Deusto
Facultad de Ingeniería
Ingeniaritza Fakultatea

Grado en Ingeniería Informática

Informatikako Ingeniaritzako gradua

Proyecto fin de grado

Gradu amaierako proiektua

Estudio sobre la privacidad en el aprendizaje federado mediante el desarrollo de un sistema de recomendación online

Ibai Guilén Pacho

Director: Diego Casado Mansilla

Bilbao, mayo de 2021

Índice general

Capítulo 1

Antecedentes y justificación

1.1. Antecedentes

Este proyecto parte del proyecto de R.Sánchez[1] donde se desarrolló un sistema de recomendación de estrategias de persuasión basado en aprendizaje activo. Para comprender los antecedentes que han incurrido en el desarrollo de este proyecto primeró hay que conocer y comprender qué son los sistemas de recomendación y después entender los problemas de seguridad y privacidad que van ligados a estos.

Los sistemas de recomendación fueron mencionados por primera vez en los años 90 y han ido evolucionando hasta estar implantados en gran parte de las empresas actuales. Estos sistemas son sistemas de filtrado de información, que partiendo de una gran cantidad de datos, tanto del usuario como de los elementos a recomendar, pueden predecir cuál será el elemento más apropiado para el usuario. Estos sistemas están estrechamente relacionados con el marketing, ya que el objetivo es conseguir recomendar un elemento que sea del agrado e interés de los usuarios suele ir ligado a fines económicos.

Hoy en día se usan en multitud de ámbitos, desde las redes sociales hasta las distribuidoras de contenido como Netflix, pasando por compañías de comercio electrónico como Amazon.

A la hora de realizar la recomendación se realizan filtrados de diferente tipo, estos no son más que la forma en la que el sistema busca correlacionar los usuarios con los elementos que estos consumen, compran o ven. Entre los métodos más comunes para relacionar esta información y obtener resultados eficientes se encuentran:

- Filtros demográficos, que recomiendan en función del sexo, edad, país, oficio, ...
- Filtros basados en contenidos, como Youtube, que recomiendan contenidos similares a los valorados por los usuarios.
- Filtrado colaborativo, que consiste en recomendar al usuario elementos valorados positivamente por usuarios similares a él.

Sin embargo, existen sistemas híbridos que utilizan varias de las estrategias de filtrado anteriores combinadas. Un ejemplo de ello es el mencionado Amazon, que tiene uno de los algoritmos de

recomendación más potentes y eficientes actualmente.

Sin embargo, existen sistemas híbridos que utilizan varias de las estrategias de filtrado anteriores combinadas. Un ejemplo de ello es el mencionado Amazon, que tiene uno de los algoritmos de recomendación más potentes y eficientes actualmente.

Esto se debe a dos cosas, en primer lugar que cuenta con una gran cantidad de información de los usuarios, tanto su edad, género, país, dirección, rutinas (si tienen Alexa), ... como los artículos que miran, compran, añaden a la lista, etc. Con todo esto el algoritmo es capaz de ofrecer recomendaciones muy precisas a los usuarios, lo que le da una gran calidad de servicio a la empresa gracias a esto.

De hecho, en Amazon, si queremos revisar las recomendaciones tenemos un apartado propio para ellas al que se puede acceder fácilmente (Fig1.1).

Fig. 1.1: Recomendaciones de Amazon (Fuente: Amazon[2])

Queda claro que un sistema de recomendación depende plenamente de la información de la que dispone, de forma que, de cuanta más información disponga de los usuarios, más precisión tendrá en sus predicciones. Para obtener esta información muchos anunciantes se han valido de diversas herramientas y utilidades, pero si se ha de destacar alguna son las cookies de terceros. Estas, son definidas por Google[**ComoBorrarHabilitar**] cómo:

> *"... son archivos que crean los sitios web que visitas para guardar información de la navegación y facilitar tu experiencia en línea. Gracias a las cookies, los sitios pueden mantener abierta tu sesión, recordar tus preferencias del sitio y proporcionarte contenido relevante en función del lugar donde te encuentres. Hay dos tipos de cookies:*
>
> - *Las cookies de origen, que las crea el sitio que visitas. El sitio se muestra en la barra de direcciones.*
> - *Las cookies de terceros, que las crean otros sitios. Parte del contenido que ves en la página web que visitas, como anuncios o imágenes, pertenece a estos sitios."*

A simple vista estas cookies parecen inofensivas, pero el problema radica en su utilización. Aunque suelen ser usadas con fines analíticos por compañías de marketing online, estas se usan para registrar el comportamiento de un usuario y crear un perfil para generar publicidad personalizada. Mediante estas se puede registrar el comportamiento de un usuario en internet, saber sus movimientos, hábitos, páginas que visita, edad, sexo, etc.

En España y en la Unión Europea existe una legislación sobre la privacidad y la protección de datos que restringe y limita los datos que estas empresas pueden recibir de nuestra navegación, el Reglamento General De Protección de Datos (RGPD). En el documento adjunto Anexo A, sección A.2 se puede consultar un breve resumen de la legislación en vigor. En lo referente a las cookies este reglamento presenta varios elementos importantes. En primer lugar, debe haber un consentimiento explicito por el usuario en el consentimiento de la política de cookies. En segundo lugar, la aceptación de las cookies de terceros no ha de ser un impedimento para el uso del servicio. En último lugar, todas las cookies y rastreadores que operen en la web del propietario deberán ser mostradas al usuario en un lenguaje sencillo.

De todos modos esta legislación no impide que se realice un perfilado del usuario, que se haga un seguimiento de este por la red o que tiempo pasa el usuario en la página. Pero con la maduración de la tecnología cada vez más gente se preocupa por el uso que le dan estas empresas a los datos y son más críticos con estas prácticas. Esto ha llegado incluso a la política, donde España esta desarrollando una carta pionera de derechos digitales, donde, entre otras cosas, recoge el derecho de los usuarios a no ser perfilados (Anexo A, sección A.3).
Hay navegadores que permiten deshabilitarlas Google las eliminará Cambio de paradigma del marketing online Respuesta de que se puede usar la msima tecnología que google para construir un sistema recomendador sin vulnerar la privacidad de los participantes

Bibliografía

[1] R. Sánchez-Corcuera, D. Casado-Mansilla, C. E. Borges y D. López-de-Ipiña, "Persuasion-based recommender system ensambling matrix factorisation and active learning models," en, *Pers Ubiquit Comput*, mar. de 2020, ISSN: 1617-4909, 1617-4917. dirección: `http://link.springer.com/10.1007/s00779-020-01382-7` (visitado 24-04-2021).

[2] *Amazon.es: compra online de electrónica, libros, deporte, hogar, moda y mucho más.* es-es. dirección: `https://www.amazon.es/` (visitado 25-04-2021).

Apéndice A

Protección de Datos en la Unión Europea y España

A.1. Preludio

Ante la llegada de las Tecnologías de la Información y la Comunicación (TIC), la Organización para la Cooperación y el Desarrollo Económicos (OCDE) decidió definir unas directrices el 23 de septiembre de 1980 que regulen *".. la protección de la privacidad y el flujo transfronterizo de datos personales ... "* [1, págs.1]. Este fue uno de los primeros intentos de regular la gestión y utilización de los datos personales.

No llegaron medidas concretas a Europa hasta el 24 de octubre de 1995, cuando el Parlamento Europeo y el Consejo Europeo publicaron la *"Directiva 95/46/CE"* [2] con el objetivo de crear un marco jurídico que garantizase la protección de los datos y la libre circulación de estos[1]. Por eso, en España se creó la *"Ley Orgánica 15/1999, de 13 de diciembre, de Protección de Datos de Carácter Personal"* [3] más conocida por sus siglas LOPD, que era la encargada de garantizar el cumplimiento de la directiva europea en España.

Ante la ausencia de leyes que garantizaran el cumplimiento de los derechos humanos en internet, algunos autores empezaron a definir los derechos digitales. Una de las primeras declaraciones de derechos fundamentales tecnológicos fue la de Robert B. Gelman, que en 1997 difundió una propuesta de *"Declaración de los Derechos Humanos en el Ciberespacio"*, basada en la Declaración Universal de los Derechos Humanos de 1948.

Sin embargo, aunque el trabajo de Robert B. Gelman fue ampliamente apoyado por una gran comunidad de expertos, no ha sido hasta la entrada en el siglo XXI (cuando las TIC ya han conseguido una gran relevancia en la sociedad) que se ha empezado a atisbar la necesidad de derechos digitales inherentes de vivir en una sociedad tecnológica. Tanto es así, que varios autores empezaron a escribir sobre la necesidad de una nueva generación de derechos humanos que incluyese el campo de la tecnología.

[1] Las directivas de la UE no son legalmente vinculantes para los ciudadanos, se dirigen a todos los estados miembros y cada uno deberá transponer la directiva a las leyes locales

La corriente pro derechos digitales llegó inclusive a la Unión Europea, que en el año 2000 redactó la *Çarta de los Derechos Fundamentales de la Unión Europea"* [4] en Niza con el objetivo de *"... reforzar la protección de los derechos fundamentales a tenor de la evolución de la sociedad, del progreso social y de los avances científicos y tecnológicos."* [4, págs.5]. El problema es que no pasó a ser jurídicamente vinculante hasta el año 2009 junto con el Tratado de Lisboa. Esta carta es especialmente innovadora en cuanto a la protección de datos de carácter personal, puesto que establecía como derecho fundamental la protección de estos en el artículo 8. En este artículo se establecía que además, estos se tratarían de modo leal y con el consentimiento de la persona afectada. Asimismo, la persona afectada tendría derecho a acceder a los datos recogidos, así como a la rectificación de los mismos.

A.2. Legislación

La legislación vigente en Europa cambió radicalmente en el 2016, el parlamento europeo adoptó nuevas medidas para la era digital.

El 5 de mayo de 2016 entró en vigor la Directiva (UE) 2016/680 [5], reconocida como la Directiva sobre protección de datos en el ámbito penal, *"... relativa a la protección de las personas físicas en lo que respecta al tratamiento de datos personales por parte de las autoridades competentes para fines de prevención, investigación, detección o enjuiciamiento de infracciones penales o de ejecución de sanciones penales, y a la libre circulación de dichos datos ..."* [5, págs.1]. La Comisión Europea define esta medida en su página web [6] como:

> *"... protege el derecho fundamental de los ciudadanos a la protección de sus datos cuando los utilicen las autoridades policiales y judiciales a efectos de aplicación de la ley. Más concretamente, la Directiva garantizará que se protejan adecuadamente los datos personales de víctimas, testigos y sospechosos de delitos, además de facilitar la cooperación transfronteriza en la lucha contra la delincuencia y el terrorismo."*

El 24 de mayo de 2016 el Reglamento (UE) 2016/679 [7], reconocido como el Reglamento General de Protección de Datos (RGPD). *"... relativo a la protección de las personas físicas en lo que respecta al tratamiento de datos personales y a la libre circulación de estos datos ..."* [7, págs.2]. La Comisión Europea define esta medida en su página web como:

> *"... es una medida esencial para fortalecer los derechos fundamentales de las personas en la era digital y facilitar la actividad económica, ya que aclara las normas aplicables a las empresas y los organismos públicos en el mercado único digital. Además, la existencia de una norma única pone fin a la fragmentación en distintos sistemas nacionales y a las cargas administrativas innecesarias."*

Esta legislación no sería aplicable hasta dos años más tarde, el 6 y 25 de mayo de 2018 respectivamente. Dos años durante los cuales cualquier empresa de la unión o cualquier empresa que tenga negocios en la Unión Europea se debía adaptar a la nueva normativa.

El objetivo principal de este reglamento era proteger a las personas físicas en cuanto al trata-

miento de sus datos personales y a la libre circulación de estos. Estos datos son de naturaleza sensible, puesto que pueden revelar la identidad de la persona, ya sea mediante el nombre completo, la dirección de su domicilio, el número de tarjeta bancaria, etc.

Entre las cuestiones más relevantes del RGPD destacan:

- La transparencia: obliga a las organizaciones a comunicar a los usuarios el tratamiento que se realiza a sus datos.
- El derecho a la rectificación y borrado: cada usuario tendrá derecho a pedir que se rectifiquen sus datos en caso de ser inexactos. Además, los usuarios tendrán derecho a que las organizaciones borren sus datos y rescindir el consentimiento de tratamiento de ellos.
- Derechos sobre el procesamiento: cada usuario podrá solicitar la limitación de procesamiento de sus datos, así como negarse a que se usen en tomas de decisiones o en perfiles automatizados, aparte, también podrá solicitar que le entreguen sus datos en formato estructurado cuando sea posible.

La directiva sobre protección de datos en el ámbito penal y el RGPD dejaron obsoleta la LOPD española, por lo cual, para adaptarse al nuevo marco jurídico europeo el 7 de diciembre de 2018 entró en vigor la *"Ley Orgánica 3/2018, de 5 de diciembre, de Protección de Datos Personales y Garantía de los Derechos Digitales"* [8] (LOPD-GDD), acorde con el RGPD.

Esta ley transpone el RGPD europeo a la legislación española, incluyendo medidas como:

- La incorporación del término privacidad desde el diseño, lo que quiere decir que deben elaborarse procesos empresariales teniendo en cuenta la LOPD-GDD desde un primer momento.
- El consentimiento, donde se obliga a las organizaciones a obtener el consentimiento explícito para el tratamiento de datos de las personas, estas personas, además, podrán solicitar la portabilidad de sus datos o la eliminación de los mismos.
- Sin embargo, la LOPD-GDD no se ciñe únicamente a adaptar el RGPD a la legislación Española, fue la primera ley europea de protección de datos en incluir explícitamente los derechos digitales de las personas en los entornos digitales.

A.3. Derechos fundamentales

Como se ha comentado en el apartado anterior, con la entrada en vigor de la LOPD-GDD España se convirtió en el primer país europeo en garantizar los derechos digitales. Estos derechos digitales llegaron en 2018 en el Título X de la LOPD-GDD, titulado *"Garantía de los derechos digitales"*, en respuesta a la Carta de los Derechos Fundamentales de la Unión Europea y a las nuevas medidas del Reglamento (UE) 2016/679 y de la Directiva (UE) 2016/680.

Varios de los derechos recogidos en la LOPD-GDD fueron:

- El derecho a la intimidad y uso de dispositivos digitales en el ámbito laboral.
- El derecho a la desconexión digital en el ámbito laboral.

- Derecho al olvido en búsquedas de Internet.
- El derecho al testamento digital.

Este paso adelante por parte del estado español incentivó a que un año después de la entrada en vigor de la LOPD-GDD, el Ilustre Colegio de la Abogacía de Barcelona (ICAB) presentase la *Çarta de Barcelona por los Derechos de la Ciudadanía en la Era Digital"* [9], apoyada por universidades y entidades de la sociedad civil.

A la vista de la relevancia e impacto del tema, el gobierno español (representado por el Ministerio de Asuntos Económicos y Transformación Digital), se inspirará en la carta del ICAB para redactar la pionera *Çarta de Derechos Digitales"* [10] de España, que contribuirá con los objetivos ya avanzados en el Título X de la LOPD-GDD. Se prevé que sea aprobada antes de 2022 para la estrategia España Digital 2025.

Esta carta cuenta con 25 puntos de alta relevancia entre los que están los recogidos en el anteriormente mencionado Título X de la LOPD-GDD, y entre otros, algunos de especial impacto en el desarrollo de este proyecto: los derechos ante la Inteligencia Artificial (Derecho XXIII) y el derecho a no ser localizado y perfilado (Derecho V).

El derecho a no ser localizado y perfilado deja claro que cada persona tiene derecho a la *"... libre autodeterminación individual (...) a no ser objeto de localización, ni a ser sometido a análisis de la personalidad o conducta que impliquen el perfilado de la persona"*, además, añade que sólo se podrán realizar *"... tratamientos de información personal con el consentimiento de la persona afectada ..."*.

En cuanto al derecho ante la Inteligencia Artificial expone como primer punto que, en lo que concierne al desarrollo y ciclo de vida de esta, *"Se deberá garantizar el derecho a la no discriminación algorítmica . . . "*. Además. en procesos de decisión automatizada las personas tienen derecho a *"Solicitar una supervisión e intervención humana"* o *Ïmpugnar las decisiones automatizadas o algorítmicas"*.

La carta, al no estar aprobada aún, no es legalmente vinculante, por lo que no es de obligado cumplimiento. Sin embargo, deja clara la orientación que van a tomar los derechos digitales en los próximos años.

Referencias

[1] OECD, *OECD Guidelines on the Protection of Privacy and Transborder Flows of Personal Data*, es. OECD, feb. de 2002, ISBN: 978-92-64-19719-0 978-92-64-19639-1. dirección: https://www.oecd-ilibrary.org/science-and-technology/oecd-guidelines-on-the-protection-of-privacy-and-transborder-flows-of-personal-data_9789264196391-en (visitado 28-03-2021).

[2] *DIRECTIVA 95/46/CE DEL PARLAMENTO EUROPEO Y DEL CONSEJO de 24 de octubre de 1995*. dirección: https://eur-lex.europa.eu/legal-content/ES/TXT/PDF/?uri=CELEX:31995L0046&from=ES (visitado 28-03-2021).

[3] *LEY ORGÁNICA 15/1999, de 13 de diciembre, de Protección de Datos de Carácter Personal*. dirección: https://www.boe.es/boe/dias/1999/12/14/pdfs/A43088-43099.pdf (visitado 28-03-2021).

[4] "Carta de los Derechos Fundamentales de la Unión Europea," pág. 17,

[5] "DIRECTIVA (UE) 2016/ 680 DEL PARLAMENTO EUROPEO Y DEL CONSEJO de 27 de abril de 2016," es, pág. 43,

[6] *La protección de datos en la UE*, es, Text. dirección: https://ec.europa.eu/info/law/law-topic/data-protection/data-protection-eu_en (visitado 27-03-2021).

[7] *REGLAMENTO (UE) 2016/679 DEL PARLAMENTO EUROPEO Y DEL CONSEJO de 27 de abril de 2016*. dirección: https://eur-lex.europa.eu/legal-content/ES/TXT/PDF/?uri=CELEX:02016R0679-20160504&from=ES (visitado 27-03-2021).

[8] M. B. Samper, *Ley Orgánica 3/2018, de 5 de diciembre, de Protección de Datos Personales y garantía de los derechos digitales*, es, 1.ª ed. Dykinson, sep. de 2020, ISBN: 978-84-13-77085-7 978-84-13-77029-1. dirección: http://www.jstor.org/stable/10.2307/j.ctv17hm980 (visitado 28-03-2021).

[9] *Carta de Barcelona por los Derechos de la Ciudadanía en la Era Digital*. dirección: https://www.icab.cat/export/sites/icab/.galleries/documents-noticies/2019/documento-carta-de-barcelona-por-los-derechos-de-la-ciudadania-en-la-era-digital-21-02-2019.pdf (visitado 01-04-2021).

[10] *Carta de Derechos Digitales*. dirección: https://portal.mineco.gob.es/RecursosArticulo/mineco/ministerio/participacion_publica/audiencia/ficheros/SEDIACartaDerechosDigitales.pdf (visitado 30-03-2021).