

Fonaments de la Ciència de Dades

PAC1: Ciència a les dades?

Font d'informació:

- [El ser humano se transformó en algoritmo. - Andy Stalman](#)
- [Tendencias en RRSS 2024 • SDi](#)
- [The Top 12 Social Media Trends for 2024 You Need to Know](#)
- [¿Cómo funcionan los algoritmos? Así deciden las redes sociales que contenido enseñarle a tus hijos – Alfabetización Digital](#)
- [La ética del dato: hacia el 'homo algorithmus'](#)
- [Entender el algoritmo de X | Juice Studio](#)
- [Elon Musk anuncia cambios en el algoritmo de X, antes Twitter, por un par de razones](#)

Pregunta 1 (10% puntuació)

Per l'actualització del nou algoritme de X:

- Creus necessària la presència de Científics de Dades?

Tenint en compte que els Científics de Dades són professionals que combinen coneixements en estadística, programació i anàlisi per a extreure informació valuosa a partir de grans volums de dades, la seva presència és important, ja que són els adequats per a utilitzar les eines d'anàlisi de X per obtenir informació important com les dades demogràfiques de l'audiència, les mètriques de participació o el rendiment dels continguts[1]. A més, s'ha de tindre en compte que l'algorisme de X és un sofisticat sistema dissenyat per seleccionar continguts adaptats a les preferències i interessos de cada usuari[1], cosa que implica, un altre cop, la necessitat de professionals que puguin gestionar aquests grans volums de dades i ajustar l'algorisme. Analitzant aquestes dades correctament es poden trobar tendències, perfeccionar l'estratègia i optimitzar l'esforç d'altres departaments perquè tinguin el màxim impacte.

- Quines tasques farien?

Les principals tasques serien:

- Anàlisi de dades d'usuari: Tenint en compte diverses mètriques com els likes, comentaris i comparticions, és poder identificar patrons de comportament i preferències[1].

- Optimització de l'algorisme: També ajustarien l'algorisme per assegurar que el contingut sigui el més rellevant i atractiu possible pels usuaris.

- Detecció i reducció del clickbait: Controlar i desenvolupar models que identifiquin continguts fraudulents per així pot reduir-ne la presència. En paraules del mateix Elon Musk, «Canvis que vindran en una nova actualització de X que oferirà more bangers i menys clickbait»[2].

- Experimentació: Provar diferents formats i estratègies, aplicant-t'ho a diferents tipus de contingut i mètodes de presentació, poden veure com afecta la participació i l'interès dels usuaris. Sempre tenint en compte que és important mantenir-los àgils i adaptables[1].

- Quins coneixements i habilitats han de tenir?

Per a poder fer la seva feina, un Científic de Dades ha de tindre una comprensió profunda del comportament dels usuaris en les xarxes socials, així com:

- Anàlisi de dades i estadística avançada: Per a poder extreure informació útil de grans conjunts de dades, identificant patrons i tendències[1].

- Programació i desenvolupament d'algorismes: Per a poder generar nous algorismes o modificar els actuals per a millorar la personalització del contingut a l'experiència de l'usuari[1]. També per identificar i minimitzar la presència de falsa informació i el clickbait[2]. I per evitar les bombolles de filtre, evitant aïllar als usuaris en les seves pròpies bombolles d'informació [3].

Pregunta 2 (25% puntuació)

Per concretar la configuració dels equips de les Xarxes Socials:

- Defineix els perfils involucrats

Per gestionar de manera efectiva les xarxes socials, s'ha de tindre en compte que es tracta d'un entorn canviant i competitiu, per tant, es necessita un equip multidisciplinari.

- Social media manager: Responsable de la planificació estratègica i coordinació de totes les activitats. Aquest perfil és vital per mantenir la xarxa social rellevant i competitiva.
- Científic de Dades: Analitzar grans volums de dades per extreure informació útil i generar models que millorin l'efectivitat de les estratègies.
- Analista de Dades: Interpreta les mètriques i dades recollides, avaluant el rendiment i suggerint millores.
- Enginyer de dades: Encarregat de la infraestructura de dades i les eines que s'utilitzaran durant el seu tractament.
- Especialista en SEO i Algoritmes: optimitza els continguts per als motors de cerca i els manté actualitzats amb els canvis en els algorismes[4]
- Especialista en Ètica de Dades: Per assegurar que es fa un ús de les dades de forma ètica i conforme a la legislació vigent.
- Especialista en visualització de dades: Crea visualitzacions precises per fer que la informació obtinguda sigui més fàcilment interpretable.
- Chief Data Officer: S'encarrega de tots els processos relacionats amb la ciència de dades.
- Dissenyador Frontend: Treballa en el disseny de la interfície de la xarxa social.
- Dissenyador Backend: Treballa en la lògica del servidor.

● Quins nivells de competència analítica han de tenir cadascun?

- Social media manager: intermedi, ja que ha de comprendre les mètriques clau per prendre decisions estratègiques.
- Científic de Dades: Molt alt, perquè la seva feina requereix habilitats altes en anàlisi de dades, estadística i programació per desenvolupar models complexos.
- Analista de Dades: Alt, perquè ha de poder interpretar dades i traduir-les a coneixement útil per a l'equip.

- Enginyer de dades: Molt alt, perquè ha de conèixer els requisits de les dades amb les quals es treballarà i programar les eines necessàries.
- Especialista en SEO i Algoritmes: Molt Alt, perquè ha de comprendre els algorismes per optimitzar el contingut i mantenir-los al dia.
- Especialista en Ètica de Dades: Alt, perquè necessita coneixements en anàlisi de dades i comprensió de les implicacions legals i ètiques.
- Especialista en visualització de dades: Molt Alt, perquè per a poder fer una bona presentació de dades és necessari comprendre-les en profunditat.
- Chief Data Officer: Molt Alt, perquè està al capdamunt de tots els processos relacionats amb la ciència de dades i ha de tindre una visió general de l'ús de les dades.
- Dissenyador Frontend: Molt alt, perquè ha de programar tota la interfície d'usuari, assegurant una experiència òptima basada en dades d'usuari i analítica.
- Dissenyador Backend: Molt alt, perquè gestiona que les dades es processin de manera eficient i segura.

● Quines bones pràctiques han de considerar com treballadors analítics?

Tenint en compte les bones pràctiques descrites a “Organitzacions orientades a la dada[5]”:

- Tindre un bon coneixement de les xarxes socials
- Comprendre la manera de pensar dels usuaris i com els models i algorismes afectaran aquesta manera de fer-ho.
- Desenvolupar una relació de treball col·laborativa amb la resta de treballadors.
- Cal fer servir un llenguatge professional per a comunicar idees i millores.
- Establir una manera ordenada i estructurada d'obtenir informació dels usuaris.

Pregunta 3 (20% puntuació)

- Descriu les diferents dades que utilitza X i classifica-les en funció de si són Simples, Estructurades, Semi Estructurades o No Estructurades.

D'acord amb la font d'informació "Entender el algoritmo de X | Juice Studio[1]", he pogut detectar les següents dades:

- Dades de participació dels usuaris: Inclouen likes, comentaris, comparticions i altres interaccions amb el contingut. Són la base de l'algorisme de X per a determinar la visibilitat i l'abast de les publicacions. Són dades Semi Estructurades, ja que, si ve la seva estructura està definida, pot variar el seu format i contingut.
- Dades de preferències i interessos dels usuaris: Són les dades que es fan servir per personalitzar el contingut en funció de les preferències individuals. També són Semi Estructurades, perquè combinen informació estructurada amb elements variables.
- Dades de rellevància del contingut: Basades en la qualitat i rellevància percebuda del contingut pels usuaris. No estructurades perquè depenen de factors subjectius i qualitius.
- Dades de freqüència: Indiquen la freqüència de les publicacions i interacció dels usuaris amb la plataforma. Aquests són simples, ja que es tracten de dades numèriques senzilles.
- Dades d'autenticitat: Aporten informació sobre la fiabilitat de les fonts i l'autenticitat del contingut. Són dades No Estructurades, perquè són de naturalesa qualitativa i variable.
- Dades demogràfiques dels usuaris: Acostumen a incloure edat, gènere, ubicació i altres característiques dels usuaris. Aquestes són Estructurades, perquè estan organitzades en camps definits.
- Dades analítiques: Es tracta de mètriques com l'abast, les impressions i el nivell de participació que permeten quantificar l'eficàcia del contingut. Són estructurades perquè són dades quantitatives organitzades.

- Descriu el Cicle de Vida de les dades involucrades a l'anterior apartat.
 1. Captura: Les dades són capturades a través de les interaccions dels usuaris.
 2. Emmagatzematge: Les dades capturades s'emmagatzemen en bases de dades de X per al seu ús posterior. Aquí hi ha les dades estructurades i no estructurades.
 3. Preprocessat: Abans de fer-ne una anàlisi, les dades es preprocessen per assegurar que estan al format adequat. Aquí poden haver-hi processos de neteja de dades, ordenació i transformació.
 4. Anàlisi: L'algorisme analitza les dades preprocessades per comprendre les preferències individuals dels usuaris i determinar per exemple, quin contingut és més rellevant per a ells.-

Pregunta 4 (20% puntuació)

- Les xarxes socials compleixen amb els criteris per ser considerades com a una Organització Orientada a la Dada?

Les xarxes socials poden ser considerades organitzacions orientades a la dada, ja que, tal i com s'explica a Organitzacions orientades a la dada[5], aquestes organitzacions utilitzen les dades per a prendre decisions i gestionar les interaccions amb els usuaris. Aquestes recullen grans volums de dades sobre els usuaris i les seves activitats (publicacions, interaccions, preferències, etc.) que són analitzades per millorar els serveis, personalitzar l'experiència de l'usuari, i optimitzar campanyes de màrqueting, entre altres aplicacions.

- Analitza el model de maduresa DELTA de l'organització

Seguint les fases descrites a "El model de maduresa DELTA[5]":

| Factors d'èxit | Fase | Justificació |
|----------------|------|--|
| Dades | 5 | S'utilitzen grans volums de dades per personalitzar l'experiència de l'usuari, |

| | | |
|--------------|---|--|
| | | optimitzar processos interns i desenvolupar nous serveis. |
| Organització | 4 | Existeix una infraestructura establerta que facilita de col·laboració entre departaments, els processos estan dissenyats per aprofitar les dades. Encara s'està treballant per aconseguir seguir la cultura analítica en tots els nivells. |
| Lideratge | 4 | El seu lideratge promou l'ús de l'analítica per a la presa de decisions, S'encoratgen mentalitats basades en dades. |
| Objectius | 4 | Els objectius de l'empresa estan alineats amb una visió analítica. L'empresa busca expandir-se i innovar gràcies a les dades. |
| Tecnologia | 5 | Es disposa de tecnologia avançada i infraestructura robusta que permet processar i analitzar grans volums de dades. Això permet implementar solucions innovadores i respondre ràpidament a les necessitats del mercat. |
| Persones | 5 | L'equip està altament capacitat en cultura analítica. Els empleats utilitzen les dades en les seves tasques diàries, |

| | | |
|--|--|--|
| | | millorant els servei base a evidències empíriques. |
|--|--|--|

Pregunta 5 (15% puntuació)

Posa't a la pell d'un membre de l'equip de visualització de dades

- Quina visualització proposaries sobre l'ús de X per millorar l'experiència, incrementar l'ús i el nombre d'usuaris?

Per exemple es podria dissenyar una visualització interactiva i dinàmica per analitzar l'ús de la plataforma.

Mapa de calor d'activitats dels usuaris: Aquest mapa mostraria l'activitat dels usuaris al llarg del dia, de la setmana, mes i any. Això permetria identificar franges temporals amb més i menys activitat per optimitzar el llançament de continguts i funcionalitats.

Es podrien identificar els punts on els usuaris troben dificultat o abandonen la plataforma, i implementar millores específiques per fer el flux de navegació més intuïtiu i agradable. De la mateixa manera, veure en quins moments del dia o en quins mesos l'activitat és menor, permetria millorar l'aplicació, incrementant l'ús i el nombre d'usuaris, ja que una experiència millorada no només manté els usuaris sinó que també n'atrau de nous.

Pregunta 6 (10% puntuació)

Destaca dues de les intervencions del primer debat.

- Què destacaries de cadascuna d'elles? Hi ha algun aspecte en el que no coincideixis?

La primera intervenció és la meva (Marc Bracons) i la resposta de Paula Barrachina Cáceres.

De la intervenció de la Paula destacaria que deixa clar els objectius dels algorismes, coincidint en el fet que l'objectiu principal és comercial. Comenta també que tenen una funció que busca mantenir enganxats als usuaris i que en tindre una constant

evolució, les marques s'han vist obligades a canviar les seves estratègies. La pregunta que fa al final és molt interessant, "creieu que les modificacions dels algoritmes són sostenibles a llarg termini, o podríem veure un rebuig dels usuaris davant d'un excés de contingut comercialitzat?".

Si bé crec que en general coincidim en la majoria dels punts, no coincideixo en la idea que les modificacions beneficien els usuaris en la mateixa mesura que les plataformes i les marques. Aquesta personalització, si no és tractada correctament pot crear filtres de bombolla, limitant l'exposició a perspectives diverses. A més, voler mantenir als usuaris enganxats pot afectar negativament el seu benestar i la seva capacitat de discernir informació de qualitat. Responent a la seva pregunta, considero que, pels motius exposats, veurem un rebuig dels usuaris.

La segona intervenció és la de l'Ana Moscardó García al post "Quin paper poden tenir els científics de dades a empreses com TikTok, Instagram o X?".

D'ella en destacaria la cerca de l'equilibri entre l'eficiència i l'ètica en el disseny d'algorismes. La preocupació per sobre qui recau la responsabilitat de l'ètica i la creació d'una àrea que s'encarregui d'aquest tema. A més comenta (amb encert en la meua opinió) que no hem d'oblidar que els Científics de Dades poden treballar en l'ètica, però sense descuidar que aquesta no és la seva principal responsabilitat.

Bibliografia

1. [juice-studio; Entender el algoritmo de X; 28 Mar 2024; <https://juice-studio.com/entender-algoritmo-x/>]
2. [sdpnoticias; Andrea Luices; 11 abril 2024; Elon Musk anuncia cambios en el algoritmo de X, antes Twitter, por un par de razones; <https://www.sdpnoticias.com/tecnologia/elon-musk-anuncia-cambios-en-el-algoritmo-de-x-antes-twitter-por-un-par-de-razones/>]
3. [Alfabetización Digital; elperiodico.com; ¿Cómo funcionan los algoritmos? Así deciden las redes sociales que contenido enseñarle a tus hijos; <https://alfabetizaciondigital.redem.org/como-funcionan-los-algoritmos-asi-deciden-las-redes-sociales-que-contenido-ensenarle-a-tus-hijos/>]

4. [SDI Digital Group; 17/01/2024; Tendencias en RRSS 2024; Tendencias en TikTok para 2024; <https://www.sdi.es/actualidad/tendencias-en-rrss-2024/>]
5. [Organitzacions orientades a la dada; Josep Curto Díaz.]