¿Algoritmos responsables? Las implicaciones éticas de los modelos de negocio basados en datos

El Desafío Ético del "Data Sharing"

En la actualidad, el "data sharing" o intercambio de datos ha sido fundamental para el desarrollo de modelos de negocio impulsados por datos. Este proceso permite a las empresas compartir, analizar y utilizar grandes volúmenes de información para optimizar sus operaciones, personalizar servicios y generar ventajas competitivas, no obstante, está práctica conlleva desafíos éticos. El intercambio de datos plantea constantes preocupaciones sobre la privacidad de los individuos, la seguridad de la información y el potencial uso indebido de datos sensibles. Esto tiene un fuerte impacto en el contexto de la inteligencia artificial y la ciencia de datos, ya que, la calidad y la cantidad de los datos son cruciales para el éxito de los modelos predictivos.

Adoptando una postura crítica, es esencial considerar los riesgos asociados con el intercambio de datos, uno de los principales desafíos éticos es la protección de la privacidad de los usuarios, a medida que las empresas recopilan y comparten grandes cantidades de datos, la posibilidad de que se vulneren los derechos de privacidad aumenta significativamente. Un caso concreto que ejemplifica este problema es el de la empresa Cambridge Analytica, que en 2018 fue acusada de recolectar datos personales de millones de usuarios de Facebook sin su consentimiento para influir en procesos electorales. Este incidente mostró cómo la falta de transparencia en la recopilación, almacenamiento y uso de datos puede llevar a la explotación de la información personal, vulnerando derechos fundamentales de las personas.

Además, existe el riesgo de que los datos compartidos sean utilizados para fines distintos a los inicialmente previstos, como la generación de perfiles que favorecen la discriminación o la implementación de decisiones automatizadas que refuerzan prejuicios preexistentes. Por ejemplo, en el sector de los seguros, algunas empresas han utilizado datos de redes sociales para ajustar las primas de seguro basándose en el estilo de vida del asegurado, lo cual podría llevar a prácticas discriminatorias.

Otro aspecto crucial del "data sharing" es la seguridad de la información. Los datos compartidos entre diferentes entidades pueden estar expuestos a ciberataques o accesos no autorizados, lo que pone en peligro no solo la privacidad de los individuos sino también la integridad de los sistemas que dependen de estos datos. Asimismo, la dispersión de la responsabilidad en la gestión de los datos compartidos puede complicar la identificación de los responsables en caso de una brecha de seguridad.

En el ámbito de los retos de ciencia de datos como el que estamos desarrollando sobre la predicción de supervivencia en el Titanic utilizando una base de datos de Kaggle, surgen preocupaciones éticas relacionadas con el "data sharing". Kaggle, una plataforma que permite a los usuarios subir y compartir datasets, requiere que quienes suben información cumplan con ciertas políticas de privacidad y permisos. Los usuarios deben asegurarse de que tienen los derechos necesarios para compartir cualquier información y que no están violando la privacidad de las personas involucradas. Sin embargo, en casos de información histórica o de bases de datos que contienen datos personales, como nombres, edades o antecedentes, el manejo ético de estos datos es fundamental.

Para subir datos a Kaggle, no se requieren permisos explícitos de los familiares de personas fallecidas en eventos históricos como el del Titanic, ya que esta información se considera de dominio público. Sin embargo, en otros contextos, especialmente cuando se trata de datos más recientes o sensibles, se necesitan permisos de los individuos involucrados, sus familiares o incluso del gobierno, dependiendo de la naturaleza de los datos. La privacidad de las personas debe ser protegida bajo leyes como el Reglamento General de Protección de Datos (GDPR) en Europa, que impone restricciones significativas sobre cómo se manejan los datos personales y exige el consentimiento explícito para su uso.

En el caso de México, la Ley Federal de Protección de Datos Personales en Posesión de los Particulares (LFPDPPP) establece que cualquier recopilación, uso, y almacenamiento de datos personales debe contar con el consentimiento explícito de los individuos, así como proporcionar información clara sobre el uso que se dará a esos datos. Esta ley también exige la protección de datos personales mediante medidas de seguridad administrativas, técnicas y físicas que eviten su daño, pérdida, alteración, destrucción, o uso no autorizado. Para subir datos a Kaggle desde México, sería necesario asegurarse de que cualquier información personal esté anonimizada o que se cuente con el consentimiento explícito de los involucrados. Además, en situaciones en las que los datos sean más recientes o involucren a personas vivas, los permisos explícitos de los individuos, sus familiares, o incluso de las autoridades gubernamentales pueden ser necesarios para cumplir con la normativa mexicana.

En términos legales, las empresas y plataformas como Kaggle están obligadas a cumplir con las normativas locales e internacionales sobre el manejo de información, lo que incluye garantizar la anonimización de los datos para evitar la identificación de individuos específicos. En este sentido, surge una cuestión ética en el uso de información personal: ¿hasta qué punto es apropiado utilizar datos personales, incluso anonimizados, para construir modelos predictivos? En el caso del Titanic, aunque los datos pueden parecer inofensivos, el uso de información personal para

análisis y predicciones debe considerar el impacto potencial en la privacidad de las personas.

En conclusión, el "data sharing" en el contexto de modelos de negocio impulsados por datos presenta desafíos éticos significativos que no pueden ser ignorados. Si bien el intercambio de datos es esencial para el avance de la inteligencia artificial y la ciencia de datos, es fundamental que se implementen medidas robustas para garantizar la protección de la privacidad y la seguridad de la información. Las plataformas como Kaggle y las empresas que utilizan datos deben adoptar un enfoque transparente y responsable en la gestión de los datos, asegurando que los usuarios estén plenamente informados sobre cómo se utilizan sus datos y que se respeten sus derechos en todo momento. De esta manera se podrá aprovechar el potencial del intercambio de datos sin comprometer los principios éticos fundamentales.

Referencias

Breidbach, C. F., & Maglio, P. (2020). *Accountable algorithms? The ethical implications of data-driven business models*. Journal of Service Management, 31(2), 163-185. https://www.researchgate.net/publication/341168881_Accountable_algorithms_The_ethical_implications_of_data-driven_business_models

European Union. (2016). *General Data Protection Regulation* (GDPR). Official Journal of the European Union, L119, 1-88.

Granville, K. (2018). *Facebook and Cambridge Analytica: What you need to know as fallout widens*. The New York Times. https://www.nytimes.com/2018/03/19/technology/facebook-cambridge-analytica-explained.html

Kaggle. (n.d.). *Datasets rules*. Kaggle. https://www.kaggle.com/datasets

Instituto Nacional de Transparencia, Acceso a la Información y Protección de Datos Personales (INAI). (2010). *Ley Federal de Protección de Datos Personales en Posesión de los Particulares (LFPDPPP)*. Diario Oficial de la Federación. https://www.diputados.gob.mx/LeyesBiblio/pdf/LFPDPPP.pdf