

Hiring Bias Gone Wrong: Amazon Recruiting Case Study

Luciano Davico

Descripción del caso

En 2014, el equipo de Research AI de Amazon creó un sistema automático de reclutamiento, el cual recibía una serie de currículums de personas que querían aplicar para sus vacantes y arrojaba lo que el sistema creía eran los mejores talentos. El sistema como tal realizaría una selección de los top N currículums más competentes para el cargo en cuestión, basado en modelos de Machine Learning cuyos datos de entrenamiento provenían de procesos de selección de Amazon en el pasado. El sistema se creó para automatizar la tarea de revisión de currículums y determinación de mejores prospectos para el cargo (Reuters, 2018).

Un año después, en 2015, se comenzó a detectar que el modelo tenía un claro sesgo que perjudicaba a los aplicantes de sexo femenino. En concreto, los rating a personas de sexo femenino en cargos de perfil más técnico, como ingeniero de software, no eran justos, en comparación a los rating de personas de sexo masculino. Dado lo anterior, la mayoría de recomendaciones que el modelo arrojaba sobre candidatos ideales eran hombres, lo que habla de un posible sesgo en los datos históricos utilizados para el entrenamiento de los sistemas recomendadores utilizados.

Principios éticos

Uno de los principios que se vulneró es el principio de transparencia. Según el artículo Amazon Scraps Secret AI Recruiting Engine that Showed Biases Against Women, los desarrolladores del modelo estaban al tanto del problema de sesgo a favor de los hombres, por lo que modificaron los algoritmos en pos de reducir este sesgo, sin considerar que otras variables del modelo finalmente también podrían introducir el mismo sesgo, por una cierta correlación entre variables de género que no fue controlada (ciertas key words en los CVs analizados permitían inferir que el aplicante era hombre o mujer). Esto induce a pensar que no se tenía un entendimiento amplio del modelo ni tampoco los desarrolladores de este trabajaron en la interpretabilidad de los resultados, lo que atenta con el principio mencionado más arriba. Este principio para este caso es fundamental, ya que los procesos de reclutamiento generan cierta expectativa y ansiedad a los participantes, por lo que se requiere seriedad, objetividad y la capacidad para poder entregar feedback a aquellos que no fueron seleccionados, por lo que basar la decisión en un modelo requiere que este pueda explicar las variables que más pesaron en la toma de decisiones y luego traspasar en simple esta decisión, de carácter sensible, al postulante.

Otro de los principios vulnerados es el principio de justicia. Este principio es más evidente: el proceso favorecía la contratación de hombres en desmedro de mujeres para cargos de área tech. En particular, en los requisitos de postulación se pedían habilidades no relacionadas con el género, por lo cual esto no debía influir en la decisión para la contratación. Sin embargo, no se aplicó justicia para las mujeres, incluso pudiendo haber sido más competentes en algún caso. Es más, el modelo a veces solo por leer un texto que aludiera a una mujer, ya asignaba menor probabilidad que a un hombre para aplicar, lo cual es claramente injusto y le resta legitimidad al proceso de selección que Amazon pretendía lanzar. Lo anterior ya es razón suficiente para no incluir estos algoritmos en el proceso.

Análisis

En este caso, estamos ante un caso en el cual hay más beneficios individuales que sociales, dado que en el caso se implementa un software de ayuda en toma de decisiones pero para individuos de una empresa, aunque generando un perjuicio a la sociedad al generar sesgo contra las mujeres en el proceso de selección. En este caso el software va en contra de tomar una postura consecuencialista, dado que no logra el mejor impacto para la mayoría de las personas, sino solo a un determinado grupo.

Por otra parte, la falta de responsabilidad ética en los datos, al no controlar los posibles sesgos que recaen en perjuicio para la sociedad tiene un impacto negativo mayor, que el beneficio para los individuos asociados al caso, el cual sería el beneficio de poder ser más productivos en el proceso de reclutamiento al automatizar las decisiones. Puede que durante el desarrollo del proceso no se haya comunicado de manera clara los riesgos de sesgo en el modelo, lo cual es un aspecto fundamental a considerar cuando se trabaja con sistemas de tomas de decisiones.

Referencias bibliográficas

- Roberto Iriondo. (2018, Octubre). *Insight – Amazon scraps secret AI recruiting tool that showed bias against women*. Medium.com. Recuperado de: <https://medium.datadriveninvestor.com/amazon-scraps-secret-ai-recruiting-engine-that-showed-biases-against-women-995c505f5c6f>
- Kat Chia. (2024). *Research Scientist. Hiring Bias Gone Wrong: Amazon Recruiting Case Study*. Cangrade.com. Recuperado de: <https://www.cangrade.com/blog/hr-strategy/hiring-bias-gone-wrong-amazon-recruiting-case-study/>