



Facultad de Ciencias Económicas y Empresariales
ICADE

**¿Representan Los Avances En Inteligencia Artificial (IA) Una Barrera
Para La Igualdad De Género En El Mercado Laboral?**

Autor: Barbara Vedri Almenar

Director: María Eugenia Fabra Florit

MADRID | 2024

Abstract

Este trabajo de fin de grado analiza el impacto de los sesgos en la inteligencia artificial (IA) en términos de género en el ámbito laboral. Se explora cómo la IA, utilizada en procesos de reclutamiento y selección, puede perpetuar y/o amplificar las desigualdades de género. A través de una revisión de la literatura y a través de una metodología cuantitativa basada en modelos de regresión logística se demuestra que las mujeres enfrentan mayores obstáculos para ser contratadas en comparación con los hombres. Los resultados subrayan la necesidad de mitigar los sesgos en los algoritmos de IA para promover la equidad en el mercado laboral.

Palabras Clave

Sesgos - Inteligencia Artificial - Igualdad de Género – Reclutamiento – Discriminación – Barreras de empleo – Algoritmos – Tecnología - Mercado Laboral - Oportunidades Laborales

Abstract (English)

This thesis analyzes the impact of biases in artificial intelligence (AI) in terms of gender in the workplace. It explores how AI, used in recruitment and selection processes, can perpetuate or amplify gender inequalities. Through a literature review and a quantitative methodologic analysis based on logistic regression models, it demonstrates that women face greater obstacles in being hired compared to men. The results highlight the need to mitigate biases in AI algorithms to promote equity in the labor market.

Keywords

Biases - Artificial Intelligence - Gender Equality - Recruitment - Discrimination - Employment Barriers - Algorithms - Technology - Labor Market - Employment Opportunities

Resumen Ejecutivo

Este estudio investiga cómo los sesgos en los datos utilizados para entrenar modelos de inteligencia artificial (IA) pueden perpetuar o incluso amplificar las desigualdades de género, afectando negativamente las oportunidades laborales para las mujeres. La metodología empleada combina análisis cualitativos y cuantitativos. En la parte cualitativa, se realiza una exhaustiva revisión de la literatura existente sobre la discriminación de género en la tecnología de IA, complementada con formaciones y asistencia a eventos relevantes para entender mejor la problemática. En la parte cuantitativa, se desarrollan modelos de regresión logística para analizar cómo los sesgos en los datos influyen en los resultados de los algoritmos de IA en los procesos de selección de personal. Se utiliza un conjunto de datos con 73,000 observaciones, incluyendo datos históricos y reales de procesos de selección de candidatos.

En primer lugar, se examina la implicación de las mujeres en su trayectoria profesional y las barreras que enfrentan para encontrar empleo. Estos obstáculos incluyen diferencias en oportunidades de estudio y trabajo, estereotipos de género que limitan la incursión en campos STEM y barreras en la educación superior que reducen la percepción de competencia y mentoría para las mujeres. En segundo lugar, se estudia el poder de decisión de la IA en las organizaciones y las repercusiones de su uso sin control. Aunque la IA está diseñada para eliminar sesgos, puede amplificar las discriminaciones existentes si se entrena con datos sesgados. Los sesgos pueden manifestarse de tres maneras:

1. **Discriminaciones Sociales en los Datos:** Las discriminaciones sociales se reflejan en los datos. El estudio descubre que, inicialmente, la probabilidad de ser contratado era un 10% menor si eres mujer en comparación con un hombre.
2. **Sesgo de Muestreo:** Los sesgos se introducen en la selección y recopilación de datos, donde las mujeres están subrepresentadas. En este caso, el conjunto de datos utilizado representa solo al 5% de mujeres, lo que distorsiona las estadísticas. Como resultado, la probabilidad general de contratación es del 53%, pero esta cifra se reduce en 9 puntos porcentuales para el género femenino y aumenta un punto perceptual para el masculino.

3. Prejuicios en el Diseño de Algoritmos: Los prejuicios de los creadores se transmiten a los algoritmos a través del diseño de características y métricas de evaluación. Por ejemplo, si el creador considera que las habilidades en C++ son más importantes que las de Java o Python, el algoritmo tenderá a elegir a hombres, ya que hay un 13% más de probabilidad de que un hombre tenga esta habilidad.

Tras el entrenamiento del modelo, los resultados de la prueba demuestran que este acentúa la discriminación en un 4,2% más, es decir, la nueva probabilidad de ser contratada es del 38,7%. Esto subraya la necesidad de mitigar los sesgos en los algoritmos de IA para promover la equidad en el mercado laboral.

Finalmente, se proponen soluciones para mitigar estos sesgos, como la eliminación de variables discriminatorias en los procesos de selección y la incorporación de equipos diversos en el desarrollo de estos algoritmos. Se destaca la importancia de intervenir y supervisar constantemente los modelos de IA para garantizar que sean justos y equitativos, evitando que la tecnología represente un retroceso en la lucha por la igualdad de género.

CONTENIDO

INTRODUCCIÓN	7
OBJETIVOS DE LA INVESTIGACIÓN	9
REVISIÓN LITERATURA/MARCO TEÓRICO	11
1. BARRERAS EDUCATIVAS Y DE BÚSQUEDA DE EMPLEO	11
2. DE OBSERVACIONES SOCIALES A REPRESENTACIONES NUMÉRICAS Y DECISIONES ALGORÍTMICAS	15
I. DEFINICIÓN Y FUNCIONAMIENTO DE LA INTELIGENCIA ARTIFICIAL	17
II. RIESGO DE SESGO Y COMO SE MANIFIESTA.....	19
III. LA IMPORTANCIA DE INTERVENIR.	22
IV. SOLUCIONES PROPUESTAS	23
METODOLOGÍA.....	26
RESULTADOS.....	32
CONCLUSIONES	49
LIMITACIONES.....	52
FUTURAS INVESTIGACIONES	52
BIBLIOGRAFIA:.....	54
ANEXO	60
APÉNDICE	65

Tabla de Contenidos de Gráficos y Tablas

<i>Tabla 1: Hoja de Cálculo Informativa sobre Conjunto de Datos.....</i>	<i>11</i>
<i>Tabla 2: Probabilidades de Ser Contratado.....</i>	<i>32</i>
<i>Tabla 3: Distribución de Nuevas Variables en Términos de Género.....</i>	<i>41</i>
<i>Tabla 4: AUC, Tasa FP y FN Matriz Hombre y Mujer Modelo 1.....</i>	<i>44</i>
<i>Tabla 5: Diferencias Porcentuales Matrices de Modelo Variables Originales y Modelo Nuevas Variable.....</i>	<i>47</i>
<i>Tabla 6: AUC, Tasa FP y FN (%) de Matrices Hombre y Mujer Modelo 2.....</i>	<i>48</i>
<i>Gráfico 1: Histograma de Habilidades Digitales Condicional a Contratado.....</i>	<i>33</i>
<i>Gráfico 2: Histograma de Habilidades Digitales Condicional al Género</i>	<i>34</i>
<i>Gráfico 3: Probabilidad de ser Contratado por Género y Nivel Educativo.....</i>	<i>36</i>
<i>Gráfico 4: Histograma de Salario Previo Condicional al Género.....</i>	<i>37</i>
<i>Gráfico 5: Histograma de Salario Previo Condicional a Contratado.....</i>	<i>37</i>
<i>Gráfico 6: Probabilidades de ser Empleado por Continente y Género.....</i>	<i>39</i>
<i>Gráfico 7: Probabilidades de Poseer la Habilidad y Ser o No Contratado.....</i>	<i>40</i>
<i>Gráfico 8: Curvas ROC Hombre y Mujer Modelo 1.....</i>	<i>44</i>
<i>Gráfico 9: AIC y BIC Modelo 1 vs Modelo 2.....</i>	<i>46</i>
<i>Gráfico 10: Curva ROC Mujer y Hombre Modelo 2</i>	<i>48</i>
<i>Matriz Confusión 1: Global.....</i>	<i>42</i>
<i>Matriz Confusión 2: Hombres.....</i>	<i>43</i>
<i>Matriz Confusión 3: Mujeres.....</i>	<i>43</i>

Introducción

La tecnología ha sido tanto aliada como enemiga en la lucha por la igualdad de género. Ha contribuido a mejorar el acceso a la información y oportunidades, pero también ha exacerbado las desigualdades existentes (Belingeri et al, 2021). A pesar de los avances, persisten importantes brechas de género en el ámbito laboral y social. De acuerdo con el Foro Económico Mundial, en 2023, ningún país ha logrado aún la igualdad de género, y se estima que aún quedan 131 años para alcanzarla a nivel global. En la vida laboral, la disparidad de género se manifiesta de diversas formas (Cleveland, 2004). La situación se agrava aún más cuando se interceptan las desigualdades de género y raza (Breda, 2020).

Actualmente, negocios, organizaciones e incluso gobiernos utilizan sistemas de IA como herramienta omnipresente en la toma de decisiones que afectan tanto a nivel individual como social. Esta tecnología ha experimentado un crecimiento exponencial y se le está otorgando la capacidad de tomar decisiones con mayor peso y responsabilidad. En los departamentos de recursos humanos, la integración de la IA ha sido significativa. Recientemente, se han observado varios ejemplos de cómo estas tecnologías en procesos de selección han resultado ser dañinas y discriminatorias hacia las candidatas de género femenino. Ser aceptado en un trabajo, la prioridad de un tratamiento médico o la calificación de un estudiante son ejemplos de decisiones determinadas por estos algoritmos. Esto ha generado inquietud sobre su sesgo discriminatorio y su repercusión en los derechos humanos (Felin, T. et al., 2021). Algunas discriminaciones son fácilmente identificables, como la brecha salarial o las políticas de faltas por maternidad. Sin embargo, otras formas de segregación, como los estereotipos de género y la falta de oportunidades de liderazgo, están tan arraigadas que pasan desapercibidas (C. Criado Perez, 2021). Estas desigualdades también se reflejan en los procesos de reclutamiento, donde la inteligencia artificial (IA) juega un papel cada vez más significativo.

Los modelos de IA tienen la capacidad de simular el intelecto humano, ingiriendo y procesando grandes cantidades de datos (Big Data), analizando patrones, aprendiendo de ellos y

proporcionando respuestas de manera similar a como lo haría un humano (Crawford, 2021). Sin embargo, los prejuicios humanos se impregnan en estos algoritmos. A través de la codificación y los datos sesgados que representan observaciones de nuestra sociedad, se perpetúan y refuerzan discriminaciones existentes por sistemas sociotécnicos complejos como el Internet. Estudios explican la capacidad que tiene la IA de desarrollar nuevas clasificaciones y criterios con un enorme potencial para nuevos tipos de sesgos (Ntoutsis, E., Fafalios, P., Gadiraju, U., et al., 2020), lo que plantea interrogantes sobre su impacto en las mujeres en el ámbito laboral.

A través de datos extraídos de diversas fuentes, se ha analizado la posibilidad de ocurrencia de tres tipos de sesgo mencionados: (1) Las discriminaciones sociales que se manifiestan a través de datos; (2) el sesgo de muestreo y la falta de representación de grupos minoritarios y (3) la impregnación de prejuicios de los creadores a través de establecer objetivos en los algoritmos. Este estudio utiliza un análisis cuantitativo para evaluar cómo los sesgos en los datos pueden influir en los resultados de los algoritmos de IA en los procesos de reclutamiento. Concretamente, se ha examinado una base de datos histórica, que refleja información relevante sobre candidatos a puestos STEM, estas variables incluyen, entre otras, sus habilidades tecnológicas, edad, salario previo y la respuesta de si fue o no empleado. A través de distintos análisis se verifica y se mide la posibilidad de sesgos de género como respuesta a la pregunta de investigación:

¿Serán las máquinas de IA nuestras aliadas para ayudar a cerrar las brechas de género o representan una amenaza adicional? Y si son una amenaza, ¿hay alguna forma de impedirlo?

La relevancia de esta pregunta en mi trabajo de fin de grado radica en su alcance transformador. Es esencial investigar las áreas laborales más afectadas y educar tanto en el conocimiento de estos sesgos como en las estrategias de corrección. Comprender cómo abordar estos problemas es crucial para evitar que la tecnología sea un paso atrás en la lucha feminista, garantizando que su influencia en el mundo laboral y la toma de decisiones sea justa y equitativa.

Objetivos de la Investigación

1. Identificar los desafíos laborales a los cuales se enfrenta la mujer y cómo estos se manifiestan en los datos, con un enfoque particular en el proceso de reclutamiento y selección.

- *Explorar las barreras y desigualdades que afectan a las mujeres en el ámbito laboral.*
- *Examinar cómo estas desigualdades se reflejan en los datos utilizados en los procesos de reclutamiento.*

2. Entender la trayectoria, el funcionamiento y el impacto de la tecnología AI en la toma de decisiones en las organizaciones, y describir el riesgo de sesgo y cómo la IA se plaga de prejuicios en ausencia de intervenciones.

- *Analizar el desarrollo y uso de la inteligencia artificial en los recursos humanos.*
- *Evaluar los potenciales sesgos en los modelos de IA y cómo estos afectan las decisiones organizacionales.*

3. Analizar y conocer la magnitud y el peso de la variable de género a la hora de elegir al candidato ideal, y si es así, cómo podría afectar a la mujer el uso de modelos de IA en las decisiones de contratación.

- *Evaluar la influencia de la variable de género en modelos estadísticos.*
- *Determinar el impacto potencial de la IA en las oportunidades laborales para las mujeres.*

Metodología

Para abordar los objetivos detallados, la investigación se divide en dos partes, extensa revisión de la literatura y un estudio cuantitativo.

Revisión de la literatura

Enfocada en la mujer en el mundo laboral y la discriminación de género en la tecnología AI, la revisión se centrará en fuentes clave, como "The Atlas of AI: Power, Politics, and the Planetary Costs of Artificial Intelligence" y "Invisible Women: Exposing Data Bias in a World

Designed for Men". El objetivo es identificar tendencias y hallazgos previos en este campo. Se explorarán los numerosos informes e investigaciones científicos sobre el impacto actual y futuro de esta nueva tecnología, sobre todo en el mercado laboral. Las palabras claves incluyen, pero no se limitan a: Sesgos en AI, Discriminaciones de género en el reclutamiento, el futuro en el trabajo, Impacto de los algoritmos, AI en los recursos humanos etc. El marco teórico se divide en tres puntos claves, primero la mujer en el mundo laboral, la IA como herramienta en el trabajo y, por último, los sesgos en estas herramientas y como afecta a la mujer en las organizaciones.

Estudio Cuantitativo

La investigación se centró en medir los sesgos de género en un conjunto de datos sobre reclutamiento en STEM recopilado por EKIMETRICS. Se utilizan las 73,000 observaciones para entrenar y probar dos modelos de regresión logística. Las variables del conjunto incluyen información demográfica, geográfica e información sobre la competencia del candidato como su experiencia o sus habilidades digitales. Previo al análisis, se llevó a cabo un preprocesamiento de los datos para eliminar valores atípicos y ordenar las variables. Se llevo a cabo un análisis estadístico y se desarrollaron dos modelos de regresión logística en R, los notebooks se encuentran accesibles en [GitHub](#). El Primer Modelo se entrena con las 12 variables originales del conjunto mientras que al segundo se le añaden cinco variables de habilidades tecnológicas. Ambos modelos usaron el 80% de los datos para entrenamiento y el 20% para pruebas. Se evaluaron los coeficientes y significancia estadística.

Por último, se evaluaron los modelos en términos de género. Utilizando matrices de confusión y curvas ROC se compara el rendimiento de los modelos divididos por género, estableciendo un umbral de clasificación del 0,05. El AIC se utilizó para comparar el ajuste de los modelos, evaluando el impacto de las nuevas variables de habilidades. Este enfoque permitió identificar y analizar posibles sesgos de género en la contratación en campos STEM.

Revisión Literatura/Marco Teórico

1. Barreras Educativas y de Búsqueda de Empleo

Las mujeres enfrentan obstáculos en el acceso y avance laboral desde el comienzo de su vida profesional (Field, 2023). Tras múltiples estudios distintos se ha demostrado que hay una brecha entre las oportunidades de estudio y laborales que obtienen hombres y mujeres, siendo una diferencia desfavorable para el género femenino (AAUW, 2016). La primera barrera a la que se enfrentan las mujeres es el estudiado 'sesgo de brillantez'. Desde los seis años, a las niñas se les imponen estereotipos arraigados que las convencen de su supuesta inferioridad en competencia/rendimiento académico (Lin Bian et al, 2017). Estos prejuicios limitan e influyen en las decisiones de elección de carrera, inclinándolas hacia campos que históricamente se perciben como menos dependientes de talento. Es más, los estudios han demostrado que, en los campos, como matemáticas, composición de música, física etc. donde se requiere a gente "brillante [brillante]", "genious [genios]" o de "raw talent [talento nato]" la proporción de mujeres es proporcionalmente baja (Criado Pérez, 2021). Además, el acceso a la educación superior presenta obstáculos específicos para el género femenino en contraste con el masculino, especialmente en términos geográficos y económicos (Breda, 2020). En la universidad, las evaluaciones y el feedback hacia las estudiantes femeninas tienden a ser más críticas, lo que resulta en una percepción de menor competencia y menos oportunidades de mentoría independientemente del género del profesor (Corinne Moss-Racusin, 2012). Una vez graduadas, las mujeres se enfrentan a barreras significativas en el proceso de reclutamiento y de promoción, lo que resulta en menos participación y un mayor 'burn out' (Cleveland, 2004). La mujer, de media, es significativamente más pobre, y se estima que quedan 169 años para llegar a una paridad salarial (Global Gender Gap Report, 2023). Desgraciadamente, estos son solo escasos ejemplos de lo que enfrenta la mujer en el mundo contemporáneo. En este estudio se argumenta que existen dos razones principales por las que es más complejo e improbable para una mujer acceder al mercado laboral y alcanzar el éxito: Por un lado, el diseño del mercado laboral, el cual está impregnado de prejuicios sexistas y estructurado para que, en la mayoría de las ocasiones, favorezca a los

hombres; y, por otro lado, se habla de la confianza como disfraz de competencia. Esta está alimentada por años de exposición a estereotipos de género y expectativas sociales, la mujer duda de su valía y capacidades, mientras que el hombre mantiene un exceso de confianza (Criado Pérez, 2021).

Estas barreras condicionan la vida laboral de la mujer y afecta enormemente a su participación en el mundo laboral, lo que agrava la segregación de género (Blau & Kahn, 2016). El 41.9% del mercado laboral se compone por el género femenino, mientras que solo ocupan un 32.2% de los puestos senior de liderazgo (Cobrerros, 2024). Además, la tasa de desempleo global femenina es del 4.5%, un 0.3% más alta que la de los hombres. Y aunque sean capaces de conseguir empleo, suelen enfrentar condiciones laborales deficientes, siendo la economía informal¹ una realidad para cuatro de cada cinco mujeres (ILO, 2023). No obstante, existe una discrepancia en la búsqueda de empleo entre géneros (Datta et al., 2015). Esta situación plantea la pregunta: ¿Por qué, a pesar de necesitar más un empleo estable y bien remunerado, las mujeres no buscan con mayor frecuencia? Esta diferencia se atribuye a varios factores como (i) las percepciones, (ii) la confianza, (iii) la falta de habilidades, (iv) el diseño de anuncios y de los canales de publicación de ofertas laborales y (v) la participación femenina en trabajos de cuidado no remunerados (Llinares-Insa, González-Navarro, et al., 2018). La brecha digital limita a las mujeres para buscar, conseguir y mantener un empleo. Una investigación mostró que las mujeres son un 25% menos propensas que los hombres a buscar trabajo en Internet y a tener habilidades básicas en tecnología digital (Collete, 2022). Y a pesar de que las plataformas si muestren oportunidades laborales, persiste una diferencia en habilidades digitales y en la sensación de estar menos preparadas para un nuevo empleo (Collete, 2022). La falta de habilidades se arraiga principalmente al coste asociado a ello, y gracias a los estudios sabemos que, de media, la mujer es más pobre (OECD, 2018). Esta disparidad se agudiza en países subdesarrollados, donde las dificultades para acceder a instalaciones públicas de TIC

¹ La economía informal constituye el conjunto diversificado de actividades económicas, empresas y trabajadorxs no reguladxs o protegidxs por el estado.

surgen por preocupaciones de seguridad en las calles, restricciones en la libertad de movimiento o simplemente por la percepción social de que su uso es ‘inapropiado’ para las mujeres (Petrie et al., 2021).

Hasta ahora ha sido el departamento de recursos humanos el cual analiza, diseña y anuncia la vacante de trabajo requerido (Drage & Mackereth, 2023). Estos puestos suelen tener requisitos y características que cumplen con mayor frecuencia hombres que mujeres, por lo que las mujeres pueden sentirse desmotivadas al presentar su candidatura (UNESCO, 2021). El lenguaje con el que se transmite el anuncio también tiene implicaciones (Llinares-Insa, González-Navarro, et al., 2018). Las vacantes suelen emplear un lenguaje sesgado hacia el género masculino, los estudios han demostrado que este sesgo influye en la participación femenina en los procesos de atracción de talento (Llinares-Insa, González-Navarro, et al., 2018). Una compañía europea observó un cambio significativo en la cantidad de candidatas presentadas al modificar un anuncio de un puesto técnico. En la primera prueba, aparecía una foto de un hombre y se destacaron palabras como *agresivo* y *competitivo*. La imagen se intercambió por la de una mujer y las palabras marcadas en este caso eran *innovación* y *entusiasmo*. El porcentaje de candidatas se disparó del 5% al 40% (Mihaljević, Müller, Dill, & Yollu-Tok, 2022). Otro gran factor que se debe considerar es el tiempo. Globalmente, el 75% del trabajo de cuidado no remunerado lo lleva a cabo la mujer (World Economic Forum, 2023). Al día, esto ocupa un promedio de tres a seis horas, (comparado a la media de los hombres de 30 minutos a dos horas lo que implica que la mujer promedia tiene una jornada laboral más larga (UNESCO, 2021). Depende del país la media varía, por ejemplo, en Portugal, son 90 minutos más al día; en Uganda, 15h comparado a 9h; en Reino Unido, los hombres tienen cinco horas más de ocio a la semana que las mujeres (Criado Perez, 2019). A lo anterior se le suma que las mujeres son menos adversas al riesgo (Samek, 2019). Por lo que, si no están seguras de tener las cualidades necesarias, y la probabilidad de ser contratadas es menor, deciden no invertir el tiempo para presentarse.

Tanto Thomas Chamorro (2018) como Caroline Criado (2021) plantean la pregunta de si el problema real no es tanto que las mujeres postulen menos, sino que los hombres lo hagan en exceso. Ambos sugieren que la confianza externa está ligada a una percepción de competencia, y que 'el exceso de confianza es una consecuencia del privilegio', la cual lleva a los hombres a considerarse más capaces para un puesto de trabajo. Asimismo, una mujer debe estar 110% segura de sus ideas para participar en un ambiente mayoritariamente masculino (Chamorro-Premuzic, 2017) y mientras que las mujeres suelen valorar su inteligencia de manera más acorde a su capacidad real, los hombres tienden a tener una autoimagen académica más elevada y subjetiva a la realidad (Reuben, 2010). Lo que conlleva a una desproporción entre candidatos femeninos y masculinos. Esto lleva a preguntarnos, una vez la mujer decide presentarse al puesto, si es cierto que está más segura y evalúa sus cualidades más objetivamente, ¿Por qué la probabilidad de contratar a una mujer es menor a la de un hombre? ¿Y por qué es aún menor si pertenecen a comunidades subrepresentadas?

Numerosas respuestas han surgido en torno a esta cuestión, ya que no se puede atribuir a una sola variable ni a un momento específico. En 2018 aun existían 106 países con leyes que incapacitaban el empleo a más de 2,7 mil millones de mujeres (Wood, 2018). Con el tiempo, especialmente en países desarrollados, las discriminaciones han disminuido y se han implementado más leyes para proteger a las mujeres (Burton, 2016). Anteriormente, las mujeres jóvenes eran frecuentemente discriminadas y no se le contrataba debido a la posibilidad de que pudieran quedar embarazadas, lo cual hoy en día es ilegal (Burton, 2016). La posibilidad de un embarazo es una discriminación fácil de identificar, sin embargo, un gran factor influyente y arraigado más profundamente son los prejuicios y estereotipos de género en los puestos de trabajo. Investigaciones recientes indican que los hombres tienden a sobrevalorar tanto sus propias habilidades como las de su mismo género (Bian, Leslie & Cimpian, 2017). Un estudio descubrió que los empleadores masculinos que se consideran imparciales a la hora de contratar son más propensos a seleccionar a un candidato de su mismo género, incluso si una candidata femenina está igualmente

cualificada (Moss-Racusin, 2012). Además, mientras que las mujeres tienden a evaluar a sus compañeros según su rendimiento real, los hombres tienden a clasificar más alto a otros hombres y por encima de las mujeres, incluso si estas tienen un desempeño superior (Cooper, 2018). Estas percepciones pueden influir en las decisiones de contratación, favoreciendo a los hombres y perjudicando a las mujeres (Drage & Mackereth, 2022), sobre todo en entornos laborales con alta concentración de hombres, como lo son los puestos STEM (Cobrerros, 2024). Por ello, cada vez son más las empresas que están implementando lo que llaman 'blind hiring [contratación a ciegas]' y cuando se lleva a cabo este método los resultados demuestran que el 60% de personas contratadas provienen de grupos subrepresentados, lo cual explica que no es que las mujeres u otros grupos minoritarios sean menos competentes si no que, se les percibe como tal (Drage & Mackereth, 2022).

2. De Observaciones Sociales A Representaciones Numéricas Y Decisiones

Algorítmicas

La era digital ha impulsado el uso de datos, los cuales representan información digital de nuestros comportamientos y acciones pasadas (Villars, 2011). En la sociedad contemporánea, estos han sido y son beneficiosos para informar decisiones, innovación, eficiencia, personalización de experiencias y predicciones futuras (Anugerah, 2022). Sin embargo, también han presentado desafíos. Las discriminaciones no son solo meras observaciones sociales, sino que también se infiltran en las bases de datos y las representaciones numéricas (Drage & Mackereth, 2022). Que las mujeres apliquen de menos y los hombres de más conduce a una amplia brecha de datos, e implica una falta de representación de la mujer en los datos almacenados para utilizar en procesos de reclutamiento (Criado Perez, 2019). LinkedIn descubrió que se mostraban puestos abiertos a más varones que mujeres simplemente porque los hombres buscaban nuevas oportunidades laborales más a menudo (Datta et al., 2015). Es más, tras un experimento donde anunciaba un puesto STEM en 191 países concluyó que el anuncio se mostró un 20% más a hombres que a mujeres. Otro estudio demostró que al usuario femenino se le mostraban empleos similares, pero de menor remuneración que al usuario masculino

(Datta et al., 2015). El sistema de LinkedIn se basa en datos extraídos de sus usuarios (MDPI, 2023), y las mujeres tienen de promedio un sueldo menor en puestos de misma responsabilidad, por lo que el algoritmo entrenado con estos mismos datos basa su decisión acorde con las variables presentadas, replicando el patrón, perpetuando la brecha financiera y mostrando oportunidades laborales con mayor frecuencia a los hombres que a las mujeres. En el sistema actual, la remuneración es una forma cuantitativa de valorar la competencia de los candidatos (Manurung, 2020), las máquinas son fácilmente capaces de identificar una correlación entre competencia y salario. Sin embargo, los algoritmos no son capaces de distinguir si existe o no una discriminación y por tanto esta falta de información tiene un alto coste para las mujeres. Esta disparidad revela cómo la utilización de datos sesgados en algoritmos puede perpetuar e incluso acentuar la brecha de género.

Hay una inmensidad de datos discriminatorios que podrían afectar a la decisión de contratación de puestos que requieren candidatos con alta competencia. Por ejemplo, en los países de la OCDE, el 5% de los chicos adolescentes desean perseguir una carrera en TIC², frente al 0,5% de las chicas. Además, el doble de chicos que chicas de quince años aspiran a una carrera en STEM³ (Collete, 2022). Lo que explica cómo es que solo el 29,2% de las mujeres ocupan estos puestos (World Economic Forum, 2023); Existe una diferencia del 25% en los bonos otorgados a hombres vs. mujeres en el mismo puesto (Criado Pérez, 2019); Más del 40% de las mujeres abandonan una empresa tecnológica después de 10 años frente al 17% de los varones (Criado Pérez, 2019). En los puestos CEO, la edad de las mujeres es de 4 años mayor de media, lo que se traduce a cuatro años más de experiencia, y, por tanto, se exige que estén 4 años más cualificadas para conseguir el mismo puesto (Chamorro-Premuzic, 2017). Tras el manifiesto de estas investigaciones, la conciencia sobre el uso, la gestión y la calidad de los datos están en el punto de mira (Hellström, 2020). Como explican los científicos, la calidad de los resultados depende de la calidad de la información que se introduce (Alihodzic et al, 2020). La contaminación de nuestras bases

² TIC - Tecnologías de la Información y la Comunicación

³ STEM- Siglas en inglés de Ciencias, Tecnología, Ingeniería y Matemáticas [Science, Technology, Engineer and Mathematics]

de datos es una preocupación destacada en el mundo del big data ya que revela como los sesgos arraigados se manifiestan en lo que se han considerado modelos ‘objetivos’ que influyen en nuestra toma de decisiones y políticas (Gao et al., 2020; Raffel et al., 2020). Estos deben estar supervisados en todo momento para garantizar fiabilidad y prevenir un uso perjudicial. La responsabilidad de que los datos sean representativos de todos es crucial para poder evitar que estos sigan moldeando percepciones y perpetuando sesgos (O’Neil, 2017); (Criado Pérez, 2019); (Buolamwini, 2018). El uso, y por tanto la responsabilidad, de datos se ha vuelto aún más significativa con la implementación de la tecnología de inteligencia artificial (IA).

i. Definición Y Funcionamiento De La Inteligencia Artificial

En el campo tecnológico, se presentan diversas definiciones sobre la Inteligencia Artificial, no obstante, todas ellas están de acuerdo en que la IA se enfoca en crear agentes inteligentes capaces de alcanzar objetivos (Ntoutsis, Fafalios, Gadiraju, et al., 2020). Demuestran inteligencia si toman acciones flexibles adecuadas a las circunstancias, a la capacidad de aprendizaje y apropiadas dentro de ciertas limitaciones perceptuales (Crawford, 2021). Según Bartneck, C., Lütge, C., Wagner, A., & Welsh, S. (2020), las máquinas deben extraer y procesar el significado correcto de los datos para poder concluir con las acciones deseadas. Todo el proceso requiere conocimiento sobre el mundo, la lógica, el razonamiento y las habilidades para aprender y adaptarse. Esta tecnología es un campo en constante evolución, "*it is a moving target*" (un objetivo en movimiento), la cual abarca diferentes enfoques y escuelas de pensamiento. El aprendizaje automático, conocido como Machine Learning (ML), es un componente fundamental que se utiliza en la mayoría de los algoritmos que replica comportamientos inteligentes. Se alega que en este campo los algoritmos responden preguntas y resuelven problemas, siendo capaces de asimilar conocimientos, razonar, actuar y adaptarse, imitando la inteligencia natural, y a su vez automatizando⁴ tareas realizadas por el ser humano (Drage & Mackereth, 2022). O’Neill

⁴ Refiriéndose al concepto de automatizar a la capacidad de una máquina para operar sin intervención humana, con su significado variando según la aplicación específica (Bartneck, C., 2020).

(2020) discute en su libro que al desarrollar un modelo se construye un "yo automatizado" el cual otros pueden implementar y mejorar en ausencia del creador original. Lo que implica que se debe decidir sobre cuáles son los aspectos relevantes para incluir en un modelo, para simplificar el contexto o el problema y desarrollar una versión más manejable y comprensible. De esta manera, la automatización optimiza la eficiencia de procesos, permite la transferencia de conocimiento y la toma de decisiones fundamentadas en ausencia del individuo original.

En términos sencillos, al igual que un modelo estadístico clásico, un modelo de inteligencia artificial es una versión simplificada de un proceso, el cual utiliza información pasada para intentar resolver problemas, a través de la evaluación de un conjunto de patrones, correlaciones y tendencias. A diferencia de los métodos tradicionales, en lugar de solo analizar datos y visualizar gráficos, estos algoritmos pueden generar soluciones razonando e interpretando los resultados de manera similar a los humanos (O'Neil, C., 2017). Además, los expertos reducen la posibilidad de error en sus algoritmos mediante el entrenamiento constante. Comparan los resultados con las predicciones y examinan los fallos, retroalimentando los sistemas con sus hallazgos para que estos 'aprendan de sus errores'. De esta manera se vuelven más inteligentes y perfeccionan la calidad y la fiabilidad de sus predicciones (Crawford, 2021). Por ejemplo, el modelo GPT-4 de OpenAI logró un aumento de 150 puntos en los resultados del examen SAT después de ser entrenado sobre el modelo anterior, GPT-3.5 (Goldman Sachs, 2023). Es importante mencionar que estos sistemas contienen complicados mecanismos internos y los expertos siguen sin comprender al 100% los aspectos de como los sistemas llagan a sus decisiones o predicciones (Ntoutsi, Fafalios, Gadiraju et al., 2020).

Esta tecnología ha sido alabada por su capacidad de ser objetiva a la hora de tomar decisiones, es más, las empresas de contratación por IA argumentan que estas herramientas eliminan el sesgo en el proceso de contratación y promueven la diversidad en la fuerza laboral (Drage & Mackereth, 2022). Sin embargo, el escritor de 'Atlas of AI',

Crawford (2021) contraargumenta las distintas definiciones existentes sobre la inteligencia artificial y explica que esta tecnología “no es ni autónoma ni racional, sino que es corporal y material, hecha de recursos naturales, combustible, trabajo humano, infraestructuras, logística, historias y clasificaciones”. Define la tecnología como un registro de poder, el cual refleja y sirve intereses de sus creadores. Y que, al igual que el ser humano, es víctima de estructuras políticas y sociales. Como explica O’Neil en ‘Weapons of Math Destruction’ (2016), los expertos a menudo carecen de información sobre los comportamientos que más les interesan. Por lo que recurren a datos sustitutos o proxis y establecen correlaciones estadísticas entre el código postal de una persona y su capacidad para pagar un préstamo o desempeñar un trabajo. Estas correlaciones son discriminatorias y ahora, ilegales (Al ACT). Como Crawford, O’Neil argumenta que, a pesar de su reputación de imparcialidad, los modelos reflejan objetivos e ideologías de los sus programadores. Los propios valores y deseos influyen en las decisiones, desde los datos que se eligen recopilar hasta las preguntas que se formulan. Los modelos son opiniones incrustadas en las matemáticas, lo que destaca la existencia de diversos sesgos.

ii. Riesgo De Sesgo Y Como Se Manifiesta

En 2012, se llevó a cabo un experimento, donde Corinne Moss-Racusin, investigadora de Yale envió currículums idénticos con nombres diferentes, John y Jennifer, a distintos profesores universitarios para evaluar la competencia del candidato. Los resultados mostraron que Jennifer recibió evaluaciones menos favorables y menos ofertas de trabajo que John, además de recibir un salario promedio inferior de \$4.000. Se descubrió que solo por ser mujer, Jennifer era percibida como menos competente (Moss-Racusin, 2012). Por lo que lleva a preguntarnos, Si la tasa de éxito durante tantos años ha dependido del género, ¿Qué pasaría si llenáramos de datos representantes del mundo laboral en un algoritmo? ¿Qué supondría para la mujer que fueran los algoritmos los que decidan quien es el candidato perfecto o cuales son los requisitos adecuados? En 2017, Amazon creó un sistema automatizado de reclutamiento cuyo propósito era evaluar a los candidatos en

función de su competencia para diversos roles. El modelo fue entrenado a base de enseñarle los currículos de aquellos trabajadores que fueron contratados, es decir, aquellos que fueron exitosos. Debido a la poca representación de mujeres en los puestos STEM de la empresa, el algoritmo supuso que el ser mujer era una variable de peso a la hora de descartar curriculums por lo que el sistema penalizaba a aquellos que contenían indicadores de que el candidato era mujer y aquellos que contenían la palabra “women’s”.

Antes de que se estableciera una ley, se experimentaron múltiples desenlaces donde los algoritmos fueron discriminatorios abiertamente, por ejemplo, los sistemas de Google clasificaban a gente de color como ‘gorilas’ (Salas, 2017), o durante la presidencia de Obama se categorizó la Casa Blanca como ‘Nigger House’ (Gibbs, 2015). Siri sabe responder medicamente a un ataque al corazón, pero se queda en blanco si se menciona el término ‘violación’ (Criado Pérez, 2019). Estos escándalos se deben al sesgo que se incrusta durante el entrenamiento de los algoritmos (Crawford, 2021). El sesgo, en este contexto, se refiere a las inclinaciones o prejuicios que pueden estar presentes en los datos utilizados para entrenar algoritmos. Según Thomas Hellström, Virginia Dignum and Suna Bensch (2020) los sesgos pueden manifestarse de manera diferente según la modalidad de los datos, como datos numéricos, textuales o de imágenes (sesgo de selección, sesgo de informe y sesgo de detección). La investigación identifica tres fuentes donde se introduce el sesgo en el aprendizaje automático: (1) Los propios datos, donde las discriminaciones en las estructuras sociales se traducen en observaciones y números, (2) la generación y la recopilación de datos, el sesgo de muestreo que no recoge información sobre grupos minoritarios y (3) El diseño de características y métricas de evaluación de los algoritmos establecidos por los equipos expertos, que introducen sus ideologías y objetivos.

Las preocupaciones de que exista discriminación en los sistemas no solo se centran en la perpetuación de este, si no que se ha demostrado que los modelos pueden realzar el sesgo (Varona & Suárez, 2022). En numerosas ocasiones, los sesgos son útiles y beneficiosos para las maquinas ya que ayudan a reducir los errores y mejorar la precisión de los resultados.

Reducen la variancia de las predicciones al introducir suposiciones propias (lo que denominamos sesgos) para aproximar la respuesta final. Este sesgo es necesario para obtener resultados óptimos, sin embargo, han demostrado ser discriminatorios (IBM Data and AI Team, 2023). Caroline Criado Perez (2019) denomina la exageración de estos sesgos como el efecto de amplificación “*the amplification effect*”. A pesar del entrenamiento y los datos integrados específicos, los algoritmos pueden agravar por si solos estas suposiciones y con ello la segregación de los datos. Durante el entrenamiento de un modelo procesador de imágenes se implementaron fotos de gente cocinando. En un 33% de los casos, las fotografías mostraban una cocina donde sus cocineros eran hombres. Tras el entrenamiento, el software creó clasificaciones donde el 84% eran mujeres y el 16% restante eran hombres, una reducción del último grupo a más de la mitad (Salas, 2017). A pesar del entrenamiento y los datos integrados específicos, los algoritmos pueden agravar por si solos estas suposiciones y con ello la segregación de los datos. Como señala O'Neil, el sesgo algorítmico puede ser mucho más peligroso y de mayor alcance que el sesgo humano.

En los procesos de reclutamiento, se espera que la IA actúe como herramienta para que el proceso se vuelva más común y sofisticado. La IA puede automatizar tareas del departamento de recursos humanos como la creación de descripciones de trabajo, la selección de candidatos y la programación de entrevistas, lo que podría aumentar la eficiencia y la precisión del proceso de contratación (Tanantong & Wongras, 2024). Además, se argumenta que la IA puede ayudar a personalizar las ofertas de trabajo para adaptarse mejor a las habilidades y preferencias de los candidatos, lo que podría mejorar la calidad de las contrataciones y la satisfacción laboral tanto para los empleadores como para los empleados (Tanantong & Wongras, 2024). No obstante, según la Unión Europea (2024) ‘La opacidad de muchos algoritmos puede crear incertidumbre y obstaculizar la aplicación efectiva de la legislación existente en materia de seguridad y derechos fundamentales.’ Por lo que se ha implementado un marco regulatorio basado en el riesgo. El acto de IA (2024) clasifica de alto riesgo las tecnologías de IA utilizadas en empleo, gestión de trabajadores y

acceso al autoempleo, destacando ejemplos como publicar anuncios de empleo específicos, analizar y filtrar solicitudes de trabajo y evaluar candidatos. La nueva ley exige que los poseedores de estos algoritmos creen un sistema de gestión de riesgos, aseguren calidad y representatividad de los datos, y mantengan una documentación técnica detallada. Además de realizar evaluaciones periódicas para identificar y mitigar posibles sesgos discriminatorios.

iii. La Importancia De Intervenir.

El poder de la inteligencia artificial es una (1) reforma de la naturaleza del trabajo (2) la transformación de industrias, y (3) cambios de estructuras económicas (Crawford, 2021). En el ámbito empresarial es fundamental, ya que permite extender el alcance e influencia de un individuo y transformarlo a nivel organizacional. Además, se ha revelado que estos modelos son capaces de mejorar y ser más inteligentes que el propio individuo en diversos campos (Goldman Sachs, 2023) (Anexo A). Esta tecnología ha impactado la naturaleza del trabajo y transformando los mercados de manera inadvertida durante años (Buchanan, M., 2019). Su importancia, interés e integración en distintos mercados no ha hecho más que incrementar. Si su evolución sigue como planeado, los sistemas tienen el potencial de perturbar significativamente (1) la economía global y (2) el mercado laboral. Desde 2013 hasta 2022, se han invertido alrededor de 400 mil millones de dólares en inteligencia artificial a nivel mundial (Anexo B). Según McKinsey (2022), existe una correlación positiva entre la inversión en el campo de AI y la adopción de este en organizaciones. Un 50% de las empresas encuestadas confirmaron haber adoptado esta tecnología de una forma u otra. Los servicios operacionales, la finanzas estratégicas y corporativas y los recursos Humanos son los departamentos más afectados y se espera que la adopción siga aumentando, lo que convierte a esta tecnología en la más revolucionaria de nuestra generación. Se estima que la integración de estos sistemas aumentará el PIB mundial en un 7%, y las estadísticas

sugieren que en la próxima década el crecimiento de productividad anual en Estados Unidos será de casi 1,5% (Goldman Sachs, 2022).

Por un lado, los cambios estructurales en el mercado laboral vendrán por la falta de necesidad de mano de obra en muchos puestos actuales. Se estima que dos tercios de trabajos actuales de Europa y Estados Unidos estarán expuestos a cierto nivel de automatización por AI, y en un futuro próximo, solamente el enfoque de la IA generativa tiene la capacidad de sustituir hasta un cuarto del trabajo actual (QuantumBlack By McKinsey, 2022). Por otro lado, varios estudios discuten que esta automatización de los puestos mejorará la calidad de vida del trabajador y su motivación, dará pie a que los puestos requieran más pensamiento y creatividad, dejando las partes repetitivas y burocráticas a las máquinas (Maslej, 2023) (Briggs & Kodnani, 2023). Además, se anticipa la creación de nuevas responsabilidades, el 60% de los trabajos actuales no existían en el 1940, lo que defiende que los mercados laborales se adaptan a las tendencias tecnológicas, cambian su naturaleza del trabajo y se forman nuevos puestos (Briggs & Kodnani, 2023) (Anexo C). Es más, desde el 2010 hemos presenciado un gran cambio en la demanda de habilidades y profesiones en el sector laboral. El crecimiento en la demanda sobre puestos y habilidades relacionadas con la IA ha incrementado drásticamente, siendo expertos en (ML) machine learning (1,03% sobre el mercado global) y Python (un incremento de 597% de puestos que requieren esta habilidad) lo más demandado (Anexo D). Esto indica que, si bien algunos trabajos pueden ser automatizados, también se están creando nuevas oportunidades laborales (Briggs & Kodnani, 2023).

iv. Soluciones Propuestas

Los expertos proponen soluciones para abordar el sesgo en los algoritmos de IA. Sugieren que se identifiquen los sesgos en los sistemas y se proporcione retroalimentación adecuada para que las máquinas puedan reaprender. Para lograrlo, es fundamental que los equipos que desarrollan estos algoritmos sean diversos, de modo que se incluyan diferentes perspectivas y necesidades (O'Neil, 2017); (Buolamwini, 2018); (Crawford, 2021). Caroline Criado-Pérez (2019) argumenta que, si una mujer hubiera formado parte del

equipo creador de Siri, probablemente habría destacado importantes problemas de seguridad para las mujeres, como el abuso sexual. Sin embargo, implementar esta solución presenta desafíos significativos debido a la falta de mujeres en el campo de la tecnología. Según la investigación de Era Dabla-Norris (2018), los cambios en el mercado laboral y la amenaza de pérdida de empleos impactarán más a las mujeres, con un 11% de riesgo frente al 9% de los hombres. Además, la IA está generando empleos principalmente en áreas STEM, donde las disparidades de género son más pronunciadas (Cobrerros, 2024). La probabilidad de que una mujer se convierta en programadora es cuatro veces menor que la de un hombre (Collete, 2022). Según una encuesta de la OECD, menos del 5% de los profesionales de IA, como Data Scientists y expertos en Machine Learning, son mujeres. Otra fuente estima que la representación femenina en este campo es del 30%, pero este aumento ha sido lento, creciendo solo un 4% desde 2016 (Anexo E) (OCED.AI, 2024). Esta disparidad de género refleja una tendencia persistente en la industria de la tecnología y la ciencia, donde históricamente ha habido una subrepresentación de mujeres. Por lo tanto, aunque la solución de incorporar más mujeres en los equipos de desarrollo de IA es ideal en teoría, su implementación práctica enfrenta desafíos significativos debido a la actual falta de diversidad en el sector tecnológico.

Como solución a la influencia del género en la selección de candidatos, se ha propuesto eliminar dicha variable junto a otras que han demostrado ser discriminatorias (Drage & Mackereth, 2022). Es una opción que muchos organismos, como el gobierno de reino unido o el partido político español han apoyado con el uso del 'blind recruitment process' (Gov.uk, 2024); (ValenciaPlaza, 2024). Empresas como Censia, una plataforma de inteligencia de talentos impulsada por IA, proporcionan este servicio donde se "oculta nombres, correos electrónicos, enlaces de perfiles, género y otros identificadores raciales para un grupo anónimo de candidatos" (Drage & Mackereth, 2022). Sin embargo, la propia organización admite que no pueden eliminar los prejuicios al 100%. Además, aunque se eliminan los datos característicos de un candidato aún se pueden implementar algunos códigos de identidad. Como, por ejemplo, los pasatiempos, los deportes y las experiencias

internacionales o de voluntariado de los candidatos pueden dar una idea sobre los candidatos y su entorno. Y el algoritmo identifica lo que se llaman 'perfiles' sin necesidad de especificar su género.

Metodología

Con R como lenguaje programable, la siguiente investigación se realizó mediante la siguiente investigación. Se garantiza acceso a cada notebook en [GitHub](#). Se ha desarrollado dos modelos de regresión logística. La elección de estos métodos se fundamenta en la necesidad de explorar la existencia de sesgos en los datos utilizando enfoques tradicionales ya que los algoritmos de IA utilizan un conjunto de ellos. A diferencia de métodos avanzados como el Deep Learning, cuyo proceso de toma de decisiones se desconoce, los métodos tradicionales permiten una mayor transparencia y comprensión en la interpretación de los resultados. Y se establece que un sesgo en estos modelos más sencillos podría significar un sesgo en los algoritmos más complejos. La inteligencia artificial, al basarse estrictamente en datos sin incorporar ética o razonamiento moral, requiere un análisis detallado para identificar posibles sesgos inherentes.

La tabla 1 contiene detalles sobre el conjunto de datos utilizado y sus variables:

Conjunto de datos con 73462 observaciones sobre 14 variables				
Variables	Data Type	Niveles	Información	Ejemplo
X	int	1-72091	Numero Candidato	1;2;3..
Employed	Factor	2	Decision de contratacion de la empresa	0;1
Gender	Factor	2	Genero candidato	Man; Woman; NonBinary
Age	Factor	2	Edad candidato: mayor o menor de 35	<35;>35
EdLevel	Factor	2	Nivel Educativo: Instituto, Grado, PostGrado, Doctorado u Otro	Master; NoHigherEd; ..
Employment	Factor	2	En posesión de un puesto de trabajo al aplicar a un puesto	No; Yes
MentalHealth	Factor	2	Registro de sufrir alguna enfermedad mental	No; Yes
MainBranch	Factor	2	Candidato aplica a un puesto STEM	Dev; NotDev
YearsCode	int	0-50	Cantidad de años que el candidato ha estado programando	1;2;3..
YearsCodePro	int	0-50	Cantidad de años que el candidato ha estado programando como profesión	1;2;3..
Country	chr	172	Pais de origen de candidato	Sweden; Spain; United States
PreviousSalary:	num	0-224000	Salario del candidato en el trabajo anterior al anunciado	51552
HaveWorkedWith:	chr	-	Habilidades tecnológicas con la que el candidato tiene experiencia	C++; Python; Git; PostgreSQL
ComputerSkills:	int	0-107	Cantidad de habilidades tecnológicas que posee el candidato	4

Tabla 1- Hoja de cálculo informativa sobre conjunto de datos

CSV:

	Age	EdLevel	Emp	Gen	Ment	Mai	Year	Years	Count	Previous	HaveWorkedWith	Comp	Employ
0	<35	Master	1	Man	No	Dev	7	4	Swede	51552.0	C++;Python;Git;P	4	0
1	<35	Undergra	1	Man	No	Dev	12	5	Spain	46482.0	Bash/Shell;HTML/	12	1
2	<35	Master	1	Man	No	Dev	15	6	Germa	77290.0	C;C++;Java;Perl;F	7	0
3	<35	Undergra	1	Man	No	Dev	9	6	Canad	46135.0	Bash/Shell;HTML/	13	0
4	>35	PhD	0	Man	No	NotI	40	30	Singap	160932.0	C++;Python	2	0
5	<35	Master	1	Man	No	Dev	9	2	France	38915.0	JavaScript;Python	5	0

Recopilación de datos

Los datos que se han utilizado en la investigación fueron recopilados a través de diversas fuentes, como portales de empleo, ferias de reclutamiento y aplicaciones en línea, por la compañía EKIMETRICS, la cual organiza distintos hackáthones centrados en concienciar sobre la responsabilidad y la ética detrás de la inteligencia artificial. Este conjunto de datos consta de alrededor de 73,000 observaciones e incluye varias variables que describen a los candidatos reales sobre puestos de STEM y la siguiente decisión de si fueron o no aceptados por la compañía. Algunas de las variables clave en el conjunto de datos son: la edad del candidato (categorizada como mayor o menor de 35 años), el nivel educativo del solicitante (como pregrado, maestría, doctorado, entre otros), su género (hombre, mujer o no binario), si el candidato proviene de un base profesional técnica, la cantidad de años que el solicitante ha estado programando, y la cantidad de años que ha estado programando en un contexto profesional, el salario anterior del candidato, el número de habilidades informáticas que posee, y su país de origen del solicitante.

Preprocesamiento de los datos

Tras un breve estudio de los datos se realizaron varios pasos de preprocesamiento de datos antes de proceder al análisis. En primer lugar, se eliminaron todos los valores nulos del conjunto de datos para garantizar la integridad y la calidad de los datos. Y tras un análisis

de las variables se eliminan los valores atípicos de las variables numéricas (ComputerSkills, PreviousSalary y YearsCode), elimina los candidatos que se clasifican como No Binarios y la variable de have Worked with ya que es una variable tipo caracteres con desorden. Además, se llevó a cabo una categorización de la variable 'Country', agrupando cada país en su continente correspondiente. Esta acción simplifica el análisis geográfico y facilita la comparación entre regiones. Como resultado de esta categorización, se creó una nueva variable llamada 'Continents'. Esta categorización también puede ser útil para visualizar e identificar patrones o tendencias relacionadas con la ubicación geográfica en el análisis posterior. Este conjunto con las variables originales se utiliza para el primer modelo. Para el segundo modelo de regresión logística, se procedió a crear nuevas variables nominales a partir de la variable que describe las habilidades de los candidatos [Have Worked With]. Estas nuevas variables se centraron en habilidades tecnológicas clave en el campo STEM, como SQL, Python, Cloud Services y Java. La creación de estas variables permite una representación más detallada y específica de las habilidades de los participantes, lo que facilita el análisis posterior. Los valores en estas nuevas variables se codificaron como 1 si el participante posee la habilidad correspondiente y como 0 si no la posee.

Conjunto datos modificado modelo 2:

72091 OBSERVACIONES CON 20 Variables				
Variables	Data Type	Niveles	Información	Ejemplo
X	int	1-72091	Numero Candidato	1;2;3..
Employed	Factor	2	Decision de contratacion de la empresa	0;1
Gender	Factor	2	Genero candidato	Man; Woman
Age	Factor	2	Edad candidato: mayor o menor de 35	<35;>35
EdLevel	Factor	2	Nivel Educativo: Instituto, Grado, PostGrado, Doctorado u Otro	Master; NoHigherEd;..
Employment	Factor	2	En posesión de un puesto de trabajo al aplicar a un puesto	No; Yes
MentalHealth	Factor	2	Registro de sufrir alguna enfermedad mental	No; Yes
MainBranch	Factor	2	Candidato aplica a un puesto STEM	Dev; NotDev
YearsCode	int	0-50	Cantidad de años que el candidato ha estado programando	1;2;3..
YearsCodePro	int	0-50	Cantidad de años que el candidato ha estado programando como profesión	1;2;3..
Country	chr	172	Pais de origen de candidato	Sweden; Spain; United States
PreviousSalary:	num	0-224000	Salario del candidato en el trabajo anterior al anunciado	51552
HaveWorkedWith:	chr	-	Habilidades tecnologicas con la que el candidato tiene experiencia	C++; Python; Git; PostgreSQL
ComputerSkills:	int	0-107	Cantidad de habilidades tecnologicas que posee el candidato	4
SQL	Factor	2	Experiencia con SQL	0;1
python	Factor	2	Experiencia con Python	0;1
java	Factor	2	Experiencia con Java	0;1
c_plus_plus	Factor	2	Experiencia con C++	0;1
cloud	Factor	2	Experiencia con programas de software en la nube	0;1
Continent	Factor	7	continente	Africa; Asia;..

Análisis Estadístico

Medias, probabilidades y distribuciones binomiales

Se realiza un análisis exploratorio de los siguientes datos, utilizando histogramas, gráficos de densidad, gráficos de barras y mapas de calor. Se examina principalmente la relación entre la variable 'Employed' y las diferencias entre géneros. Se investigan las probabilidades basadas en nuestro conjunto de datos y se exploran las correlaciones entre variables. Asumiendo que las variables son independientes, para las variables categóricas se hace una prueba estadística de X^2 , mientras que para las continuas se hacen pruebas de distribución binomial normal. Dado que existe una amplia brecha de datos entre mujeres y hombres el intervalo de confianza se establece en 95%, por lo que el nivel de satisfacción es de 5%. Si la hipótesis nula H_0 es cierta, la variable es irrelevantes en cuanto a si tienen un efecto sobre la decisión de ser contratado o no. A esto le sigue un examen del peso de la evidencia y el valor informativo para identificar qué tan importante y relevante es la variable.

Modelo de Regresión Logística

Para esta investigación se crean dos modelos:

- (1) El primero modelo contiene las 12 variables originales del conjunto de datos extraído de Ekimetrics.
- (2) Un segundo modelo con las mismas variables, pero se le añaden 5 variables categóricas sobre las habilidades informáticas del candidato en estas entran la utilización de: Python, SQL, Java, C++ o Cloud Computing.

Para ambos modelos, la variable independiente es Contratado [Employed = 1], mientras que el resto de las variables son dependientes y se centra el foco en la variable de género. Los datos se dividen en dos para el análisis:

- (1) el 80% de los datos se utiliza para entrenar los modelos y se crea un resumen con el p-valor de cada variable dependiente

(2) el 20% restante se utiliza para evaluar el rendimiento de los modelos

Datos de entrenamiento (Training Set)

El resumen del modelo proporciona información sobre los coeficientes estimados y su significancia estadística en el modelo de regresión logística. Los coeficientes representan el efecto de cada variable sobre la probabilidad de ser contratado ($\text{Employed} = 1$). El valor del coeficiente indica la magnitud del cambio en la probabilidad de la variable dependiente ($\text{Employed} = 1$) por una unidad de cambio en la variable independiente, manteniendo constantes las demás variables. Un coeficiente positivo indica que un aumento en la variable independiente se asocia con un aumento en la probabilidad de la variable dependiente, mientras que un coeficiente negativo indica lo contrario. Para establecer si esta variable es o no relevante se observó el valor p , el cual indica la extremidad de la probabilidad. Un valor p menor al nivel de significancia establecido (0,05) indica que es poco probable que el efecto observado sea coincidencia, y se considera que la variable dependiente tiene un efecto estadísticamente significativo sobre la variable independiente.

Datos de prueba (Test Set)

Tras la realización de la prueba de evaluación se ilustran los resultados en tres matrices de confusión: una matriz global y unas matrices para cada género. De esta manera se evalúan los valores de la sensibilidad y precisión del modelo en base a los distintos géneros. Para la verificación de si existe o no sesgo de género se establece un umbral de clasificación de 0,05, Si la probabilidad asignada aleatoriamente es mayor que 0,05, la predicción es "empleado" (clase 1). De lo contrario, la predicción es "no empleado" (clase 0).

Además, para establecer un rendimiento de referencia para la comparación con el modelo de regresión logística, se utilizó un clasificador aleatorio. Este clasificador realiza predicciones para el resultado binario (empleado o no empleado) de forma totalmente aleatoria. La curva ROC para el clasificador aleatorio se construye analizando los verdaderos positivos, los falsos positivos, la sensibilidad y la especificidad en diferentes umbrales de clasificación. A medida que el umbral varía de 0 a 1, se calculan la tasa de

verdaderos positivos (TPR) y la tasa de falsos positivos (FPR). La curva ROC representa la TPR (también conocida como sensibilidad) en el eje y contra la FPR (1 - especificidad) en el eje x. La curva ROC del clasificador aleatorio sirve como punto de referencia. Un clasificador perfecto alcanzaría un AUC (Área bajo la curva) de 1.0, lo que indica una separación perfecta entre las clases positiva y negativa. Por el contrario, un clasificador aleatorio tendría un AUC de 0.5, lo que significa que no es mejor que adivinar al azar. Al comparar el AUC del modelo de regresión logística con el AUC del clasificador aleatorio, podemos evaluar qué tan bien funciona el modelo para distinguir entre individuos empleados y no empleados.

Modelo de Regresión Logística 2 (Adición de Variables Habilidades)

Se investiga cuáles son y se establecen las habilidades de STEM más importantes a la hora de contratar como variables adicionales al modelo. Se cree que estas variables ayudaran a predecir mejor si el candidato debe o no ser contratado dado que son las habilidades más buscadas. Las variables se extraen de la columna original 'Ha Trabajado Con [Have Worked With]', es decir, si el candidato ha añadido SQL, MySQL o PostgreSQL se clasifica como Si (SQL = 1), mientras que si el candidato no menciona la habilidad se marca como No (SQL = 0). El mismo proceso se repite simultáneamente con las 4 variables restantes. Por último, se repite el mismo análisis del modelo 1. El Criterio de Información de Akaike (AIC) compara modelos de regresión logística, proporcionando una evaluación global del rendimiento del modelo. Un AIC más bajo indica un mejor ajuste del modelo, considerando su capacidad de explicar la varianza de la variable dependiente como su simplicidad. Se utilizó esta última métrica para comprar ambos modelos y se comparan los resultados de las matrices en el rendimiento para entender el impacto de introducir nuevas variables de habilidades.

Resultados

I. Genero Y Contratado [Gender & Employed]

El objetivo de esta investigación es demostrar si la variable de genero ha sido o no decisiva a la hora de contratar en el pasado, y por tanto si un modelo regresivo puede o no estar sesgado. Es importante mencionar que existe una brecha de datos sobre el género importante ya que, en esta muestra, solamente el 5% de las observaciones pertenecen a la mujer. Por lo tanto, existe un sesgo inicial debido a la desproporción en la representación de los géneros. En base a la muestra la probabilidad de que un candidato sea contratado es de 53,6%. En términos de género, existe una discrepancia entre hombre y mujeres; la probabilidad de que un hombre sea contratado es de un 54,1% (un 0,5% por encima de la media) mientras que la de la mujer es de un 44,8% (un 8,8% por debajo de la media).

Probability <chr>	Probability_Value <dbl>
P(Employed)	0.5362805
P(Woman and Employed)	0.4488346
P(Man and Employed)	0.5407668

Tabla 2- Probabilidades de ser Contratado

Los resultados del test Chi-cuadrado indican una asociación estadísticamente significativa entre el género y la probabilidad de ser contratado (Chi-cuadrado = 113.36, $p < 0.001$). La hipótesis nula se descarta ya que es poco probable que la diferencia observada en las tasas de contratación de hombres y mujeres se deba al azar. El Peso de la Evidencia (WOE) indica la fuerza de predicción de cada género. Para los hombres, el WOE es bajo y positivo (0.01805), sugiriendo que su género aporta poca información sobre la probabilidad de ser contratados. En cambio, las mujeres tienen un WOE negativo mayor (-0.35076), lo que indica que ser mujer disminuye significativamente la probabilidad de ser contratado. Sin embargo, el Valor Informativo (IV) de la variable "género" se acerca a cero (0.0063), categorizándola como no predictiva. Esto podría deberse a la poca información disponible sobre las mujeres en el conjunto de datos. Por lo que, aunque existe una diferencia en el WOE entre hombres y mujeres, el género no parece ser un predictor fuerte del

estado de empleo en este conjunto de datos. Es probable que otras variables sean más relevantes para la predicción del empleo. Este análisis evalúa la variable de forma independiente. Es posible que, en conjunto con otras variables, el género sí tenga un papel predictivo importante.

II. Variables Originales

i. Habilidades Informáticas [Computer Skills]

La media de habilidades que posee un candidato es de 13.42 y la moda es de 11. El gráfico muestra la distribución condicional de las habilidades informáticas (computerskills) de los individuos, separada por su estado de empleo. Se ve una posición positivamente sesgada para ambas distribuciones. No obstante, se ve una diferencia entre los dos grupos. A medida que aumenta el número de habilidades, la proporción de individuos empleados incrementa notablemente. Esto

sugiere que los reclutadores de talento prefieren candidatos con un rango de 10-20 habilidades. Además, aquellos candidatos que poseen más de 30 habilidades son altamente propensos a ser contratados. Por otro lado,

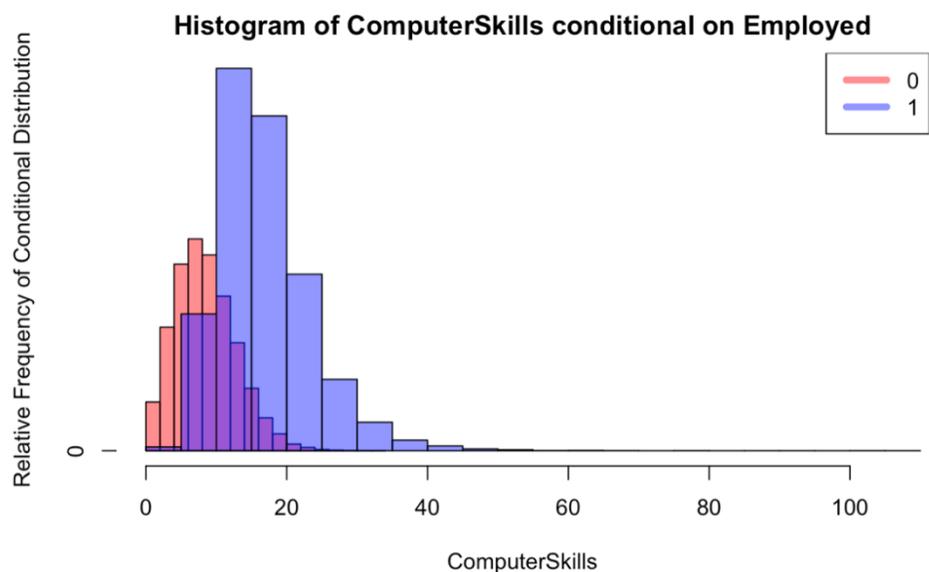


Gráfico 1- Histograma de Habilidades Digitales Condicional a Contratado

los candidatos no contratados tienden a tener una concentración de habilidades entre 5 y 10, y es menos probable que sean empleados si poseen menos de 5 habilidades. En resumen, se puede concluir que las habilidades informáticas juegan un papel importante en la empleabilidad del candidato. Aquellos con más habilidades tienen mayores probabilidades de ser empleados, mientras que la falta de estas reduce significativamente las oportunidades de empleo.

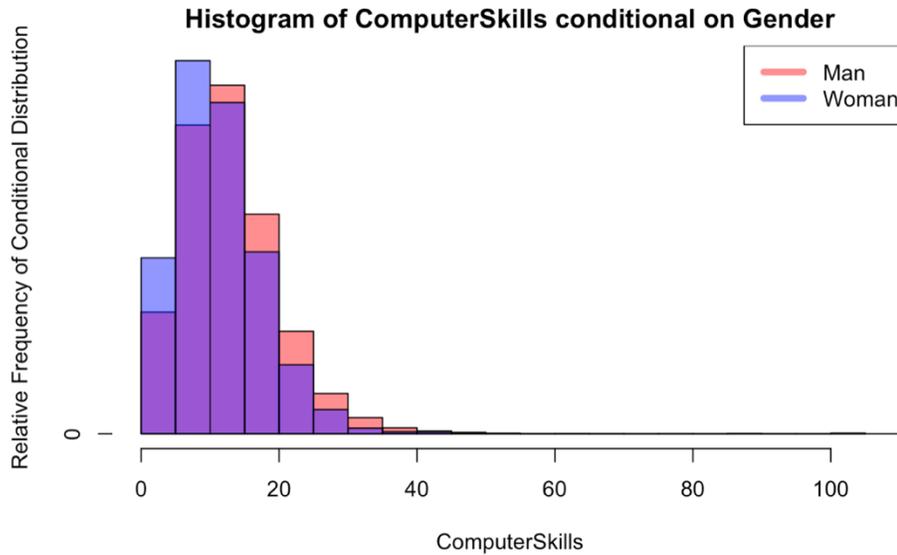


Gráfico 2- Histograma de Habilidades Digitales Condicional al Genero

El grafico 2 muestra la diferencia entre las distribuciones de mujeres y hombres, ambos géneros están sesgados positivamente sin embargo es notable que las mujeres tienden a tener menos habilidades y su distribución esta más concentrada donde el número de habilidades es menor. Los hombres tienen una distribución más amplia de habilidades informáticas, con algunos individuos alcanzando hasta 40 o más habilidades. Mientras que las candidatas femeninas tienden a concentrarse más en el rango de 5 a 15 habilidades, y pocas tienen más de 20 habilidades. Esto sugiere que tener más habilidades informáticas no solo está asociado con una mayor probabilidad de empleo, sino también que los hombres tienden a reportar más habilidades informáticas que las mujeres. Lo que implica que Los hombres, al tener una distribución más amplia y mayor cantidad de habilidades, podrían tener una ventaja en términos de empleabilidad en roles que requieren múltiples habilidades informáticas. El análisis de los dos gráficos muestra que las habilidades informáticas son un factor clave en la empleabilidad y que existe una diferencia de género en la distribución de estas habilidades.

ii. Rama Principal [Main Branch]

Esta variable define a los candidatos de puestos STEM como desarrolladores de código o no desarrolladores. Los desarrolladores (Dev) tienen una mayor proporción de empleo (0,55) en comparación con los no desarrolladores (NotDev) (0,35); sin embargo, estos

últimos componen solo un 5,3% del conjunto total. La categoría "Dev" tiene un WOE positivo, indicando una mayor probabilidad de empleo, mientras que "NotDev" tiene un WOE negativo, sugiriendo una menor probabilidad de empleo. El IV sugiere que la categoría de Dev (0.004192734) aporta poca información adicional para predecir la contratación y para "NotDev" el valor IV es más alto (0.051840784), lo que indica que esta categoría tiene una mayor capacidad predictiva respecto a la probabilidad de empleo. Las mujeres tienen una representación ligeramente mayor en la categoría de no desarrolladores.

Se examina la relación entre el género (hombres y mujeres) y la categoría de MainBranch (Desarrolladores y No Desarrolladores) con el Test de Chi-cuadrado. Dado que el valor p es mucho menor que 0,05, podemos rechazar la hipótesis nula de que no existe una relación entre el género y la categoría de MainBranch. Esto significa que hay una asociación estadísticamente significativa entre el género y si un individuo es un desarrollador (Dev) o no (NotDev). Simultáneamente, la prueba Chi-cuadrado de esta variable contra el ser contratado también concluye que existe una asociación significativa entre el estado de empleo y el tipo de rol (Desarrollador vs. No Desarrollador).

iii. Nivel Educativo [EdLevel]

Esta variable establece el nivel educativo de los participantes. Los datos indican que la probabilidad de ser contratado es mayor si se tiene un título universitario (53%), seguido por un máster (23%) y otros niveles educativos (16%). Por otro lado, la probabilidad de ser contratado sin poseer una educación superior es del 5,5%, y si se tiene un doctorado, del 1,9%; ya que solo el 10% del conjunto no tiene educación superior y el 7,3% posee un doctorado. Sin embargo, si se analiza la probabilidad de ser contratado dentro de estas categorías específicas, se observa que tener un doctorado conlleva una probabilidad de contratación del 29%, mientras que aquellos candidatos que no tienen una educación superior tiene una probabilidad del 0,59 de ser contratados. En términos absolutos y porcentuales, un mayor número de hombres no han tenido una educación superior mientras que en términos porcentuales las mujeres son más propensas en tener un

doctorado. El grafico 3 representa las probabilidades de ser contratado dependiendo del nivel educativo y el género. Esto sugiere que la probabilidad de ser contratado siendo una

con mujer un

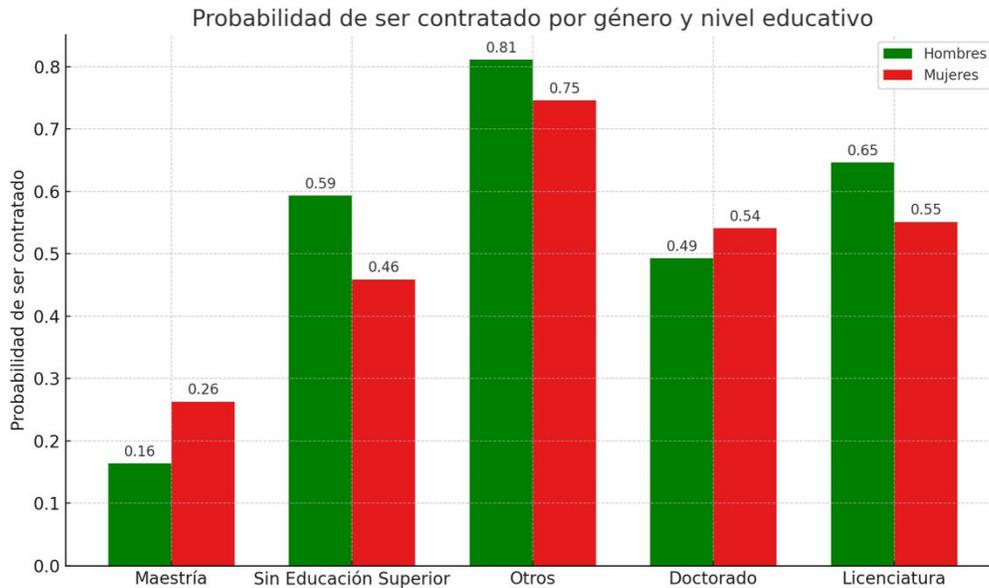


Gráfico 3- Probabilidad de ser Contratado por Genero y Nivel Educativo

doctorado es aproximadamente del 54%. Mientras que la probabilidad de ser contratado siendo un hombre sin educación superior es aproximadamente 59,3% (vs. un 45,9% de la mujer sin educación superior).

No obstante, el análisis de los niveles educativos y su relación con la empleabilidad muestra que aquellos con un Doctorado (PhD) o un grado (Undergraduate) tienen una ligera ventaja en términos de empleabilidad, reflejada por un WOE positivo. Los individuos con un postgrado (Master) o sin educación superior (No Higher Education) tienen una ligera desventaja, reflejada por un WOE negativo. La categoría de Doctorado y Grado también tiene el mayor valor de información (IV), lo que indica que esta categoría es más predictiva para la empleabilidad en comparación con las otras. En resumen, un mayor nivel educativo parece estar asociado con una mayor probabilidad de empleo.

Salario Previo [Previous Salary]:

Los histogramas A y B muestran que la mayoría de las personas con salarios bajos están empleadas, y la proporción de contratados disminuye ligeramente a medida que el salario aumenta. A medida que el salario aumenta, hay una mayor proporción de no empleados, pero los empleados siguen siendo mayoría en todos los rangos de salario.

En cuanto al género, los hombres tienen una mayor representación en todos los rangos de salario, especialmente en los salarios más altos. Las mujeres están presentes en todos los rangos de salario, pero su proporción disminuye a medida que el salario aumenta, especialmente en los rangos superiores a 100,000. Estos resultados sugieren que tanto el género como el estado de empleo pueden influir en la distribución del Salario previo, con diferencias notables en la representación en diferentes rangos de salario.

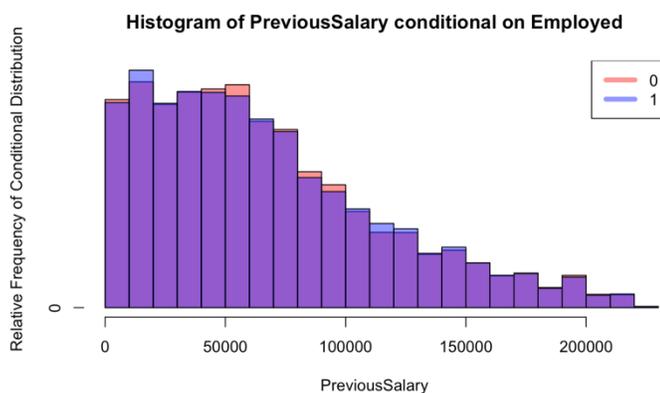


Gráfico 5- Histograma de Salario Previo Condicional en ser Contratado

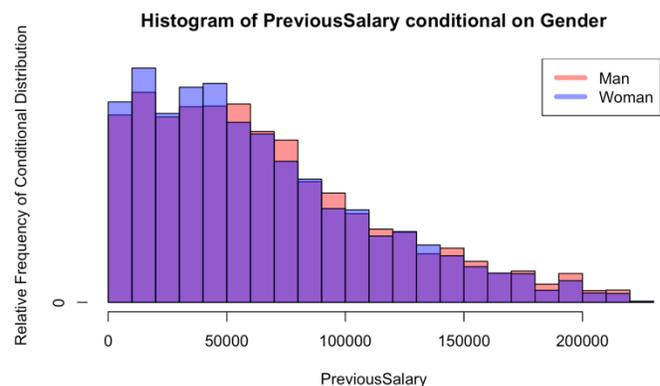


Gráfico 4 - Histograma de Salario Previo Condicional al Genero

La propuesta de rango de valores (bins) para el WOE y VI se basa en la distribución observada en un histograma de 10 bins sobre el salario previo, tiene como objetivo capturar mejor la variabilidad en los datos de PreviousSalary. Ningún rango muestra una fuerte asociación con la probabilidad de ser contratado, ya que todos los valores de WOE están cercanos a cero. Los valores de IV son también muy bajos, indicando que PreviousSalary aporta poca información adicional para la predicción de la probabilidad de empleo en este modelo. Esto sugiere que PreviousSalary no es un predictor fuerte de la empleabilidad en

este conjunto de datos, y la segmentación actual en bins no proporciona una gran diferenciación en términos de probabilidad de ser contratado.

Continentes

En términos proporcionales, el análisis de las frecuencias relativas de empleo por continente revela que África tiene la mayor proporción de contratados (0,61) en comparación con los demás continentes. Asia, Europa, Norteamérica, Oceanía y Sudamérica muestran una ligera mayoría de candidatos empleados sobre los desempleados, con proporciones de contratados entre 0,52 y 0,56. La probabilidad de ser contratado si te encuentras en categoría "Other" es de 0,51. En cuanto al conjunto, Europa y Norteamérica son los continentes con la mayor representación en ambos estados de empleo, sugiriendo una alta presencia en la muestra total. En cambio, continentes como Oceanía, África y Other tienen una representación significativamente menor, lo que podría influir en la variabilidad de los datos y en el análisis predictivo basado en la variable "Employed".

Se observan diferencias significativas en la capacidad predictiva de las distintas categorías. Aunque la categoría correspondiente a Europa tiene el mayor número de observaciones (28.903), su bajo valor de WOE (-0,06) y Valor Informativo de 0,003 indica que su capacidad predictiva es limitada. En contraste, la categoría "Other", con el menor número de observaciones (2.150), muestra un alto valor de WOE (0,32), lo que sugiere una mayor capacidad predictiva a pesar de su baja frecuencia. El conjunto de África, Oceanía y Asia presenta el valor más alto de VI (0.005), lo que indica que es la más predictiva en términos generales.

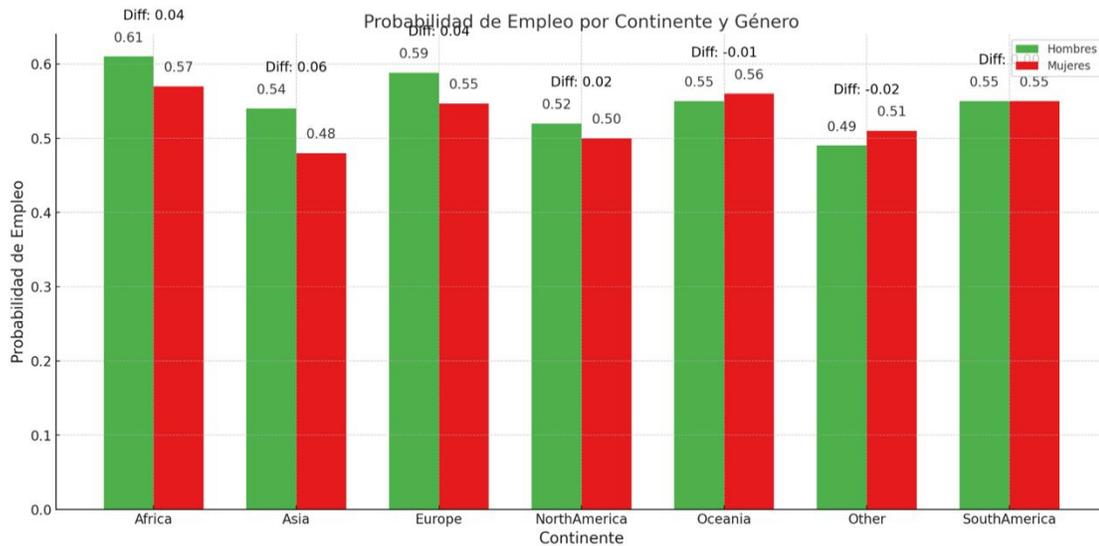


Gráfico 6- Probabilidades de ser Empleado por Continente y Genero

Al observar la distribución de género en estos continentes las diferencias más notables se encuentran en los continentes con mayor número de observaciones, en Europa y Norteamérica la frecuencia relativa entre hombres y mujeres tiene una brecha de 0,05 y de 0,10 respectivamente, habiendo una mayor representación de hombres en estos datos. En otros continentes, las diferencias son menores (menos de 0,03), lo que sugiere una distribución más equilibrada entre géneros. El gráfico muestra la probabilidad de empleo por continente y género. En general, en la mayoría de los continentes, los hombres tienen una probabilidad ligeramente mayor de estar empleados en comparación con las mujeres, excepto en Oceanía y en la categoría "Otros", donde las mujeres tienen una ligera ventaja.

Edad [Age]

Esta variable se presenta como una variable categórica donde los participantes se clasifican en <35 o >35. Se observa que las personas menores de 35 años tienen una mayor probabilidad de contratación en comparación con las personas mayores de 35 años.

La diferencia en la probabilidad de empleo entre los dos grupos de edad es del 3% (55% y 52%). En los datos hay una mayor representación de jóvenes menores de 35. Sin embargo, la proporción de menores de 35 años es ligeramente mayor entre los empleados (66%) en

comparación con los no empleados (64%). Estos gráficos indican que la edad es un factor en la empleabilidad, con una ligera ventaja para los menores de 35 años en términos de probabilidad de estar empleados. Sin embargo, en general, los menores de 35 años representan la mayoría tanto entre los empleados como entre los no empleados. De las mujeres candidatas el 73% de ellas pertenecen al grupo de menor edad, comparado al 65%.

III. Nuevas variables

El gráfico 7 muestra el porcentaje de empleados y no empleados que poseen las habilidades específicas (ej. SQL = 1). Observamos que tener o no la habilidad de Python no es significativa a la hora de contratar ya que de los grupos de gente que fueron contratados o no el 42% de los candidatos poseen esta habilidad. Sin embargo, si el candidato posee el resto de las habilidades digitales la probabilidad de ser contratado es mayor. Siendo Java y SQL las más diferenciadas.

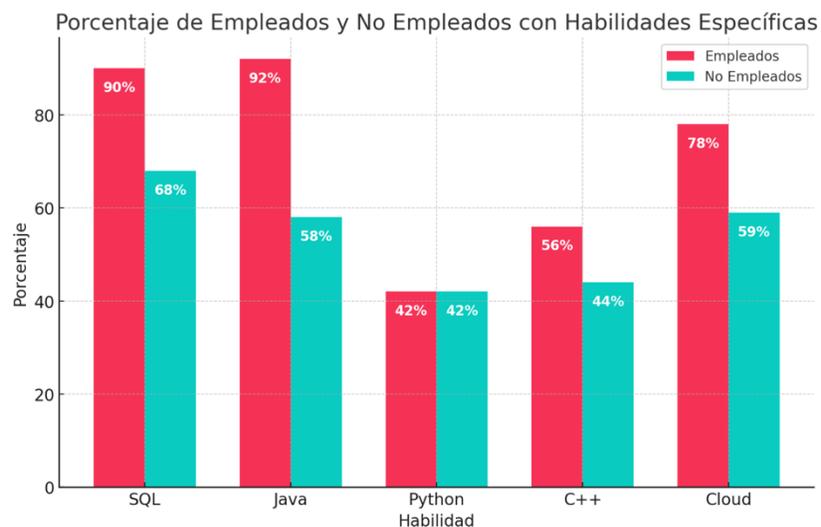


Gráfico 7- Probabilidades de poseer la habilidad y ser o no contratado

Es más, la tabla 2 muestra las distribuciones de las nuevas variables en términos de género. En términos relativos, el promedio de mujeres que han trabajado con estas tecnologías es de 57,44% mientras que el promedio de hombres utilizando estas tecnologías es de 62,62% (diferencia de 5,18%). La proporción de mujeres que controlan Java es un 0,5 superior a la de los hombres, mientras que el resto de las habilidades las dominan mayoritariamente los

hombres con una diferencia de entre 3-12%. En términos absolutos son 2,029 mujeres contra 42,931 hombres.

	Mujer	Hombre	Diferencias
	Si (%)	Si (%)	Si (M - H) %
SQL	74,8	79,8	-5
Python	39,3	42,4	-3.1
Java	76,9	76,3	+0.6
C++	32,8	45,2	-12.4
Cloud Computing	63,4	69,4	-6
AVERAGE	57,4	62,6	-5.2

Tabla 3 - Distribución de Nuevas Variables en Términos de Género

Modelo Regresión Múltiple 1 – Variables Originales

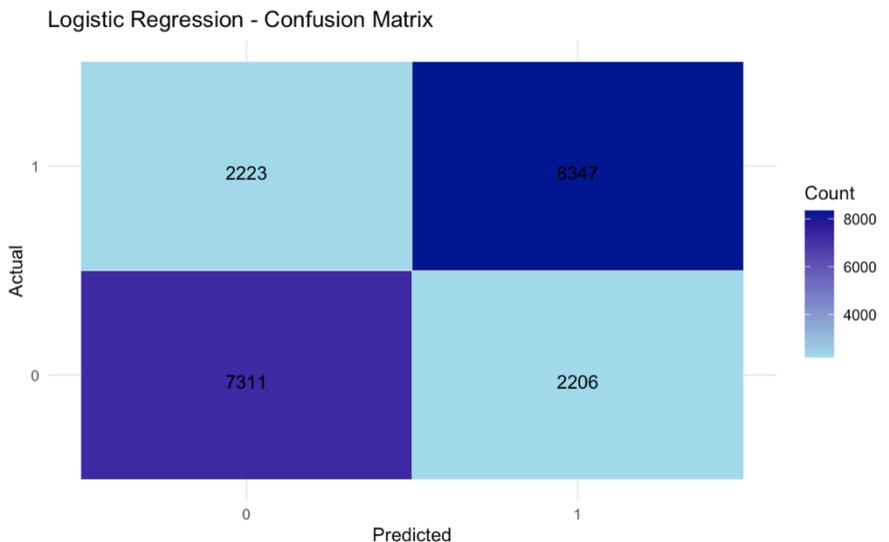
El intercepto tiene un valor de -4.070 con un valor $p < 2e-16$, lo que indica que es altamente significativo. Esto representa el logaritmo de las probabilidades (log-odds) de ser contratado cuando todas las variables independientes son iguales a cero. Las variables más significativas, dado que el valor $p < 0.05$, indican que factores como ser desarrollador, el salario previo, las habilidades informáticas y el nivel educativo tienen un impacto considerable en las probabilidades de la variable dependiente. Tras la prueba, el modelo predice que el 38,41% de las mujeres y el 53,28% de los hombres deben ser contratados. En comparación a las probabilidades preanálisis la diferencia de probabilidad de contratación es de -0,7% para el hombre, pero un 6,47% para la mujer. Lo que indica que el modelo amplifica el sesgo inicial reduciendo la probabilidad de que una mujer sea contratada aún más. Lo que indica que el uso de este modelo podría perjudicar al género femenino.

El coeficiente de habilidades informáticas es de 0.3251, lo que significa que cuantas más habilidades posea un candidato, mayores serán las probabilidades logarítmicas de ser contratado. En comparación con no tener educación superior, los niveles educativos de posgrado (0.227), título universitario (0.138) y otros (0.184) tienen un impacto positivo en la variable dependiente, mientras que un doctorado (-0.47) tiene un impacto negativo. Además, los resultados sugieren que el continente de residencia puede influir en las

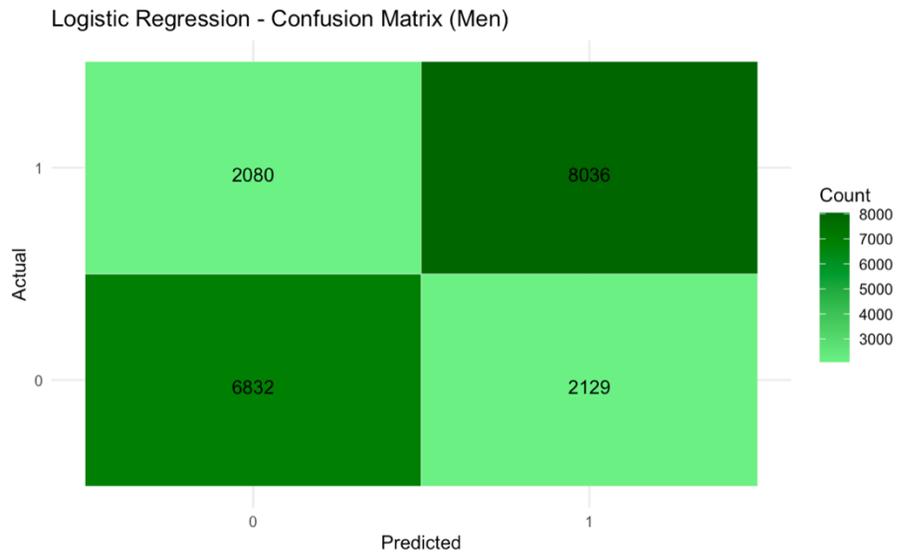
probabilidades de ser contratado. En comparación con el nivel de referencia (Norteamérica), África y Asia parecen tener una leve significancia positiva (valores $p = 0.037$ y 0.067), mientras que residir en Oceanía disminuye significativamente las probabilidades de ser contratado. Curiosamente, ser desarrollador (Dev) reduce significativamente las probabilidades logarítmicas de ser contratado en comparación con no ser desarrollador. Finalmente, estar empleado actualmente y los años de experiencia profesional en codificación también influyen significativamente en la decisión de contratar al candidato. El resumen indica que la variable "Género" resultó ser estadísticamente significativa para predecir la probabilidad de ser contratado. El coeficiente estimado fue de -0.1356 ($p = 0.014313$), en igualdad de condiciones, las mujeres (Gender1 = 1) tienen un 13.56% menos de probabilidades de ser contratadas que los hombres (Gender1 = 0). La influencia del género no es independiente del efecto de otras variables, como la experiencia laboral o las habilidades que poseen los candidatos. Esto significa que el efecto del género puede variar dependiendo de los valores de estas otras variables.

Utilizando el umbral de 0.5, la Matriz de confusión 1 (global), muestra el rendimiento general del modelo para todos los individuos. La Matriz 2 muestra el rendimiento del modelo para los hombres y por último la matriz 3 muestra el rendimiento del modelo para predecir si una mujer será contratada o no. A partir de estas, se calcularon métricas como la precisión (y margen de error), la sensibilidad y la especificidad.

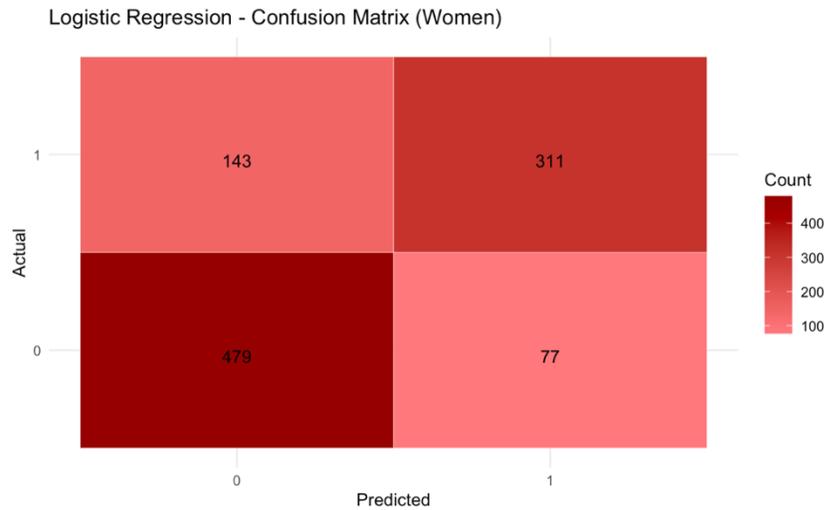
1.1 – Matriz Confusión Global



1.2 – Matriz Confusión Hombres



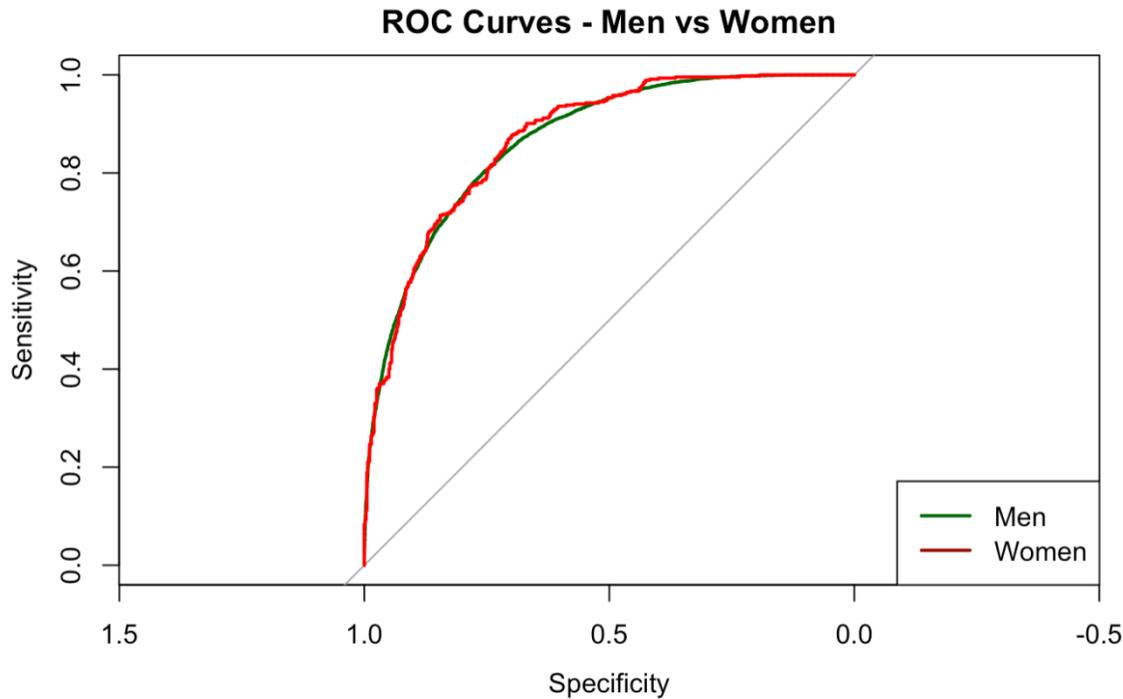
1.3 – Matriz Confusión Mujeres



En general, el modelo presenta un rendimiento similar para hombres y mujeres, con una precisión cercana al 78% (Anexo x). Sin embargo, se observan diferencias leves en las métricas. En términos de Sensibilidad, la capacidad del modelo para identificar correctamente a los individuos que sí deben ser contratados es ligeramente mayor para los hombres (79.44%) que para las mujeres (68.5%). Es alrededor de un 10% más probable que el modelo clasifique erróneamente a las mujeres como "no contratadas", incluso cuando deberían ser contratadas. Al mismo tiempo, la diferencia entre la métrica de especificidad entre ambos géneros es del 10%, mayor para las mujeres (86.2%) que para los hombres (76.2%). Por lo que el modelo es un más probable que clasifique erróneamente a los

hombres como "contratados", incluso cuando no deberían serlo. El gráfico x muestra las curvas ROC para hombres y mujeres, mientras que la tabla x muestra las Tasas de Falsos Positivos (FP) y de Falsos negativos (FN) además del valor del área bajo la curva (AUC).

Gráfico 8- Curvas ROC Hombre y Mujer Modelo 1



	Tasa FP (%)	Tasa FN (%)	AUC
Hombres	23,76	13,85	0,8672244
Mujeres	20,56	31,50	0,8704878

Tabla 4- AUC, Tasa FP y FN Matriz Hombre y Mujer Modelo 1

Según el AUC, ambos modelos tienen buena capacidad para distinguir los empleados de los no empleados ($AUC < 0,8$). Curiosamente el área debajo de la curva para las mujeres es ligeramente más alta, por lo que esto implica que la probabilidad de clasificar correctamente a una candidata es mayor. Adentrándonos más a fondo en el porqué de estos resultados observamos que el modelo predijo que de aquellos hombres que se debían contratar el 23,76% fue incorrecto, un error de un 3% más que para las mujeres. Mientras que de aquellas mujeres que el modelo decidió no contratar, el 31,5% de las mujeres fueron contratadas, una tasa de FN de 17.65% mayor a la del hombre. Aunque el

modelo presenta una capacidad similar para predecir el empleo tanto en hombres como en mujeres, existe un sesgo evidente en las tasas de error de predicción. El modelo tiende a cometer más errores falsos positivos en hombres y más errores falsos negativos en mujeres. Estos sesgos pueden indicar que el modelo, no trata a los géneros por igual en términos de precisión de clasificación.

Modelo de Regresión Múltiple 2 – Nuevas Variables Habilidades

El intercepto del segundo modelo tiene una línea base más baja ($4.880e+00$) para las probabilidades logarítmicas de ser contratado cuando todas las variables independientes son iguales a cero.

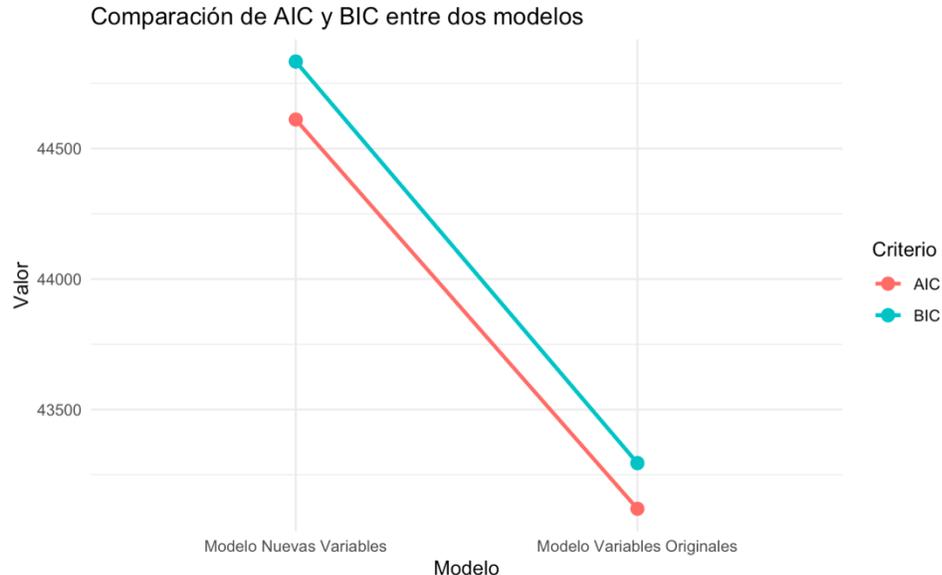
Tras añadir las 5 variables categóricas binomiales al modelo el coeficiente de género es mayor en magnitud (-0.9636 frente a -0.1356), lo que indica un impacto mucho más fuerte en la disminución de la probabilidad de ser contratado para las mujeres. Sin embargo, la significancia de la variable de género disminuye, de hecho, el valor p es > 0.5 ($0,074$) por lo que se encuentra en el intervalo de confianza y no se puede descartar la hipótesis nula. Y, por tanto, se concluye que para el modelo 2 se la variable no es relevante a la hora de decidir (ANEXO). Esto sugiere que la relación entre el género y la probabilidad de pertenecer a $Employed = 1$ se ve afectada por la inclusión de variables adicionales en el nuevo modelo, haciendo que el efecto del género sea más pronunciado, pero menos estadísticamente significativo.

Las variables de habilidades técnicas incluidas en el modelo (SQL, cloud, Python, Java, y C++) muestran diferentes impactos en la probabilidad de ser contratado. Los coeficientes de Java (1.004) y C++ (1.198) se consideran altos y tienen un impacto significativo en la probabilidad de ser contratado. El coeficiente de Python (-0.8427) también es alto, pero tiene un impacto negativo. Los coeficientes de SQL (-0.1782) y Cloud (-0.3471) son moderados y tienen un menor impacto comparativo. Los coeficientes negativos indican que poseer conocimientos en estas tecnologías disminuye la log-odds de ser contratado. Por otro lado, los coeficientes positivos sugieren que tener habilidades en estos lenguajes de

programación aumenta las log-odds. En resumen, tener conocimientos en Python y Cloud Computing tiene un efecto negativo, mientras que tener dominio de Java y C++ incrementa significativamente la probabilidad de que un individuo sea contratado. En este nuevo modelo, la edad gana importancia, siendo más probable que te contraten si tienes menos de 35 años. En cuanto a la educación y los antecedentes académicos, tanto el grado como el postgrado y los años de experiencia en codificación ya no son significativos en el nuevo modelo. El doctorado y el ser desarrollador reducen su impacto negativo a pesar de que se mantienen significativos. El resto muestran resultados parecidos.

Si comparamos los valores de AIC y BIC de ambos modelos, estos son menor en el modelo, lo que sugieren que el modelo con las nuevas variables es más complejo, lo que podría resultar en un sobreajuste. Esto tiene sentido ya que se han añadido 5 variables más a considerar.

Gráfico 9- AIC y BIC Modelo 1 vs Modelo 2



Para el modelo 2 se repite la examinación de las tres matrices: global, mujeres y hombres. Las probabilidades del modelo son las siguientes: $P(\text{Employed} = 1) = 52,7\%$; $P(\text{Employed} = 1 | \text{Gender} = 0) = 53,3$; $P(\text{Employed} = 1 | \text{Gender} = 1) = 41,3$. El Modelo 2 muestra mejoras

significativas en sensibilidad y especificidad para ambos géneros, reduciendo las brechas de equidad observadas en el Modelo 1 (tabla x). El Modelo 2 mejora notable en sensibilidad con un aumento de 7.1% para mujeres y 2.7% para hombres y se reduce la brecha entre géneros un 4,4%, mejorando la equidad en la detección entre géneros. Además, el segundo modelo también muestra una mejora en especificidad, con un aumento de 7.3% para mujeres y 3.2% para hombres.

Reduciendo la brecha entre géneros del 10% a un 6,1%. La precisión mejora ligeramente y se mantiene equitativa entre géneros. En general, el Modelo 2 proporciona un rendimiento más equilibrado y preciso.

Diferencia entre Modelo 1 y 2:

Verde → Impacto Positivo

Rojo → Impacto Negativo

DIFERENCIAS	MUJERES (%)	HOMBRES (%)	Différence (M – H) (%)
SENSIBILIDAD	+ 7,1	+ 2,7	- 4,4
ESPECIFICIDAD	+ 7,3	+ 3,2	- 3,9
PRECISIÓN	+ 3,0	+ 2,9	+ 0,001

Tabla 5- Diferencias Porcentuales Matrices de Modelo Variables Originales y Modelo Nuevas Variables

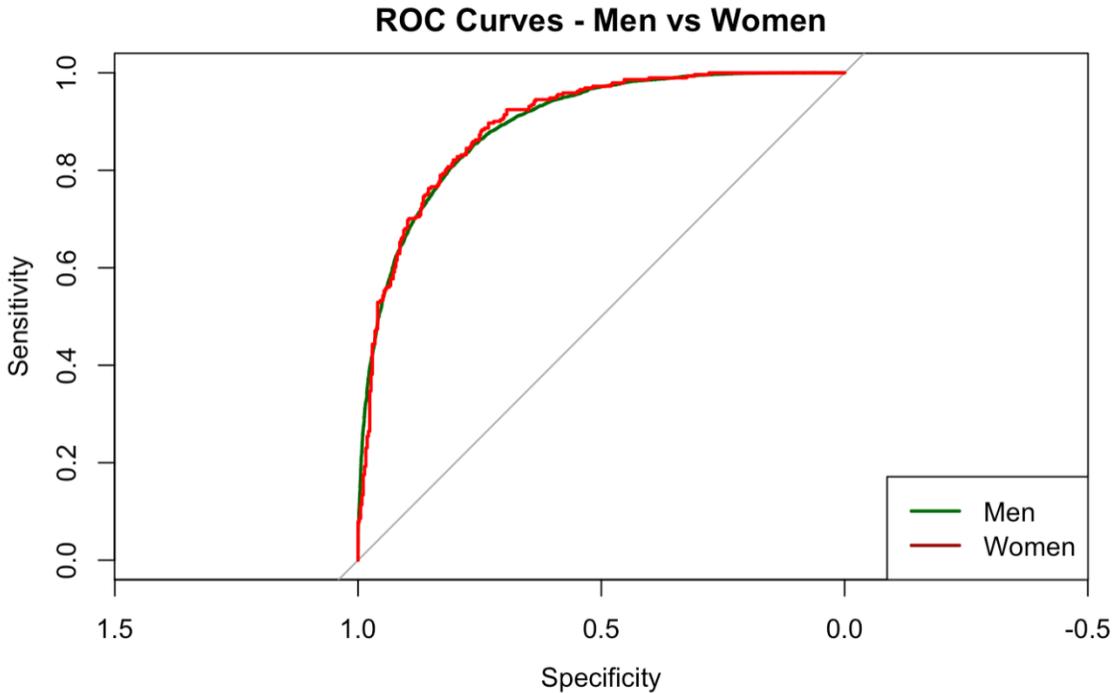


Gráfico 10- Curva ROC Mujer y Hombre Modelo 2

	Tasa FP (%)	Tasa FN (%)	AUC
Hombres	20.65%	14.48%	0.8936581
Mujeres	17.85%	24.40%	0.8954884

Tabla 6- AUC, Tasa FP y FN (%) de Matrices Hombre y Mujer Modelo 2

En términos de rendimiento, observamos mejoras en el Modelo 2. Comparando, la tasa de falsos positivos es menor para ambos géneros, y una mayor disminución para el hombre, lo que indica una mejor capacidad del modelo para reducir las clasificaciones incorrectas como positivas ($FPR(H) = - 3,11 \%$ y $FPR(M) = - 2,71$). La tasa de falsos negativos es significativamente menor para las mujeres ($FNR(M) = -7,1$), lo que sugiere una mejora considerable en la detección de casos positivos entre las mujeres. El AUC es mayor para ambos géneros, lo que indica un mejor rendimiento global del Modelo 2 en comparación con el Modelo 1.

Conclusiones

Como hemos establecido, el mundo en el que vivimos está marcado por una discriminación de género omnipresente. Los modelos de inteligencia artificial se basan en datos históricos, y estas bases de datos están impregnadas de sesgos que reflejan las barreras existentes en los procesos de contratación y las percepciones de confianza tanto del candidato como del empleador. Se ha demostrado, que no es coincidencia que la probabilidad de ser contratada siendo una mujer sea considerablemente más baja que la de un hombre. Los datos (históricos y ya sesgados) introducidos muestran como la mujer tiene de base inicial un 10% menos de probabilidades que un hombre de ser contratada. Esta disparidad no solo es evidente en los resultados de los modelos estadísticos, sino que estos modelos amplifican y ensanchan la brecha a una diferencia del 14,9% entre hombres y mujeres.

Además, el rendimiento de un modelo se mide en función de su capacidad para predecir correctamente los resultados deseados. Sin embargo, cuando solo el 5% de los datos son representativos y la probabilidad de contratar a una mujer es menor, los resultados de la prueba reflejan que el rendimiento de un modelo varía según el género. Curiosamente, siendo este rendimiento más eficiente hacia el género femenino, al predecir correctamente si una mujer va a ser contratada o no. En otras palabras, el modelo se equivoca más a menudo al ofrecer erróneamente más puestos a hombres que al no darle el puesto a una mujer. Esto refleja la teoría de Thomas Chamorro, quien sostiene que el error radica en que se percibe al hombre como más competente debido a su exceso de confianza, más que que se perciba a la mujer como menos competente por falta de ella. Esto ocurre porque los modelos introducen suposiciones matemáticas para precisar sus resultados replicando y acentuando patrones. Para las máquinas, el género es una variable más con unos y ceros y estas no son conscientes de su significado e impacto, ni capaces de detectar estos errores y cambiarlos. Por ello está en manos de sus creadores detectar y mitigar estos sesgos. La Ley de IA establece que los algoritmos deben auditarse y comprenderse; sin embargo, no se especifica que estas auditorías deben ser realizadas por profesionales con antecedentes diversos. No es suficiente que un hombre blanco diseñe el algoritmo y otro hombre blanco

lo audite. La responsabilidad cae tanto en las organizaciones privadas como públicas en garantizar la diversidad en sus equipos para reducir el riesgo de errores.

Introducir variables relacionadas con las habilidades y capacidades del candidato puede mitigar el riesgo de sesgo de género, como así lo demuestra nuestro segundo modelo. Pero estos algoritmos no son capaces de distinguir entre el bien y el mal; carecen de ética y se limitan a los datos que se les proporcionan. Por lo que es crucial auditar estos sesgos, ya que estos se introducen de distintas maneras. Aunque agregar más variables relevantes puede mitigar la discriminación de género, es esencial considerar las variables introducidas ya que podrían generar nuevos sesgos, como hemos observado con la edad. El segundo modelo muestra que la variable de género ya no es relevante cuando se añaden habilidades esenciales para el sector STEM, y la decisión de contratación se basa más en cualidades y habilidades, independientemente del género. Sin embargo, al observar las variables relacionadas a habilidades digitales más detenidamente se observa que los hombres adquieren estas habilidades con más frecuencia y mayor cantidad. Lo que indica que las carreras STEM siguen siendo ambientes masculinos. Aunque es cierto que también se observa un incremento en las habilidades y la motivación para aplicar a puestos de trabajo entre las mujeres más jóvenes.

Para eliminar el riesgo de sesgo, es necesario contar con equipos diversificados que establezcan diferentes necesidades. Las mujeres deben estar implicadas en la creación de modelos y algoritmos. Sin embargo, dado el contexto en el que nos encontramos esto podría no ser factible, y más aún si se establecen modelos sesgados a la hora de contratar; Si se quiere llegar a una paridad de género en las carreras de STEM el cambio debe provenir de la educación, mitigando el sesgo de brillantez y empoderando a las mujeres. Sin embargo, un análisis más profundo y sesgado por género de la variable 'Nivel de educación' muestra una realidad preocupante: un hombre sin educación superior tiene más probabilidades de ser contratado que una mujer con un doctorado. Esto evidencia que, a pesar de los avances en la preparación y motivación de las mujeres jóvenes, persisten

fuerres sesgos de género que favorecen a los hombres en el mercado laboral, independientemente de su nivel educativo. Por lo que se considera que no se debería empoderar solamente a la mujer, pero comunicar y crear conciencia entre el género masculino y evitar que caigan en los estereotipos.

Eliminar la variable de género mediante el blind hiring no es la solución definitiva. Como mencionado en el marco teórico, muchos puestos están diseñados en base a habilidades y competencias típicamente asociadas a hombres, y son pocas las mujeres que aplican a estos puestos. Como resultado, los algoritmos pueden rechazar perfiles femeninos y obtener el resultado opuesto al deseado. Sabemos que es crucial la incorporación de mujeres en este campo, y por ello, utilizar la variable de género puede ayudar a empoderar, monitorear y controlar la diversidad de la plantilla, así como a identificar la necesidad de contratar a alguien específico para aportar una perspectiva diversa. La variable de género debe ser un factor significativo que favorezca a las mujeres y a los grupos minoritarios hasta que se alcance una paridad. Por ejemplo, al establecer un rango salarial para cubrir un puesto senior, se podría fijar que el salario previo debe oscilar entre 30,000 y 60,000 euros. Sin embargo, esto podría descalificar a mujeres que están capacitadas en cubrir el rol pero que reciben menos paga debido a la disparidad salarial. Si la variable de género existe, se le podría instruir al algoritmo que ajuste el rango para las mujeres (por ejemplo, de 25,000 a 55,000 euros); ya que aplicar un mayor rango (25,000 – 60,000) a todo el conjunto sería un error, porque permitiría que hombres con menos experiencia sean aceptados en detrimento de candidatas femeninas más cualificadas.

En conclusión, hemos establecido que el mundo en el que vivimos está marcado por una discriminación de género omnipresente. Los modelos de inteligencia artificial, basados en datos históricos ya sesgados, amplifican estas disparidades, resultando en que las mujeres tienen significativamente menos probabilidades de ser contratadas. Sin embargo, creando conciencia y cambiando políticas se puede evitar que esta tecnología se convierta de enemigo a aliado de las mujeres. La introducción de variables relacionadas con habilidades

y capacidades del candidato puede mitigar el riesgo de sesgo de género, no obstante, persisten desafíos fundamentales. Es esencial auditar estos modelos de IA con diversidad en los equipos y considerar cuidadosamente las variables introducidas para evitar la introducción de nuevos sesgos. La solución radica no solo en mejorar los algoritmos, sino también en abordar las desigualdades desde la raíz mediante la educación y crear conciencia en el género masculino sobre lo perjudicial que es (especialmente en el área STEM). La responsabilidad de garantizar la equidad en el uso de la IA recae tanto en las organizaciones privadas como en las públicas, promoviendo la diversidad y realizando auditorías con profesionales de diferentes antecedentes.

Limitaciones

Esta investigación enfrenta a varias limitaciones. En primer lugar, (1) se basa en datos históricos que ya contienen sesgos de género, lo que puede limitar la precisión de los hallazgos. (2) Los resultados pueden no ser generalizables a todos los contextos laborales y culturales debido a la diversidad de industrias y regiones, y se desconoce la localización específica de los puestos de trabajo, lo que impide comparar si existe más discriminación en empresas o países determinados. (3) La investigación se centra en la discriminación de género, pero otras formas de discriminación, como la basada en raza, etnicidad o discapacidad, no se abordan en profundidad. (4) Las técnicas de análisis pueden tener limitaciones en su capacidad para identificar y medir los sesgos presentes en los algoritmos de IA. Esto se debe a que la IA establece un conjunto de modelos y análisis; el resultado podría ser distinto si se utilizan otros métodos, como los árboles de decisión u otros enfoques. (5) La implementación y evaluación de intervenciones para mitigar estos sesgos en la tecnología de IA están fuera del alcance del estudio.

Futuras Investigaciones

Para futuras investigaciones, se podría repetir el modelo incrementando proporcionalmente el número de mujeres en la muestra de datos. Esto permitiría conocer si se observa la una menor, misma, o mayor disparidad que en el estudio si los datos fueran más representativos. Al mismo tiempo, este método ayudaría medir y cuantificar los sesgos que se producen al recolectar datos

(una de las tres maneras comentadas por la que se introducen los sesgos). Al tener una representación más equitativa de mujeres, se podría evaluar con mayor precisión cómo los sesgos de género influyen en los resultados de los algoritmos de inteligencia artificial y en los procesos de toma de decisiones en el ámbito laboral. Al mismo tiempo, al ser uno de los métodos por donde se introduce el sesgo, se podría medir el impacto de este método de infiltración en comparación a los otros dos para poder establecer una prioridad de enfoque.

Otra línea de investigación interesante sería realizar un análisis de clúster para observar si los algoritmos crean perfiles significativamente diferentes entre hombres y mujeres. Este enfoque permitiría identificar patrones y categorías dentro de los datos que podrían revelar cómo los algoritmos diferencian entre géneros y si estas diferencias son perjudiciales o discriminatorias. Concluiríamos con más certeza la validez del uso del reclutamiento a ciegas. Además, al comprender mejor estas diferencias, sería posible desarrollar estrategias para ajustar los algoritmos y reducir los sesgos de género, promoviendo así una mayor equidad en los procesos de selección y otras decisiones organizacionales.

Bibliografía:

AAUW. (2016). *Barriers and Bias: Status of Women in Leadership*. Washington DC.

Alihodzic, D., Broeker, A., Baehr, M., Kluge, S., Langebrake, C., & Wicha, S. G. (2020). Impact of inaccurate documentation of sampling and infusion time in model-informed precision dosing. *Frontiers in Pharmacology*, 11, 172.

Allen, G. (2020). *Understanding AI Technology*. Joint Artificial Intelligence Center, Department of Defence. JAIC.

Anugerah, A. R., Muttaqin, P. S., & Trinarningsih, W. (2022). Social network analysis in business and management research: A bibliometric analysis of the research trend and performance from 2001 to 2020. *Heliyon*, 8(4).

Banco Mundial. (2022). Datos Banco Mundial. Retrieved from Población, Mujeres (%): <https://datos.bancomundial.org/indicador/SP.POP.TOTL.FE.ZS>

Bartneck, C., Lütge, C., Wagner, A., & Welsh, S. (2020). *An Introduction to Ethics in Robotics and AI*. Springer International Publishing AG.

Bian, L., Leslie, J., & Cimpian, A. (2017). Gender stereotypes about intellectual ability emerge early and influence children's interests. *Science*. <https://doi.org/aah6524>

Blau, F. D., & Kahn, L. M. (2016). The Gender Wage Gap: Extent, Trends, and Explanations. *Quarterly Journal of Economics*, 137(4), 2163-2194. doi:10.1093/qje/qjw019

Briggs, J., & Kodnani, D. (2023, March 26). The Potentially Large Effects of Artificial Intelligence on Economic Growth. *Global Economics Analyst*. Goldman Sachs Global Investment Research.

Burton, T. I. (2016, January). The Feminist History of Prohibition. *JSTOR Daily*.

Clementine Collete, G. N. (2022). *Los efectos de la IA en la vida laboral de las mujeres*. Francia: UNESCO.

Corinne A. Moss-Racusin, J. F. (2012, October 9). Science faculty's subtle gender biases favor male students. *PNAS*, 109(41).

Crawford, K. (2021) *The Atlas of AI: Power, Politics, and the Planetary Costs of Artificial Intelligence*. New Haven: Yale University Press.

Dahlbom, P., Siikanen, N., Sajasalo, P., & Jarvenpää, M. (2020). Big data and HR analytics in the digital era. *Baltic Journal of Management*, 15(1), 120-138.

Datta, A., Tschantz, M. C., & Datta, A. (2015). Automated Experiments on Ad Privacy Settings: A Tale of Opacity, Choice, and Discrimination [Article]. *Proceedings on Privacy Enhancing Technologies*, 92-112

Drage, E., Mackereth, K. Does AI Debias Recruitment? Race, Gender, and AI's "Eradication of Difference". *Philos. Technol.* **35**, 89 (2022). <https://doi.org/10.1007/s13347-022-00543-1>

Emily Field, A. K. (2023). *Women in the Workplace 2023*. McKinsey & Company.

Era Dabla-Norris, K. K. (2018, Noviembre). Las mujeres, la tecnología y el futuro del trabajo. IMF.

Ernesto Reuben, P. R.-B. (2010). The Emergence of Male Leadership in Competitive Environments. *Forschungsinstitut zur Zukunft der Arbeit Institute for the Study of Labor, Discussion Paper (5300)*.

Jeanette N. Cleveland, T. K.-F. (2004). Gender Discrimination in Organizations. In A. C.

Hellström, T. (2020, April 1). *Bias in Machine Learning -- What is it Good for?* arXiv.org. <https://arxiv.org/abs/2004.00686>

IBM Data and AI Team. (2023, October 16). Shedding light on AI bias with real world examples. IBM.

International Labour Organization (ILO). (2023). Women and the World of Work: Statistical Overview. International Labour Organization. Retrieved from <https://www.ilo.org/resource/news/ilo-expects-global-unemployment-drop-slightly-2024-slow-progress-reduce>

Gao, L., et al. (2021). The Pile: An 800GB Dataset of Diverse Text for Language Modeling. Retrieved from <https://ar5iv.labs.arxiv.org/html/2101.00027>.

Gibbs, S. (2015, May). Google says sorry over racist Google Maps White House search results. *The Guardian*.

Global Education Monitoring Report Team. (2023). SDG 4 mid-term progress review: progress since 2015 has been far too slow. UNESCO.

Gov.uk. (2024). *Recruitment Support*. From Civil Services Career: https://www.civil-service-careers.gov.uk/departments/working_for_ipo/ipo-recruitment-support/#:~:text=We%20ask%20you%20provide%20your,it's%20called%20name%20blind%20recruitment.

Katelyn M. Cooper, A. K. (2018). Who perceives they are smarter? Exploring the influence of student characteristics on student academic self-concept in physiology. *Advances in Physiology Education*, 42(2), 200-208.

Lin Bian et al. (2017). Gender stereotypes about intellectual ability emerge early and influence children's interests. *Science* 355, 389-391 DOI:10.1126/science.aah6524

Llinares-Insa, L. I., González-Navarro, P., Córdoba-Iñesta, A. I., & Zacarés-González, J. J. (2018). Women's job search competence: A question of motivation, behavior, or gender. *Frontiers in Psychology*, 9, 137. <https://doi.org/10.3389/fpsyg.2018.00137>

Lucía Cobreros, J. G. (2024). Mujeres en STEM Con el apoyo de : Desde la educación básica hasta la carrera laboral. Esade. EsadeEcPol - Center for Economic Policy.

Manurung, E. F. (2020). The effects of transformational leadership, competence and compensation on work motivation and implications on the performance of lecturers of Maritime College in DKI Jakarta. *International Journal of Multicultural and Multireligious Understanding*, 7(6). Retrieved from <https://ijmmu.com/index.php/ijmmu/article/view/1385>

Mihaljević, H., Müller, I., Dill, K., & Yollu-Tok, A. (2022). Towards gender-inclusive job postings: A data-driven comparison of augmented writing technologies. *PLOS ONE*, 17(9), e0274312. <https://doi.org/10.1371/journal.pone.0274312>

MDPI. (2023). Using LinkedIn endorsements to reinforce an ontology and machine learning-based recommender system to improve professional skills. *Electronics*. Retrieved from <https://www.mdpi.com/journal/electronics>

Nestor Maslej, L. F. (2023). The AI Index 2023 Annual Report. Institute for Human-Centered AI. Stanford: Stanford University.

Ntoutsis, E, Fafalios, P, Gadiraju, U, et al. (2020) Bias in data-driven artificial intelligence systems—An introductory survey. *WIREs Data Mining Knowl Discov*. p.1-14, vol 10

OECD. (2024, March 2). Live data. Retrieved from [oecd.ai: https://oecd.ai/en/data?selectedArea=ai-demographics&selectedVisualization=ai-demographics-by-gender](https://oecd.ai/en/data?selectedArea=ai-demographics&selectedVisualization=ai-demographics-by-gender)

OECD. (2023). Bridging the Digital Gender Divide. OECD. Retrieved from <https://www.oecd.org/digital/bridging-the-digital-gender-divide.pdf>.

O’Neil, C. (2017). *Weapons of Math Destruction*. Penguin Books.

ONU Mujeres. (2015). El progreso de las mujeres en el mundo. Transformar las economías para realizar los derechos. Mexico: ONU.

Open AI. (2024, Feb 13). Openai. Retrieved from What is ChatGPT?: <https://help.openai.com/en/articles/6783457-what-is-chatgpt>

Petrie, Christopher. García-Millán, Clara. Mateo-Berganza Díaz, María Mercedes. (2021). Spotlight: Habilidades del siglo XXI en América Latina y el Caribe HundrED Research. <https://hundred.org/en/research>

QuantumBlack AI by McKinsey. (2022). The state of AI in 2022—and a half decade in review. McKinsey.

Randstad. (2024). the labor market and AI. Randstad. Retrieved from Workforce Insights.

Raffel, C., Shazeer, N., Roberts, A., Lee, K., Narang, S., Matena, M., ... & Liu, P. J. (2020). Exploring the limits of transfer learning with a unified text-to-text transformer. *Journal of Machine Learning Research*, 21(140), 1-67.

Robert L. Dipboye, *Discrimination at Work the Psychological and Organizational Bases*. New York.

Salas, J. (2017, Octubre). A woman's place is in the home – at least according to AI... ElPais.

Samek, A. (2019). Gender Differences in Job Entry Decisions: A University-Wide Field Experiment. *Management Science*, 65(7), 3272-3281.

Susanne Hupfer, A. B. (2021, December). Women in the tech industry: Gaining ground but facing new headwinds. Deloitte Insights.

Tanantong, T., & Wongras, P. (2024). A UTAUT-based framework for analyzing users' intention to adopt artificial intelligence in human resource recruitment: A case study of Thailand. *Systems*, 12(1), 28. <https://doi.org/10.3390/systems12010028>

Thomas Breda, E. J. (2020, December). Gender stereotypes can explain the gender-equality paradox. *PNAS*, 117(49), 31063–31069.

ValenciaPlaza. (2024, May 9). El PP propone en el Congreso el currículum ciego que defiende Díaz para evitar discriminaciones. *ValenciaPlaza*.

Varona, D., & Suárez, J. L. (2022). Discrimination, Bias, Fairness, and Trustworthy AI. *Applied Sciences*, 12(12), 5826. <https://doi.org/10.3390/app12125826>

Villars, R. L., Olofson, C. W., & Eastwood, M. (2011). Big data: What it is and why you should care. *White paper, IDC*, 14, 1-14.

Work Bank Group. (2024, April). Woman, Business and the Law. Retrieved from The World Bank: https://wbl.worldbank.org/en/data/exploretopics/wbl_sj

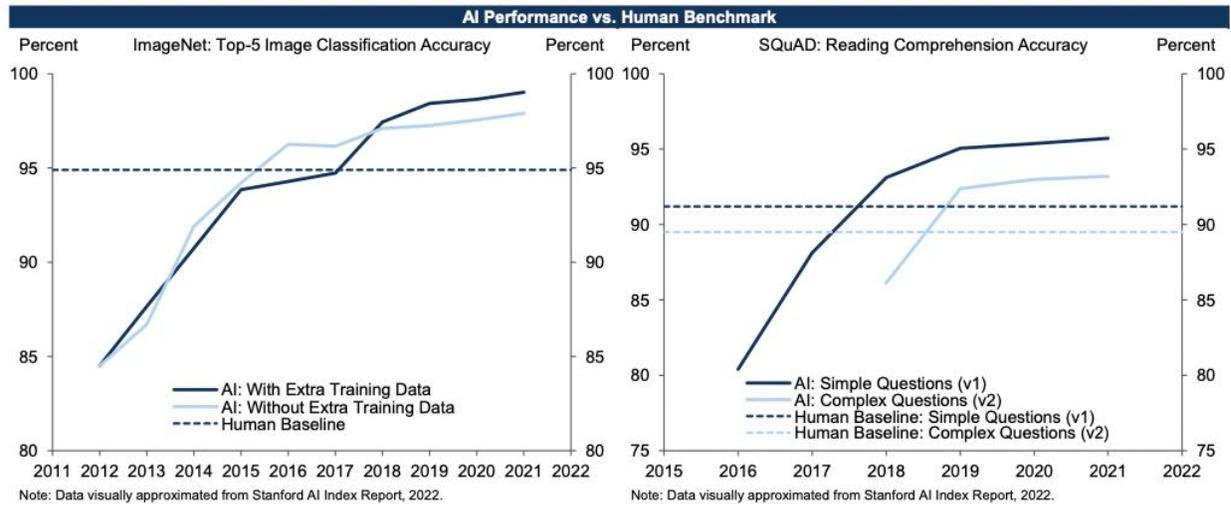
World Economic Forum. (2021). Global Gender Gap Report. Geneva.

World Economic Forum. (2023). Global Gender Gap Report. Switzerland: World Economic Forum.

Wood, J. (2018). Existen leyes en 104 países que impiden que las mujeres trabajen en algunos empleos. World Economic Forum.

Anexo

3.1 - La IA Supera Rápidamente los Puntos de Referencia Humanos (AI Increasingly Outperforms Humans Benchmarks)



Source: Stanford Institute for Human-Centered Artificial Intelligence, Goldman Sachs Global Investment Research

3.2 Inversión Privada en IA por Área Geográfica, 2013-22 (Suma)

Private Investment in AI by Geographic Area, 2013–22 (Sum)

Source: NetBase Quid, 2022 | Chart: 2023 AI Index Report

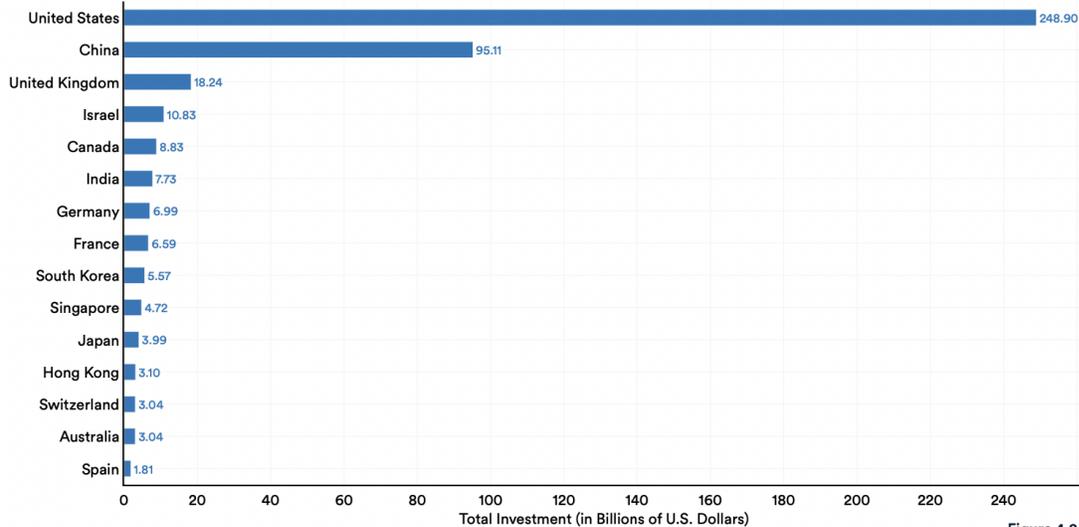


Figure 4.2.11

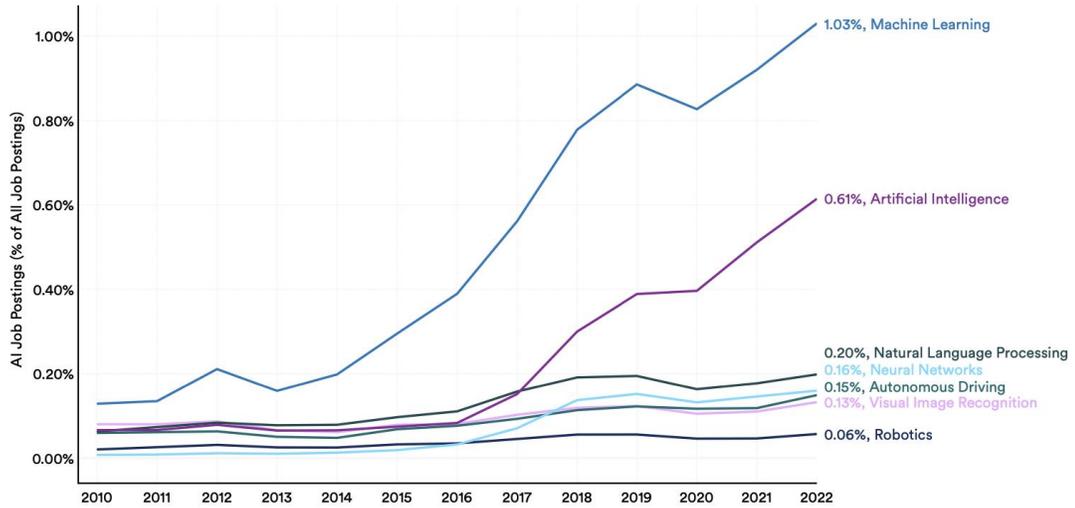
3.3- La Innovacion Tecnologica Conlleva a la Creación de Nuevos Puestos de Empleo

Exhibit 10: Technological Innovation Leads to the Creation of New Occupations That Account for the Bulk of Employment Growth



Source: Autor et al. (2022), Goldman Sachs Global Investment Research

3.4- Anuncios de Trabajo en IA (% en EEUU por Clúster de Habilidad), 2010-22



3.5- Top 10 Habilidades Especializadas en IA en 2022 en Anuncios de Trabajo en EEUU por Habilidad

Top Ten Specialized Skills in 2022 AI Job Postings in the United States by Skill Share, 2010–12 Vs. 2022

Source: Lightcast, 2022 | Chart: 2023 AI Index Report

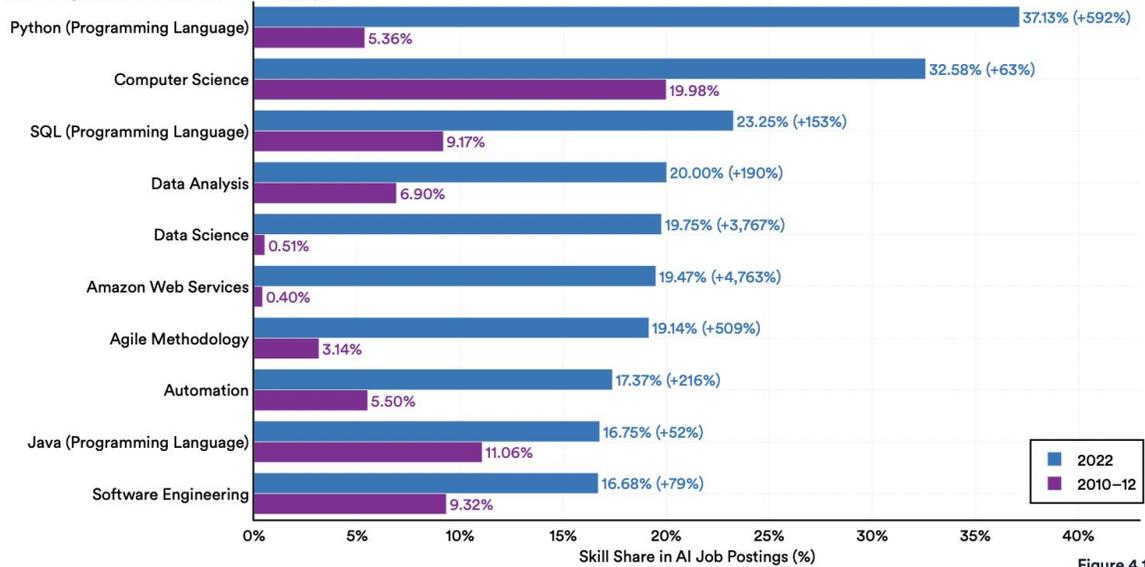
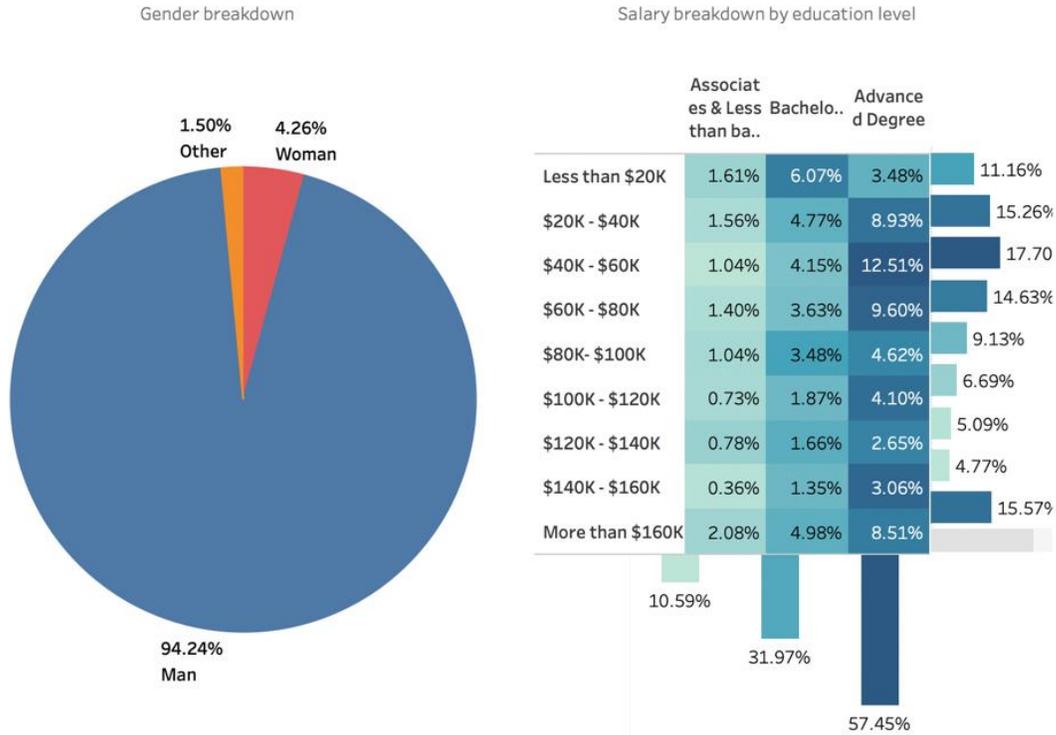


Figure 4.1.4

3.6- Informacion Demografica sobre los Encuestados



Note: Aggregate demographic information from survey respondents is leveraged to build indicators and identify trends related to the profession, country, salary, education, gender, and age of AI developers. Please see [methodological note](#) for more information.

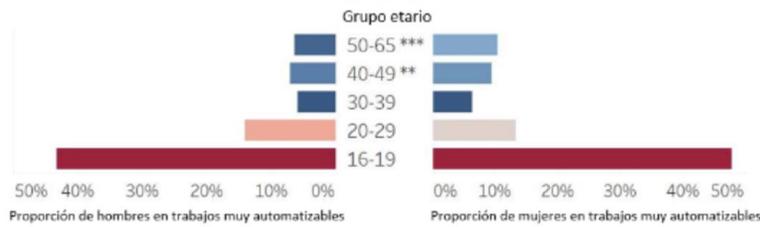
Source of data: Stackoverflow Survey
 Please cite as: OECD.AI (2023), visualisations powered by Tableau using data from Stackoverflow, accessed on www.oecd.ai. Supported by the [Patrick J. McGovern foundation](#).

3.7 - Proporción de Hombres y Mujeres con alto riesgo de automatización de sus empleos, en porcentaje por grupo etario

¿Será reemplazado por un robot?

Las mujeres trabajadoras mayores de 40 años corren un riesgo mucho mayor que los hombres en el mismo grupo etario de que sus trabajos se automaticen. El máximo riesgo de automatización del empleo corresponde a los trabajadores de 16 a 19 años.

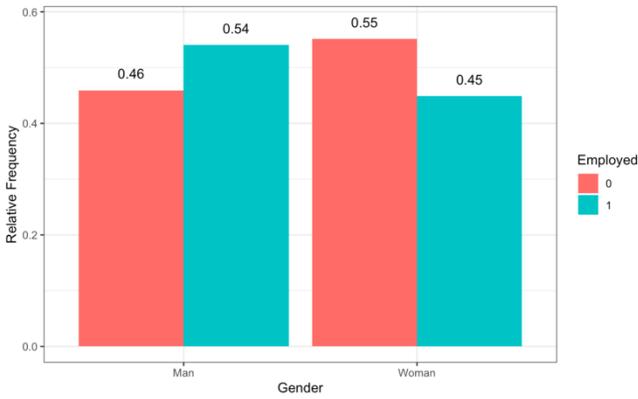
(Proporción de hombres y mujeres con alto riesgo de automatización de sus empleos, en porcentaje por grupo etario)



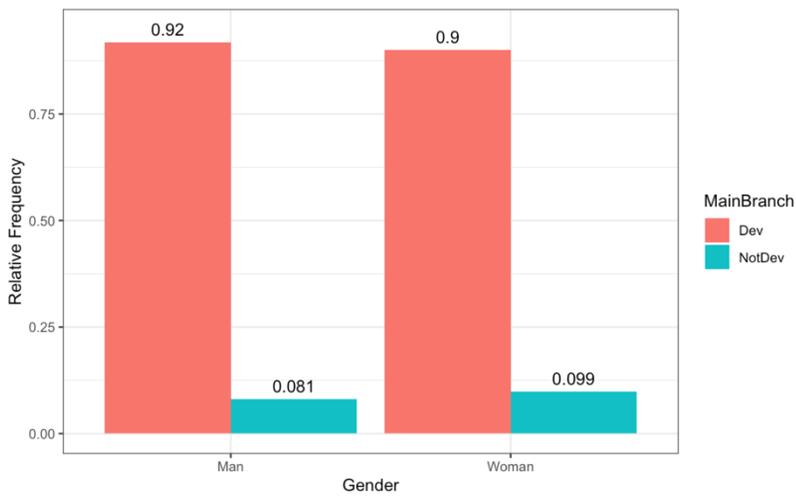
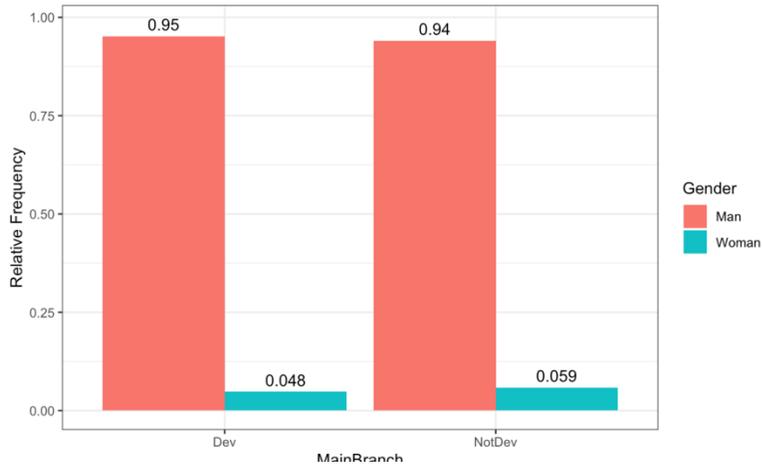
Fuentes: Frey y Osborne (2017); encuesta PIAAC; y estimaciones del personal técnico del FMI.
 Nota: Los detalles sobre la metodología y las variables constan en "Gender, Technology, and the Future of Work", IMF Staff Discussion Note 18/07, anexo III (2018).

Apéndice

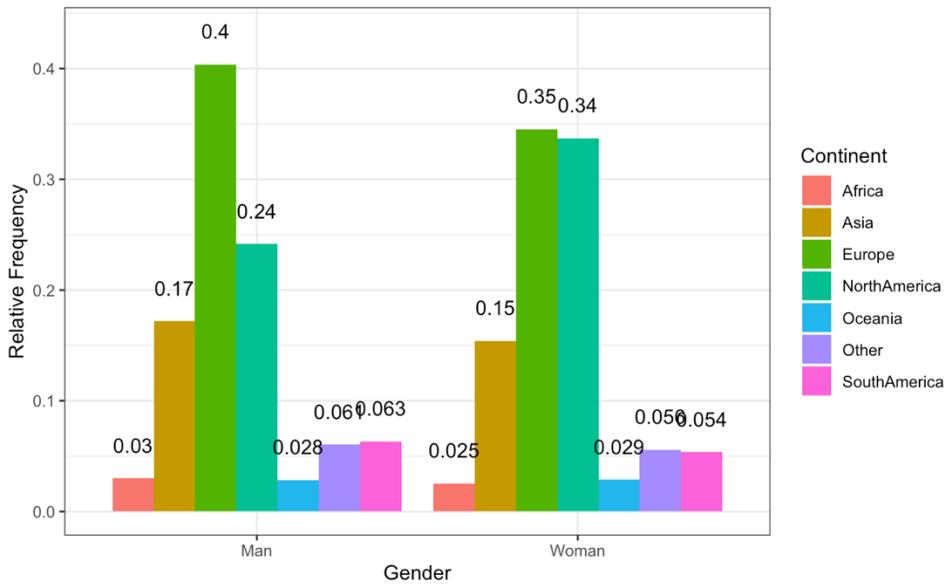
Apéndice 1- Gráfico de Barras $P(\text{Employed}|\text{Gender})$



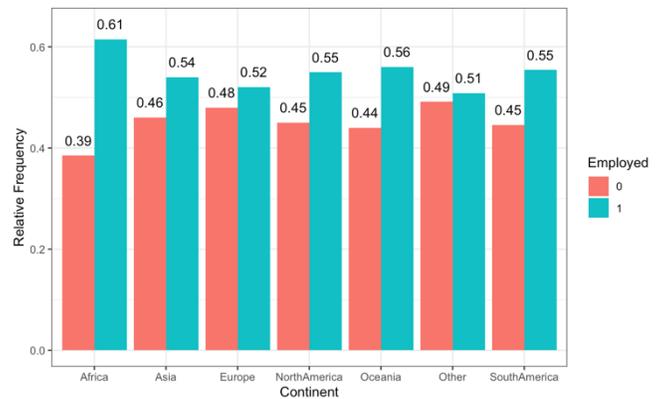
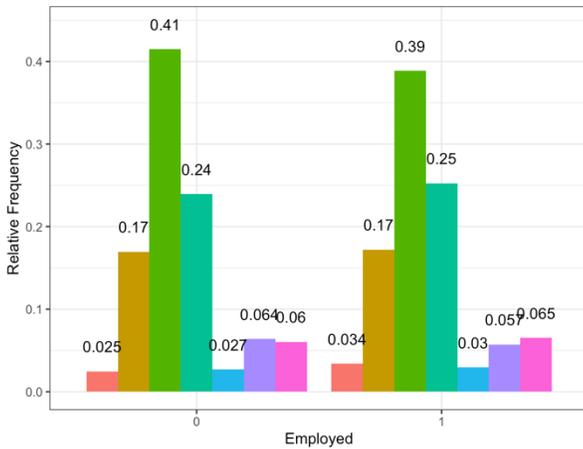
Apéndice 2- Gráfico de Barras $P(\text{Empleo} | \text{Genero})$ y $P(\text{Genero} | \text{Empleo})$



Apendice 3- Gráfico de Barras P(Genero | Continente)



Apendice 4 - Gráfico de Barras P(Empleo | Continente) y P(Genero | Continente)



Apendice 5- Tabla P(Empleo | Genero n Continente)

Probabilidad_de_Empleo_por_Continente_y_G_nero

Continente	Employed_Men	Unemployed_Men	Employed_Women	Unemployed_Women	Probabilidad_Empleo_Hombres	Probabilidad_Empleo_Mujeres	Diferencia
Africa	0.61	0.39	0.57	0.43	0.61	0.57	0.040000000000000036
Asia	0.54	0.46	0.48	0.52	0.54	0.48	0.060000000000000005
Europe	0.4	0.28	0.35	0.29	0.5882352941176471	0.546875	0.04136029411764708
NorthAmerica	0.52	0.48	0.5	0.5	0.52	0.5	0.020000000000000018
Oceania	0.55	0.45	0.56	0.44	0.55	0.56	-0.010000000000000009
Other	0.49	0.51	0.51	0.49	0.49	0.51	-0.020000000000000018
SouthAmerica	0.55	0.45	0.55	0.45	0.55	0.55	0.0

Apendice 6- Resumen Modelo Logistico Regresivo 1

Call:

```
glm(formula = class ~ ., family = binomial, data = trainSet2)
```

Coefficients:

	Estimate	Std. Error	z value	Pr(> z)	
(Intercept)	-4.070e+00	9.122e-02	-44.618	< 2e-16	***
Age1	5.405e-02	3.616e-02	1.495	0.134969	
EdLevel1	2.277e-01	5.867e-02	3.881	0.000104	***
EdLevel2	1.382e-01	3.993e-02	3.461	0.000538	***
EdLevel3	-4.700e-01	7.844e-02	-5.992	2.07e-09	***
EdLevel4	1.843e-01	3.009e-02	6.124	9.10e-10	***
Employment1	1.989e-01	3.868e-02	5.142	2.71e-07	***
Gender1	-1.356e-01	5.536e-02	-2.449	0.014313	*
MentalHealth1	2.083e-02	2.972e-02	0.701	0.483284	
MainBranch1	-6.035e-01	4.827e-02	-12.502	< 2e-16	***
YearsCode	6.841e-03	3.197e-03	2.140	0.032382	*
YearsCodePro	8.942e-03	3.940e-03	2.270	0.023232	*
PreviousSalary	-4.272e-06	3.664e-07	-11.661	< 2e-16	***
ComputerSkills	3.251e-01	3.096e-03	105.031	< 2e-16	***
Continent1	-1.418e-01	7.460e-02	-1.901	0.057261	.
Continent2	2.274e-02	7.273e-02	0.313	0.754499	
Continent3	1.628e-01	7.819e-02	2.082	0.037338	*
Continent4	1.853e-01	1.011e-01	1.833	0.066807	.
Continent5	-1.035e-01	8.376e-02	-1.236	0.216526	
Continent6	-3.507e-01	8.365e-02	-4.192	2.76e-05	***

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Apendice 7- Tabla Sensibilidad, Especificidad y Precisión de Matrices Mujer y Hombre Modelo 1

MODELO 1	MUJERES	HOMBRES	Diferencia (M – H)
SENSIBILIDAD	0,685	0,794	- 0,109
ESPECIFICIDAD	0, 862	0, 762	+ 0,1
PRECISIÓN	0,782	0,779	+ 0,003

Apendicee 8- Modelo Regresion Logistico 2

```
glm(formula = class ~ ., family = binomial, data = trainSet)
```

Coefficients:

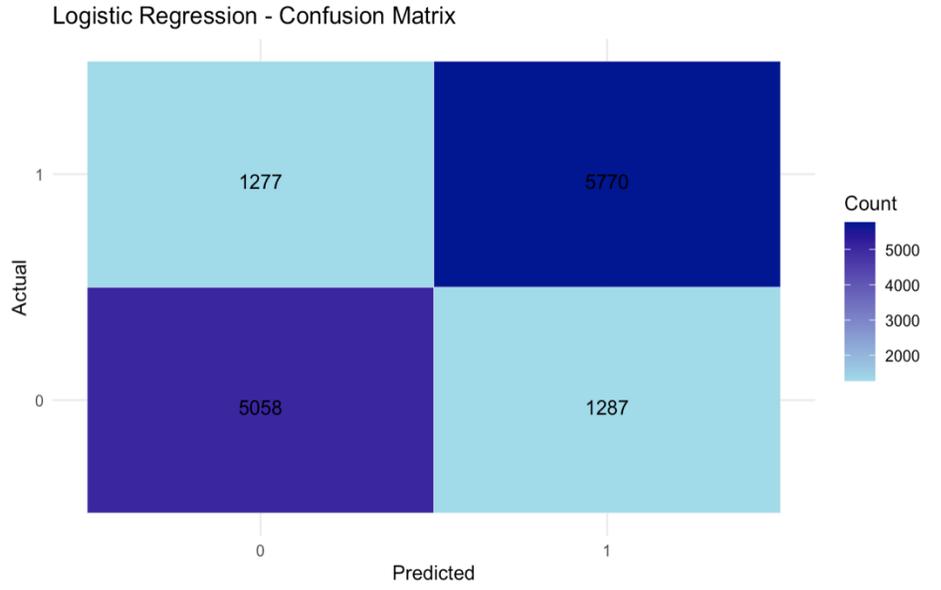
	Estimate	Std. Error	z value	Pr(> z)	
(Intercept)	-4.880e+00	9.507e-02	-51.331	< 2e-16	***
Age1	8.425e-02	3.569e-02	2.361	0.018249	*
EdLevel1	2.645e-02	5.785e-02	0.457	0.647537	
EdLevel2	1.210e-02	3.970e-02	0.305	0.760613	
EdLevel3	-2.763e-01	7.832e-02	-3.528	0.000418	***
EdLevel4	8.847e-02	2.997e-02	2.952	0.003158	**
Employment1	1.968e-01	3.807e-02	5.170	2.34e-07	***
Gender1	-9.636e-02	5.395e-02	-1.786	0.074074	.
MentalHealth1	4.098e-02	2.945e-02	1.391	0.164107	
MainBranch1	-2.806e-01	4.871e-02	-5.760	8.39e-09	***
YearsCode	-1.352e-04	3.180e-03	-0.043	0.966086	
YearsCodePro	1.415e-03	3.921e-03	0.361	0.718250	
PreviousSalary	-2.017e-06	3.677e-07	-5.485	4.15e-08	***
ComputerSkills	3.462e-01	3.763e-03	92.005	< 2e-16	***
SQL1	-1.782e-01	3.200e-02	-5.569	2.56e-08	***
python1	-8.427e-01	2.670e-02	-31.568	< 2e-16	***
java1	1.004e+00	3.305e-02	30.387	< 2e-16	***
c_plus_plus1	1.198e+00	2.558e-02	46.827	< 2e-16	***
cloud1	-3.471e-01	2.775e-02	-12.508	< 2e-16	***
Continent1	-7.755e-02	7.419e-02	-1.045	0.295879	
Continent2	4.794e-02	7.235e-02	0.663	0.507569	
Continent3	9.360e-02	7.791e-02	1.201	0.229610	
Continent4	1.989e-01	1.002e-01	1.985	0.047164	*
Continent5	-9.016e-02	8.330e-02	-1.082	0.279110	
Continent6	-1.776e-01	8.315e-02	-2.135	0.032738	*

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

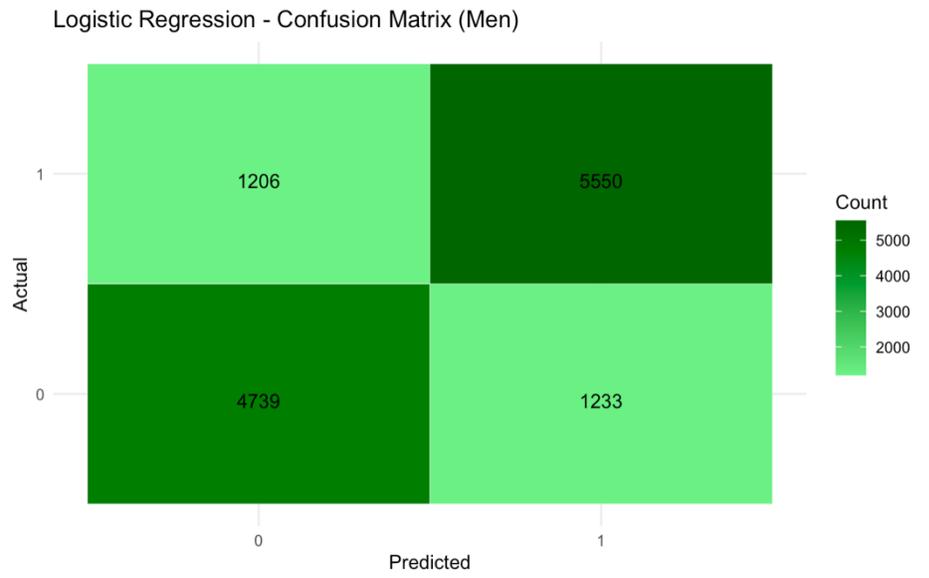
(Dispersion parameter for binomial family taken to be 1)

Null deviance: 74111 on 53565 degrees of freedom
 Residual deviance: 44561 on 53541 degrees of freedom
 AIC: 44611

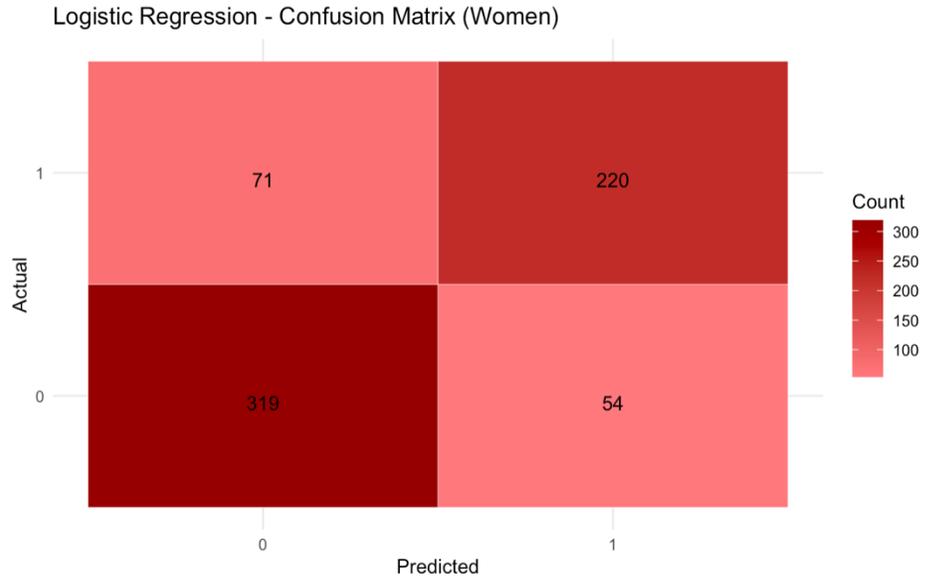
Apendice 9- Matriz Global Modelo 2



Apendice 10- Matriz Hombres Modelo 2



Apendice 11- Matriz Mujeres Modelo 2



Apendice 12- Tabla Sensibilidad, Especificidad y Precisión de Matrices Mujer y Hombre Modelo 2

MODELO 2	MUJERES	HOMBRES	Diferencia (M – H)
SENSIBILIDAD	0,756	0,821	- 0,065
ESPECIFICIDAD	0,855	0,794	+ 0,061
PRECISION	0,812	0,808	+ 0,004

