

La inteligencia artificial, los sesgos del algoritmo y la discriminación en las relaciones laborales

Artificial intelligence, algorithmic bias, and discrimination in labor relations

MILAGROS ALEJANDRA SALINAS ATENCIO*

Universidad Nacional Mayor de San Marcos / Universidad de San Martín de Porres
Contacto: salinas.atencio.mili@gmail.com
<https://orcid.org/0009-0003-6951-5934>

CHRISTIAN GUZMÁN NAPURÍ**

Pontificia Universidad Católica del Perú
Contacto: guzmannapuri@hotmail.com
<https://orcid.org/0000-0001-9929-0068>

RESUMEN: Los autores analizan el impacto de la inteligencia artificial (IA) en las relaciones laborales, destacando cómo los sesgos algorítmicos pueden generar discriminación en procesos de selección, evaluación y despidos. Expone que la falta de regulación y supervisión adecuada puede perpetuar desigualdades, afectando principalmente a grupos vulnerables. Además, examinan estrategias para mitigar estos sesgos, como auditorías, supervisión humana y regulaciones específicas. También mencionan la importancia de la transparencia en la toma de decisiones automatizadas y el derecho de los trabajadores a impugnar decisiones injustas, finalmente subrayando la necesidad de un marco normativo que garantice un uso ético e inclusivo de la IA en el ámbito laboral.

ABSTRACT: The authors analyzes the impact of artificial intelligence (AI) on labor relations, highlighting how algorithmic biases can lead to discrimination in hiring, evaluation, and dismissal processes. She argues that the lack of regulation and proper oversight can perpetuate inequalities, disproportionately affecting vulnerable groups. Additionally, she examines strategies to mitigate these biases, such as audits, human

* Lic. Administración de Negocios Internacionales. Maestrante de Finanzas Corporativas en UPN. CEO Maviliari SAC. Estudiante de Derecho en la Universidad Nacional Mayor de San Marcos (cursando 4to año), Conciliadora Extrajudicial. Especializaciones en Marketing Digital, Inteligencia Artificial, Gestión Pública, Contrataciones con el Estado y Finanzas Avanzadas. Presidente de la Asociación Jóvenes del Milenio. Escritora y pintora. Autora de Poemas Naturales, Convivencia Inteligente y Diccionario Retail 2025

** Magíster en Derecho, con mención en Derecho Constitucional, por la PUCP y doctorando en la misma universidad. Profesor titular del Departamento de Derecho de la PUCP. Director de la Maestría de Derecho Administrativo Económico y de Derecho Constitucional de la Universidad Continental. Ha sido funcionario de diversas entidades públicas, entre ellas presidente y vocal del Tribunal de Susalud. Árbitro en materia de contratación administrativa. Socio de GNS Guzmán Napurí & Segura Abogados.

supervision, and specific regulations. The authors also emphasize the importance of transparency in automated decision-making and the right of workers to challenge unfair decisions. Finally, she underscores the need for a regulatory framework that ensures the ethical and inclusive use of AI in the workplace.

PALABRAS CLAVE: Inteligencia Artificial, Sesgo Algorítmico, Discriminación Laboral, Gestión Del Talento, Regulación Y Normativa, Transparencia Algorítmica

KEYWORDS: Artificial Intelligence, Algorithmic Bias, Workplace Discrimination, Talent Management, Regulation And Standards, Algorithmic Transparency

Recibido: 20/02/2025

Aceptado: 24/03/2025

Publicado en línea: 31/03/2025

SUMARIO: I. Introducción. II. Fundamentos Teóricos. III. Aplicación de la IA en la Gestión Laboral. IV. Estrategias para Mitigar los Sesgos Algorítmicos en el Ámbito Laboral. 1. Auditorías y transparencia algorítmica. 2. Uso de datos representativos. 3. Supervisión humana y control ético. 4. Explicabilidad y derechos de los trabajadores. 5. Regulación y Políticas Públicas. 6. Capacitación y sensibilización sobre la IA. 7. Modelos alternativos de IA ética. V. Evaluaciones de Desempeño y Promociones Asistidas por IA. VI. Riesgos de Sesgo en la Automatización de Despidos. VII. Impacto de los Sesgos Algorítmicos en la Discriminación Laboral. VIII. ¿Cómo se Manifiestan los Sesgos en la Toma de Decisiones?. 1. Sesgo de representación. 2. Sesgo de confirmación. 3. Discriminación indirecta. 4. Transparencia limitada. IX. Consecuencias en la Equidad de Oportunidades Laborales. 1. Perpetuación de desigualdades preexistentes. 2. Menos diversidad en las empresas. 3. Dificultad en la impugnación de decisiones. 4. Desconfianza en la automatización. X. Marco Normativo y Regulaciones Aplicables. 1. Legislación laboral peruana y su relación con la IA. 2. Normativas internacionales sobre IA y sesgos algorítmicos. A) Reglamento General de Protección de Datos (GDPR) de la Unión Europea. B) Propuesta de Ley de Inteligencia Artificial de la UE (AI Act.). C) Fair Hiring Algorithms Act. (EE.UU). D) Declaración de la OIT sobre el Futuro del Trabajo (2019). 3. Propuestas para una regulación más equitativa en el Perú. A) Regulación de la toma de decisiones automatizada. B) Creación de un ente regulador de IA en el trabajo. C) Incorporación de auditorías algorítmicas obligatorias. D) Fomento de la educación en IA y ética laboral. XI. Estrategias para Mitigar los Sesgos Algorítmicos en el Ámbito Laboral. 1. Modelos de auditoría y supervisión de IA. 2. Estrategias para llevar a cabo las auditorías. A) Auditorías internas. B) Auditorías externas. C) Supervisión continua. XII Implementación de Principios de Equidad y Trasperencia. 1. Medidas para garantizar la implementación de los principios. B) Uso de datos representativos. C) Desarrollo de mecanismos de corrección. XIII. Rol de los Sindicatos y Reguladores en la Vigilancia del Uso de la IA en el Ámbito Laboral. 1. Casos de auditorías exitosas. XIV. El Contexto Regulador en América Latina y Perú. 1. Propuestas regulatorias para el Perú. A) Obligatoriedad de informes de impacto de IA. B) Creación de un ente supervisor. XV. Defensa del Derecho a Impugnar Decisiones Automatizadas. 1. Principales preocupaciones sobre la IA y el debido proceso laboral. A) Desconocimiento del criterio de decisión. B) Dificultad para apelar decisiones. C) Falta de supervisión humana. 2.

Medidas para fortalecer los derechos de los trabajadores ante la IA. A) Revisión humana. B) Comités de IA y equidad laboral. C) Garantía del derecho a la explicación. XVI. Conclusión. Referencias.

◆

I. INTRODUCCIÓN

En años más recientes, la inteligencia artificial (IA) ha tomado un papel crucial en los diversos aspectos de la toma de decisiones en múltiples industrias. En el área laboral, estos sistemas automatizados se volvieron cruciales al optimizar la selección, valoración y administración del talento humano, no obstante, su uso ha generado preocupaciones marcadas en temas de equidad y transparencia, particularmente por los sesgos algorítmicos que se generan en este proceso de automatización y que pueden resultar en ambientes discriminatorios que pueden manifestarse en la preferencia por ciertos perfiles o en la plena exclusión de determinados grupos vulnerables, suscitando retos para la justicia laboral y la protección de derechos de los trabajadores dado que la falta de regulación perpetúa las desigualdades existentes.

El presente artículo, entonces, analizará cómo estos sesgos algorítmicos impactan directamente en la perpetuación de la discriminación laboral y explorará diversas medidas para garantizar un uso justo de la inteligencia artificial. Al abordar los fundamentos teóricos, las aplicaciones en la gestión laboral, el marco normativo y las estrategias para mitigar la discriminación desde un enfoque multidisciplinario. A través de este análisis se busca aportar a la discusión sobre la implementación responsable en el mercado laboral, con énfasis en el contexto peruano, mismo en el que su experiencia a través del proceso de incorporación de la digitalización genera un entorno en el cual se vuelve crucial garantizar el uso inclusivo y transparente de la IA.

En la gestión laboral, uno de los principales usos de la inteligencia artificial se presenta en la automatización de los procesos de selección de personal. Estos algoritmos analizan una cantidad elevada de currículums en segundos a través de la identificación de patrones de experiencia y habilidades, no obstante, estos tienen un alto índice de reproducción de sesgos implícitos si los datos utilizados para calibrar los algoritmos reflejan prácticas discriminatorias, afectando, en la

mayoría de situaciones, a mujeres, minorías y otros grupos vulnerables. Otro uso frecuente lo encontramos en la evaluación de desempeño y promoción de empleados, en este caso, el entendimiento de factores contextuales se vuelve esencial para brindar una decisión objetiva, sin embargo, los algoritmos muchas veces ignoran este factor, centrándose en perfiles específicos en base a los perfiles usados en su calibración, excluyendo a trabajadores que no se encuentren dentro de ese marco de ascensos o beneficios sin brindar una justificación objetiva de acuerdo a cada caso.

Siguiendo la misma línea, en los países que utilizan la IA en procesos de despido y rotación laboral mediante algoritmos de modelo predictivo basados en información histórica de métricas de desempeño y la retroalimentación de supervisores también presentan problemas graves en cuando a la consideración de factores contextuales, sin la supervisión adecuada, en ambos casos, los sistemas refuerzan patrones discriminatorios.

Al emplear la IA, pese a que represente un avance significativo en el proceso de digitalización en el ámbito laboral, se vuelve innegable admitir que la controversia alrededor de su requerimiento de datos históricos puede, y ha, resultado en la marginación sistemática en el acceso a oportunidades de crecimiento profesionales hacia empleados con perfiles específicos, lo que complica su uso a largo plazo.

II. FUNDAMENTOS TEÓRICOS

Para comprender el impacto de la IA en las relaciones laborales, es esencial conocer sus principios fundamentales y la problemática que representa el sesgo algorítmico. La inteligencia artificial es un campo de la informática que busca crear sistemas capaces de realizar tareas que normalmente requieren de la inteligencia humana, como el aprendizaje, la toma de decisiones y el reconocimiento de patrones (Russell y Norvig 2021). Dentro del ámbito laboral, se ha convertido en una herramienta clave para la optimización de procesos, desde la selección de personal hasta la evaluación de desempeño y gestión de talento (Kaplan y Haenlein 2019). Sin embargo, el impacto de estas tecnologías no está exento de problemáticas, principalmente debido a los sesgos algorítmicos que pueden derivar en discriminación laboral, los sesgos algorítmicos pueden definirse como desviaciones sistemáticas en los resultados producidos por un sistema de IA,

debido a la forma en que se diseñan los modelos o a los datos utilizados para su entrenamiento (Barocas, Hardt y Narayanan 2019). Uno de los principales problemas con los sesgos en IA es que pueden replicar y amplificar desigualdades estructurales preexistentes en la sociedad. Por ejemplo, un sistema de selección de personal basado en IA podría preferir candidatos de cierto género o grupo étnico si los datos históricos utilizados para su entrenamiento reflejan prácticas discriminatorias en contrataciones previas (Binns 2020).

A partir del punto de vista jurídico, la discriminación laboral basada en IA enfrenta múltiples desafíos regulatorios, aún cuando ya existen normativas internacionales que promueven la transparencia y equidad en su uso, su aplicación en el contexto laboral aún es limitada. En el Perú, es crucial establecer lineamientos que garanticen que la inteligencia artificial se utilice de forma responsable y sin perpetuar desigualdades estructurales, este debate sobre la regulación de la misma en el trabajo también involucra principios éticos fundamentales, como la aplicabilidad y la responsabilidad algorítmica. Los empleadores, en este caso, deben asegurarse de que sus sistemas sean auditables y comprensibles, permitiendo a los trabajadores conocer los criterios utilizados en la toma de decisiones automatizadas al mismo tiempo en el que la supervisión humana sigue siendo esencial para prevenir, corregir fallos y garantizar la equidad en el proceso.

En el contexto laboral, la discriminación algorítmica puede manifestarse en diversas etapas del proceso de empleo, incluyendo la contratación, promoción y despidos (Eubanks 2018). Investigaciones han demostrado que algunos algoritmos de contratación han penalizado de manera sistemática a candidatas mujeres, simplemente porque los datos históricos reflejaban una menor proporción de mujeres en roles de liderazgo dentro de ciertas industrias (Dastin 2018). Del mismo modo, sistemas de IA utilizados en el sector financiero para evaluar la confiabilidad de los trabajadores pueden discriminar a ciertos grupos con base en variables aparentemente neutrales, pero que terminan favoreciendo a individuos de determinados estratos socioeconómicos (O'Neil 2016).

Para mitigar estos problemas, es entonces fundamental comprender los principios de equidad en el diseño de algoritmos y el proceso de implementación de estrategias de auditoría y corrección de sesgos, para esto existen diversos enfoques para lograr una inteligencia artificial más imparcial, por un lado, se propone establecer metodologías de “*debiasing*” o eliminación de sesgos en datos

y modelos utilizando técnicas como la recolección de datos balanceados y el ajuste de pesos en los modelos de aprendizaje automático (Mehrabi et al. 2021). Por el otro se destaca la importancia de la supervisión humana en la toma de decisiones automatizadas, garantizando que los algoritmos sean herramientas complementarias y no sustituyen la evaluación de expertos en recursos humanos (Rahwan et al. 2019).

El debate sobre la ética del uso de inteligencia artificial también se ha vuelto un foco de atención significativo, lo que se refleja en el aumento de iniciativas que buscan promover principios de transparencia y responsabilidad en el desarrollo de estas tecnologías (Floridi et al. 2018). En el ámbito jurídico, diversas normativas han comenzado a abordar la problemática de los sesgos algorítmicos, como el “Reglamento General de Protección de Datos” (GDPR) de la Unión Europea, que exige explicabilidad en los procesos de toma de decisiones automatizadas (Wachter, Mittelstadt y Floridi 2017).

III. APLICACIÓN DE LA IA EN LA GESTIÓN LABORAL

74

La implementación de la IA ha transformado significativamente la gestión laboral en todos sus niveles. Su implementación ha permitido agilizar procesos, reducir costos operativos y mejorar la precisión en la toma de decisiones en las empresas (Kaplan y Haenlein 2019). Sin embargo, su uso también plantea riesgos relacionados con la transparencia y la equidad en la administración del talento humano (Binns 2020).

Como ya se ha mencionado, uno de los principales usos de la inteligencia artificial en la materia implica la automatización del proceso de selección de personal. Los algoritmos están en capacidad de analizar grandes volúmenes de currículums en un muy corto periodo de tiempo, identificando patrones en la experiencia y habilidades de los candidatos, este avance, sin embargo, no elimina los sesgos históricos de discriminación que los sistemas replican en base a la información histórica utilizada en ellos, excluyendo de manera sistemática a ciertos grupos demográficos. Por ejemplo, un estudio realizado por Dastin (2018) evidenció cómo un sistema de contratación basado en IA de Amazon penaliza automáticamente los currículums de mujeres debido a la predominancia histórica de hombres en roles tecnológicos.

Su segundo ámbito de aplicación es la evaluación del desempeño de los trabajadores, por ejemplo, empresas del rubro de IBM han implementado sistemas de IA que analizan métricas de productividad y generan informes sobre el rendimiento de los empleados (Brynjolfsson y McAfee 2017). Sin embargo, estos sistemas pueden generar problemas cuando las métricas utilizadas no toman en cuenta factores contextuales, siendo los más comunes la carga laboral desigual o las diferencias en el acceso a recursos dentro de la empresa. Ante esto, O'Neil (2016) advierte que los algoritmos utilizados en la evaluación del desempeño podrían reforzar desigualdades preexistentes si no son debidamente supervisados por humanos. Empresas de distribución y logística han desarrollado sistemas automatizados para optimizar los horarios de trabajo y asignar turnos de acuerdo a patrones de productividad, sin embargo, investigaciones recientes han demostrado que estos sistemas pueden priorizar la eficiencia operativa sobre las necesidades humanas, resultando en horarios poco flexibles que afectan la calidad de vida de los trabajadores (Eubanks 2018).

Otro aspecto clave se encuentra en la predicción de rotación de empleados, por ejemplo, empresas como *Workday* han desarrollado algoritmos que predicen qué empleados podrían estar considerando renunciar en base a sus datos de rendimiento, satisfacción y comportamiento (Rahwan et al. 2019). Pese a que esta tecnología permite a las empresas implementar estrategias de retención proactivas, también plantea graves preocupaciones en cuanto a la privacidad de los empleados y la posibilidad de uso indebido de estos datos para justificar despidos anticipados sin una evaluación justa.

La Implementación de esta tecnología en las áreas de promoción y desarrollo profesional y vigilancia y monitoreo también ha generado debates, en el caso del primero, los algoritmos diseñados para recomendar ascensos pueden favorecer a empleados con perfiles similares a aquellos que, de acuerdo a su base histórica de datos, han alcanzado posiciones de liderazgo, excluyendo a grupos menos representados en la empresa (Floridi et al. 2018). Mehrabi et al. (2021) sostienen que la falta de diversidad en los conjuntos de datos de entrenamiento puede resultar en decisiones sesgadas que perpetúan la desigualdad en las oportunidades de crecimiento laboral. En cuanto al segundo, algunas empresas han implementado sistemas de seguimiento basados en IA para analizar la productividad en tiempo real, de la mano del uso de cámaras de reconocimiento facial y software de seguimiento de actividad digital. Si bien estos sistemas pueden aumentar la

eficiencia, también han sido criticados por violaciones a la privacidad y generar entornos de trabajo altamente controlados y estresantes (Wachter, Mittelstadt y Floridi 2017).

IV. ESTRATEGIAS PARA MITIGAR LOS SEGOS ALGORÍTMICOS EN EL ÁMBITO LABORAL

1. Auditorías y Transparencia Algorítmica

La implementación de auditorías, diseñadas específicamente para evaluar cómo los algoritmos procesan los datos, permite detectar posibles discriminaciones durante el proceso de selección de candidatos y en las evaluaciones de desempeño (Raji et al. 2020). De la misma manera, las empresas deben desarrollar herramientas con un formato amigable que permita a los trabajadores y reguladores entender cómo se toman las decisiones automatizadas (Wachter, Mittelstadt y Floridi 2017).

2. Uso de Datos Representativos

Otro factor importante que contribuye a los sesgos algorítmicos es el uso de conjuntos de datos no representativos. Mehrabi et al. (2021) sostienen que una solución efectiva es recopilar datos de manera equitativa, asegurando que los modelos incluyan información de diferentes grupos demográficos y socioeconómicos. La inclusión de estos datos balanceados y el ajuste constante de pesos en los modelos de aprendizaje automático contribuiría significativamente a reducir la discriminación algorítmica (Binns 2020).

3. Supervisión Humana y Control Ético

La capacidad masiva de los algoritmos al analizar grandes cantidades de información no implica la eliminación de la supervisión humana como herramienta base para garantizar la equidad. Al respecto, Rahwan et al. (2019) argumenta que los sistemas de inteligencia artificial deben ser herramientas de apoyo en la toma de decisiones, no sustitutos del juicio humano, para lo cual, las empresas estarían en la obligación de establecer comités de ética que supervisen y validan las decisiones automatizadas.

4. Explicabilidad y Derechos de los Trabajadores

El derecho a la explicabilidad de las decisiones algorítmicas es crucial en la mitigación de sesgos (Floridi et al. 2018). Los trabajadores deben tener acceso a explicaciones claras sobre cómo se toman decisiones automatizadas en su entorno laboral además del establecimiento de mecanismos de apelación para impugnar decisiones que se consideren injustas o discriminatorias.

5. Regulación y Políticas Públicas

La regulación gubernamental juega un papel clave en la mitigación de los sesgos algorítmicos. Legislaciones como el “Reglamento General de Protección de Datos” (GDPR) de la Unión Europea establecen principios de equidad y explicabilidad en la toma de decisiones automatizadas (Wachter, Mittelstadt y Floridi 2017).

6. Capacitación y Sensibilización sobre la IA

Las empresas, al emplear la inteligencia artificial, deben invertir en la capacitación de sus empleados para que comprendan cómo funciona la automatización de decisiones y cómo pueden identificar sesgos en sus procesos de selección. Brynjolfsson y McAfee (2017) sugieren que la alfabetización en IA debe ser parte de la formación de los profesionales de recursos humanos, garantizando que puedan interpretar correctamente los resultados de los algoritmos y tomar decisiones justas.

7. Modelos Alternativos de IA Ética

Un enfoque innovador para mitigar los sesgos algorítmicos es el desarrollo de modelos de inteligencia artificial diseñados con principios de equidad desde su concepción, técnicas tales como la regularización de equidad en el aprendizaje automático han demostrado ser efectivas para reducir la discriminación en la toma de decisiones laborales (Dastin 2018). Dichos modelos ajustarán automáticamente los pesos de los algoritmos para evitar la prevalencia de decisiones discriminatorias.

V. EVALUACIONES DE DESEMPEÑO Y PROMOCIONES ASISTIDAS POR IA

Empresas que han iniciado el proceso de implementación de sistemas avanzados automatizados en las evaluaciones de desempeño y concursos de promoción analizan métricas de productividad, interacción con sistemas internos y retroalimentación de supervisores para determinar niveles de rendimiento y recomendar promociones (Brynjolfsson y McAfee 2017). Este enfoque reduce la subjetividad en la toma de decisiones y promueve una mayor eficiencia en la gestión del talento, esto sin embargo no significa que la dependencia exclusiva de la IA para estas decisiones genere los problemas ya mencionados. O'Neil (2016) advierte que, si los datos históricos reflejan tendencias discriminatorias, como la falta de promoción de mujeres o minorías, los algoritmos pueden reforzar estas prácticas. De la misma forma en la que un informe de la Organización Internacional del Trabajo (OIT) sugiere que la supervisión humana sigue siendo esencial para evitar decisiones sesgadas y garantizar la equidad en las evaluaciones de desempeño (ILO 2021).

78

VI. RIESGOS DE SESGO EN LA AUTOMATIZACIÓN DE DESPIDOS

Con el tiempo, el uso de la inteligencia artificial en la toma de decisiones sobre despidos se ha incrementado en los últimos años, particularmente en empresas que buscan optimizar su fuerza laboral a través de análisis predictivos (Rahwan et al. 2019). Utilizando algoritmos que identifican patrones de baja productividad, ausentismo o insatisfacción laboral y sugerir qué empleados deberían ser considerados para desvinculación laboral en base a esos resultados.

Sin embargo, este enfoque también presenta serios desafíos, tanto éticos como legales, un estudio de Eubanks (2018) expone casos en los que empleados fueron despedidos injustamente debido a decisiones automatizadas que no consideraban factores contextuales, como problemas de salud o circunstancias personales. Wachter, Mittelstadt y Floridi (2017) advierten que la falta de transparencia en estos sistemas puede dificultar la impugnación de decisiones injustas por parte de los empleados.

Teniendo en mente que el objetivo final es mitigar estos riesgos, las empresas deben garantizar que sus sistemas sean auditables y proporcionar explicaciones

claras sobre las decisiones de despido tras un análisis humano de los motivos destacados por los algoritmos. La introducción de mecanismos de apelación y la intervención humana en la última etapa del proceso de toma de decisiones pueden contribuir a la equidad y a la protección de los derechos de los trabajadores (ILO 2021).

VII. IMPACTO DE LOS SEGOS ALGORÍTMICOS EN LA DISCRIMINACIÓN LABORAL

En Estados Unidos, los efectos de los sesgos algorítmicos en el ámbito laboral han sido documentados, por ejemplo, el sistema de selección de personal de Amazon fue descontinuado después de que se evidenciara que su algoritmo penaliza automáticamente a las candidatas mujeres, debido a que su base de datos fue entrenada con datos históricos de contratación dominados por hombres (Dastin 2018). Este caso puso en evidencia cómo los sesgos preexistentes en la sociedad pueden ser amplificadas por la inteligencia artificial si estos no se diseñan con mecanismos adecuados para corregirlos.

En Europa, un estudio del Consejo de Europa (2021) evidenció que las herramientas de IA utilizadas en procesos de selección laboral en grandes empresas multinacionales presentaban sesgos contra personas de ciertos orígenes étnicos. Se detectó que los algoritmos tendían a favorecer a candidatos cuyos perfiles coincidían con los empleados preexistentes, lo que perpetuaba la falta de diversidad en el lugar de trabajo (Wachter et al. 2017).

A nivel latinoamericano, un informe del Banco Interamericano de Desarrollo (BID) mostró cómo en Brasil y México ciertos sistemas de contratación digital excluían involuntariamente a candidatos de zonas rurales o con menor acceso a educación superior, lo que reducía significativamente sus oportunidades de empleo (BID, 2021). En el caso de Perú, aún hay pocos estudios específicos sobre el impacto de la IA en la contratación, pero se ha identificado que la falta de regulación clara puede generar desigualdades en la automatización de procesos de selección (OIT, 2021).

VIII. ¿CÓMO SE MANIFIESTAN LOS SESGOS EN LA TOMA DE DECISIONES?

Los sesgos algorítmicos en la IA laboral pueden manifestarse de diversas maneras, dependiendo del tipo de datos utilizados y del diseño del modelo algorítmico, dentro de las manifestaciones más comunes podemos encontrar:

1. Sesgo de representación

Cuando los datos de entrenamiento de un algoritmo no son suficientemente diversos se desarrollan patrones discriminatorios. Por ejemplo, si un sistema ha sido entrenado con datos de empleados predominantemente masculinos considera que ciertos rasgos o experiencias laborales femeninas no son relevantes para ciertos puestos (Bogen 2019), como en la anterior mención del caso de Amazon.

2. Sesgo de confirmación

Ocurre cuando el algoritmo refuerza patrones históricos de contratación en favor de perfiles similares a los ya existentes dentro de la fuerza labora de la empresa, acción que prevalece desigualdades en el acceso a oportunidades laborales e impede la diversidad en el lugar de trabajo (Mehrabi et al. 2021).

3. Discriminación indirecta

El uso de variables, aparentemente neutrales, como el código postal o la universidad de procedencia, pueden correlacionarse con factores socioeconómicos o étnicos, lo que, de ser parte importante de la información usada para calibrar la IA, genera efectos discriminatorios al excluir automáticamente a ciertos grupos de candidatos sin que la variable en sí esté explícitamente relacionada con su capacidad para el puesto (Raji et al. 2020).

4. Transparencia limitada

Debido a que muchas empresas no comprenden en su totalidad el cómo funcionan los modelos de inteligencia artificial que utilizan para la selección de personal hace que los sesgos en los sistemas sean más difíciles de detectar y corregir (Floridi et al. 2018).

IX. CONSECUENCIAS EN LA EQUIDAD DE OPORTUNIDADES LABORALES

1. Perpetuación de desigualdades preexistentes

Si los algoritmos replican patrones históricos de discriminación, grupos marginados tendrán menos oportunidades de acceso a empleos bien remunerados y de calidad (Eubanks 2018).

2. Menor diversidad en las empresas

La falta de inclusión de perfiles diversos en los procesos de selección automatizados puede afectar la capacidad de innovación y competitividad de las empresas (ILO 2021).

3. Dificultad en la impugnación de decisiones

La opacidad en los sistemas de IA dificulta que los trabajadores afectados por decisiones algorítmicas puedan impugnar su validez o exigir una revisión justa (Wachter et al. 2017).

4. Desconfianza en la automatización

Cuando los trabajadores perciben que los procesos de IA son injustos o sesgados, pueden generar resistencia a la adopción de nuevas tecnologías en el entorno laboral (Rahwan et al. 2019).

Para mitigar estas consecuencias, es fundamental que las empresas adopten principios de transparencia y supervisión en el uso de IA en la contratación y gestión de personal al mismo tiempo en el que los gobiernos deben desarrollar regulaciones que aseguren que la IA se utilice de manera equitativa y responsable en el ámbito laboral.

X. MARCO NORMATIVO Y REGULACIONES APLICABLES

1. Legislación laboral peruana y su relación con la IA

En el Perú, la regulación sobre el uso de la inteligencia artificial en el ámbito laboral aún es incipiente. Actualmente, no existe una ley específica que regule el uso de algoritmos en la contratación y gestión del talento humano. Sin embargo, algunas disposiciones del Código del Trabajo y la Ley de Protección de Datos Personales (Ley N° 29733) pueden aplicarse a ciertos aspectos del uso de algoritmos de inteligencia artificial en las relaciones laborales (Congreso del Perú 2011).

El artículo 2 de la Constitución Política del Perú ya establece el derecho a la igualdad y no discriminación, lo que implica que cualquier tecnología utilizada en la gestión de empleados no debe generar prácticas discriminatorias (Tribunal Constitucional del Perú 2020). Otro recurso legal es lo dispuesto en la Ley de Protección de Datos Personales, misma que exige a las empresas que garanticen el consentimiento informado de los trabajadores en el uso de sus datos privados, un aspecto crucial cuando entendemos que los sistemas de inteligencia artificial recopilan y analizan dicha información.

No obstante, la falta de regulaciones específicas sobre el uso de esta tecnología genera vacíos legales que pueden derivar en la aplicación de algoritmos sin supervisión adecuada, lo que resalta la necesidad de desarrollar un marco regulatorio que garantice la transparencia y equidad en su uso.

2. Normativas internacionales sobre IA y sesgos algorítmicos

A nivel internacional, diversas regulaciones han comenzado a abordar los desafíos planteados, por ejemplo:

- A) Reglamento General de Protección de Datos (GDPR) de la Unión Europea: Marco normativo que establece la obligación de las empresas en cuanto a la proporción de explicaciones claras sobre las decisiones automatizadas y que, a su vez, garantiza el derecho de los trabajadores a impugnar decisiones tomadas exclusivamente por esta tecnología (Wachter, Mittelstadt y Floridi 2017).
- B) Propuesta de Ley de Inteligencia Artificial de la UE (AI Act.): Reglamento, aún en proceso de aprobación, que busca establecer restricciones sobre el uso de la inteligencia artificial en materias sensibles, incluyendo el empleo, donde se prohíben prácticas que puedan generar discriminación laboral automatizada (Consejo de la UE 2021).
- C) Fair Hiring Algorithms Act. (EE.UU.): En algunos estados de EE.UU., se han impulsado legislaciones que obligan a las empresas a evaluar la equidad de sus algoritmos de contratación y proporcionar informes sobre su impacto en la diversidad laboral (Barocas, Hardt y Narayanan 2019).
- D) Declaración de la OIT sobre el Futuro del Trabajo (2019): La Organización Internacional del Trabajo ha enfatizado la necesidad de es-

tablecer marcos normativos que regulen el uso de IA en la gestión de empleo, asegurando la protección de los trabajadores contra posibles sesgos algorítmicos (OIT 2019).

3. Propuestas para una regulación más equitativa en el Perú

- A) Regulación de la toma de decisiones automatizada: A través del establecimiento de normativas que exijan a las empresas evaluar la equidad de sus algoritmos y proporcionar explicaciones claras a los trabajadores sobre cómo se toman las decisiones que emplean este método.
- B) Creación de un ente regulador de IA en el trabajo: Utilizando de ejemplo los modelos de la UE y EE.UU., se podría desarrollar una entidad encargada de supervisar el impacto de la adaptación tecnológica en su propio mercado laboral de tal forma que se garantice su uso transparente y equitativo.
- C) Incorporación de auditorías algorítmicas obligatorias: Diseñado para las empresas que utilicen sistemas automatizados en la contratación y gestión de empleados deberían someterse a auditorías periódicas para detectar y corregir posibles sesgos discriminatorios.
- D) Fomento de la educación en IA y ética laboral: Resalta la necesidad de establecer sistemas que eduquen tanto a empleadores como a trabajadores sobre los riesgos y beneficios del uso de algoritmos en el ámbito laboral.

XI. ESTRATEGIAS PARA MITIGAR LOS SESGOS ALGORÍTMICOS EN EL ÁMBITO LABORAL

1. Modelos de auditoría y supervisión de IA

Según Raji et al. (2020), la auditoría algorítmica permite evaluar el desempeño de los sistemas de inteligencia artificial, identificando y corrigiendo sesgos antes de que generen impactos negativos en los trabajadores.

2. Estrategias para llevar a cabo las auditorías:

- A) Auditorías internas: Utilizando el ejemplo de empresas como IBM y Microsoft, que han desarrollado equipos especializados en ética algo-

rítmica que supervisan la implementación de estos sistemas, asegurando que cumplan con principios de equidad y transparencia (Binns 2020).

- B) Auditorías externas: A través de organizaciones independientes que evalúen los algoritmos utilizados en procesos de selección y evaluación laboral para detectar posibles sesgos. Un estudio de Sandvig et al. (2014) sugiere que la supervisión externa es clave para garantizar la imparcialidad de los sistemas.
- C) Supervisión continua: Se recomienda una evaluación constante para monitorear el desempeño y ajustar los modelos en caso de que surjan desviaciones discriminatorias para garantizar la funcionalidad de las medidas anteriores (Wachter, Mittelstadt y Floridi 2017).

XII. IMPLEMENTACIÓN DE PRINCIPIOS DE EQUIDAD Y TRANSPARENCIA

Según Floridi et al. (2018), la equidad en el uso de inteligencia artificial implica que los algoritmos no favorezcan de forma injusta a ciertos grupos por sobre otros, mientras que la transparencia requiere que los procesos de toma de decisiones sean comprensibles para los trabajadores en su totalidad.

1. Medidas para garantizar la implementación de los principios

- A) Explicabilidad de los algoritmos: Los trabajadores deben ser capaces de comprender cómo funcionan los modelos utilizados que afectan sus oportunidades laborales. La UE, por ejemplo, ha impulsado regulaciones que exigen la explicabilidad en la toma de decisiones automatizadas, como el Reglamento General de Protección de Datos (GDPR) (Wachter et al. 2017).
- B) Uso de datos representativos: Mehrabi et al. (2021) argumentan que los conjuntos de datos utilizados para entrenar algoritmos deben ser diversos e inclusivos para minimizar los riesgos de discriminación.
- C) Desarrollo de mecanismos de corrección: aspecto fundamental para que los algoritmos puedan ser ajustados lo más temprano posible en caso de que se detectan sesgos, de tal forma en la que se asegure

que las decisiones sean justas y equitativas para todos los trabajadores (Rahwan et al. 2019).

XIII. ROL DE LOS SINDICATOS Y REGULADORES EN LA VIGILANCIA DEL USO DE IA EN EL ÁMBITO LABORAL

La intervención de estos elementos, luego de analizar todos los problemas ya mencionados, se vuelve crucial para garantizar que la automatización no se traduzca en prácticas injustas ni en la vulneración de derechos laborales. Los sindicatos y las organizaciones de trabajadores desempeñan un papel clave en la supervisión de los sistemas algorítmicos utilizados en la contratación, evaluación y despidos dado que se encuentran en una posición más ventajosa para exigir auditorías independientes y transparentes es una de las estrategias más efectivas.

1. Casos de auditorías exitosas

Utilizando el ejemplo de HireVue, una plataforma de selección de personal basada en sistemas de inteligencia artificial que, tras múltiples cuestionamientos sobre sesgos de género y raza, fue sometida a una auditoría independiente en 2021. Como resultado de este proceso, la empresa se vió en la obligación de modificar sus modelos y mejorar la transparencia en la toma de decisiones automatizadas (Bogen y Rieke 2020).

Los sindicatos tienen un papel importante en el proceso para lograr que las auditorías, internas o externas, se conviertan en un requisito obligatorio en los procesos de contratación y evaluación que empleen este tipo de tecnología. Al mismo tiempo que promueven la creación de estándares éticos y técnicos para asegurar que los sistemas automatizados sean supervisados regularmente y evitar la dependencia única en la automatización.

XIV. EL CONTEXTO REGULADOR EN AMÉRICA LATINA Y PERÚ

En América Latina, el avance de medidas preventivas y regulatorias en torno a los sistemas de inteligencia artificial en el ámbito laboral ha sido más lento a comparación de Estados Unidos o Europa. En Perú, aún no existe una legislación específica que regule el uso de IA en la contratación y evaluación de empleados. Sin embargo, los sindicatos pueden impulsar regulaciones que se

alineen con estándares internacionales y exigir que las empresas implementen mecanismos de transparencia.

1. **Propuestas regulatorias para el Perú**

- A) **Obligatoriedad de informes de impacto de IA:** Las empresas que utilicen IA en procesos laborales deberían presentar informes que evidencien cómo sus algoritmos afectan la contratación y promoción de trabajadores.
- B) **Creación de un ente supervisor:** Similar al modelo europeo, el Perú podría establecer un organismo encargado de vigilar el uso de IA en el sector laboral y garantizar la equidad en la automatización del empleo.

XV. DEFENSA DEL DERECHO A IMPUGNAR DECISIONES AUTOMATIZADAS

Uno de los mayores riesgos en la automatización del empleo es la falta de mecanismos para que los trabajadores impugnen decisiones injustas tomadas por algoritmos. Los sindicatos pueden desempeñar un papel clave en la promoción de derechos digitales en el trabajo.

86

1. **Principales preocupaciones sobre la IA y el debido proceso laboral**

- A) **Desconocimiento del criterio de decisión:** Muchos trabajadores no saben cómo funcionan los sistemas de IA que evalúan su desempeño o deciden sobre su permanencia en la empresa.
- B) **Dificultad para apelar decisiones:** En algunos casos, las empresas confían exclusivamente en los algoritmos y no permiten que los empleados impugnen decisiones de despido o promoción.
- C) **Falta de supervisión humana:** Si la IA toma decisiones sin intervención humana, se pueden generar despidos injustos basados en métricas poco confiables.

2. **Medidas para fortalecer los derechos de los trabajadores ante la IA**

- A) Revisión humana: Ningún sistema de inteligencia artificial debería tomar decisiones laborales sin la supervisión de un ser humano.
- B) Comités de IA y equidad laboral: Integrar equipos de expertos en tecnología, derecho laboral y representantes sindicales para evaluar el impacto de la IA en el empleo.
- C) Garantía del derecho a la explicación: Las empresas se ven en la obligación de proporcionar la información detallada detrás del proceso de la toma de decisiones basadas en algoritmos, así como permitir a los empleados cuestionarlas si es que consideran que han sido injustas de alguna forma.

XVI. CONCLUSIÓN

En síntesis, podemos establecer que la inteligencia artificial tiene un enorme potencial para transformar la administración del trabajo, al verse en capacidad de ofrecer soluciones revolucionarias en sectores fundamentales, como lo son las áreas de selección de personal, distribución de responsabilidades y valoración del rendimiento. Sin embargo, es crucial que su implementación en el entorno laboral sea controlada y supervisada adecuadamente por todos los medios posibles para evitar que se perpetúen las desigualdades existentes y la discriminación debido a los reportes históricos con los que se calibra su uso. De no ser manejada con cautela, se podría fortalecer prejuicios previos a su implementación en decisiones laborales actuales impactando negativamente a determinados grupos e intensificando las desigualdades sociales y económicas históricas. Es debido a estos factores que se vuelve crucial el establecimiento de marcos normativos correctamente desarrollados e implementados que garanticen una correcta implementación de esta tecnología.

Una de las principales medidas para lograr esto se refiere a la implementación de auditorías periódicas de los algoritmos empleados, particularmente en procesos de reclutamiento, que permitan identificar y corregir errores en los modelos antes de que causen impactos negativos significativos contra las oportunidades laborales, además, resulta crucial que exista claridad en la operación de los algoritmos, tanto con las compañías como con los empleados para que estos últimos comprendan a profundidad cómo se realizan las decisiones automatizadas. Otro punto clave se refiere a contar con datos representativos y variados, compren-

diendo que los algoritmos aprenden de los datos que reciben, con mayor énfasis si estos reflejan desigualdades pasadas, se prioriza la implementación de medidas para que los modelos no repliquen los mismos sesgos. Por ejemplo, si los datos históricos de contratación favorecen a un grupo demográfico determinado, los algoritmos podrían también inclinarse a seleccionar a personas de ese mismo grupo, excluyendo a otros sin justificación. Por lo que se vuelve necesario por parte de las empresas asegurarse de que los datos sean inclusivos y representen de manera equilibrada a todas las personas, sin discriminación por género, raza, edad o cualquier otro diferencial.

Ante esto, la supervisión humana retoma una posición fundamental, se entiende que los algoritmos son herramientas poderosas, pero la toma de decisiones finales debe recaer en los seres humanos, dado que tienen la capacidad de interpretar los resultados de los algoritmos teniendo en cuenta un contexto más amplio en función de los valores sociales y éticos.

Finalmente, la normativa también tiene un rol crucial para garantizar un uso responsable de la tecnología. Se vuelve imprescindible que tanto leyes como políticas se ajusten rápidamente a los progresos tecnológicos, definiendo normativas precisas en torno a la salvaguarda de datos, la privacidad y la no discriminación. La colaboración entre empresas, reguladores y expertos en ética de implementación de inteligencia artificial se ha vuelto crucial para diseñar políticas que permitan a las empresas innovar sin comprometer los derechos de sus trabajadores.

En síntesis, aunque la IA puede transformar positivamente la gestión laboral, su implementación debe ser cuidadosamente controlada a través de auditorías regulares, la transparencia, los datos representativos, la supervisión humana y una regulación eficaz. Solo a través de un enfoque colaborativo entre diversos actores se podrá integrar la tecnología de manera que beneficie a la sociedad en su conjunto.

REFERENCIAS

Banco Interamericano de Desarrollo (BID). 2021. *AI in Latin America: Ethical and Social Implications*.

- Barocas, Solon, Moritz Hardt, y Arvind Narayanan. 2019. *Fairness and Machine Learning: Limitations and Opportunities*. Cambridge: MIT Press.
- Binns, Reuben. 2020. "Human Judgement in Algorithmic Recommendations." *Harvard Journal of Law & Technology* 33(1): 55-98.
- Bogen, Miranda y Aaron Rieke. 2020. *Help Wanted: An Examination of Hiring Algorithms, Equity, and Bias*. Upturn.
- Bogen, Miranda. 2019. "All the Ways Hiring Algorithms Can Introduce Bias." *Harvard Business Review*.
- Brynjolfsson, Erik, y Andrew McAfee. 2017. *Machine, Platform, Crowd: Harnessing Our Digital Future*. New York: W. W. Norton & Company.
- Congreso del Perú. 2011. *Ley de Protección de Datos Personales, Ley N° 29733*.
- Consejo de la Unión Europea. 2021. *Artificial Intelligence Act: Regulating the Use of AI in the EU*.
- Dastin, Jeffrey. 2018. "Amazon Scraps Secret AI Recruiting Tool that Showed Bias Against Women." *Reuters*, October 10, 2018.
- Eubanks, Virginia. 2018. *Automating Inequality: How High-Tech Tools Profile, Police, and Punish the Poor*. New York: St. Martin's Press.
- Floridi, Luciano, et al. 2018. "AI4People—An Ethical Framework for a Good AI Society." *Mind & Machine* 28(4): 689-707.
- International Labour Organization (ILO). 2019. *The Future of Work Report*.
- International Labour Organization (ILO). 2021. *Artificial Intelligence and the Future of Work*. ILO Report.
- Mehrabi, Ninareh, et al. 2021. "A Survey on Bias and Fairness in Machine Learning." *ACM Computing Surveys* 54(6): 1-35.
- O'Neil, Cathy. 2016. *Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy*. New York: Crown Publishing.
- Rahwan, Iyad, et al. 2019. "Machine Behaviour." *Nature* 568(7753): 477-486.

Raji, Inioluwa, et al. 2020. "Closing the AI Accountability Gap: Defining an End-to-End Framework for Internal Algorithmic Auditing." *Conference on Fairness, Accountability, and Transparency*.

Sandvig, Christian, Kevin Hamilton, Karrie Karahalios, y Cedric Langbort. 2014. "Auditing Algorithms: Research Methods for Detecting Discrimination on Internet Platforms." *Data and Discrimination: Collected Essays*.

Tribunal Constitucional del Perú. 2020. *Sentencia sobre igualdad y no discriminación en el empleo*.

Wachter, Sandra, Brent Mittelstadt, y Luciano Floridi. 2017. "Why a Right to Explanation of Automated Decision-Making Does Not Exist in the General Data Protection Regulation." *International Data Privacy Law* 7(2): 76-99.