de l'autor

En el nostre cas hem decidit utilitzar una funció exponencial perquè creiem que la diferència d'anys entre pel·lícules és un valor que, si es fa molt gran, és molt rellevant. No és el mateix una pel·lícula de ciència-ficció dels anys 70 que dels anys 90. D'aquesta manera, obtenim que quan la diferència d'anys és 0, l'índex de similitud és 1, i com major sigui la diferència d'anys menor serà l'índex de similitud.

Ara, haurem de crear la funció que, a partir d'aquests índexs de similitud, crearà la predicció de valoració que tindrà cada pel·lícula segons l'usuari que fem la recomanació. Ho farem mitjançant el mètode k-nearest-neighbours.

La funció, resumidament, tindrà aquests passos.

1. Calcular l'índex de similitud entre totes les pel·lícules que l'usuari ha valorat i la pel·lícula que vulguem calcular-hi la predicció.
2. Calcular les k pel·lícules (en el nostre cas 50) amb un major índex de similitud a la pel·lícula a què volem predir la valoració.

3. Fer la mitjana ponderada de l'índex de similitud d'aquestes 30 pel·lícules amb la pel·lícula que volem calcular-hi la predicción, ponderant segons la valoració que l'usuari hagi donat de cadascuna de les pel·lícules

Sembla un procés complex, però realment no ho és tant. Com l'usuari no haurà valorat totes aquestes 30 pel·lícules, realment només es ponderaran a la mitjana final els índexs de similitud de les pel·lícules que sí han estat valorades, fent que quedi un càlcul simple.

Aplicat al codi, ens quedaria així:

```
def estimate(self, u, i):
    if not (self.trainset.knows_user(u) and self.trainset.knows_item(i)):
        raise PredictionImpossible('User and/or item is unknown.')

    # Crear índice de similitud entre pel·lícula [i] i tots els elements que ha valorat l'usuari
    neighbors = []
    for rating in self.trainset.ur[u]:
        genreSimilarity = self.similarities[i,rating[0]]
        neighbors.append( (genreSimilarity, rating[1]) )

    # Extreure les valoracions més similars (top-k)
    k_neighbors = heapq.nlargest(self.k, neighbors, key=lambda t: t[0])

    # Calcular la mitjana ponderada dels top-k neighbours segons les valoracions de l'usuari
    simTotal = weightedSum = 0
    for (simScore, rating) in k_neighbors:
        if (simScore > 0):
            simTotal += simScore
            weightedSum += simScore * rating

    if (simTotal == 0):
        raise PredictionImpossible('No neighbors')

    predictedRating = weightedSum / simTotal
    return predictedRating
```

Il·lustració 26: K-Nearest-Neighbours
 Font: Codi i Imatge de l'autor

Mitjançant un bucle, anem per cada pel·lícula que l'usuari hagi valorat i omplim la llista “neighbours” amb l'índex de similitud entre aquesta pel·lícula i la pel·lícula que volem predir-hi la valoració. Com he esmentat abans, els índex de similitud els podem calcular fàcilment amb les funcions que el mòdul “Surprise” ens proporciona.

Després, ordenem mitjançant la funció heapq.nlargest aquesta llista per tornar-la en les top-K pel·lícules amb el major índex de similitud, per a finalment computar la mitjana ponderada d'aquestes top-k pel·lícules segons la valoració de cada pel·lícula segons l'usuari. Aquest valor serà la predicción de la valoració de la pel·lícula.

Aquests trets són els que diferencien aquest motor de recomanació amb l'altre, però encara no hem ensenyat com es computa en conjunt. Dit això, anem a veure les parts del nostre codi.

Abans de tot, cal parlar de la classe AlgoBase. Aquesta és una classe⁵¹ del mòdul “Surprise” que ajuda a poder entrenar qualsevol algorisme amb les funcions que el mòdul incorpora. Dona igual quin tipus de subclasse utilitzis, com en el nostre cas és KNN, que si deriva d'una arquitectura Algo-Base podrà ser entrenada per funcions del mòdul “Surprise”.

A trets generals, en aquest tros de codi primerament definim la k, que és el nombre de “neighbours” que volem utilitzar al procés explícit anteriorment de predicción de valoracions. Després implementem una funció “fit” que posteriorment serà cridada per la llibreria “Surprise” quan entrenem l'algorisme.

També, el que estem fent aquí és crear una matriu 2D que serveix com a cerca de la puntuació de similitud basada en el contingut entre dues pel·lícules qualsevol. Això triga un temps a executar-se, així que cada 100 pel·lícules processades mostrem per pantalla el progrés.

Després, calculem els índex de similitud per gènere i any per cada parell possible i els multipliquem junts per a relacionar aquests dos puntuacions. Sabem que la similitud entre les pel·lícules A i B és la mateixa que entre les pel·lícules B i A, així que només calculem aquesta puntuació una vegada i la copiem en ambdós casos a la matriu. Aquesta funció, enllaç de utilitzar una altra funció preestablerta pel mòdul “Surprise” a l'hora d'elegir un criteri de creació d'índex de similituds, utilitza les dos funcions que hem creat abans sobre el gènere i l'any. Aquest arxiu de codi també conté la funció *Estimate* que hem explicat abans, que resumidament és la que aplica el mètode *k-neighbours* per a convertir aquests índexs de similitud en una verdadera predicción de valoració.

⁵¹ Les classes proveeixen una forma d'empaquetar dades i funcionalitat junts. En crear una classe nova, es crea un tipus d'objecte nou i permet crear noves instàncies d'aquest tipus.

Font: [Python.org](https://www.python.org)

Posteriorment, tenim 4 arxius dins del directori del motor de recomanació que van junts, així que els detallarem en conjunt. Igual que ha passat amb l'arxiu anterior, també deriven de la classe *AlgoBase*. Aquests són els arxius que s'encarreguen de valorar la prediccio del sistema de recomanació. Realment, podríem simplement executar els resultats que rebem amb el submòdul KNN que acabem d'explicar, però creiem que és important valorar si de veritat és útil una recomanació.

En primer lloc hem de definir unes mètriques d'entrenament, i això és el que fa l'arxiu Recommender Metrics, que trobem a l'annex 7.

Bàsicament la funció d'aquest arxiu és definir tots els tipus d'estratègies d'avaluació de prediccions que li aplicarem a les prediccions que generem amb el primer arxiu. Entre elles es troben totes aquelles esmentades a la part teòrica transformades en codi. Moltes d'aquestes funcions també ens són facilitades pel mòdul “Surprise”, o per individuals a la xarxa. Com ja he dit, l'explicació més extensa del funcionament d'aquests criteris d'avaluació està a la part teòrica, però anem a repassar-la.

Si volem entrenar un algorisme de recomanació, com és obvi hem d'utilitzar les dades de l'usuari que ja disposem. Primerament, dividim aquestes dades en diferents paquets, cosa que farà una altra classe. Aquests paquets inclouran sets d'entrenament i sets de pràctica (train/test sets). L'objectiu és que, amb els patrons que l'algorisme detecti en el paquet d'entrenament endevini les valoracions del paquet de pràctica. Bàsicament, el que volem amb aquest procés és veure si el motor de recomanació s'apropa per ell sol als valors reals dels quals ja disposem. Si sabem que un usuari ha posat 5 estrelles a una pel·lícula, però l'algorisme entrenant diu que n'hauria posat 3, aquestes estratègies d'avaluació de prediccions ens ensenyen l'error que ha tingut. El mateix passa amb una llista de prediccions. El sistema entrenant presenta el que ell creu que seria la llista d'elements (pel·lícules) preferits per l'usuari, i aquestes estratègies d'avaluació mostren si està molt o molt poc equivocat.

En primer lloc, tenim el RMSE, o l'arrel de l'error quadràtic mitjà, que juntament amb el MAE o error quadràtic mitjà mostren la diferència entre els valors obtinguts i els reals. Després està el “hit rate”, que mesura la freqüència amb què el nostre sistema recomana l'element correcte al model d'aprenentatge “Leave One Out”. Entre altres, també trobem

valors com la diversitat, que mesura la “similitud” entre ítems recomanats; o la popularitat, que mesura la popularitat dels ítems recomanats. Segons cada motor de recomanació els valors desitjats canvien, ja que potser per un servei interessa un valor de diversitat que a un altre no. Ja tenint això, necessitem una manera d’aplicar aquestes mètriques d’avaluació a l’algorisme que hem creat a l’anterior arxiu KNN, i per això tenim la classe anomenada “EvaluatedAlgorithm”. La definirem en un altre arxiu del codi.

Aquesta classe conté un algorisme de “Surprise”, anomenat “innit” però introduceix una nova funció anomenada “Evaluate” que executa totes les mètriques a RecommenderMetrics en aquest algorisme, fet per nosaltres. Per tant, aquesta classe facilita la mesura de la precisió, la cobertura, diversitat i tota la resta d'un algorisme determinat, que en aquest cas és el de filtratge basat en contingut.

Ara, per avaluar un sistema de recomanació de diverses maneres diferents, hem de dividir les nostres dades d'entrenament en seccions dividides de diferents maneres. Per això serveix la classe “EvaluationData”, també definida a part dins del codi.

Aquesta classe pren una sèrie de “Data-Sets” d’entrenament que provenen de la base de dades “MovieLens” i crea tots els “Train/Test Splits⁵²” necessaris per a la nostra classe “EvaluatedAlgorithm”, que és la que vertaderament entrena l’algorisme. En aquest cas en crea molts tipus, un per cada tipus d’avaluació. Per exemple, crea un “Train/Test Split” de 75/25, que utilitza el 75% de les dades per entrenar i l’últim 25% per practicar, o també crea un set de dades per generar matrius de similitud. D'aquesta manera, gràcies a la cohesió entre pràctica i entrenament, la intel·ligència artificial, és a dir, el motor de decisió del motor de recomanació es perfecciona, de manera que recomani més eficientment. Això, tècnicament, es realitza fent que la màquina estableixi patrons entre

⁵² Els Train/Test splits són un mètode que divideix les dades de valoracions passades de l’usuari en disposició del motor en dos grups, un que es farà servir per entrenar (“train”) i un altre per practicar (“test”), de manera que el motor de decisió de la intel·ligència artificial sigui més eficient (estableixi més patrons i relacions entre productes).

productes, de manera que a l'hora de recomanar prengui en compte tots els exemples vists anteriorment i els posi en pràctica

En resum, la classe “central” “EvaluatedAlgorithm” utilitza els data sets creats per “Evaluation Data” per posar a prova les mesures definides per “Recommendation Metrics”.

Per últim està la classe i arxiu “Evaluator”, que resumidament compara algorismes. En aquest cas, només compararà l'algorisme KNN principal amb un recomanador aleatori, com hem esmentat anteriorment, però no deixa de ser interessant la seva funció. Quan executem el codi, aquesta classe determinarà quin usuari és el que utilitzarem, a la foto el 85, quantes recomanacions es faran, en aquest cas 10, i també és el que preserva la sintaxi del que es mostrarà per pantalla. Aquesta classe extreu les dades de la base de dades, en aquest cas de MovieLens, i mitjançant la funció “AddAlgorithm” crea un “EvaluatedAlgorithm” per cada algorisme que vulguem comparar internament. Això fa que si es vol provar nous algorismes no s'hagi de crear una classe “EvaluatedAlgorithm” o “EvaluationData” per cada algorisme, sinó que simplement s'haurà d'afegir a “Evaluator”.

L'arxiu amb què executarem el motor de recomanació és aquest, “ContentRecs”, que crida les funcions necessàries per trobar les dades que volem i processar-les.

3.1.2 Resultats

Ara només queda veure que ens retorna el codi executat. Primer ho farem amb l'usuari 85.

13186	85,25,0,837511784
13187	85,3,2,0,837512420
13188	85,5,3,0,837512493
13189	85,10,5,0,837507143
13190	85,19,3,0,837512280
13191	85,21,4,0,837511743
13192	85,23,3,0,837512169
13193	85,44,2,0,837512134
13194	85,58,1,0,837506990
13195	85,110,5,0,837506903
13196	85,153,4,0,837506803
13197	85,158,1,0,837513025
13198	85,160,3,0,837512403
13199	85,161,4,0,837507117
13200	85,165,4,0,837506796
13201	85,170,3,0,837512681
13202	85,172,3,0,837513025
13203	85,173,4,0,837512148
13204	85,177,2,0,837512454
13205	85,181,10,837512701
13206	85,185,3,0,837511998
13207	85,186,3,0,837512987
13208	85,188,1,0,837511493
13209	85,196,3,0,837511986
13210	85,203,3,0,837512599
13211	85,208,3,0,837511797
13212	85,216,3,0,837512325
13213	85,227,4,0,837512325
13214	85,230,2,0,837512519
13215	85,231,4,0,837507101
13216	85,234,4,0,837512212
13217	85,247,1,0,837512599
13218	85,253,3,0,837512030
13219	85,255,5,0,837512635
13220	85,256,4,0,837513025
13221	85,261,1,0,837513012
13222	85,275,3,0,837512134
13223	85,277,2,0,837512482
13224	85,282,1,0,837512962
13225	85,288,4,0,837512114
13226	85,291,1,0,837512611
13227	85,292,4,0,837507117
13228	85,293,5,0,837511431
13229	85,296,5,0,837506719

Il·lustració 27: Full d'excel amb valoracions

Font: Movie Lens

Abans d'executar, però, seria adient donar un cop d'ull a l'historial de valoracions d'aquest usuari, per a valorar qualitativament si les recomanacions són adients.

movielid	title	genres
1	Toy Story (1995)	Adventure Animation Children Comedy Fantasy
2	Jumanji (1995)	Adventure Children Fantasy
3	Grumpier Old Men (1995)	Comedy Romance
4	Waiting to Exhale (1995)	Comedy Drama Romance
5	Father of the Bride Part II (1995)	Comedy
6	Heat (1995)	Action Crime Thriller
7	Sabrina (1995)	Comedy Romance
8	Tom and Huck (1995)	Adventure Children
105	"Bridges of Madison County, The (1995)"	Drama Romance
107	Muppet Treasure Island (1996)	Adventure Children Comedy Musical
108	Catwalk (1996)	Documentary
110	Braveheart (1995)	Action Drama War
111	Taxi Driver (1976)	Crime Drama Thriller
112	Rumble in the Bronx (Hont faan kui) (1995)	Action Adventure Comedy Crime
293	L.A. Confidential: The Professional (A.k.a. The Professional) (L.A. Confidential) (1994)	Action Crime Drama Thriller
294	"Perez Family, The (1995)"	Comedy Romance
295	Pyromaniac's Love Story, A (1995)	Comedy Romance
296	Pulp Fiction (1994)	Comedy Crime Drama Thriller

Il·lustració 28: Pel·lícules Correspondents a cada ID

Font: MovieLens

A la foto de l'esquerra podem observar una sèrie de valors extrets de la base de dades MovieLens que ens indiquen, en primer lloc, l'usuari, després l'ID de la pel·lícula que l'usuari ha valorat i posteriorment la valoració com a tal. Amb aquesta informació podem buscar quines pel·lícules corresponen a cada ID de pel·lícula que ha valorat positivament (5 estrelles).

No és difícil adonar-se que a les valoracions d'aquest usuari predominen les pel·lícules de Drama i acció, així que, basant-nos en el gènere, li haurien d'agradar pel·lícules amb atributs similars. A més, si ens fixem en la data aquestes pel·lícules són totes dels 90, així que seria rar que una recomanació s'allunyés molt d'aquests termes.

En executar, el programa ha mostrat això per pantalla:

```
Carregant valoració de pel·lícules...
Computant el rànking de popularitat de les pel·lícules epr mesurar la popularitat del resultat...
Estimating biases using als...
Computing the cosine similarity matrix...
Done computing similarity matrix.
Avaluant ContentKNN ...
Avaluant precisió...
Computant la matriu de similitud per contingut...
0 of 8211
100 of 8211
200 of 8211
300 of 8211
400 of 8211
500 of 8211
600 of 8211
700 of 8211
800 of 8211
```

Il·lustració 29: Execució part 1

Font: Imatge i Codi de l'autor

En començar el procés el primer que ha fet és reproduir iteracions per entrenar l'algorisme. Traduït al que hem treballat a l'apartat anterior, bàsicament avaluua la qualitat de l'algorisme de la classe KNN a partir de “data sets” d’entrenament i pràctica, fent-li les proves de “Recommender Metrics”. A més a més, ho repeteix amb un algorisme que simplement agafa pel·lícules aleatòries, per a utilitzar-ho com a grup control.

```
...fet
...fet
Anàlisi complet.
Avaluant Random ...
Avaluant precisió...
Anàlisi complet.

Algorithm RMSE MAE
ContentKNN 0.9380 0.7263
Random 1.4385 1.1478

Legend:
RMSE: Root Mean Squared Error. Valors petits signifiquen més precisió.
MAE: Mean Absolute Error. Valors petits signifiquen més precisió.

Usant recomanador ContentKNN

Constraint model de recomanació...
Computant la matriu de similitud per contingut...
0 of 9066
100 of 9066
200 of 9066
300 of 9066
400 of 9066
500 of 9066
```

Il·lustració 30: Execució Part 2

Font: Codi i Imatge de l'autor

Aquí mostrem per pantalla les puntuacions de RMSE i MAE que ha obtingut l'algorisme KNN i l'aleatori. Efectivament, podem afirmar que els algorismes de similitud entre ítems que hem programat a la classe KNN són efectius, ja que els valors de RMSE i MAE són significativament menors. Això vol dir que la relació que el nostre motor de recomanació fa entre els atributs de les pel·lícules (gènere i any) i les pel·lícules valorades per cada usuari és certament eficient, encara que, òbviament, no perfecte. No oblidem que aquest

motor de recomanació només té en compte un factor de vegades tan irrelevants com és el gènere d'una pel·lícula, així que no pot ser comparat als centenars de dades implícites que utilitzen motors com Netflix o Amazon.

```

8900 of 9066
9000 of 9066
...fet
Computant recomanacions..

Recomanem:
Romeo Must Die (2000) 3.9495359023237966
Licence to Kill (1989) 3.945765099237318
Raiders of the Lost Ark: The Adaptation (1989) 3.945765099237318
Die Hard 2 (1990) 3.9452414915214012
Hunt for Red October, The (1990) 3.9452414015214012
Living Daylights, The (1987) 3.9435774883384176
Rambo: First Blood Part II (1985) 3.931062071636673
View to a Kill, A (1985) 3.931062071636673
Die Hard (1988) 3.918913648648398
Dead Pool, The (1988) 3.918913648648398

Usant recomanador Random

Constraint model de recomanació...
Computant recomanacions..

```

Il·lustració 31: Execució Part 3

Font: Imatge i Codi de l'autor

Després, com també observem a la imatge anterior, torna a fer una sèrie d'iteracions, però per una raó diferent. En lloc d'entrenar l'algorisme (quelcom irrelevants pel consumidor) aquesta vegada el nostre motor de recomanació ha creat una llista de les 10 pel·lícules que més poden agradar al consumidor segons les dades donades, mitjançant l'algorisme de “k-nearest neighbours” esmentat anteriorment, produint una llista “Top-N”. Aquesta primera llista, com ja es pot suposar, és el de l'algorisme aleatori, ja que com podem veure el número al costat de la pel·lícula (la valoració que prediu el nostre motor) no és encertat (5 és el màxim).

```

Constraint model de recomanació...
Computant recomanacions..

Recomanem:
Sleepers (1996) 5
Beavis and Butt-Head Do America (1996) 5
Fear and Loathing in Las Vegas (1998) 5
Happiness (1998) 5
Summer of Sam (1999) 5
Bowling for Columbine (2002) 5
Babe (1995) 5
Birdcage, The (1996) 5
Carlito's Way (1993) 5
Wizard of Oz, The (1939) 5

In [33]:

```

Il·lustració 32: Execució part 4

Font: Codi i Imatge de l'autor

Finalment, l'execució acaba mostrant per pantalla les pel·lícules que el nostre algorisme ha determinat com més probables a què rebin una puntuació positiva de l'usuari 85.

Hem de tenir en compte que no només utilitza el factor dels atributs (gènere i any) sinó que també ha classificat les dades segons la diversitat, popularitat i altres factors que hem determinat al codi anteriorment. Això és el que realment varia la recomanació.

Com podem veure els resultats, qualitativament, no estan mal encaminats, encara que no estan encertats del tot. La majoria de títols recomanats preserven els atributs predominants en l'usuari 85 (drama i acció), però hi ha alguns, com Babe, que no tenen res a veure. Això és molt comú als sistemes de recomanació basats en contingut, ja que de vegades hi ha molts factors que no poden ser etiquetats amb un atribut. En canvi, els de filtratge col·laboratiu, si disposen de moltes dades tendeixen a donar millors resultats, puix que les decisions que pren són fruit de patrons humans.

A continuació editarem un tros molt petit del codi, i executarem el mateix procés. Per a fer més clar i evident el resultat de la nostra hipòtesi, simplement canviarem 2 valors arbitràriament dins d'algunes funcions del codi. Básicament, un canvi que sembla inofensiu, però si realment fa variar les recomanacions finals que se'n mostren per pantalla deixarà de ser-ho.

Aquí estan els canvis fets:

```
def __init__(self, k=30, sim_options={}):
    pass

def __init__(self, k=15, sim_options={}):
    pass
```



Il·lustració 33: Canvis realitzats al codi. Part 1

Font: Codi i Imatge de l'autor

En primer lloc, canviem la `k` de la funció KNN, dins del mòdul “ContentKNNAlgorithm”, la qual defineix el nombre de “neighbours” que se li assignarà a la pel·lícula a què li volem predir la valoració.

```
self.trainSet, self.testSet = train_test_split(data, test_size=.25, random_state=1)
self.trainSet, self.testSet = train_test_split(data, test_size=.45, random_state=1)
```



Il·lustració 34: Canvis Realitzats al codi. Part 2

Font: Codi i Imatge de l'autor

lel

“Data-Set” que anirà destinat a entrenament o a pràctica, fent que agafi d’exemples, en lloc del 75% del total de les dades, només un 55%.

Anem a veure com afecten aquests canvis a l’execució del codi. Gran part de l’execució l’hem obviat, ja que és bàsicament el mateix (la part aleatòria i les iteracions), així que saltarem directament a aquell apartat variable.

```
7000  of  7333
7100  of  7333
7200  of  7333
7300  of  7333
...fet
Anàlisi complet.
Avaluant Random ...
Avaluant precisió...
Anàlisi complet.

Algorithm RMSE      MAE
ContentKNN 0.9491    0.7326
Random      1.4337    1.1445

Legend:
RMSE:      Root Mean Squared Error. Valors petits signifiquen més precisió.
MAE:       Mean Absolute Error. Valors petits signifiquen més precisió.
```

Il·lustració 35: Segona execució (després del canvi), Part 1

Font: Imatge i Codi de l'autor

El primer que observem és que el nombre d’iteracions que ha de fer per “entrenar” és menor, per això el que abans era 8281 ara s’ha convertit en 7333, però això és irrelevat. Un factor que si és rellevant és la precisió que ha tingut aquest motor. Com es pot considerar, els valors de RMSE i MAE han augmentat lleugerament. Això vol dir que, a l’hora d’entrenar el motor, els valors que ha donat l’algorisme al set de pràctica s’allunyen més dels reals que amb el primer motor.

Això, però, realment és irrelevat per la nostra hipòtesi. Hem volgut posar-ho perquè és cabdal per a entendre el funcionament del motor i el seu criteri d’autoavaluació, però com l’entrenament no s’acaba reflectint en un motor de recomanació real, ja que l’usuari només veu les recomanacions, no és el factor principal que defensarà la nostra hipòtesi.

```
Recomanem:  
Dumbo (1941) 5  
Cinema Paradiso (Nuovo cinema Paradiso) (1989) 5  
Ben-Hur (1959) 5  
Brothers McMullen, The (1995) 5  
First Knight (1995) 5  
Paper, The (1994) 5  
Remains of the Day, The (1993) 5  
Trainspotting (1996) 5  
Finding Nemo (2003) 5  
Dark Knight, The (2008) 5  
  
In [6]:
```

Il·lustració 36: Segona execució (després del canvi), part 2

Font: Codi i Imatge de l'autor

Les recomanacions obtingudes després d'aquest lleuger canvi han estat aquestes. Podem observar que, com ens plantejàvem, són completament diferents de les assolides al primer motor de recomanació. De fet, el canvi és tan gran que no hi ha cap pel·lícula que hagi coincidit. Per fer una anàlisi concreta d'aquestes dades, podríem dir que, encara que la variació entre anys s'ha mantingut, els gèneres de les pel·lícules obtingudes (que podem consultar a la base de dades MovieLens) és més adient, puix que s'adqua sense errors dràstics als gèneres de les pel·lícules ben valorades per l'usuari. Com a apunt, també cal dir que ha trobat resultats amb un índex de popularitat major, ja que algunes de les pel·lícules són conegeudes.

Aquest és un exemple clar d'una realitat bastant present al món de la programació, i és que quan pensàvem que els resultats serien pitjors atès que la valoració de l'entrenament ha estat pitjor, passa el contrari.

Per a validar la nostra hipòtesi hem reproduït aquest procés amb molts altres usuaris, i el resultat també ha estat el mateix.

Per tant, podem afirmar que la nostra hipòtesi és certa, i que les recomanacions que un sistema de recomanació produeix són molt sensibles al codi que el regeix. Cada línia de codi, seguint aquesta lògica, afecta les recomanacions que veiem, i això repercuteix en el nostre consum i conducta. Això és fàcilment comprovable, ja que com ens diu la doctora Eva Martin, tendim a confiar molt en les recomanacions. Com molts estudis (i la pròpia experiència de cadascú) proven, no ens anem a, per exemple, la vintena pàgina d'Amazon

per comprar un producte, sinó que tendim a escollir els primers que veiem. Per tant, si per aquesta immediatesa tendim a consumir el primer que veiem, i allò que veiem primer ho dicta el motor de recomanació, consumim el que ens mostra el motor de recomanació. Això seria inofensiu excepte pel fet que ara sabem que, efectivament, els resultats que mostren els motors de recomanació son susceptibles, i per tant manipulables, per aquells que els creen.

Sabent això, que les recomanacions siguin tan variables i susceptibles als desitjos del productor del sistema de recomanació poden esdevenir un gran estri de manipulació tant econòmica com ideològica. Com hem argumentat a la part teòrica, l'estat d'ànim de moltes persones (sobretot adolescents) està directament relacionat amb les xarxes socials, i ara ja sabem que allò que apareix a les xarxes socials pot ser manipulable. Per tant, així com hem observat amb l'algorisme de Booking o amb el fenomen de l'"Echo Chamber", tenim menys control sobre nosaltres mateixos del que pensem, i això és un problema.

3.2 Contextualització de l'entorn social

El fet d'investigar tant sobre el rerefons social que tenen les tecnologies com els motors de recomanació ens ha fet pensar sobre què pensa la gent sobre aquest tema. Més concretament, a preguntar-nos quin és el nivell de conscienciació que els usuaris d'aquest sistema socioeconòmic, o ara podríem dir ja víctimes del mateix, els consumidors, tenen envers la situació actual. Hem parlat sobre com totes aquestes tecnologies actuen de mirall a molts aspectes que defineixen el postmodernisme occidental, com bé és la cultura de la immediatesa, la sobrecàrrega d'estímuls o el context neoliberalista; però és la gent conscient d'això?

El format d'investigació va sorgir d'una simple pregunta. No enteníem com en un món on tothom té un accés massiu a la informació, hi pot haver a la vegada tanta manipulació.

Per esbrinar-ho, hem volgut preparar unes simples preguntes que ens serviran de mostra, de tastet, de l'opinió i posició popular sobre aquest sistema tan llibertari però controlador. Cal recalcar que els resultats no determinen una mostra significativa de la nostra societat,

i que per a res aquest apartat té intenció de ser considerat una enquesta, sinó més aviat un apropament al nostre entorn que crida a una reflexió.

Com hem parlat anteriorment als apartats socials, el monitoratge massiu de les nostres dades és una característica clau pel sosteniment de les noves tecnologies, ja que, sense tenir dades amb què operar tecnologies com els motors de recomanació, serien inoperatives. Amb aquestes dades, aquests sistemes ens indueixen a un consumisme personalitzat que es manté gràcies a la “falsa felicitat” que ens transmeten a través de petits pics de dopamina, i així en bucle.

Per aquesta raó hem volgut encarar aquestes preguntes a aquest apartat de la nostra societat, el mercat de dades, tenint també en compte que és un tema de què mínimament es coneix, a diferència d'altres factors potser més profunds del neoliberalisme.

Per a realitzar aquesta petita “enquesta” varem anar per diferents zones de Lleida preguntant a 23 persones menors de 18 anys, 22 persones entre 25 i 60 anys i 18 persones majors de 60 anys sobre aquest tema. Per a no fer-ho llarg, cosa que afectaria negativament a les respostes, l’”enquesta” constà de 2 preguntes senzilles (i l’edat), que si haguéssim de formalitzar prendrien aquesta forma:

1. Quina és la teva edat?
2. Ets conscient que, utilitzant una xarxa social o qualsevol servei de venda “online” s’utilitzen les teves dades i es manipulen els resultats que obtens?
3. Disposar d’aquest coneixement afecta o afectaria (en el cas de resposta negativa a la pregunta anterior) la teva manera d’interaccionar amb aquests serveis?

Els resultats obtinguts després d'un temps preguntant, obviant les respostes que ometien la pregunta (unes quantes), es troba a la següent pàgina. Aquestes varen ser anotades in situ, ja que les preguntes eren orals, així que no hi ha cap mena de formulari que hagi estat explícitament transmès als participants.

Els resultats, en forma de gràfica, varen ser els següents:

Entre les persones de menys de 18 anys:



Il·lustració 40: Gràfic -18, P.1 (15 SI 8 NO)

Font: Recull de dades i Gràfic de l'autor

Entre les persones entre 18 i 60 anys:



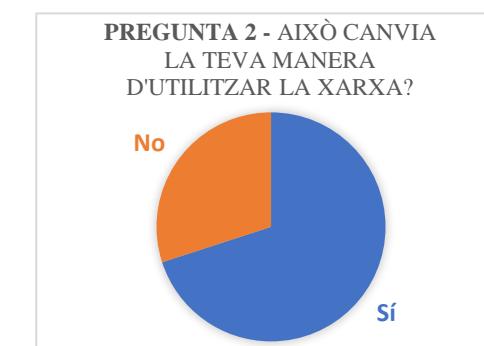
Il·lustració 39: Gràfic 18-60, P.1 (13 SI 8 NO)

Font: Recull de dades i Gràfic de l'autor



Il·lustració 37: Gràfic -18, P.2 (19 SI 4 NO)

Font: Recull de dades i Gràfic de l'autor



Il·lustració 38: Gràfic 18-60, P.2 (15 SI 6 NO)

Font: Recull de dades i Gràfic de l'autor

Entre les persones de més de 60 anys:



Il·lustració 42: Gràfic +60, P.1 (7 SI 11 NO)

Font: Recull de dades i Gràfic de l'autor



Il·lustració 41: Gràfic +60, P.2 (15 SI 3 NO)

Font: Recull de dades i Gràfic de l'autor

Primer de tot, el que voldria recalcar és el general desconeixement del tema fora de la carcassa. La majoria de persones, aproximadament un 80%, no comprenen sobre quina cosa estaven sent preguntats fins que els vam començar a parlar de “cookies”, algorismes intel·ligents, o altres idees més comunes.

Per tant, com a primer raonament extret d'aquest tastet, trobo que hi ha una gran desinformació general entre la gent sobre això. És sorprenent que, amb l'auge que hi ha de les tecnologies i el gairebé infinit portal d'informació que internet ens proporciona, hi hagi tan poc coneixement sobre un tema que tant ens afecta a tots.

Potser, com diu Nicholas Carr al seu llibre *Superficiales: ¿Qué está haciendo internet con nuestras mentes?* o també esmenta Roberto Casati a *El Elogio del papel*, si deixéssim de retroalimentar aquest sistema que ens condueix a la ignorància crònica deixaríem de ser una societat superficial i immobilista.

Això va lligat, en la meva opinió, amb el segon punt a comentar. Sobretot en l'apartat d'adolescents podem observar com molta gent que coneix els sistemes que utilitzen serveis com les xarxes socials, després afirmen que no els fa variar la seva manera d'utilitzar-los. De fet, més d'un 20% dels adolescents ens van dir que sabien que era quelcom “dolent” però que no els deia res, no volien fer res al respecte. Això ens va fer pensar que el problema d'avui dia no només és la ignorància, sinó el conformisme. Com diu el llibre *Inmediatez, capitalismo y vides aceleradas*, de Borja Muntadas Figueras, tendim a una societat on ja no només deixarem de tenir essència humana, sinó que tampoc voldrem recuperar-la.

Tendim a un conformisme que, com diu John Tomilson al seu llibre *La cultura de la rapideza*, ve derivat d'una sobrecàrrega d'estímuls i distraccions que ens porten cap a la manca de creativitat i a la ignorància.

4. CONCLUSIONS

Durant tot el treball, els motors de recomanació ens han fet virar entre la tecnologia i la sociologia amb el propòsit de donar resposta a la nostra hipòtesi. Concretament, hem arribat a aquesta contesta mitjançant la consecució dels tres objectius primaris, i això ens ha deixat amb els següents punts clau.

En primer lloc, a la part teòrica, el concepte de motor de recomanació ha estat explicat des de diferents punts de vista, mitjançant un desglossament en les diferents parts que el componen. A més a més, a la part pràctica, s'ha elaborat un motor de recomanació que ha permès mostrar tot el procés tecnològic que hi ha darrere d'aquesta eina informàtica, de manera que s'ha pogut aproximar el tipus de procés que usen les grans empreses per a recomanar els seus productes.

Gràcies a l'acompliment d'aquest objectiu hem pogut acotar l'estratosfèric nombre de processos que es duen a terme en una acció tan petita com mirar el mòbil, i hem observat que el terme central de tota recomanació és les nostres dades. D'aquesta manera, fem una crida a deixar de treure-li importància a aquesta acció tan quotidiana i a donar-nos compte que, a la xarxa, tot està calculat al mil·límetre.

En segon lloc, mitjançant la part pràctica tecnològica, s'ha demostrat que un petit canvi en el codi d'un motor de recomanació es tradueix en la modificació completa de les recomanacions resultants. Això s'ha fet demostrant que, prenent un mateix motor de recomanació i executant-lo una vegada abans de canviar un valor al codi i una després de fer-ho, les seves recomanacions varien.

L'assoliment d'aquest objectiu també permet afirmar la hipòtesi principal, ja que els motors de recomanació es poden manipular molt fàcilment. De la mateixa manera que canviant uns valors arbitraris totes les recomanacions han canviat, una empresa destinada a la venda pot predeterminar un cert tipus de recomanacions que ens facin consumir una cosa o una altra. Si consumim allò que veiem primer (per pur efecte d'immediatesa), i allò que veiem primer va determinat pels motors de recomanació, podem afirmar una cosa: aquestes tecnologies tenen relació directa amb el nostre consum i ideologia, i ambdues coses poden ser manipulades per aquells que creen els sistemes recomanadors.

Per exemple, una empresa pot fer que el motor de recomanació mostri recomanacions molt semblants a allò prèviament consumit per l'usuari, fent que es produueixi efecte “Echo Chamber” i que, consegüentment, es polaritzin i radicalitzin opinions. O també, com va fer Booking, pot fer que certs productes se’ns mostrin abans que altres per acord econòmic, fent que aquells que acabaran consumint els usuaris estiguin ja preestablerts.

Per últim, mitjançant l'apartat dels motors de recomanació a la societat s'ha donat una visió holística de què suposen aquests algorismes a la nostra vida quotidiana, així com també de la seva traducció en fenòmens que descriuen la nostra era, com la immediatesa, la sobrecàrrega d'estímuls, o l'esmentat “Echo Chamber”.

D'aquesta manera hem pogut tocar de peus a terra i veure que els motors de recomanació no són només un grapat de codi. Hem vist com uns “simples” algorismes, efectivament, tenen molt a veure amb les tendències de la nostra societat, i això no és quelcom per prendre's a la lleugera.

Per tant, mitjançant tots aquests punts que han anat progressivament afirmant la hipòtesi i acomplint els diferents objectius, arribem a una reflexió que ho tanca.

Els motors de recomanació, com a hores d'ara ja no podem negar, són estris gairebé imprescindibles dins del nostre dia a dia. Són una peça angular dins de molts dels serveis que usem en línia, i com ja hem esmentat, donen un ordre a aquest abocador d'informació que és internet. Són indispensables, i no només pel seu ús rutinari, sinó per la importància que tenen dins d'aquest gran engranatge que és el consumisme, ja que el benefici de moltes empreses es veuria afectat en la seva absència.

No obstant, en un món on s'ha desvirtuat l'ètica i on el valor humà està en decadència, és bastant il·lusori pensar que aquests estris només siguin utilitzats amb aquesta finalitat. Tenint en compte que el sistema socioeconòmic en què vivim pren el capital com a centre de tot, els motors de recomanació acaben sent una simple víctima, i es converteixen en un estri de fer diners, i malauradament, manipular.

A més a més, hem observat que aquests estris ens fan tendir a un conformisme que atenua encara més aquest sentiment de canvi, fent que ens endinsem en un cercle viciós que mai acaba. I si segueix tot igual, com ara exemplificarem, mai acabarà. Com hem vist, els motors de recomanació són conseqüència del sistema socioeconòmic actual; un sistema

que, entre altres coses, es manté per una necessitat d'immediatesa que ens predisposa a ser superficials. Però a més, aquesta necessitat d'immediatesa es nodreix dels mateixos sistemes de recomanació, ja que el contingut que mostren ens genera pics de dopamina que ens adhereixen encara més al producte que consumim i a la seva consegüent gratificació instantània. Com podem veure, tot és un peix que es mossega la cua.

Tot i això, no ens hem d'oblidar de la dualitat d'aquests motors, com passa amb totes les coses. La veritat és que, en un context social ideal, un motor de recomanació podria usar les dades mèdiques de tots els humans i, per exemple, crear una base de dades clínica immensa amb l'objectiu de predir el tractament més adient per a cada persona. D'aquesta manera, els avenços de la medicina personalitzada serien gegants.

És precisament aquí on ens agradarà introduir una proposta futura que, encara que ambiciosa, no està lluny dels avenços tecnològics que veiem dia rere dia. A més a més, ens presenta l'esmentada dualitat que tenen aquests estris. Bàsicament, mitjançant un motor de recomanació de filtratge col·laboratiu, seria possible crear una base de dades que prengui l'historial clínic de totes les persones d'x circumscriptió. Després, establint connexions entre pacients similars es podrien determinar les tendències patològiques de cada un d'aquests, i posteriorment, decidir un medicament o tractament adient segons l'experiència d'altres pacients.

Com hem vist amb el potencial que presenten els motors de recomanació, és poc adient culpar-los de la situació en què es troba la societat, però tampoc s'ha de negar que són un dels engranatges que mouen el bucle. El bucle cap a la inhumanitat.

Precisament per això, dins d'una societat on se'ns predisposa a un cercle vicios on l'essència humana cada vegada té menys cabuda, cal fer que aquestes eines de manipulació es converteixin en eines de cohesió i ajuda humana. Cal fer que aquest conformisme es converteixi en un anhel de canvi.

Cada vegada tenim més informació al nostre abast, i fins i tot les tesis i bases del sistema que, com hem vist, ens controla, estan a un sol clic. Encara així, en seguim sent víctimes, i la manipulació és una realitat massificada i normalitzada. Som realment lliures o la llibertat ens empresona?

Motors de recomanació, en efecte, una arma de doble fil.

5. ANNEXOS

ANNEX 1 - Origen dels algorismes computacionals

Al segle XIX, el matemàtic Charles Babbage va dissenyar el primer ordinador primigeni, una màquina analítica dissenyada específicament per construir taules de logaritmes i funcions trigonomètriques evaluant polinomis per aproximació.⁵³

Aquesta tenia dispositius d'entrada basats en les targetes perforades de Jacquard, un processador aritmètic que calculava nombres, una unitat de control que determinava quina tasca realitzar, un mecanisme de sortida i una memòria on els nombres podien ser emmagatzemats fins ser processats. Malauradament, aquesta màquina mai va poder ser creada per raons d'índole política, ja que hi va haver detractors per un possible ús bèlic de la màquina.⁵⁴

Diagram for the computation by the Engine of the Numbers of Bernoulli. See Note G. (page 729 of rev.)												Result Variables.
Statement of Results.	Working Variables.											Result Variables.
	W ₁	W ₂	W ₃	W ₄	W ₅	W ₆	W ₇	W ₈	W ₉	W ₁₀	W ₁₁	
1 X Y ₁ × Y ₂ Y ₃ Y ₄ Y ₅ Y ₆	Y ₁ Y ₂ Y ₃ Y ₄ Y ₅ Y ₆	—	—	—	—	—	—	—	—	—	—	—
2 — Y ₁ - Y ₂ Y ₃ —	Y ₁ Y ₂ Y ₃ Y ₄ Y ₅ Y ₆	—	—	—	—	—	—	—	—	—	—	—
3 — Y ₁ + Y ₂ Y ₃ —	Y ₁ Y ₂ Y ₃ Y ₄ Y ₅ Y ₆	—	—	—	—	—	—	—	—	—	—	—
4 — Y ₁ × Y ₂ Y ₃ —	Y ₁ Y ₂ Y ₃ Y ₄ Y ₅ Y ₆	—	—	—	—	—	—	—	—	—	—	—
5 — Y ₁ + a Y ₂ Y ₃ —	Y ₁ Y ₂ Y ₃ Y ₄ Y ₅ Y ₆	—	—	—	—	—	—	—	—	—	—	—
6 — Y ₁ × Y ₂ Y ₃ Y ₄ —	Y ₁ Y ₂ Y ₃ Y ₄ Y ₅ Y ₆	—	—	—	—	—	—	—	—	—	—	—
7 — Y ₁ + Y ₂ Y ₃ Y ₄ —	Y ₁ Y ₂ Y ₃ Y ₄ Y ₅ Y ₆	—	—	—	—	—	—	—	—	—	—	—
8 — Y ₁ × Y ₂ Y ₃ Y ₄ Y ₅ —	Y ₁ Y ₂ Y ₃ Y ₄ Y ₅ Y ₆	—	—	—	—	—	—	—	—	—	—	—
9 — Y ₁ + Y ₂ Y ₃ Y ₄ Y ₅ —	Y ₁ Y ₂ Y ₃ Y ₄ Y ₅ Y ₆	—	—	—	—	—	—	—	—	—	—	—
10 — Y ₁ × Y ₂ Y ₃ Y ₄ Y ₅ Y ₆	Y ₁ Y ₂ Y ₃ Y ₄ Y ₅ Y ₆	—	—	—	—	—	—	—	—	—	—	—
11 — Y ₁ + a Y ₂ Y ₃ Y ₄ Y ₅ —	Y ₁ Y ₂ Y ₃ Y ₄ Y ₅ Y ₆	—	—	—	—	—	—	—	—	—	—	—
12 — Y ₁ × Y ₂ Y ₃ Y ₄ Y ₅ —	Y ₁ Y ₂ Y ₃ Y ₄ Y ₅ Y ₆	—	—	—	—	—	—	—	—	—	—	—
13 — Y ₁ + Y ₂ Y ₃ Y ₄ Y ₅ —	Y ₁ Y ₂ Y ₃ Y ₄ Y ₅ Y ₆	—	—	—	—	—	—	—	—	—	—	—
14 — Y ₁ × Y ₂ Y ₃ Y ₄ Y ₅ Y ₆	Y ₁ Y ₂ Y ₃ Y ₄ Y ₅ Y ₆	—	—	—	—	—	—	—	—	—	—	—
15 — Y ₁ + a Y ₂ Y ₃ Y ₄ Y ₅ Y ₆	Y ₁ Y ₂ Y ₃ Y ₄ Y ₅ Y ₆	—	—	—	—	—	—	—	—	—	—	—
16 — Y ₁ × Y ₂ Y ₃ Y ₄ Y ₅ —	Y ₁ Y ₂ Y ₃ Y ₄ Y ₅ Y ₆	—	—	—	—	—	—	—	—	—	—	—
17 — Y ₁ + Y ₂ Y ₃ Y ₄ Y ₅ —	Y ₁ Y ₂ Y ₃ Y ₄ Y ₅ Y ₆	—	—	—	—	—	—	—	—	—	—	—
18 — Y ₁ × Y ₂ Y ₃ Y ₄ Y ₅ Y ₆	Y ₁ Y ₂ Y ₃ Y ₄ Y ₅ Y ₆	—	—	—	—	—	—	—	—	—	—	—
19 — Y ₁ + a Y ₂ Y ₃ Y ₄ Y ₅ Y ₆	Y ₁ Y ₂ Y ₃ Y ₄ Y ₅ Y ₆	—	—	—	—	—	—	—	—	—	—	—
20 — Y ₁ × Y ₂ Y ₃ Y ₄ Y ₅ —	Y ₁ Y ₂ Y ₃ Y ₄ Y ₅ Y ₆	—	—	—	—	—	—	—	—	—	—	—
21 — Y ₁ + Y ₂ Y ₃ Y ₄ Y ₅ —	Y ₁ Y ₂ Y ₃ Y ₄ Y ₅ Y ₆	—	—	—	—	—	—	—	—	—	—	—
22 — Y ₁ × Y ₂ Y ₃ Y ₄ Y ₅ Y ₆	Y ₁ Y ₂ Y ₃ Y ₄ Y ₅ Y ₆	—	—	—	—	—	—	—	—	—	—	—
23 — Y ₁ + a Y ₂ Y ₃ Y ₄ Y ₅ Y ₆	Y ₁ Y ₂ Y ₃ Y ₄ Y ₅ Y ₆	—	—	—	—	—	—	—	—	—	—	—
24 — Y ₁ × Y ₂ Y ₃ Y ₄ Y ₅ —	Y ₁ Y ₂ Y ₃ Y ₄ Y ₅ Y ₆	—	—	—	—	—	—	—	—	—	—	—
25 — Y ₁ + Y ₂ Y ₃ Y ₄ Y ₅ —	Y ₁ Y ₂ Y ₃ Y ₄ Y ₅ Y ₆	—	—	—	—	—	—	—	—	—	—	—
Here follows a sequence of operations thrown to memory.												—
26 — Y ₁ × Y ₂ Y ₃ Y ₄ Y ₅ Y ₆	Y ₁ Y ₂ Y ₃ Y ₄ Y ₅ Y ₆	—	—	—	—	—	—	—	—	—	—	—
27 — Y ₁ + a Y ₂ Y ₃ Y ₄ Y ₅ Y ₆	Y ₁ Y ₂ Y ₃ Y ₄ Y ₅ Y ₆	—	—	—	—	—	—	—	—	—	—	—
28 — Y ₁ × Y ₂ Y ₃ Y ₄ Y ₅ —	Y ₁ Y ₂ Y ₃ Y ₄ Y ₅ Y ₆	—	—	—	—	—	—	—	—	—	—	—
29 — Y ₁ + Y ₂ Y ₃ Y ₄ Y ₅ —	Y ₁ Y ₂ Y ₃ Y ₄ Y ₅ Y ₆	—	—	—	—	—	—	—	—	—	—	—
30 — Y ₁ × Y ₂ Y ₃ Y ₄ Y ₅ Y ₆	Y ₁ Y ₂ Y ₃ Y ₄ Y ₅ Y ₆	—	—	—	—	—	—	—	—	—	—	—
31 — Y ₁ + a Y ₂ Y ₃ Y ₄ Y ₅ Y ₆	Y ₁ Y ₂ Y ₃ Y ₄ Y ₅ Y ₆	—	—	—	—	—	—	—	—	—	—	—
32 — Y ₁ × Y ₂ Y ₃ Y ₄ Y ₅ —	Y ₁ Y ₂ Y ₃ Y ₄ Y ₅ Y ₆	—	—	—	—	—	—	—	—	—	—	—
33 — Y ₁ + Y ₂ Y ₃ Y ₄ Y ₅ —	Y ₁ Y ₂ Y ₃ Y ₄ Y ₅ Y ₆	—	—	—	—	—	—	—	—	—	—	—
34 — Y ₁ × Y ₂ Y ₃ Y ₄ Y ₅ Y ₆	Y ₁ Y ₂ Y ₃ Y ₄ Y ₅ Y ₆	—	—	—	—	—	—	—	—	—	—	—
35 — Y ₁ + a Y ₂ Y ₃ Y ₄ Y ₅ Y ₆	Y ₁ Y ₂ Y ₃ Y ₄ Y ₅ Y ₆	—	—	—	—	—	—	—	—	—	—	—
36 — Y ₁ × Y ₂ Y ₃ Y ₄ Y ₅ —	Y ₁ Y ₂ Y ₃ Y ₄ Y ₅ Y ₆	—	—	—	—	—	—	—	—	—	—	—
37 — Y ₁ + Y ₂ Y ₃ Y ₄ Y ₅ —	Y ₁ Y ₂ Y ₃ Y ₄ Y ₅ Y ₆	—	—	—	—	—	—	—	—	—	—	—
38 — Y ₁ × Y ₂ Y ₃ Y ₄ Y ₅ Y ₆	Y ₁ Y ₂ Y ₃ Y ₄ Y ₅ Y ₆	—	—	—	—	—	—	—	—	—	—	—
39 — Y ₁ + a Y ₂ Y ₃ Y ₄ Y ₅ Y ₆	Y ₁ Y ₂ Y ₃ Y ₄ Y ₅ Y ₆	—	—	—	—	—	—	—	—	—	—	—
40 — Y ₁ × Y ₂ Y ₃ Y ₄ Y ₅ —	Y ₁ Y ₂ Y ₃ Y ₄ Y ₅ Y ₆	—	—	—	—	—	—	—	—	—	—	—
41 — Y ₁ + Y ₂ Y ₃ Y ₄ Y ₅ —	Y ₁ Y ₂ Y ₃ Y ₄ Y ₅ Y ₆	—	—	—	—	—	—	—	—	—	—	—
42 — Y ₁ × Y ₂ Y ₃ Y ₄ Y ₅ Y ₆	Y ₁ Y ₂ Y ₃ Y ₄ Y ₅ Y ₆	—	—	—	—	—	—	—	—	—	—	—
43 — Y ₁ + a Y ₂ Y ₃ Y ₄ Y ₅ Y ₆	Y ₁ Y ₂ Y ₃ Y ₄ Y ₅ Y ₆	—	—	—	—	—	—	—	—	—	—	—
44 — Y ₁ × Y ₂ Y ₃ Y ₄ Y ₅ —	Y ₁ Y ₂ Y ₃ Y ₄ Y ₅ Y ₆	—	—	—	—	—	—	—	—	—	—	—
45 — Y ₁ + Y ₂ Y ₃ Y ₄ Y ₅ —	Y ₁ Y ₂ Y ₃ Y ₄ Y ₅ Y ₆	—	—	—	—	—	—	—	—	—	—	—

Il·lustració 43: Els nombres de Bernoulli

Font: Alamy

patrons algebraics, les que varen ser essencials per programar els primers ordinadors a mitjans del segle XX. A les seves notes també va anticipar conceptes com el de la

Motors de Recomanació, Una Arma de Doble Fil

Matías Mora

⁵³ (Bromley & Bromley, 1982)

⁵⁴ (Máquina Diferencial – Maquina Analítica / IDIS, n.d.)

subrutina⁵⁵ o el bucle⁵⁶ informàtic, i és una de les raons per la que se la considera la mare de la programació computacional.⁵⁷

El que la va portar a la història, però, va ser com va agafar com a referència els nombres de Bernoulli⁵⁸ per desenvolupar allò que ara coneixem com algorisme informàtic. Tot i que mai va poder materialitzar els coneixements anotats (300 dibuixos i 2200 pàgines descriptives) perquè la màquina analítica no va ser acabada, van servir d'inspiració per milers de programadors posteriors.⁵⁹ Actualment és un referent per milers de persones arreu del món, i la seva valentia i resiliència per superar les normes socials d'aquella època anteposant-se a les idees que tots els altres podien aconseguir d'una via més fàcil és admirable.

⁵⁵ En computació, una **subrutina** o **subprograma**, com a idea general, es presenta com un subalgorisme que forma part de l'algorisme principal, el qual permet resoldre una tasca específica.

Font: Wikipedia

⁵⁶ Un **bucle** o **cicle**, en programació, és una sentència que es realitza diverses vegades a una secció aïllada del codi, fins que la condició assignada al bucle deixa de complir-se.

Font: Wikipedia

⁵⁷ (*Ada Lovelace y El Algoritmo Informático* - Agencia B12, n.d.)

⁵⁸ En matemàtiques, els **Nombres de Bernoulli**, són una seqüència de nombres racionals amb connexions profundes amb la teoria de nombres. Els nombres de Bernoulli apareixen a l'expansió en sèrie de Taylor de les funcions tangent i tangent hiperbòlica, en les fórmules per la suma de potències dels primers nombres naturals, a la fórmula d'Euler–Maclaurin i a l'expressió de certs valors de la funció zeta de Riemann.

Font: Wikipedia

⁵⁹ (*Ada Lovelace, Pionera de La Informática* - Fundación Aquae, n.d.).

ANNEX 2 – Origen dels Ordinadors

Per això, ens hem de traslladar uns quants segles enrere, on alguns teoremes matemàtics i concepcions filosòfiques van establir la base d'aquest concepte.

Aristòtil va ser el primer a descobrir de manera estructurada un conjunt de regles, els anomenats sil·logismes⁶⁰, que descriuen una part del funcionament de la ment humana i que, en seguir-les pas a pas, produueixen conclusions racionals a partir de les premisses.

El 250 aC, Ctesibi d'Alexandria va construir la primera màquina autocontrolada. Es tractava d'un regulador de l'aigua que en modificava el funcionament, però sense un raonament previ.

Fent un salt a la història, el 1637, un dels filòsofs més influents del segle XVII va predir la possibilitat de crear màquines que pensessin per si mateixes. Aquesta figura era René Descartes, encara que Alan Turing, anys després és el que presenta la idea com a viable.

⁶¹

$$P(A|B) = \frac{P(B|A) * P(A)}{P(B)}$$

Il·lustració 44: Representació matemàtica del Teorema de Bayes

Font: Wikipedia

El teorema de Bayes, al 1763, va definir la probabilitat de que un esdeveniment passi basant-se en el coneixement de les condicions prèvies que es relacionen amb l'esdeveniment en qüestió.⁶²

⁶⁰ El sil·logisme és un mètode lògic creat per Aristòtil, a través del qual s'obté una conclusió mitjançant dues premisses: premissa major, que inclou el seu predicat (P), i premissa menor, que inclou el seu subjecte (S). És possible que les premisses siguin ambdues veritables o bé que una sigui veritable i l'altra falsa (mai no poden ser falses totes dues), però la conclusió ha de ser necessàriament veritable o falsa.

Font: Wikipedia

⁶¹ (*Origen Del Concepto de Inteligencia Artificial*, n.d.)

⁶² (*Te Contamos Qué Es El “machine Learning” y Cómo Funciona*, n.d.)

Si ens regim a una definició més matemàtica, diem que aquest teorema expressa la probabilitat condicionada⁶³ d'un esdeveniment aleatori **A** donat **B** en termes de la distribució de probabilitat condicionada de l'esdeveniment **B** donat **A** i la distribució de probabilitat marginal⁶⁴ de només **A**. És a dir, per exemple, ens diu que sabent la probabilitat de tenir un mal de cap atès que es té grip, es podria saber la probabilitat de tenir grip si es té un mal de cap tenint en compte la probabilitat de tenir grip sense considerar cap factor (com podria ser el mal de cap).⁶⁵

Després d'anys d'aturada en aquest sentit, el 1847, el matemàtic George Boole va posar un component més a aquesta història, establint que el raonament lògic podria sistematitzar-se, de la mateixa manera que es resol una equació matemàtica.

Com hem vist, juntament amb molts altres teoremes i raonaments matemàtics i filosòfics, fa anys que es va establir la idea primitiva del que després nosaltres coneixeríem com Machine learning, però no va ser fins la invenció de l'ordinador electrònic que es va realment aproximar el potencial d'aquesta idea.

És difícil definir què separa un ordinador del que no és un ordinador, però hi ha molts pensadors que també es van fer la mateixa pregunta. Si fem servir la definició d'Alan

⁶³ La **probabilitat condicionada** és la probabilitat que hi hagi un succès *A*, sabent que també succeeix un altre succès *B*. La probabilitat condicional s'escriu $P(A|B)$, i es llegeix «la probabilitat de *A* donada *B*».

Font: Wikipedia

⁶⁴ Dins la teoria de probabilitats, donades duesvariables aleatòries juntes *A* & *B*, la **distribució marginal** de *A* és simplement la distribució de probabilitat de *A* fent cas omís de la informació referent a *B*.

Font: AulaFacil

⁶⁵ (*Teorema de Bayes - Wikipedia, La Enciclopedia Libre*, n.d.)

Turing i la seva màquina universal de Turing, un ordinador seria una màquina programable, de propòsit general i completament automàtica.⁶⁶

Anteriorment els ordinadors eren artefactes que mecanitzaven un càlcul, però gràcies a la proposta d'Alan Turing l'ordinador va passar a ser un computador de programa emmagatzemat; és a dir, que a la memòria no només emmagatzemava dades, sinó també els programes a executar.



Il·lustració 45: La màquina ENIGMA

Font: *Wikipedia*

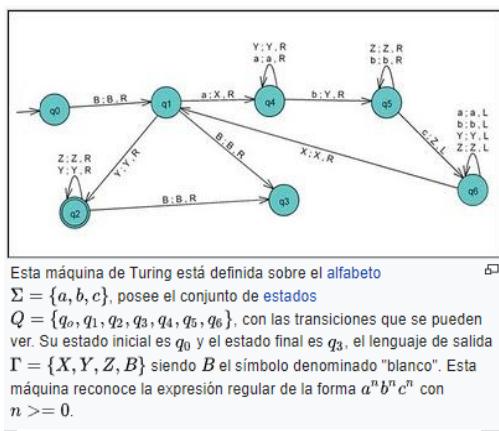
Alan Turing va ser un matemàtic britànic, que a part de la seva cabdal contribució a la computació que posteriorment explicaré, va ser decisiu per la caiguda de l'exèrcit nazi i per definir el concepte de intel·ligència artificial, entre altres coses. La Segona Guerra Mundial va oferir un marc insospitat d'aplicació pràctica de les seves teories

matemàtico-informàtiques, en sorgir la necessitat de desxifrar els missatges codificats que la Marina alemanya feia servir per enviar instruccions als submarins que fustigaven els combois d'ajuda material enviats des dels Estats Units. Turing, al comandament d'una divisió de la intel·ligència britànica, va dissenyar tant els processos com les màquines que, capaços d'efectuar càlculs combinatoris molt més ràpidament que qualsevol ésser humà, van ser decisius en la ruptura final del codi. Entre elles podem trobar la més famosa; Bombe, que va desxifrar Enigma.

A més a més, Turing va definir un mètode teòric per deidir si una màquina és capaç de pensar com un home analitzant una “conversació” textual entre aquests dos, la qual ha de ser avaluada per un interrogador. No comença això a apropar-se a la nostra realitat contemporània? Tenint en compte que el 60% de les persones han interactuat amb un xatbot en els darrers 12 mesos (segons dades de GetVoip) i que els mateixos gestionen el 60% de les consultes d'atenció al client, Turing fa 70 anys no anava mal encaminat.

⁶⁶ (*Máquina de Turing Universal - Wikipedia, La Enciclopedia Libre*, n.d.; Rogozhin & Rogozhin, 1996a)

El que el va fer passar a la història com a pare de la computació moderna, però, és el concepte que anteriorment he mencionat; la màquina universal de Turing, ideada el 1937. Molts personatges fonamentals de la computació moderna com John von Neumann la van utilitzar posteriorment, ja que aquesta màquina va ser la primera en ser considerada un computador de programa emmagatzemat⁶⁷. Gràcies a això, von Neumann va ser capaç de crear la coneguda arquitectura de von Neumann⁶⁸, la qual comparteixen els ordinadors actuals. Per a prendre consciència de la magnitud d'aquest concepte ideat pel matemàtic britànic, hem de saber que de fet és el model computacional pel qual es basen tots els ordinadors que ara coneixem.⁶⁹



Il·lustració 46: Esquema explícit de la màquina de Turing

Font: Wikipedia

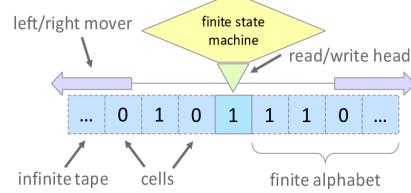
⁶⁷ Un ordinador de programa emmagatzemat és el que emmagatzema les instruccions del programa a la memòria electrònica. (“Computador De Programa Almacenado.” Wikipedia, Wikimedia Foundation, 20 Apr. 2021, https://es.wikipedia.org/wiki/Computador_de_programa_almacenado.)

⁶⁸ L'arquitectura de von Neumann és un disseny que utilitza una memòria per emmagatzemar instruccions i dades. Aquest tipus de màquina és la implementació d'una màquina de Turing i la visió d'una arquitectura seqüencial en lloc de paral·lela. (*Modelo De Von Neumann - Universidad Veracruzana*. https://www.uv.mx/rrojano/arquitectura_uno/clase4-1.pdf.)

⁶⁹ (*Máquinas de Turing. ¿Qué Es Una Máquina de Turing? / by Bootcamp AI / Medium*, n.d.)

Bàsicament, una màquina de Turing és una màquina hipotètica (un estudi teòric) destinada a simular qualsevol algorisme informàtic, sense importar-ne la complexitat.

Aquesta consta d'una cinta dividida en caselles on s'emmagatzema la memòria, on es poden escriure símbols com poden ser 1 o 0. A banda d'això també es compon d'un capçal capaç de moure's per la cinta d'esquerra a dreta i escriure símbols en cada casella. En últim terme, la constitueix un programa que digui al capçal què és el que ha de fer, el qual pot estar escrit a la cinta en forma de 1 i 0.



Il·lustració 47: Procés Màquina de Turing

Font: ResearchGate

Per continuar amb la nostra exploració de la Màquina de Turing convindrà que associem el concepte amb la idea de funció matemàtica. Una funció és una regla que associa un parell de valors (x, y). A x l'anomenem argument⁷⁰, i ens ofereix un i només un valor y per a cada valor de x . De la mateixa manera, una Màquina de Turing particular ofereix un i només un output quan introduïm un input. El resultat dependrà tant de l'input introduït com de les instruccions definides per la màquina, convertint-la en una funció matemàticament parlant.

Traslladant-ho a la mateixa màquina, anomenem T a una Màquina de Turing arbitrària, i m a la cinta. Simbolitzarem l'acció de la màquina sobre aquesta cinta com la funció $T(m)$. Aquesta acció ens tornarà un resultat p . $T(m) = p$, així que com en aquest cas ens torna un output, diem que $T(m)$ és equivalent a una funció computable⁷¹.

⁷⁰ En matemàtiques, l'argument d'una funció és un valor proporcionat per obtenir el resultat de la funció. Per exemple, la funció binària: $f(x, y) = x^2 + y^2$ té dos arguments, (x i y)

⁷¹ Les funcions computables són els objectes bàsics d'estudi en la teoria de la computabilitat.

Les funcions computables són l'anàleg formalitzat de la noció intuïtiva d'algorismes , en el sentit que una funció és computable si hi ha un algorisme que pot fer la feina de la funció, és a dir, donada una entrada del domini de funció, pot tornar la sortida corresponent.

Font: Hmong

No content amb això, just després de teoritzar la màquina de Turing, ell mateix va albirar el concepte tergiversador de la màquina universal de Turing. La diferència entre ambdues rau en que la màquina universal va establir les bases dels ordinadors moderns gràcies a la seva capacitat de dur a terme múltiples processos i d'executar la funció de qualsevol màquina de Turing. La màquina universal essencialment aconsegueix això mitjançant la lectura de tant la descripció de la màquina que volem que sigui simulada com també l'entrada de la seva pròpia cinta.⁷² En paraules més fàcils, una màquina universal de Turing és la que pren com a programa una màquina de Turing com a tal, és a dir, que pot emular el comportament de qualsevol màquina de Turing. Amb aquesta màquina, Alan Turing concep teòricament la manera de fer que es puguin resoldre tots els problemes matemàtics que puguin expressar-se mitjançant un algorisme, i per tant que poden ser resoltes per una màquina de Turing.⁷³

Per exemplificar-ho, imaginem que l'entrada que li fiquem a la màquina és “2+2”. Dos símbols de la cinta representen aquests números, i un altre símbol indica l'operador de suma. El funcionament de la màquina fa que per cada símbol llegit la reacció del capçal sigui escriure o moure's. La combinació de símbols per realitzar aquesta suma donarà com a resultat un nou símbol que serà la definició de la suma d'aquests dos nombres.

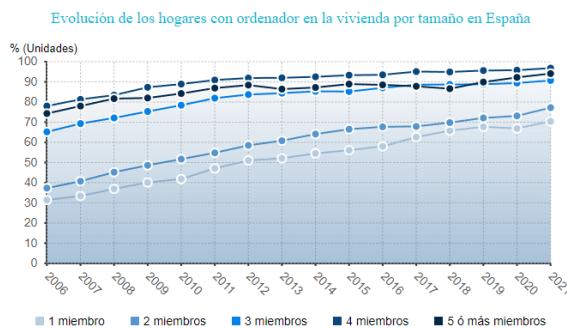
Aquest procés en un context purament teòric (com bé és la idea primigènia de la màquina universal de Turing) pot resoldre, com hem esmentat abans, qualsevol tipus d'operació matemàtica, i precisament per això se l'anomena la base dels processos computacionals. És important entendre això, ja que Turing, amb la seva teoria, estableix també que hi ha tasques que no són computables, ja que no poden ser representades amb algorismes (recordem la seva definició), i per tant cap màquina de Turing (això també inclou als ordinadors actuals) mai les podrà fer.

Alan Turing, tot i tenir una gran adversitat en contra seva com bé és ser homosexual a la seva època, ha estat una figura fonamental pel desenvolupament d'una tecnologia que molts pensem que no podríem viure sense ella. Gràcies a avenços que parteixen del seu

⁷² (Herken, 1995; Rogozhin & Rogozhin, 1996b)

⁷³ (Arora & Boaz, 2009; *La Máquina Universal de Turing • Un Artículo de La Máquina Oráculo*, n.d.).

teorema, actualment un 81,4% de les llars espanyoles amb persones d'entre 16 i 74 anys vivint-hi té un dispositiu electrònic, i l'estadística va creixent. Que el seu model computacional encara sigui axiomàtic en el desenvolupament d'ordinadors i altres dispositius fins 100 anys després de la seva creació és simplement admirable.



Il·lustració 48: Evolució de llars amb ordinador a Espanya

Font: Epdata

Ara, tenint com a premissa el que podem definir com a ordinador electrònic, tractarem de determinar quin dels diversos candidats en disputa va tenir l'honor de ser el primer ordinador de la història.

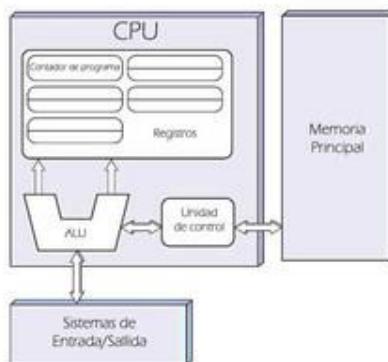
Per això, hem de conèixer a aquell qui va donar sentit a l'estructura dels ordinadors (no al seu procés computacional). Aquest va ser l'esmentat anteriorment Jon von Neumann, amb la seva encara present arquitectura de von Neumann, originada el 1947. Sigui quin sigui l'ordinador que utilitzes, consola de videojocs, telèfon mòbil o fins i tot SmartTV, disposa el hardware organitzat de la següent manera.

En primer lloc la CPU (unitat central de processament). És el circuit digital que s'encarrega de fer les instruccions d'un programa. Se'l denomina també processador. Aquesta conté l'ALU (realització d'operacions aritmètiques i lògiques), la unitat de control (control de l'ALU, la memòria i els dispositius entrada/sortida, moure dades i programes des de i cap a la memòria) i un conjunt de registres (àrees d'emmagatzematge d'alta velocitat pel qual totes les dades passen abans de processar-se).

En segon lloc trobem la memòria, la qual pot controlar dades, així com el programa que les utilitza (instruccions). En els ordinadors moderns aquesta memòria és la RAM, i és ràpida i accessible directament per la CPU.

En tercer lloc trobem l'entrada i sortida, composada pels elements que s'encarreguen de prendre dades com a entrada (com pot ser un teclat), mostrar-los en alguna sortida (com pot ser un monitor) o comunicar-se amb altres sistemes.

Per últim, trobem els busos, els quals són les vies de comunicació que permeten moure la informació entre els elements de l'arquitectura. N'hi ha de dades (transport de dades), de direccions (indicació de l'origen i destí de dades) i de control (senyals per coordinar).⁷⁴



Il·lustració 49: Arquitectura de Von Neumann
 Font: Wikipedia

Aquesta és una explicació molt ràpida de com es compon l'estructura dels ordinadors moderns, ja que no és el tema principal d'aquest treball, però una aproximació era necessària per prendre contacte amb l'immens món computacional, i així entendre el seu origen.

Durant la Segona Guerra Mundial, els artillers de l'exèrcit dels Estats Units apuntaven les seves armes usant taules de tir (dades), que contenien les trajectòries que podien seguir els projectils en funció del tipus d'arma, la zona geogràfica o la direcció i la velocitat del vent. Aquestes taules es confeccionaven manualment a la base militar d'Aberdeen Proving Grounds, a Pennsilvània. Cada taula contenia unes tres mil trajectòries i cada trajectòria

⁷⁴ (*Arquitectura de Von Newmann — Plone Site*, n.d.; *Arquitectura von Neumann: Origen, Modelo, Cómo Funciona*, n.d.; *John von Neumann, El Genio Detrás Del Ordenador Moderno*, n.d.)

necessitava uns 750 càlculs. A causa de la gran quantitat de càlculs que cal fer, l'exèrcit dels Estats Units va establir un conveni amb la Universitat de Pennsilvània, on ja treballaven amb versions primitives d'ordinadors, perquè els ajudessin a elaborar aquestes taules de tir d'artilleria.

Així, el 1943, es va iniciar la construcció del primer ordinador de propòsit general basat en circuits electrònics, l'ENIAC (acrònim d'Electronic Numerical Integrator And Computer), inventat per J. Presper Eckert i John Mauchly. (*¿Cuál Fue El Primer Ordenador de La Historia?*, n.d.)

L'ENIAC es va acabar de construir el 1945 i es va presentar al públic el 15 de febrer de 1946. Era una gegant bèstia de 27 tones de pes, que ocupava una superfície de 167 m² i que comptava amb 17.500 vàlvules de buit⁷⁵, 7.200 díodes⁷⁶ de vidre, 1.500 relés⁷⁷, 70.000 resistències, 10.000 condensadors i cinc milions de soldadures. Les vàlvules de buit de l'època no eren especialment duradores i, quan una es fonia, deixava l'ENIAC totalment inoperatiu. Com que la majoria d'aquestes errades passaven durant els encesos o apagats, es va prendre la decisió de no apagar-lo mai.

⁷⁵ Una vàlvula de buit és un component electrònic tancat dins d'un tub de vidre, utilitzat per amplificar, commutar, o modificar un senyal elèctric mitjançant el control del moviment dels electrons en un espai "buit" a molt baixa pressió, o en presència de gasos especialment seleccionats

Font: Wikipedia

⁷⁶ En electrònica un díode és un dispositiu electrònic no lineal i polaritzat format per dos èlectrodes actius. Un díode només deixa circular el corrent elèctric en un únic sentit i el bloqueja en el sentit contrari restringint el moviment dels electrons.

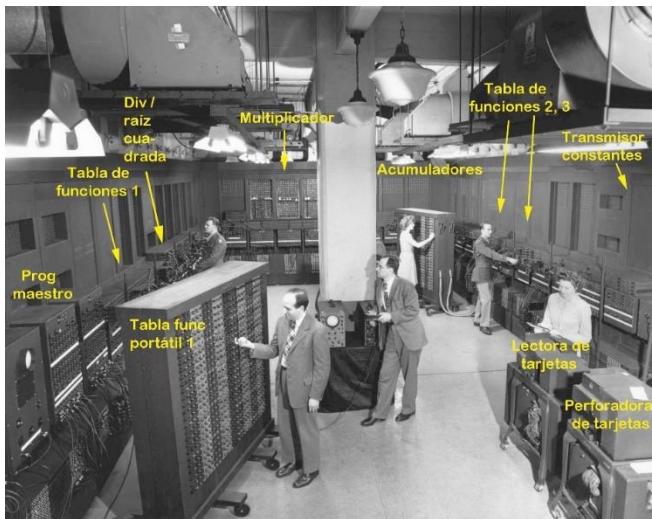
Font: Wikipedia

⁷⁷ Un relé és un dispositiu electromagnètic que permet modificar l'estat d'un commutador elèctric gràcies a l'electricitat.

Font: Wikipedia



Amb aquest senzill truc van aconseguir que les fallades es reduïssin a només una vàlvula cada dos dies, però a costa de provocar freqüents apagades a la ciutat de Filadèlfia, on estava ubicat l'ENIAC, ja que el seu consum era de 160.000 W.



Il·lustració 50: Estructura ENIAC

Font: Museu Informàtica UPV

L'ENIAC era programable per realitzar qualsevol tipus de càlcul numèric (sumes, restes, multiplicacions, divisions i arrels quadrades), arribant a fer 5.000 sumes i 300 multiplicacions per segon, cosa revolucionària per a l'època.

No tenia sistema operatiu, ni cap programa emmagatzemat, només emmagatzemava els números que feia servir en les seves operacions. Usava el sistema de numeració decimal, en lloc del sistema binari actual i podia arribar a calcular amb nombres de fins a 20 xifres.

En els 10 anys que va estar en funcionament, amb ell es van realitzar des dels càlculs esmentats per a les taules de tir d'artilleria, fins a complexos càlculs físics sobre la bomba d'hidrogen.⁷⁸

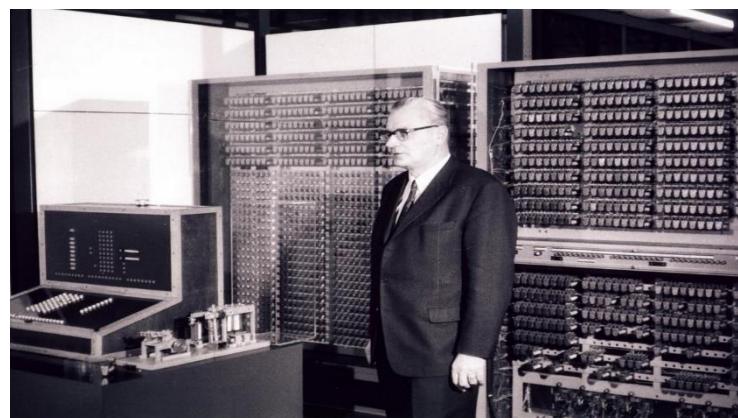
Hi ha altres interpretacions i concepcions sobre quin ordinador (prenent la definició dita) fou el primer, ja que com la motivació per crear-lo fou purament militar (com molts altres invents) no hi ha res a ciència certa, ja que hi havia moltes operacions encobertes. Per

⁷⁸ (*Computer - The First Computer / Britannica*, n.d.; *When Was the First Computer Invented?*, n.d.; de Solla Price, 1984; Leiner et al., 2009; O'Regan, n.d.)

exemple, podríem haver parlat de Mark I (acabat al 1944), el primer ordinador electromecànic d'EEUU ideat per Howard Aiken (basat en la màquina analítica de Babbage). També és acceptada la teoria de que el primer fou Colossus, creat per un grup de científics entre els que es trobaven Alan Turing o Thomas H. Flowers. Aquest era un projecte ultrasecret dirigit per Winston Churchill amb fins de desxifrar missatges nazis, i de fet va ser decisiu per la victòria dels aliats.

Per últim, segons els últims estudis el primer ordinador considerat de propòsit general (no militar ni criptogràfic), automàtic i completament programable que mai va existir va ser Z3, creat per Konrad Zuse, un enginyer alemany. Encara que la seva contribució a la història és menor a la de l'ENIAC, i per això tothom considera aquest últim el vertader pioner dels ordinadors moderns, si ens posem exactes hauríem de prendre en consideració aquest model creat per l'alemany.

Durant els estudis d'enginyeria, Zuse havia de fer molts càlculs rutinars a mà, cosa que li resultava molt avorrit. Això el va condir a pensar en una màquina que pogués fer aquests càlculs. El 1938 va crear el seu primer prototip, el Z1, el qual comptava amb tecnologies



Il·lustració 51: Konrad Zuse i el Z3

Font: ParcelaDigital

extraordinàriament avançades per l'epoca, ja que era una calculadora mecànica binària operada amb electricitat, encara que amb programabilitat limitada. Malauradament, aquest model mai va funcionar del tot bé atès a problemes mecànics.

Després de rebre el permís per treballar en els seus models computacionals enllloc de fer el servei militar, va crear el Z2, que va presentar a la mateixa casa dels seus pares, i disposava de millors com utilitzar relés telefònics enllloc de llàmines de metall. Amb els fons que li van donar va poder crear la seva empresa, Zuse KG.

D'aquesta manera, el 1941 es va concloure la construcció del Z3, una màquina programable i completament automàtica, és a dir, el primer ordinador construït totalment operatiu. El Z3 era bàsicament una calculadora binària basada en relés telefònics que podia realitzar les operacions de suma, resta, multiplicació, divisió i arrel quadrada, amb aritmètica en coma flotant purament binària. El Z3, de tecnologia electromecànica, pesava 1 tona. Estava construït amb 2.300 relés, tenia una freqüència de rellotge de 5,33 Hz, i una longitud de paraula de 22 bits. Permetia programació amb bucles, comptava amb memòria per a 64 paraules i amb una unitat de càlcul.

Perquè ens fem una idea del revolucionari de l'invent de Zuse només cal fer un breu repàs a algunes de les seves característiques:

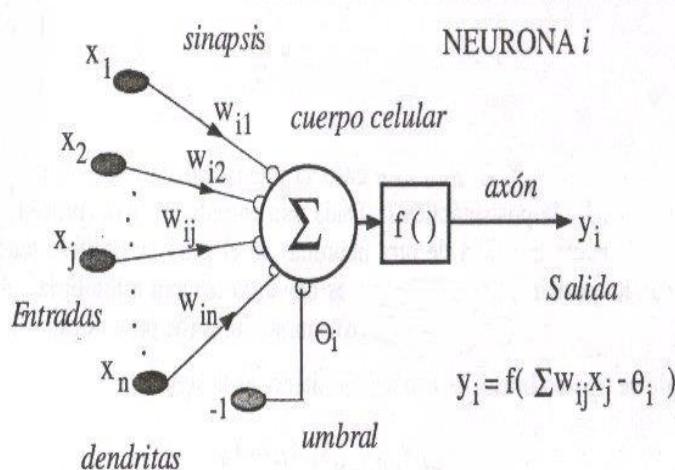
- 1- El primer disseny d'un ordinador programable va ser realitzat per Charles Babbage a mitjans del segle XIX, però no es va poder dur a terme en aquells moments. En part perquè era decimal, i per tant molt complicat, no pas binari i simple com el Z3.
- 2- Si Ada Lovelace va ser la primera programadora teòrica, escrivint programes per a una màquina que no existia, llavors Konrad Zuse va ser el primer programador pràctic.
- 3- L'ENIAC va ser completat 4 anys després que el Z3. Mentre que l'ENIAC feia servir vàlvules de buit, el Z3 feia servir relés. L'ENIAC encara era decimal i el Z3 ja era binari.
- 4- Fins al 1948 programar l'ENIAC significava tornar a soldar els cables. El Z3 llegia els programes de targetes perforades.
- 5- Avui dia els computadors estan basats en transistors en comptes de vàlvules o relés, però la seva arquitectura interna és més semblant al Z3 que a l'ENIAC.
- 6- El Z3 era una màquina universal de Turing. Era possible construir bucles al Z3, però no hi havia cap instrucció de salt condicional (encara que hagués estat senzill afegir-ne una). Tot i això, hi ha una manera d'implementar una màquina de Turing en un Z3 (assumint una capacitat d'emmagatzematge infinita), com va ser demostrat el 1998 per Raúl Rojas. És una manera estranya, però la pròpia màquina de Turing és estranya, estant dissenyada per ser simple i universal, no eficient.



Konrad Zuse i el seu treball van estar silenciats pels aliats atès al seu origen alemany, fins al punt de que el seu treball no és considerat per tothom l'ordinador electrònic primigeni. El seu taller va ser bombardejat durant la guerra, i les seves condicions inicials de treball no eren les millors, però la seva genialitat ens ha portat un invent que va marcar la diferència.

ANNEX 3 - Origen de la Intel·ligència Artificial

Anys després de la creació del primer ordinador, concretament a la dècada del 1950, diferents científics van començar a investigar com aplicar la biologia de les xarxes neuronals del cervell humà per intentar crear les primeres màquines intel·ligents. Això és gràcies a una idea concebuda per Alan Turing al veure el potencial de la seva màquina universal, i pel que va crear el seu esmentat test de Turing. La pregunta fou simple, però revolucionària: les màquines poden pensar? Ell va ser el primer científic en qüestionar-s'ho, i així ho va constar al seu article *Computing machinery and intelligence*. (Alan Turing, *El Padre de La Inteligencia Artificial* | Ministerio de Cultura, n.d.)

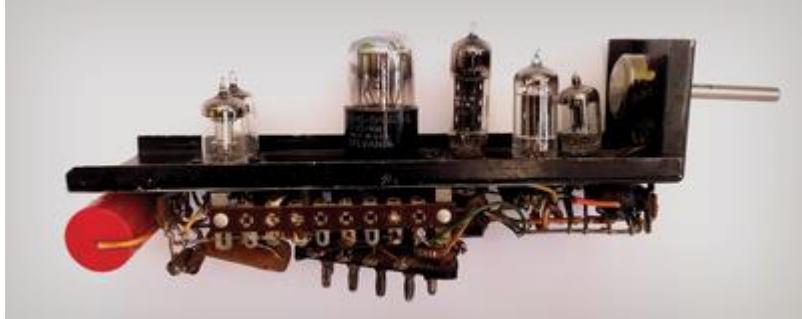


Il·lustració 52: Neurona McCulloch-Pitts

Font: Ibiblio

També va ser necessària l'aportació de Warren McCulloch i Walter Pitts al 1943, els quals van presentar el seu model de neurones artificials, i encara que es considera el primer treball del camp de la intel·ligència artificial encara no s'havia concebut la idea. Les fonts de la seva investigació varen ser els coneixements sobre la fisiologia bàsica i el

funcionament de les neurones, l'anàlisi formal de la lògica proposicional de Russell i Whitehead i la teoria de la computació de Turing.⁷⁹



Il·lustració 53: L'SNARC
Font: *The Scientist*

La idea va derivar en la creació de les xarxes neuronals artificials, un model computacional inspirat en la manera com les neurones transmeten la informació entre elles a

través d'una xarxa de nodes⁸⁰ interconnectats. Un dels primers experiments en aquest sentit el van fer Marvin Minsky i Dean Edmonds, científics de l'Institut Tecnològic de Massachusetts (MIT) al 1951. Tots dos van aconseguir crear un programa informàtic capaç d'aprendre de l'experiència per sortir d'un laberint. Aquesta neuro-computadora es va dir SNARC (Stochastic Neural Analog Reinforcement Calculator), i estava composada per 40 "neurones" informàtiques (fetes amb 3000 vàlvules de buit) que imitaven a una rata. A través d'unes llums es podia monitoritzar el moviment de les rates (composades per les neurones), i després de diversos intents casuals, les rates van començar a "pensar" sobre una base lògica gràcies a les eleccions que prèviament elegien, les quals emmagatzemaven a una memòria catalogant-se com bones o dolentes, com ara funciona una intel·ligència artificial.⁸¹

⁷⁹ (3.-*Historia de La IA - Inteligencia Artificial IA*, n.d.)

⁸⁰ Un node és un punt d'intersecció / connexió dins d'una xarxa. En un entorn on tots els dispositius són accessibles a través de la xarxa, aquests dispositius són tots considerats nodes.

Font: Wikipedia

⁸¹ (Crevier, 1993; *Marvin Minsky's SNARC, Possibly the First Artificial Self-Learning Machine : History of Information*, n.d.; *SNARC / History of AI*, n.d.; *Stochastic Neural Analog Reinforcement Calculator - Wikipedia*, n.d.; *This Week in The History of AI at AIWS.Net – Marvin*

Aquesta va ser la primera màquina capaç d'aprendre per si mateixa a resoldre una tasca sense haver estat programada de manera explícita, sinó que ho feia tan sols aprenent a partir dels exemples proporcionats inicialment. L'èxit va significar un canvi de paradigma respecte al concepte més ampli d'intel·ligència artificial.

D'aquesta manera va néixer el terme intel·ligència artificial com a tal, encara que se li va acabar de posar nom quan John McCarthy va definir la paraula com la ciència i enginyeria de “fer màquines intel·ligents”, al 1956.

Un altre pilar en la creació de xarxes neuronals artificials fou Frank Rosenblatt, un psicòleg estatunidenc que va crear la idea de la neurona artificial Perceptró. Aquesta és un algorisme capaç d'aprendre mitjançant un sistema de xarxa nerviosa que simula processos cerebrals, i és el primer model que estrictament segueix els principis neurològics que això requereix.⁸² Es considera un classificador binari (o discriminador lineal⁸³) el qual genera una predicción basant-se en un algorisme combinat amb el pes de les entrades.

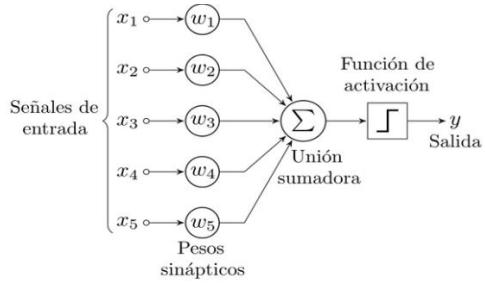
Bàsicament, el seu funcionament es basa en operacions de multiplicació entre dos components importants, comportant-se com una funció matemàtica: les entrades de dades (input) representades per x , i el pes d'aquestes (la seva rellevància en el càcul de sortida) representat w . La suma d'aquesta multiplicació es transmet a una funció d'activació, que determina un valor binari de 0 o 1, que determina si s'activa o no la neurona. (*Perceptron : Qu'est-Ce Que c'est et à Quoi Ça Sert ?, n.d.*) Això permet classificar les dades, i aquest és el seu gràfic (semblant a la neurona McCulloch-Pitts, ja que està inspirat en ella).

Minsky and Dean Edmonds Built SNARC, the First Artificial Neural Network / AIWS.Net, n.d.; Russell et al., n.d.)

⁸² . (*The History of Neural Networks and AI: Part I - Open Data Science*, n.d.)

⁸³ L'anàlisi Discriminant Lineal és una generalització del discriminant lineal de Fisher, un mètode utilitzat en estadística, reconeixement de patrons i aprenentatge de màquines per trobar una combinació lineal de trets que caracteritzen o separen dues o més classes d'objectes o esdeveniments.

Font: Wikipedia



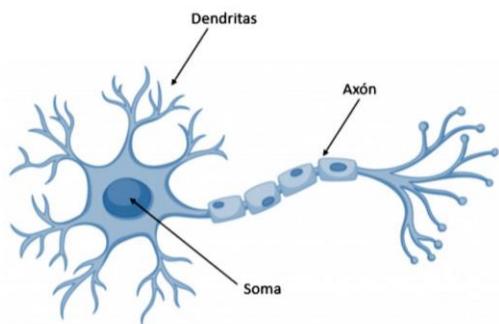
Il·lustració 54: El Perceptró

Font: [Wikipedia](#)

El Perceptró estava destinat a ser realment una màquina en comptes d'un algorisme i per això finalment es va construir el Mark I Perceptró, basat en les idees del Perceptró de Frank Rosenblatt.

Tot i això, una neurona perceptró és inútil si no es junta amb altres, creant un sistema o xarxa neuronal artificial, i per això és cabdal entendre el seu funcionament, i ho explicarem des del punt de vista que a tots ens sona al sentir la paraula neurona, el biològic.

Les xarxes neuronals artificials estan basades en el funcionament de les xarxes de neurones biològiques. Les neurones que tots tenim al nostre cervell estan compostes de dendrites, el soma i l'axó: Les dendrites s'encarreguen de captar els impulsos nerviosos que emeten altres neurones. Aquests impulsos, es processen al soma i es transmeten a través de l'axó que emet un impuls nerviós cap a les neurones contigües.



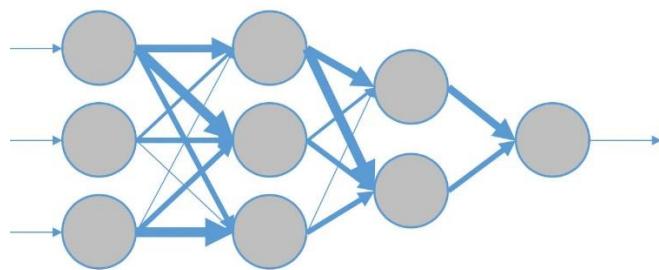
Il·lustració 55: Parts d'una Neurona

Font: [Xeridia](#)

En el cas de les neurones artificials, com ja hem dit, la suma de les entrades multiplicades pels pesos associats determina l'“impuls nerviós” que rep la neurona. Aquest valor es

processa a l'interior de la cèl·lula mitjançant una funció d'activació que torna un valor que s'envia com a sortida de la neurona. De la mateixa manera que el nostre cervell està compost per neurones interconnectades entre si, una xarxa neuronal artificial està formada per neurones artificials connectades entre si i agrupades en diferents nivells que anomenem capes.

A la següent imatge podem veure una xarxa amb quatre capes:



Il·lustració 56: Xarxa neuronal amb 4 capes

Font: Xeridia

Com hem introduït abans, la manca de disponibilitat de dades i la manca de potència de càlcul de l'època feien que aquests sistemes no tinguessin prou capacitat per resoldre problemes complexos. Això va derivar en l'arribada de l'anomenat primer hivern de la intel·ligència artificial, esdeveniment que va durar fins a la dècada dels 90. Això fou una sèrie de dècades durant les quals la manca de resultats i avenços va fer que el món acadèmic perdés esperança respecte a aquesta disciplina, ja que les expectatives que havien estat posades en l'aprenentatge automàtic no s'adherien amb la realitat.

Tot i així, durant aquests anys, els avenços no varen ser nuls. Es van començar a desenvolupar màquines a partir de tecnologia de reconeixement de patrons, que actualment és un dels majors camps dins del Machine Learning.⁸⁴

Poc a poc, els científics van començar a aplicar l'aprenentatge automàtic a la mineria de dades, el programari adaptatiu, les aplicacions web, l'aprenentatge de textos i

⁸⁴ (*La Historia Del Machine Learning - ELearning Actual*, n.d.-b)

l'aprenentatge d'idiomes a la dècada de 1990. Els científics, d'aquesta manera crearen programes informàtics que podien analitzar quantitats massives de dades i treure conclusions o aprendre dels resultats. El terme “Aprendentatge Automàtic” (Machine Learning) es va encunyar quan els científics van ser finalment capaços de desenvolupar programes informàtics de manera que poguessin aprendre i millorar per ells mateixos, sense necessitat d'intervenció humana. Així va néixer la dècada d'or de la intel·ligència artificial, on la gent va tornar a creure en les possibilitats que aquesta tecnologia pot donar.

85



Il·lustració 57: Garry Kasparov vs Deep Blue
 Font: RTVE

Entre d'altres descobriments i aplicacions descobertes aquesta dècada “d'or” per a la intel·ligència artificial, en podríem destacar una de clara: aquella que va marcar el renéixer d'aquella idea fa anys abandonada, el gairebé pleoasme d'aquest treball Machine Learning. El 1997 una cèlebre fita

va canviar la percepció de tot el món sobre la capacitat i abast de les màquines: el sistema d'IBM Deep Blue. Aquest era un algorisme entrenat a base de veure milers de partides exitoses d'escacs, i de fet va aconseguir derrotar el màxim campió mundial d'aquest joc, Garry Kasparov, succés que la gent no creia real.

Aquesta fita va ser possible aleshores –i no fa 30 anys– gràcies a l'augment de la disponibilitat de dades amb què entrenar el model. Com ens diu l'expert Yann LeCun, “el que feia aquest sistema és calcular estadísticament quin moviment té més probabilitats de fer-li guanyar la partida basant-se en milers d'exemples de partides vistes prèviament”. De fet, aquesta tècnica és també utilitzada ara per entrenar intel·ligències artificials mitjançant un bombardeig de dades.

⁸⁵ (*A Short History of Machine Learning -- Every Manager Should Read*, n.d.; *A Timeline of Machine Learning History*, n.d.; *La Historia Del Machine Learning - ELearning Actual*, n.d.-a)

Amb això arribem al que més o menys és la nostra tecnologia actual, i per tant el que componen els fonaments dels dispositius i programaris que utilitzem el nostre dia a dia.



Il·lustració 58: Tecnologia de pagament automàtic Amazon Go (en sortir)
 Font: Xataka



Il·lustració 59: Cotxe Autònom Tesla
 Font: Xataka

L'aprenentatge automàtic és una de les tecnologies actuals d'avantguarda que ens ha ajudat a millorar no només els procediments industrials i professionals, sinó també la vida quotidiana.

Actualment, mitjançant mètodes estadístics barrejats amb sistemes informàtics capaços d'aprendre experiències anteriors podem plantejar-nos possibilitats multidisciplinàriament immenses, i de fet, ja es fa servir en diversos àmbits i sectors. El diagnòstic mèdic, el processament d'imatges, la predicció, tot tipus de classificació de dades i cotxes autònoms només són algunes aplicacions.

Actualment, hi ha intel·ligències artificials basades en Machine Learning capaces de negociar, de reconèixer facialment, i de interactuar amb humans sense aquests enterar-se de que la persona darrere de la pantalla no és ben bé una persona (poden passar el test de Turing!).



La lucha por liderar el mercado de la Inteligencia Artificial

Operaciones corporativas de fusión o adquisición relacionadas con la IA.



Il·lustració 60: Progrés d'operacions econòmiques relacionades amb IA

Font: BBVA

A més a més, i centrant-nos en el tema principal del treball, l'aprenentatge automàtic és utilitzat per tota empresa per conèixer les tendències dels consumidors i els patrons operatius, així com per a la creació de nous productes. És inviable dir que aquest sector no ha esdevingut un diferenciador competitiu, i l'alça de l'enginyeria d'intel·ligència artificial i la seva demanda és una clara evidència.⁸⁶

ANNEX 4 -Tipus de Machine Learning

En primer lloc tenim el tipus menys utilitzat actualment per la seva poca efectivitat pràctica, i aquest és l'aprenentatge supervisat. Si un ordinador estigués aprenent a identificar fruites en un entorn d'aprenentatge supervisat, se li donarien imatges d'exemple de fruites etiquetades, d'això se'n diuen dades d'entrada. Per exemple, les etiquetes dirien que els plàtans són llargs, corbs i grocs, que les pomes són rodones i vermelles, mentre que una taronja és esfèrica, d'aspecte cerós i ataronjada. Després d'un temps convenient, la màquina hauria de poder identificar amb seguretat quina fruita és quina, basant-se en

⁸⁶ (9 Applications of Machine Learning from Day-to-Day Life / by Daffodil Software / App Affairs / Medium, n.d.; Applications of Machine Learning - Javatpoint, n.d.; Las 9 Aplicaciones de Machine Learning Que Deberías Conocer, n.d.; ¿Qué Es Machine Learning y Qué Aplicaciones Tiene?, n.d.; Top 10 Machine Learning Applications and Examples in 2022, n.d.)

aquestes descripcions preconcebudes (no extretes per la màquina). Si se li presenta una poma, per exemple, podria dir amb seguretat que no és de color taronja; per tant, no és una taronja, però també que no és groga i llarga, per tant, no és un plàtan. Aleshores, serà una poma pel fet que és rodona i vermella.

Parlant en termes més tècnics, l'aprenentatge de tipus supervisat es refereix a un model específic de Machine Learning on el procés de generació de coneixement es realitza amb un grup d'exemples o dades prèviament etiquetades, i on els resultats i tipus de resultats que dona l'operació són coneguts prèviament. Aquest tipus de model aprèn d'aquests resultats i incorpora ajustaments als paràmetres interiors per poder adaptar-se a dades noves que ingressen al sistema. En l'àmbit dels motors de recomanació s'utilitza en els basats en contingut, que després profunditzarem

Gràcies a l'aprenentatge desenvolupat per aquests models supervisats s'alimenta un conjunt de resultats que permet fer prediccions adequades del comportament de dades noves que encara no han estat processades, és a dir, classificació. Aquest tipus d'aprenentatge és el que s'incorpora en aplicacions tecnològiques com filters detectors de spam en correus electrònics, detectors d'imatges en "Captcha"s o en aplicacions de reconeixement de veu o escriptura. Això, és clar, és molt efectiu per reconèixer certs patrons, però si aquests han de ser prèviament especificats pels científics de dades que programen l'algorisme, perd molta viabilitat en camps on la recerca no s'incloguin aquestes facilitats.

Per contra, l'aprenentatge no supervisat és quan no hi ha cap categorització o etiquetatge de les dades. La màquina no tindrà idea del concepte de fruita, per tant no podrà etiquetar els objectes. No obstant això, podrà agrupar-los segons els colors, les mides, les formes i les diferències. La màquina agruparà les coses d'acord amb les similituds, trobant estructures i patrons ocults en dades sense etiquetar. No hi ha un camí correcte o incorrecte, ni tampoc un mestre. No hi ha resultats, només una anàlisi pur de les dades.

Si el Machine Learning fos un nen que aprengué a anar en bicicleta, l'aprenentatge supervisat és el pare que corre darrere la bicicleta i la sosté en posició vertical, mentre que l'aprenentatge no supervisat consisteix en lliurar la bicicleta, picar de mans al cap al nen i dir "bona sort".

En aquest tipus s'inclouen conjunts de dades sense etiquetar, on no es coneix prèviament l'estructura que aquests posseeixen. En aquest tipus d'aprenentatge es busca obtenir informació clau o important sense conèixer prèviament la referència de les variables de sortida, explorant l'estructura de les dades que no estan etiquetades.

Com que la màquina no sap que hi ha una resposta 'correcta', permetre que les decisions es prenguin sobre les dades basant-se únicament en la informació (és a dir, sense biaixos per part del científic) permet a l'algorisme trobar estructures interessants o ocultes a les dades que abans no eren visibles per als científics de dades, anomenades vectors de característiques.

Els algorismes d'aprenentatge no supervisats també permeten tasques de processament més complexes. Un cop més, l'absència d'etiquetatge vol dir que es poden mapejar relacions entre dades que un humà no podria mai fer, i ignora les idees (potser errònies) preconcebudes de qualsevol concepte que s'investiga, ja que es parteix des de 0.

Per tant, el millor moment per utilitzar aquest tipus d'aprenentatge és quan no hi ha dades preexistents sobre els resultats preferits, com per exemple passa quan una empresa necessita determinar el mercat objectiu d'un producte nou.

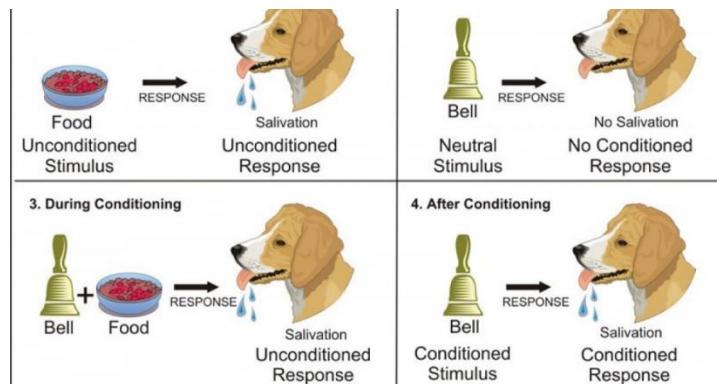
Per últim tenim el també bastant conegut aprenentatge reforçat; un model diferent dels tipus explicats anteriorment. L'aprenentatge reforçat té com a finalitat construir models que augmentin el rendiment prenent com a base el resultat o la recompensa que es genera per cada interacció realitzada. Aquesta recompensa és el producte d'una acció correcta, que deriva en una mesura específica dins del comportament per a accions futures.⁸⁷



Il·lustració 61:
 Tecnologia StockFish
 amb aprenentatge reforçat
 Font: Wikipedia

⁸⁷ (*Machine Learning / Qué Es, Tipos, Ejemplos y Cómo Implementarlo*, n.d.; *¿Qué Es El Aprendizaje No Supervisado? / TIBCO Software*, n.d.; *Types of Machine Learning - Javatpoint*, n.d.)

El “Reinforcement Learning” llavors, intentarà fer aprendre la màquina basant-se en un esquema de “premis i càstigs” –com amb el gos de Pavlov⁸⁸ (imatge a la següent pàgina)– en un entorn on cal prendre accions i que està afectat per múltiples variables que canvien amb el temps. A més, aquest sistema d’aprenentatge automàtic té la capacitat de modificar el seu entorn, de manera que pugui aplicar-se com a agent actiu a processos quotidians.



Il·lustració 62: El gos de Pavlov

Font: *Dogalize*

De fet, aquest sistema d’aprenentatge automàtic recupera un concepte aparentment obsolet en la informàtica, i això és el “bruteforce”, i consisteix en, alhora de resoldre un problema, utilitzar totes les combinacions possibles.⁸⁹

La realitat és que per fer-li aprendre a un cotxe autònom a conduir, l’hem de fer xocar, accelerar, conduir contramà i cometre tota mena d’infraccions per dir-li “això està malament, et trec els punts” i per això, cal fer que executi milers i milers de vegades en un entorn de simulat. L’experiència és molt important.

⁸⁸ Pavlov va fer un experiment on feia sonar una campana en donar menjar a uns gossos, i va demostrar que els gossos després salivaven només en escoltar la campana, cosa que indicava que havien associat el so de la campana amb el menjar.

Font: *Elcorreo*

⁸⁹ (*Aprendizaje Por Refuerzo / Aprende Machine Learning*, n.d.; *Reinforcement Learning - GeeksforGeeks*, n.d.; *Reinforcement Learning 101. Learn the Essentials of Reinforcement... | by Shweta Bhatt / Towards Data Science*, n.d.; *What Is Reinforcement Learning? The Complete Guide - Deepsense.Ai*, n.d.)

L'aprenentatge per reforç pot ser usat en robots, per exemple en braços mecànics on enlloc d'ensenyar instrucció per instrucció a moure's, podem deixar que faci intents a cegues i anar recompensant quan ho fa bé. També es pot fer servir en ambients que interactuen amb el món real, com en un altre tipus de maquinària industrial i per al manteniment predictiu, però també en l'ambient financer, per exemple per decidir com conformar una cartera d'inversió sense intervenció humana.

Si ens posem a pensar, les possibilitats són infinites.

ANNEX 5 - Entrevista Eva Martín

Aquí es troba el codi QR que porta a l'entrevista feta el 14/7/2022 amb Eva Martín, qui va fer la tesi sobre recomanació d'hotels.



ANNEX 6 - Entrevista Josep Forné

Aquí es troba el codi QR que porta a l'entrevista feta el 22/8/2022 amb el filòsof Josep Maria Forné.



ANNEX 7 - Codi del programa

Aquest és el codi QR que conduceix a un repositori de github on es troba tot el codi que hem anat referenciant



6. REFERÈNCIES

6.1 WEBGRAFIA

- *Sundog Education with Frank Kane.* (n.d.). Retrieved November 14, 2022, from
<https://www.sundog-education.com/>
- ▷ *Algoritmos Cotidianos » Qué son, Utilidad, Tipos y Ejemplos.* (n.d.). Retrieved September 29, 2022, from <https://www.cinconoticias.com/algoritmos-cotidianos/>
- 3 Ways to Calculate the Mean Absolute Error (MAE) in R [Examples].* (n.d.). Retrieved November 14, 2022, from <https://www.codingprof.com/3-ways-to-calculate-the-mean-absolute-error-mae-in-r-examples/>
- 3.-Historia de la IA - Inteligencia Artificial IA.* (n.d.). Retrieved September 7, 2022, from
<https://sites.google.com/site/iainacap/home/usos-de-la-ia>
- 9 Applications of Machine Learning from Day-to-Day Life / by Daffodil Software / App Affairs / Medium.* (n.d.). Retrieved September 30, 2022, from <https://medium.com/app-affairs/9-applications-of-machine-learning-from-day-to-day-life-112a47a429d0>
- 13 características del neoliberalismo - Significados.* (n.d.). Retrieved November 1, 2022, from
<https://www.significados.com/caracteristicas-del-neoliberalismo/>
- 25 años del duelo Kasparov - Deep Blue.* (n.d.). Retrieved November 14, 2022, from
<https://www.rtve.es/deportes/20210210/ajedrez-kasparov-deep-blue-aniversario/2074420.shtml>
- A Short History of Machine Learning -- Every Manager Should Read.* (n.d.). Retrieved September 29, 2022, from <https://www.forbes.com/sites/bernardmarr/2016/02/19/a-short-history-of-machine-learning-every-manager-should-read/?sh=2956f6a115e7>
- A systematic review and research perspective on recommender systems / Journal of Big Data / Full Text.* (n.d.). Retrieved November 14, 2022, from
<https://journalofbigdata.springeropen.com/articles/10.1186/s40537-022-00592-5>
- A Timeline of Machine Learning History.* (n.d.). Retrieved September 29, 2022, from
<https://www.techtarget.com/whatis/A-Timeline-of-Machine-Learning-History>

Ada Lovelace, pionera de la informática - Fundación Aquae. (n.d.). Retrieved September 6, 2022, from <https://www.fundacionaqua.org/wiki/ada-lovelace-madre-de-la-programacion/>

Ada Lovelace y el algoritmo informático - Agencia B12. (n.d.). Retrieved September 6, 2022, from <https://agenciab12.com/noticia/ada-lovelace-algoritmo-informatico>

Adams, J., Green, J., & Milazzo, C. (2012). Has the British Public Depolarized Along With Political Elites? An American Perspective on British Public Opinion. *Comparative Political Studies*, 45(4), 507–530. <https://doi.org/10.1177/0010414011421764>

Aditya, P. H., Budi, I., & Munajat, Q. (n.d.). *A Comparative Analysis of Memory-based and Model-based Collaborative Filtering on the Implementation of Recommender System for E-commerce in Indonesia : A Case Study PT X.*

Advantages and Disadvantages of Technology Advances by Lynda Moultry Belcher, Demand Media. (n.d.). Retrieved November 1, 2022, from <https://ininet.org/advantages-and-disadvantages-of-technology-advances-by-lynda-m.html>

Alan Turing / Biography, Facts, Computer, Machine, Education, & Death / Britannica. (n.d.). Retrieved September 6, 2022, from <https://www.britannica.com/biography/Alan-Turing>

Alan Turing, biografía del vencedor del Código Enigma - La Mente es Maravillosa. (n.d.). Retrieved September 6, 2022, from <https://lamenteesmaravillosa.com/alan-turing-biografia-del-vencedor-del-codigo-enigma/>

Alan Turing: El descifrado de la máquina Enigma >> El Año de Turing >> Blogs EL PAÍS. (n.d.). Retrieved November 14, 2022, from <https://blogs.elpais.com/turing/2013/06/alan-turing-el-descifrado-de-la-maquina-enigma.html>

Alan Turing, el padre de la inteligencia artificial / Ministerio de Cultura. (n.d.). Retrieved September 7, 2022, from <https://www.cultura.gob.ar/alan-turing-el-padre-de-la-inteligencia-artificial-9162/>

Algorisme - Viquipèdia, l'enclopèdia lliure. (n.d.). Retrieved September 5, 2022, from <https://ca.wikipedia.org/wiki/Algorisme>

Algoritmo de Euclides - Wikipedia, la encyclopédie libre. (n.d.). Retrieved November 14, 2022, from https://es.wikipedia.org/wiki/Algoritmo_de_Euclides

Alpaydin, E. (2020). *Introduction to Machine Learning* (Fourth). MIT.

Applications of Machine Learning - Javatpoint. (n.d.). Retrieved September 30, 2022, from
<https://www.javatpoint.com/applications-of-machine-learning>

Aprendizaje por Refuerzo / Aprende Machine Learning. (n.d.). Retrieved October 3, 2022, from
<https://www.aprendemachinelearning.com/aprendizaje-por-refuerzo/>

Archivo:Diagram for the computation of Bernoulli numbers.jpg - Wikipedia, la enciclopedia libre. (n.d.). Retrieved November 14, 2022, from
https://es.m.wikipedia.org/wiki/Archivo:Diagram_for_the_computation_of_Bernoulli_numbers.jpg

Arora, S., & Boaz, B. (2009). Computational Complexity: A Modern Approach Appendix A: Mathematical Background. *Cambridge University Press.*
http://www.technoeng.ro/files/publications/2_0_book1.pdf

Arquitectura de Von Newmann — Plone site. (n.d.). Retrieved September 7, 2022, from
<https://elpuig.xeill.net/Members/vcarceler/c1/didactica/apuntes/ud2/na1>

Arquitectura Von Neumann, características y limitaciones. (n.d.). Retrieved November 14, 2022, from <https://hardzone.es/tutoriales/rendimiento/von-neumann-limitaciones/>

Arquitectura von Neumann: origen, modelo, cómo funciona. (n.d.). Retrieved September 7, 2022, from <https://www.lifeder.com/arquitectura-von-neumann/>

Barberá, P., Jost, J. T., Nagler, J., Tucker, J. A., & Bonneau, R. (2015). Tweeting From Left to Right: Is Online Political Communication More Than an Echo Chamber? *Psychological Science*, 26(10), 1531–1542. <https://doi.org/10.1177/0956797615594620>

Bermúdez de Castro: “Nuestro cerebro ha cambiado muy poco desde hace 200.000 años” - Valencia Plaza. (n.d.). Retrieved November 1, 2022, from
<https://valenciaplaza.com/bermudez-de-castro-nuestro-cerebro-ha-cambiado-muy-poco-desde-hace-200000-anos>

Biografía de Alan Turing. (n.d.). Retrieved September 6, 2022, from
<https://www.biografiasyvidas.com/biografia/t/turing.htm>

Bokde, D., Girase, S., & Mukhopadhyay, D. (2015). Matrix Factorization model in Collaborative Filtering algorithms: A survey. *Procedia Computer Science*, 49(1), 136–146. <https://doi.org/10.1016/j.procs.2015.04.237>

- Boström, P., Filipsson, M., Skolan, K., Datavetenskap, F., & Kommunikation, O. (2017). *Comparison of User Based and Item Based Collaborative Filtering Recommendation Services.*
- Bromley, A. G., & Bromley, A. G. (1982). Charles Babbage's Analytical Engine, 1838. *IEEE Annals of the History of Computing*, 4(3), 197–217.
<https://doi.org/10.1109/mahc.1982.10028>
- Cinelli, M., de Francisci Morales, G., Galeazzi, A., Quattrociocchi, W., & Starnini, M. (2021). The echo chamber effect on social media. *Proceedings of the National Academy of Sciences of the United States of America*, 118(9), e2023301118.
https://doi.org/10.1073/PNAS.2023301118/SUPPL_FILE/PNAS.2023301118.SAPP.PDF
- Collaborative Filtering Vs Content-Based Filtering for Recommender Systems.* (n.d.). Retrieved October 4, 2022, from <https://analyticsindiamag.com/collaborative-filtering-vs-content-based-filtering-for-recommender-systems/>
- Computational model of a Turing machine / Download Scientific Diagram.* (n.d.). Retrieved November 14, 2022, from https://www.researchgate.net/figure/Computational-model-of-a-Turing-machine_fig1_341817215
- computer - The first computer / Britannica.* (n.d.). Retrieved September 7, 2022, from <https://www.britannica.com/technology/computer/The-first-computer>
- Condicionamiento clásico: definición y experimentos.* (n.d.). Retrieved November 14, 2022, from <https://psicologiyamente.com/psicologia/condicionamiento-clasico-experimentos>
- confirmation bias - The Skeptic's Dictionary - Skepdic.com.* (n.d.). Retrieved November 1, 2022, from <https://skepdic.com/confirmbias.html>
- Cookies de terceros: el rastro digital bajo la lupa.* (n.d.-a). Retrieved November 13, 2022, from <https://www.ionos.es/digitalguide/hosting/cuestiones-tecnicas/cookies-de-terceros/>
- Cookies de terceros: el rastro digital bajo la lupa.* (n.d.-b). Retrieved November 14, 2022, from <https://www.ionos.es/digitalguide/hosting/cuestiones-tecnicas/cookies-de-terceros/>
- Cosine similarity formula / Download Scientific Diagram.* (n.d.). Retrieved November 14, 2022, from https://www.researchgate.net/figure/Cosine-similarity-formula_fig2_345471138
- Crevier, D. (1993). *AI: The Tumultuous Search for Artificial Intelligence*. 386.
- ¿Cuál fue el primer ordenador de la historia?* (n.d.). Retrieved September 6, 2022, from <https://www.atsistemas.com/es/blog/cul-fue-el-primer-ordenador-de-la-historia>

de Solla Price, D. (1984). A History of Calculating Machines. *IEEE Micro*, 4(1), 22–52.

<https://doi.org/10.1109/MM.1984.291305>

El Filtrado colaborativo y los motores de recomendación – Juan Barrios. (n.d.-a). Retrieved

November 3, 2022, from <https://www.juanbarrios.com/el-filtrado-colaborativo-y-los-motores-de-recomendacion/>

El Filtrado colaborativo y los motores de recomendación – Juan Barrios. (n.d.-b). Retrieved

November 14, 2022, from <https://www.juanbarrios.com/el-filtrado-colaborativo-y-los-motores-de-recomendacion/>

El fin del neoliberalismo y el renacimiento de la historia / Economía / EL PAÍS. (n.d.).

Retrieved November 14, 2022, from

https://elpais.com/economia/2019/11/13/actualidad/1573640730_606639.html

“El sueño americano es una farsa”: qué es la “trampa de la meritocracia” (y cómo afecta tanto a pobres como a ricos en Estados Unidos) - BBC News Mundo. (n.d.). Retrieved

November 1, 2022, from <https://www.bbc.com/mundo/noticias-50258187>

El Sueño Americano o la nueva forma de esclavitud moderna. (n.d.). Retrieved November 14,

2022, from <https://www.sopitas.com/noticias/el-sueno-americano-esclavitud-moderna-estados-unidos-migrantes-mexico-derechos-humanos-laborales/>

Eli Pariser: Eli Pariser: cuidado con la “burbuja de filtros” en la red / TED Talk. (n.d.).

Retrieved November 1, 2022, from

https://www.ted.com/talks/eli_pariser_beware_online_filter_bubbles?language=es

Evolución de los hogares con ordenador en la vivienda por tamaño en España. (n.d.). Retrieved

November 14, 2022, from <https://www.epdata.es/evolucion-hogares-ordenador-vivienda-tamano/8c7ebe64-8e6a-4adc-9cd8-89211f12468e>

File:Perceptrón 5 unidades.svg - Wikimedia Commons. (n.d.). Retrieved November 14, 2022,

from https://commons.wikimedia.org/wiki/File:Perceptr%C3%B3n_5_unidades.svg

Filter bubbles in internet search engines - BBC News. (n.d.). Retrieved November 1, 2022, from

<https://www.bbc.com/news/av/technology-13874209>

GraphEverywhere https://www.grapheverywhere.com/grafos-que-son-tipos-orden-y-

herramientas-de-visualizacion/ - Buscar con Google. (n.d.). Retrieved November 14,

2022, from

<https://www.google.com/search?q=GraphEverywhere+https%3A%2F%2Fwww.grapheverwhere.com%2Fgrafos-que-son-tipos-orden-y-herramientas-de->

visualizacion%2F&oq=GraphEverywhere+https%3A%2F%2Fwww.grapheverywhere.co
m%2Fgrafos-que-son-tipos-orden-y-herramientas-de-
visualizacion%2F&aqs=chrome..69i57.255j0j7&sourceid=chrome&ie=UTF-8

Harper, F. M., & Konstan, J. A. (2015). The MovieLens Datasets. *ACM Transactions on Interactive Intelligent Systems (TiiS)*, 5(4). <https://doi.org/10.1145/2827872>

Herken, Rolf. (1995). *The universal Turing machine : a half-century survey*. 611.

Historia del sueño americano y sus consecuencias - EEUU - Internacional - ELTIEMPO.COM.
(n.d.). Retrieved November 1, 2022, from <https://www.eltiempo.com/mundo/eeuu-y-canada/historia-del-sueno-americano-y-sus-consecuencias-178542>

How Politics Got So Polarized / The New Yorker. (n.d.). Retrieved November 2, 2022, from
https://www.newyorker.com/magazine/2022/01/03/how-politics-got-so-polarized?utm_campaign=falcon&utm_medium=social&utm_brand=tny&utm_source=twitter&mbid=social_twitter&utm_social_type=owned&s=03

How to explain machine learning in plain English / The Enterprisers Project. (n.d.). Retrieved September 6, 2022, from <https://enterprisersproject.com/article/2019/7/machine-learning-explained-plain-english>

Introduction to recommender systems - Things Solver. (n.d.-a). Retrieved November 14, 2022, from <https://thingsolver.com/introduction-to-recommender-systems/>

Introduction to recommender systems - Things Solver. (n.d.-b). Retrieved November 14, 2022, from <https://thingsolver.com/introduction-to-recommender-systems/>

John von Neumann, el genio detrás del ordenador moderno. (n.d.). Retrieved September 7, 2022, from https://www.eldiario.es/tecnologia/diario-turing/john-neumann-revolucionando-computacion-manhattan_1_2705516.html

Kaminskas, M., & Bridge, D. (2016). Diversity, serendipity, novelty, and coverage: A survey and empirical analysis of beyond-Accuracy objectives in recommender systems. *ACM Transactions on Interactive Intelligent Systems*, 7(1). <https://doi.org/10.1145/2926720>

La batalla por liderar la inteligencia artificial / BBVA. (n.d.). Retrieved November 14, 2022, from <https://www.bbva.com/es/batalla-inteligencia-artificial/>

La cámara de eco o cómo la Red te muestra sólo lo que quieres ver - Beers&Politics. (n.d.). Retrieved November 2, 2022, from <https://beersandpolitics.com/la-camara-de-eco-o-como-la-red-te-muestra-solo-lo-que-quieres-ver>

La computadora Z3 de Konrad Zuse. (n.d.). Retrieved November 14, 2022, from
<https://parceladigital.com/articulo/la-maquina-z3-de-zuse>

La cultura de la inmediatez - MetamorfosisMetamorfosis. (n.d.). Retrieved November 1, 2022,
from <https://blogs.lavanguardia.com/metamorfosis/la-cultura-de-la-inmediatez-79017>

La gran pregunta del coche autónomo: cuánta ventaja le lleva Tesla a la competencia. (n.d.).
Retrieved November 14, 2022, from <https://www.xataka.com/vehiculos/gran-pregunta-coche-autonomo-cuanta-ventaja-le-lleva-tesla-a-competencia>

La historia del Machine Learning - eLearning Actual. (n.d.-a). Retrieved September 29, 2022,
from <https://elearningactual.com/la-historia-del-machine-learning/>

La historia del Machine Learning - eLearning Actual. (n.d.-b). Retrieved September 25, 2022,
from <https://elearningactual.com/la-historia-del-machine-learning/>

La historia del Machine Learning - eLearning Actual. (n.d.-c). Retrieved November 14, 2022,
from <https://elearningactual.com/la-historia-del-machine-learning/>

La inmediatez en la sociedad actual / Blog / Centro Cristina Andrades. (n.d.). Retrieved
November 1, 2022, from <https://centrocristinaandrades.com/blog/problema-inmediatez-sociedad-actual/>

La Máquina Universal de Turing • Un artículo de La Máquina Oráculo. (n.d.). Retrieved
September 6, 2022, from <https://lamaquinaoraculo.com/computacion/la-maquina-universal-de-turing/>

La neurona artificial. (n.d.). Retrieved November 14, 2022, from
https://www.ibiblio.org/pub/linux/docs/LuCaS/Presentaciones/200304curso-glisa/redes_neuronales/curso-glisa-redes_neuronales-html/x38.html

Las 9 aplicaciones de machine learning que deberías conocer. (n.d.). Retrieved September 30,
2022, from <https://nexusintegra.io/es/las-9-aplicaciones-de-machine-learning-que-deberias-conocer/>

Leiner, B. M., Cerf, V. G., Clark, D. D., Kahn, R. E., Kleinrock, L., Lynch, D. C., Postel, J.,
Roberts, L. G., & Wolff, S. (2009). A brief history of the internet. *ACM SIGCOMM Computer Communication Review*, 39(5), 22–31.
<https://doi.org/10.1145/1629607.1629613>

Los siete hábitos de la gente altamente efectiva - Wikipedia, la enciclopedia libre. (n.d.).

Retrieved November 14, 2022, from

https://es.wikipedia.org/wiki/Los_siete_h%C3%A1bitos_de_la_gente_altamente_efectiva

Machine Learning / Qué es, tipos, ejemplos y cómo implementarlo. (n.d.). Retrieved September 30, 2022, from <https://www.grapheverywhere.com/machine-learning-que-es-tipos-ejemplos-y-como-implementarlo/>

Machine Learning. Qué, cómo, cuándo, dónde. (n.d.). Retrieved September 29, 2022, from

<https://blogthinkbig.com/machine-learning-que-como-cuando-donde>

Máquina de Turing - Wikipedia, la enciclopedia libre. (n.d.). Retrieved November 14, 2022, from https://es.wikipedia.org/wiki/M%C3%A1quina_de_Turing

Máquina de Turing universal - Wikipedia, la enciclopedia libre. (n.d.). Retrieved September 6, 2022, from https://es.wikipedia.org/wiki/M%C3%A1quina_de_Turing_universal

Máquina diferencial – Maquina analítica / IDIS. (n.d.). Retrieved September 5, 2022, from <https://proyectoidis.org/maquina-diferencial-maquina-analitica/>

Máquinas de Turing. ¿Qué es una máquina de Turing? / by Bootcamp AI / Medium. (n.d.). Retrieved September 7, 2022, from <https://bootcampai.medium.com/m%C3%A1quinas-de-turing-c329ccc270f>

Marvin Minsky's SNARC, Possibly the First Artificial Self-Learning Machine : History of Information. (n.d.). Retrieved September 7, 2022, from <https://www.historyofinformation.com/detail.php?entryid=4343>

McPherson, M., Smith-Lovin, L., & Cook, J. M. (2003). Birds of a Feather: Homophily in Social Networks. [Https://Doi.Org/10.1146/Annurev.Soc.27.1.415](https://doi.org/10.1146/annurev.soc.27.1.415), 27, 415–444.
<https://doi.org/10.1146/ANNUREV.SOC.27.1.415>

Model-Based Recommendation System / Towards Data Science. (n.d.). Retrieved November 4, 2022, from <https://towardsdatascience.com/how-to-build-a-model-based-recommendation-system-using-python-surprise-2df3b77ab3e5>

Neoliberalism - Overview, History, Characteristics, and Schools of Thought. (n.d.). Retrieved November 1, 2022, from <https://corporatefinanceinstitute.com/resources/economics/neoliberalism/>

Neoliberalism: What It Is, With Examples and Pros and Cons. (n.d.). Retrieved November 1, 2022, from <https://www.investopedia.com/terms/n/neoliberalism.asp>

Neoliberalismo: definición, ventajas y características. (n.d.). Retrieved November 1, 2022, from <https://humanidades.com/neoliberalismo/>

Neoliberalismo: qué es, características y representantes principales - Toda Materia. (n.d.). Retrieved November 1, 2022, from <https://www.todamateria.com/neoliberalismo/>

Nilsson, N. J. (1998). *Artificial Intelligence : a new synthesis.* 513.

Online Personalization Creates Echo Chamber to Affirm Biases - The New York Times. (n.d.). Retrieved November 14, 2022, from <https://www.nytimes.com/2011/05/29/technology/29stream.html>

O'Regan, G. (Cornelius G. (n.d.). *A brief history of computing.* 359.

Origen del concepto de Inteligencia Artificial. (n.d.). Retrieved September 7, 2022, from <https://agenciab12.com/noticia/origen-concepto-inteligencia-artificial>

(PDF) *Model-based approach for Collaborative Filtering.* (n.d.). Retrieved October 5, 2022, from https://www.researchgate.net/publication/321753015_Model-based_approach_for_Collaborative_Filtering

Per una dissecció de la postmodernitat. (n.d.). Retrieved November 1, 2022, from <http://webs.racocatala.cat/eltalp/neolib.htm>

Perceptron : qu'est-ce que c'est et à quoi ça sert ? (n.d.). Retrieved September 11, 2022, from <https://datascientest.com/es/perceptron-que-es-y-para-que-sirve>

¿Por qué el neoliberalismo provoca desigualdad? (n.d.). Retrieved November 1, 2022, from <https://www.nuevatribuna.es/articulo/economia/neoliberalismo-provoca-desigualdad/20150222132252112779.html>

¿Por qué me aparece publicidad de lo que hablo? Esta es la respuesta. (n.d.). Retrieved November 13, 2022, from <https://pymesworld.com/porque-publicidad-hablo/>

Proyecto ENIAC – Historia de la Informática. (n.d.). Retrieved November 14, 2022, from <https://histinf.blogs.upv.es/2011/12/05/proyecto-eniac/>

Publicidad misteriosa: ¿nuestros móviles nos escuchan? (n.d.). Retrieved November 13, 2022, from <https://www.nobbot.com/pantallas/publicidad-misteriosa-moviles-nos-escuchan/>

¿Qué es el algoritmo de k vecinos más cercanos? / IBM. (n.d.). Retrieved November 14, 2022, from <https://www.ibm.com/es-es/topics/knn>

¿Qué es el aprendizaje no supervisado? / TIBCO Software. (n.d.). Retrieved September 30, 2022, from <https://www.tibco.com/es/reference-center/what-is-unsupervised-learning>

¿Qué es el scroll infinito y por qué deberían utilizarlo los editores? / Clickio. (n.d.). Retrieved November 14, 2022, from <https://blog.clickio.com/es/que-es-el-scroll-infinito-y-por-que-deberian-utilizarlo-los-editores/>

¿Qué es Machine Learning y qué aplicaciones tiene? (n.d.). Retrieved September 30, 2022, from <https://www.ibertech.org/que-es-machine-learning-y-que-aplicaciones-tiene-en-nuestro-dia-a-dia-2/>

¿Qué es un algoritmo? ¿De dónde surge? – Programa en Línea. (n.d.). Retrieved September 6, 2022, from <https://www.programaenlinea.net/que-es-un-algoritmo-de-donde-surge/>

Qué influye más en lo que somos, ¿la genética o la crianza? - The New York Times. (n.d.). Retrieved November 1, 2022, from <https://www.nytimes.com/es/2018/08/22/espanol/naturaleza-crianza-gemelos.html>

Qué significa aceptar las cookies al entrar en una página web. (n.d.). Retrieved November 13, 2022, from <https://www.redeszone.net/tutoriales/internet/que-es-aceptar-cookies-web/>

Real World Applications of Recommendation Engine - Muvi One. (n.d.). Retrieved November 13, 2022, from <https://www.muvi.com/blogs/real-world-applications-of-recommendation-engine.html>

Recommendation Systems: Applications and Examples in 2022. (n.d.). Retrieved November 13, 2022, from <https://research.aimultiple.com/recommendation-system/>

Redes Neuronales artificiales: Qué son y cómo se entranan / [site:name]. (n.d.-a). Retrieved November 14, 2022, from <https://www.xeridia.com/blog/redes-neuronales-artificiales-que-son-y-como-se-entrenan-parte-i>

Redes Neuronales artificiales: Qué son y cómo se entranan / [site:name]. (n.d.-b). Retrieved November 14, 2022, from <https://www.xeridia.com/blog/redes-neuronales-artificiales-que-son-y-como-se-entrenan-parte-i>

Reinforcement learning - GeeksforGeeks. (n.d.). Retrieved October 3, 2022, from <https://www.geeksforgeeks.org/what-is-reinforcement-learning/>

Reinforcement Learning 101. Learn the essentials of Reinforcement... | by Shweta Bhatt | Towards Data Science. (n.d.). Retrieved October 3, 2022, from <https://towardsdatascience.com/reinforcement-learning-101-e24b50e1d292>

- Rogozhin, Y., & Rogozhin, Y. (1996a). Small Universal Turing Machines. *Theoretical Computer Science*, 168(2), 215–240. [https://doi.org/10.1016/S0304-3975\(96\)00077-1](https://doi.org/10.1016/S0304-3975(96)00077-1)
- Rogozhin, Y., & Rogozhin, Y. (1996b). Small Universal Turing Machines. *Theoretical Computer Science*, 168(2), 215–240. [https://doi.org/10.1016/S0304-3975\(96\)00077-1](https://doi.org/10.1016/S0304-3975(96)00077-1)
- Root-Mean-Square Error in R Programming - GeeksforGeeks*. (n.d.). Retrieved November 14, 2022, from <https://www.geeksforgeeks.org/root-mean-square-error-in-r-programming/>
- Roy, D., & Dutta, M. (2022). A systematic review and research perspective on recommender systems. *Journal of Big Data*, 9(1), 1–36. <https://doi.org/10.1186/S40537-022-00592-5/TABLES/3>
- Russell, S. J. (Stuart J., Norvig, P., & Canny, John. (n.d.). *Artificial intelligence : a modern approach*. 1081.
- Sistemas de recomendación / ¿Qué es el filtrado colaborativo?* (n.d.). Retrieved October 5, 2022, from <https://www.grapheverywhere.com/sistemas-de-recomendacion-que-es-el-filtrado-colaborativo/>
- Sistemas de recomendación de contenido con Machine Learning – Cleverdata*. (n.d.). Retrieved October 4, 2022, from <https://cleverdata.io/sistemas-recomendacion-machine-learning/>
- SNARC / History of AI*. (n.d.). Retrieved September 7, 2022, from <https://historyof.ai/snarc>
- SOBRE EL PROBLEMA DE LA INMEDIATEZ O LA NECESIDAD DE GRATIFICACIÓN INMEDIATA:** “*Su relación con la autodestrucción*” – Centro Psicológico SELF. (n.d.). Retrieved November 1, 2022, from <https://www.psicologosalamancaself.com/blog/sobre-el-problema-de-la-inmediatez-o-la-necesidad-de-gratificacion-inmediata-%E2%80%9Csu-relacion-con-la-autodestruccion%E2%80%9D/>
- Stochastic neural analog reinforcement calculator - Wikipedia*. (n.d.). Retrieved September 7, 2022, from https://en.wikipedia.org/wiki/Stochastic_neural_analog_reinforcement_calculator
- Stockfish - Wikipedia, la enciclopedia libre*. (n.d.). Retrieved November 14, 2022, from <https://es.wikipedia.org/wiki/Stockfish>
- Sundog Education by Frank Kane / Founder, Sundog Education. Machine Learning Pro / Udemy*. (n.d.-a). Retrieved November 14, 2022, from <https://www.udemy.com/user/frankkane/>

Sundog Education by Frank Kane / Founder, Sundog Education. Machine Learning Pro /

Udemy. (n.d.-b). Retrieved November 14, 2022, from
<https://www.udemy.com/user/frankkane/>

Te contamos qué es el “machine learning” y cómo funciona. (n.d.). Retrieved September 5, 2022, from <https://www.bbva.com/es/machine-learning-que-es-y-como-funciona/>

Teorema de Bayes - Viquipèdia, l'enciclopèdia lliure. (n.d.). Retrieved November 14, 2022, from https://ca.wikipedia.org/wiki/Teorema_de_Bayes

Teorema de Bayes - Wikipedia, la encyclopédie libre. (n.d.). Retrieved September 5, 2022, from https://es.wikipedia.org/wiki/Teorema_de_Bayes

Teoría de Autómatas y Computación: Marvin Minsky y la computación neuronal. (n.d.). Retrieved November 14, 2022, from <http://taycalbacete.blogspot.com/2014/05/normal-0-21-false-false-false-es-x-none.html>

The Difference Between Implicit and Explicit Data for Business / by Tim Clayton / Blog / Mirumee. (n.d.). Retrieved November 3, 2022, from <https://blog.mirumee.com/the-difference-between-implicit-and-explicit-data-for-business-351f70ff3fbf>

The History of Neural Networks and AI: Part I - Open Data Science. (n.d.). Retrieved September 11, 2022, from <https://opendatascience.com/the-history-of-neural-networks-and-ai-part-i/>

The problem of living inside echo chambers. (n.d.). Retrieved November 14, 2022, from <https://theconversation.com/the-problem-of-living-inside-echo-chambers-110486>

The Psychology of Instant Gratification and How It Will Revolutionize Your Marketing Approach. (n.d.). Retrieved November 1, 2022, from <https://www.entrepreneur.com/growing-a-business/the-psychology-of-instant-gratification-and-how-it-will/235088>

This week in The History of AI at AIWS.net – Marvin Minsky and Dean Edmonds built SNARC, the first artificial neural network / AIWS.net. (n.d.). Retrieved September 7, 2022, from <https://aiws.net/the-history-of-ai/this-week-in-the-history-of-ai-at-aiws-net-marvin-minsky-and-dean-edmonds-built-snarc-the-first-artificial-neural-network/>

Top 10 Machine Learning Applications and Examples in 2022. (n.d.). Retrieved September 30, 2022, from <https://www.simplilearn.com/tutorials/machine-learning-tutorial/machine-learning-applications>

Train-Test Split for Evaluating Machine Learning Algorithms. (n.d.). Retrieved November 4, 2022, from <https://machinelearningmastery.com/train-test-split-for-evaluating-machine-learning-algorithms/>

Types of Machine Learning - Javatpoint. (n.d.). Retrieved September 30, 2022, from <https://www.javatpoint.com/types-of-machine-learning>

Various Implementations of Collaborative Filtering / by Prince Grover / Towards Data Science. (n.d.). Retrieved October 5, 2022, from <https://towardsdatascience.com/various-implementations-of-collaborative-filtering-100385c6dfe0>

What Content-Based Filtering is & Why You Should Use it / Upwork. (n.d.). Retrieved October 4, 2022, from <https://www.upwork.com/resources/what-is-content-based-filtering>

What is a Content-based Recommendation System in Machine Learning?/ Analytics Steps. (n.d.). Retrieved October 4, 2022, from <https://www.analyticssteps.com/blogs/what-content-based-recommendation-system-machine-learning>

What is a Third-Party Cookie? (n.d.). Retrieved November 13, 2022, from <https://www.techtarget.com/whatis/definition/third-party-cookie>

What is collaborative filtering? - Definition from WhatIs.com. (n.d.). Retrieved October 5, 2022, from <https://www.techtarget.com/whatis/definition/collaborative-filtering>

What is content-based Filtering? (n.d.). Retrieved November 14, 2022, from <https://www.educative.io/answers/what-is-content-based-filtering>

What is implicit data? - Definition from WhatIs.com. (n.d.). Retrieved November 3, 2022, from <https://www.techtarget.com/whatis/definition/implicit-data>

What is Machine Learning? / IBM. (n.d.). Retrieved September 6, 2022, from <https://www.ibm.com/cloud/learn/machine-learning>

What Is Machine Learning - ML - and Why Is It Important? / NetApp. (n.d.). Retrieved November 3, 2022, from <https://www.netapp.com/artificial-intelligence/what-is-machine-learning/>

What is reinforcement learning? The complete guide - deepsense.ai. (n.d.). Retrieved October 3, 2022, from <https://deepsense.ai/what-is-reinforcement-learning-the-complete-guide/>

When Was the First Computer Invented? (n.d.). Retrieved September 7, 2022, from <https://www.computerhope.com/issues/ch000984.htm>

6.2 BIBLIOGRAFIA

- TURING, Alan, The Essential Turing: Seminal Writings in Computing, Logic, Philosophy, Artificial Intelligence, and Artificial Life plus The Secrets of Enigma, Copeland J., 2004
- SCHRAGE, Michael, Recommendation Engines, PaperBag, 2020
- FISHER Mark, Realismo Capitalista- No hay alternativa?, Caja Negra 2016
- LOWEN Alexander, El narcicismo, la enfermedad de nuestro tiempo, Paidos, 2014
- BAUDRILLARD Jean, La ilusión del fin, la huelga de los acontecimientos, Anagrama 2006
- BACH Richard, Juan Salvador Gaviota, B de Bolsillo 2019
- TOMLISON John, The culture of Speed: The coming of Immediacy, Sage 2007
- CASATI Roberto, Elogio del Papel: contra el colonialismo digital, Ariel 2015
- CARR Nicholas, Superficiales: ¿Qué está haciendo internet con nuestras mentes?, DEBOLSILLO 2018
- MUNTADAS FIGUERAS Borja, Inmediatez, capitalismo y vides aceleradas, Chiado 2016
- LYOTARD Jean François, La Condición postmoderna, Ediciones Cátedra 2006
- JAPPE Anselm, La societat autófaga, Pepitas de calabaza 2019
- JAMESON Fredric, Ensayos sobre el posmodernismo, Ediciones Imago Mundi

1991

- HUXLEY Aldous, Un Món Feliç, Educaula 2019
- ORWELL George, 1984, DEBOLSILLO 2013
- MINGARDI Alberto, La Verdad sobre el Neoliberalismo
- BAUMAN, Zygmunt, Vida de Consumo 2007
- BAUFRILLARD, Jean, La Sociedad de consumo. Sus mitos, sus estructuras 2009
- CAÑAS, José Luis, SOREN KIERKEGAARD: Entre la Immediatez y la Relación
- BAUMAN, Zygmunt, Amor Líquido: Acerca de la Fragilidad de Los Vínculos Humanos 2003

6.3 ÍNDEX D'IL·LUSTRACIONS

Il·lustració 1: Funció "Et Recomanem", de Netflix	5
Il·lustració 2: Cerca Usuari 1: "Juego" Font: Twitter	6
Il·lustració 3: Cerca Usuari 2: "Juego" Font: Twitter	6
Il·lustració 4: Funcionament i Tipus de Cookies	23
Il·lustració 5: Agrupació de tipus de motors de recomanació	26
Il·lustració 6: Filtratge basat en contingut.....	27
Il·lustració 7: Matriu de 2 dimensions entre ítems i atributs.....	28
Il·lustració 8: Diagrama KNN Font: IBM.....	35
Il·lustració 9: Funcionament d'un sistema de recomanació amb filtratge col·laboratiu basat en l'usuari	35
Il·lustració 10: Funció Amazon (Compra-ho amb); si primerament elegeix comprar un comandament per a videojocs.	36
Il·lustració 11: Funcionament Sistema de recomanació amb filtratge col·laboratiu basat en l'ítem	36
Il·lustració 12: Comparació entre els dos grans tipus de motors de recomanació	38
Il·lustració 13: Funcionament del mètode "Train/Test split"	39

II·l·lustració 14: Fórmula de l'Error Absolut Mitjà	39
II·l·lustració 15: Fórmula de l'arrel de l'error quadràtic mitjà (RMSE)	40
II·l·lustració 16: Scroll Infinit Font: Clickio	Error! No s'ha definit el marcador.
II·l·lustració 17: Funció Amazon's Choice	48
II·l·lustració 18: Sistema de Grafs i nodes	53
II·l·lustració 19: Exemple gràfic del funcionament de l'Echo Chamber Font: The New York Times.....	55
II·l·lustració 20: Exemple gràfic del funcionament de l'Echo Chamber Font: TheConversation..	55
II·l·lustració 21: Funcionament similitud cosinus	61
II·l·lustració 22: Representació de pel·lícules per gènere	62
II·l·lustració 23: Fórmula de la similitud del cosinus	62
II·l·lustració 24: Funció similitud cosinus explicada en codi.	63
II·l·lustració 25: Funció similitud entre anys codificada	64
II·l·lustració 26: K-Nearest-Neighbours Font: Codi i Imatge de l'autor	65
II·l·lustració 27: Full d'excel amb valoracions	70
II·l·lustració 28: Pel·lícules Corresponents a cada ID	70
II·l·lustració 29: Execució part 1	71
II·l·lustració 30: Execució Part 2	71
II·l·lustració 31: Execució Part 3	72
II·l·lustració 32: Execució part 4	72
II·l·lustració 33: Canvis realitzats al codi. Part 1.....	73
II·l·lustració 34: Canvis Realitzats al codi. Part 2	74
II·l·lustració 35: Segona execució (després del canvi), Part 1.....	74
II·l·lustració 36: Segona execució (després del canvi), part 2	75
II·l·lustració 37: Gràfic -18, P.2 (19 SI 4 NO).....	78
II·l·lustració 38: Gràfic 18-60, P.2 (15 SI 6 NO).....	78
II·l·lustració 39: Gràfic 18-60, P.1 (13 SI 8 NO).....	78
II·l·lustració 40: Gràfic -18, P.1 (15 SI 8 NO).....	78
II·l·lustració 41: Gràfic +60, P.2 (15 SI 3 NO).....	78
II·l·lustració 42: Gràfic +60, P.1 (7 SI 11 NO).....	78
II·l·lustració 43: Els nombres de Bernoulli.....	83
II·l·lustració 44: Representació matemàtica del Teorema de Bayes.....	85
II·l·lustració 45: La màquina ENIGMA	87
II·l·lustració 46: Esquema explícit de la màquina de Turing	88

II·lustració 47: Procés Màquina de Turing.....	89
II·lustració 48: Evolució de llars amb ordinador a Espanya <i>Font: Epdata</i>	91
II·lustració 49: Arquitectura de Von Neumann <i>Font: Wikipedia</i>	92
II·lustració 50: Estructura ENIAC.....	94
II·lustració 51: Konrad Zuse i el Z3 <i>Font: ParcelaDigital</i>	95
II·lustració 52: Neurona McCulloch-Pitts	97
II·lustració 53: L'SNARC <i>Font: The Scientist</i>	98
II·lustració 54: El Perceptró <i>Font: Wikipedia</i>	100
II·lustració 55: Parts d'una Neurona	100
II·lustració 56: Xarxa neuronal amb 4 capes.....	101
II·lustració 57: Garry Kasparov vs Deep Blue <i>Font: RTVE</i>	102
II·lustració 58: Tecnologia de pagament automàtic Amazon Go (en sortir) <i>Font: Xataka</i>	103
II·lustració 60: Cotxe Autònom Tesla <i>Font: Xataka</i>	103
II·lustració 61: Progrés d'operacions econòmiques relacionades amb IA <i>Font: BBVA</i>	104
II·lustració 62: Tecnologia StockFish amb aprenentatge reforçat <i>Font: Wikipedia</i>	106
II·lustració 63: El gos de Pavlov	107