



LOS LIBERTADORES
FUNDACION UNIVERSITARIA

Modelo de estimación de perfiles laborales exitosos. Caso de estudio: Entidad financiera.

Juan Camilo Cifuentes Ruiz, jccifuentesr@libertadores.edu.co, Fundación Universitaria Los Libertadores

Jonathan Camilo Flórez Díaz, jcflorezd@libertadores.edu.co, Fundación Universitaria Los Libertadores

Edgar Alejandro Valcárcel Gómez, eavalcarcelg@libertadores.edu.co, Fundación Universitaria Los Libertadores

José John Fredy González Veloza, jjgonzalezv02@libertadores.edu.co, Fundación Universitaria Los Libertadores

RESUMEN

El siguiente trabajo se realizó con la necesidad de crear una herramienta que permita predecir posibles perfiles exitosos. Se realizó por métodos de estimaciones no lineales con métodos de machine learning como lo son las regresiones logísticas o Random Forest. Se construyó un modelo que nos permitiera tener mayor objetividad al momento de elegir un perfil para la fuerza comercial de cierto banco. Los resultados arrojados por el modelo muestran una forma más objetiva de elegir perfiles laborales, así como de obtener las variables que son significativas a la hora de realizar la elección del candidato.

Palabras clave: Perfiles, modelo, variables, herramienta.

ABSTRACT

The following work was carried out with the need to create a tool that would allow us to predict possible successful profiles. It was carried out by non-linear estimation methods with machine learning methods as logistics regressions or random forest. A model built that allowed us to be more objective when choosing a profile for the commercial strength of a certain bank. The results

produced by the model show us a more objective way of choosing job profiles, as well as obtaining the variables that are significant when choosing the candidate.

Keywords: Profiles, model, variables, tool.

INTRODUCCIÓN

El siguiente trabajo se dio con la necesidad de implementar una herramienta que nos ayude a predecir perfiles exitosos para fuerza comercial de cierta entidad financiera es por esto que se pretende evaluar el mejor modelo que nos permita predecir los perfiles comerciales con alto desempeño, para que a través de la herramienta anteriormente mencionada se pueda de una manera más óptima y objetiva predecir perfiles exitosos al momento de la selección por parte del equipo de recursos humanos.

REFERENTES TEÓRICOS

Actualmente lo que busca la Inteligencia Artificial (IA) es hacer eficientes las labores que los humanos realizan y la gestión de los Recursos Humanos (RR. HH) no es la excepción (Park, 2018) hoy en día los RR. HH tienen un papel fundamental en la planificación y conducción estratégica de las organizaciones (KUMAR & NAGRANI, 2020). La IA puede definirse como un algoritmo programado que tiene la capacidad de aprender y tiene como objetivo parecerse a una labor humana (Park, 2018). En los RR.HH esta labor empieza a cobrar relevancia ya que reduce los tiempos de contratación y por ende ayuda a aumentar la productividad de los recursos humanos (KUMAR & NAGRANI, 2020) además de fomentar un reconocimiento en tiempo real y bonificación, o promoción o cambio de rol y responsabilidad, para tener éxito en las estrategias de negocios se debe fomentar unos recursos humanos más ágiles y creativos y de igual forma que sean más rápidos y eficaces (Park, 2018).

El desempeño se comprende como la serie de habilidades, conocimientos y aportes que impactan positivamente a las necesidades de la organización, este puede ser medido de muchas formas, desde competencias, indicadores de cumplimiento y conceptos cualitativos por parte de los líderes encargados de los procesos (Montejo 2001). En general la evaluación del desempeño se ha realizado de muchas formas, sin embargo existe un ciclo que puede generalizarse en:

- Evaluación
- Conceptualización
- Desarrollo
- Implementación
- Retroalimentación

En este proceso se hace una evaluación sobre los aspectos primarios que componen el desempeño de los empleados, no evalúan a los empleados, sino las políticas y metas organizacionales para conceptualizar los arquetipos que determinan un buen desempeño, desarrollan las estrategias para extraer, medir y comprender los datos, para finalmente realizar la implementación inicial de una evaluación del desempeño dónde se busca retroalimentar tanto el proceso como a las personas que requieren alcanzar las metas organizacionales propuestas. (Latorre 2011), (Montejo 2001).

El desempeño cobra mucha relevancia y las organizaciones son consistentes de que este es un elemento fundamental e importante a el día de hoy (Guartán, Torres, & Ollague, 2019). El recurso laboral es uno de los pilares fundamentales dentro de una organización ya que contribuye a su crecimiento y alcance de los objetivos de la misma, por eso a medida que ha evolucionado la gestión de los recursos humanos se han ido creando modelos de evaluación del desempeño (Henríquez & Calderón, 2012). Dado lo anterior en una organización en la cual se presentan

actividades del servicio de personal, se necesita del manejo de información sobre las características individuales del trabajador y reconocimiento del nivel de desempeño (Palmar & Valero, 2014).

En revisiones del estado del arte en cuanto al desempeño se encuentran diversas investigaciones dónde se hallan diferencias en la percepción y medición del desempeño entre hombres y mujeres, dónde al parecer se percibe en general que las mujeres y hombres tienden a obtener los mismos resultados en las evaluaciones de sus roles personales y organizacionales aún cuando históricamente se percibe que las mujeres no obtienen el mismo desempeño que los hombres (Godoy, 2009 & Todaro, 2002). Entre las variables consideradas una de las más relevantes en diversas investigaciones es el género o sexo, sin embargo existen diversas investigaciones que encuentran relaciones entre el desempeño y la satisfacción laboral, motivación y bienestar, siendo estas variables psicológicas fundamentales para la comprensión del desempeño en los últimos años, en lo que en general cuando aumenta el bienestar, satisfacción y motivación tenderá a aumentar el desempeño, en cambio indicadores cómo el ausentismo se encuentra que puede llegar a ser una variable predictora sobre el desempeño, ya que a su vez, relacionando lo encontrado en diversas investigaciones el ausentismo puede a su vez indicar una baja motivación y bienestar laboral las que son proporcionales positivamente al desempeño, convirtiendo esta variable de ausentismo en una variable indispensable para probar nuestro modelo. (Cueva, Díaz Ruiz, 2017 & Paredes, Urquia, 2020 & Ledesma, Crespo, Castro, 2018).

METODOLOGÍA

La investigación se desarrolló en 4 fases. En la fase 1 se realizó un análisis exploratorio, donde se tomaron los datos principales de las bases de datos internas de cierto Banco con las descripciones contempladas en la Tabla 1. De los datos se asumen dos posibles variables de entrenamiento que corresponden a indicadores comerciales y evaluación de desempeño (EVD).

Sexo	Sexo de la persona (Hombre-Mujer)
Fecha de Nacimiento	Fecha de nacimiento de las personas
Edad	Edad de la persona (redondeada hacia abajo)
Fecha de Ingreso	Fecha de ingreso a sus labores en el Banco
Antigüedad	Años de antigüedad en sus labores del Banco (redondeado hacia abajo)
Vp	Estructura en el nivel 1 (Vicepresidencia a la que pertenece)
Dir	Estructura en el nivel 2 (Dirección / Gerencia a la que pertenece)
Oficina	Estructura en el nivel 3 (Oficina en la que trabaja la persona)
Ciudad	Ciudades en las que se encuentra ubicada la oficina y la persona
Salario	Salario que devenga la persona
Categoría	Nivel de responsabilidad de la persona en números romanos, siendo I el más bajo y X el más alto
Tipo de contrato	Tipo de contrato de la persona
Procesos disciplinarios	Indicador si la persona esta reportada en procesos disciplinarios
Ausentismo	Días de ausentismo en el 2021
Aplica a conv interna	Indicador si ha aplicado a convocatorias internas
Cantidad de aplicaciones conv	Cantidad de aplicaciones a convocatorias internas
Jefe	Código que identifica al jefe de la persona
EVD	Evaluación de desempeño donde E es la calificación más alta y A la más baja
Indicadores comerciales 2021	Porcentaje de cumplimiento frente a las ventas esperadas, el 100% representa el cumplimiento tope, sin embargo, se pueden superar las ventas esperadas y obtener un porcentaje mayor a este

Tabla 1. Variables consideradas en el estudio.

En la fase 2 se procede a realizar un análisis descriptivo, exploratorio y multivariado para determinar relaciones y comportamientos entre el grupo de variables mediante las bibliotecas Sweetviz 2.1.3, Scikit-learn 1.0.1, Seaborn 0.11.2 y Matplotlib 1.6.2. En la fase 3 se crean y evalúan los modelos para determinar cual se ajusta más al objetivo de predecir los perfiles exitosos, en este se utilizan las bibliotecas PyCaret 2.3, Scikit-learn 1.0.1. Por último, en la fase 4 se realiza creación e implementación de la herramienta a través de un formulario en Google Colab.

RESULTADOS

Se realizó la prueba con los datos de 1029 personas de los cuales el 24% son hombres (300) y el 76% son mujeres (909), estas personas tienen una edad promedio de 33 años, siendo la persona más joven de 19 años y la mayor de 60 años. En promedio se encuentra una antigüedad de 7.7 años, teniendo el más nuevo 0 años de antigüedad y el más antiguo 41 años. En cuanto a los salarios devengados por las personas comerciales se encuentra un promedio de 1.727.821 COP, siendo el mínimo 1.193.866 COP y el máximo 2.911.500 COP. Al revisar las variables de entrenamiento se encontró que hay 291 personas con un bajo desempeño y 918 personas con un buen desempeño, en cuanto a los indicadores comerciales se encontraron que en promedio se obtiene un cumplimiento del 60%, teniendo el mayor cumplimiento comercial con 140% y el menor cumplimiento con 0%.

Dada las revisiones hechas a las variables se observó que el grupo no tiene un comportamiento de normalidad univariada dado que al aplicar la prueba de Kolmogorov se obtiene un p-valor menor a 0.05 lo que nos indica que se rechaza la hipótesis de normalidad.

En la Figura 1 observamos la correlación entre la edad, salario y antigüedad, allí destacamos el hecho que a mayor antigüedad dentro del Banco mayor será su edad y salario, se encuentra que el salario tiende a no incrementar cuando la antigüedad se encuentre entre 15 y 35 años aproximadamente.

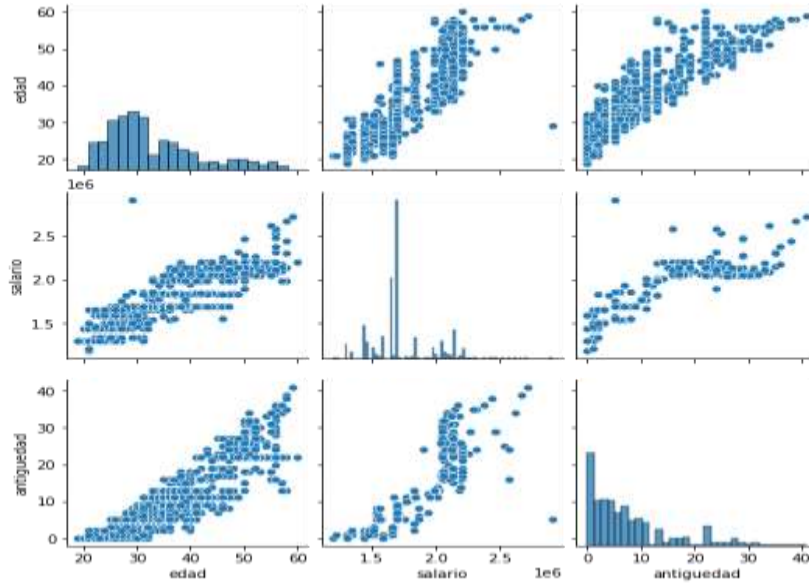


Figura 1

Como se observa en la Figura 2 no se ve de manera bivariada relación entre las variables y los indicadores comerciales, por lo que se realiza una prueba de correlación entre las variables con la EVD y se ve la existencia de posibles relaciones para entrenar un modelo.

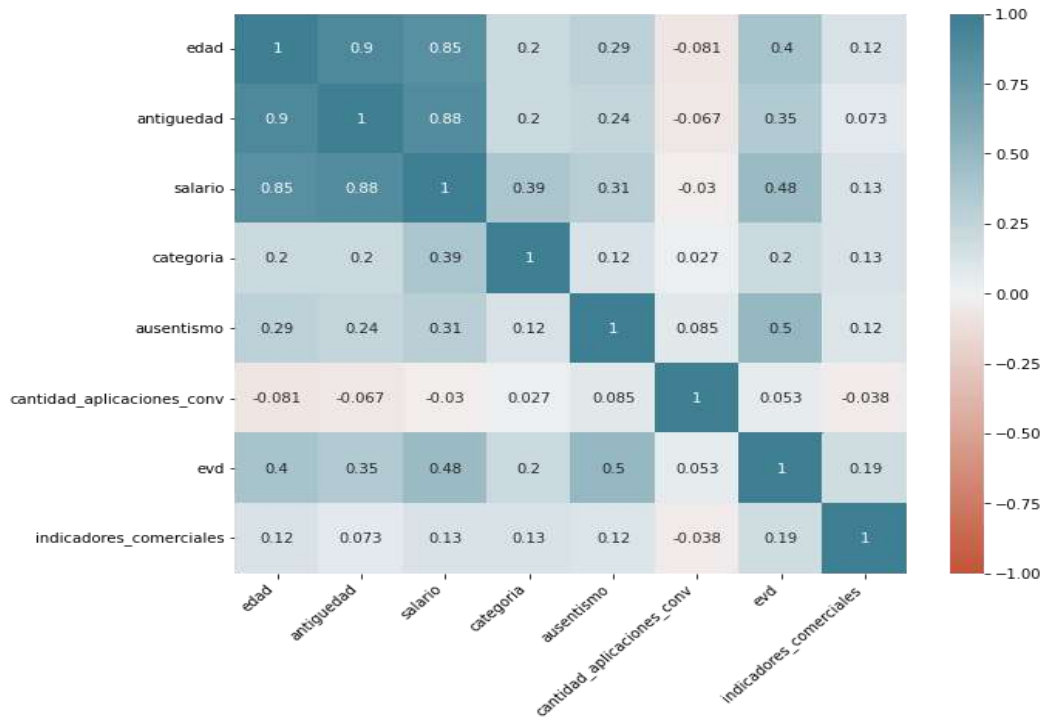


Figura 2

En la Ilustración 2 la correlación se encuentra entre 0 y 0.5, donde al parecer la edad, antigüedad y el salario tienden a estar correlacionadas fuertemente, con relaciones superiores a 0.85, por lo que podemos asumir que cuando aumenta la edad también la antigüedad y por ende el salario. La categoría sostiene correlaciones muy débiles con las otras variables, sin embargo, al parecer se relaciona un poco más con el salario, por lo que a mayor responsabilidad se encuentra un mayor salario. El ausentismo tiene una correlación positiva con la evaluación de desempeño (evd), donde se relaciona con un $r = 0.5$, lo que nos lleva a pensar que cuando aumentan los ausentismos tiende a aumentar la calificación en la evaluación de desempeño. En cuanto a las variables de entrenamiento se encontró una relación entre la evaluación de desempeño y diversas variables como lo son la edad, el salario y el ausentismo ($r = 0.4$, $r = 0.48$, $r = 0.5$), no obstante, la variable de indicadores comerciales se relaciona de manera débil con las demás variables lo que implica que el modelo de entrenamiento puede ajustarse más al entrenarse con la evaluación de desempeño. Al realizar un análisis de componentes principales (PCA) se decide eliminar la variable antigüedad para evitar colinealidad con la variable edad.

Por la naturaleza del objetivo del trabajo decidimos escoger el modelo Random Forest Classifier dado sus métricas y comportamiento de aprendizaje ya que el K Neighbors Classifier está basado en agrupaciones (clusters) por lo que al clasificar con la base de testeo cumplía con las características del “clasificador tonto”, clasificando a todas las personas con alto desempeño, por lo que debido al desbalanceo de las clases los puntajes de ajuste son bastante altos.

Modelo	Accuracy	F1 Score	Precisión
Random Forest Classifier	0.9374	0.9608	0.9304
Logistic Regression	0.7683	0.8690	0.7683
K Neighbors Classifier	0.9433	0.9646	0.9348

Tabla 2. Métricas de los modelos.

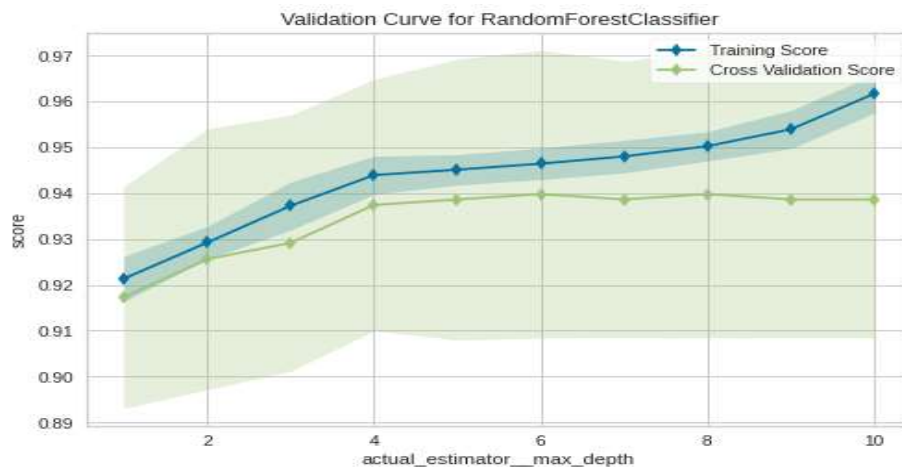


Figura 3

Como se ve en la Figura 3 a partir de la profundidad 6 empieza a existir un sobreajuste, aunque las métricas muestran y respaldan que el modelo tiene un buen comportamiento de clasificación y exactitud de los perfiles comerciales.

DISCUSIÓN DE RESULTADOS

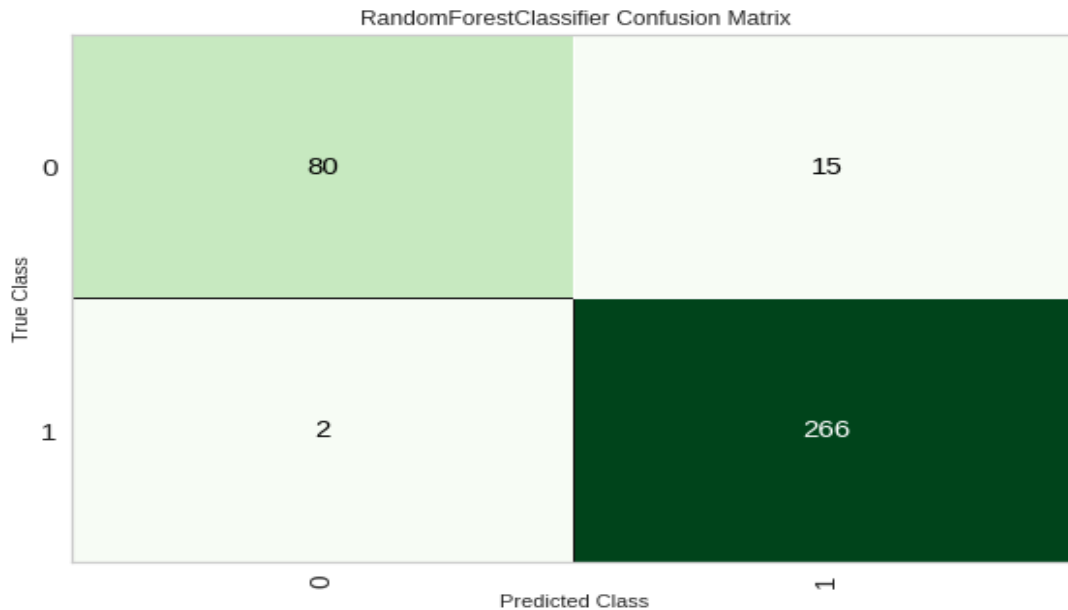


Figura 4

Teniendo en cuenta que el modelo arroja un Accuracy del 93% y un F1score del 96% y ver la clasificación que arroja el modelo se deduce que el modelo cumple con la precisión necesaria para ser implementado ya que de un total de 363 usuarios en la base de validación solo el 5% está quedando mal clasificado lo que indica que un 95% esta asignándose correctamente a los perfiles, por otro lado, los juicios quedan reducidos a cualquier elección subjetiva. Además, el modelo presenta coherencia con el comportamiento de las variables y lo descrito por los autores revisados en el estado del arte, como lo menciona Godoy y Togaro en la que el sexo es indiferente para explicar el desempeño de un usuario, además como se mostraba en el gráfico de correlación una persona que devengue más salario y posea una mayor edad tendrá mayor desempeño.

En la Figura 5 se observa que las variables de edad y salario tienen un peso en el modelo de 41% y 28% respectivamente se consideran las variables más importantes en el modelo, además el ausentismo con una representación del 13% contrasta con lo encontrado por Cueva, Diaz (2017) en donde el ausentismo implicaba una reducción directa en el desempeño ya que mostraba un bajo interés en sus funciones y una pérdida en las horas productivas, sin embargo en este modelo en

cuando los días de ausentismo se encuentran entre 1 y 7 aumenta la probabilidad de que el usuario tenga alto desempeño. Por otro lado, cuando se supera este umbral, el desempeño disminuye, lo que abre la discusión a una nueva hipótesis de que si el ausentismo se traduce en días de descanso el usuario tendrá un mejor desempeño en periodos cortos de trabajo.

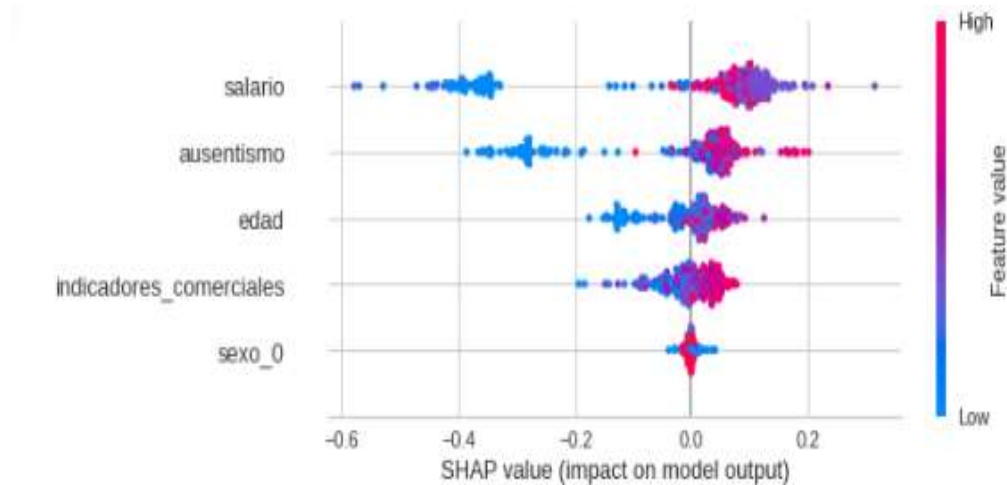


Ilustración 5

CONCLUSIONES

1. El modelo basado en Random Forest predice correctamente el 95% de los perfiles comerciales exitosos y no exitosos de la entidad financiera.
2. El modelo nos permite tener una visión más objetiva a la hora de elegir perfiles comerciales y de esta forma evitar juicios subjetivos.
3. Las variables más significativas para explicar que una persona tenga alto desempeño es la edad y el salario.
4. La variable sexo es indiferente para explicar si una persona tendrá o no un alto desempeño, así como lo detalla Godoy y Togaro en sus respectivos trabajos.

5. La variable ausentismo deja de ser representativa después de que un individuo se ausenta más de 8 días y menos de 1 día, basado en esto podríamos evaluar una hipótesis en la que una persona tendría un alto desempeño si se ausenta entre 1 y 7 días.

REFERENCIAS BIBLIOGRÁFICAS

Chacaltana, J., & Yamada, G. (2009). Calidad del empleo y productividad laboral en el Perú (No. 691). Working Paper.

Cueva Gonzales, K., & Díaz Ruiz, M. A. (2017). Satisfacción laboral y su relación con el desempeño laboral de los colaboradores de Mibanco, 2017.

Godoy, L., & Mladinic, A. (2009). Estereotipos y roles de género en la evaluación laboral y personal de hombres y mujeres en cargos de dirección. *Psyche (Santiago)*, 18(2), 51-64.

Guartán, A., Torres, K., & Ollague, J. (2019). La evaluación del desempeño laboral desde una perspectiva integral de varios factores. *Digital Publisher*, 13-26.

Henríquez, J. S., & Calderón, V. C. (2012). Diseño del proceso de evaluación del desempeño del personal y las principales tendencias que afectan su auditoría. *pensamiento y gestión*, N° 32, 54-82.

KUMAR, S. P., & NAGRANI, K. (2020). ARTIFICIAL INTELLIGENCE IN HUMAN RESOURCE MANAGEMENT. *Novateur Publication's JournalNX- A Multidisciplinary Peer Reviewed Journal*, 106-118.

Latorre Navarro, M. F. (2011). La gestión de recursos humanos y el desempeño laboral.

Ledesma, C. R. D., Crespo, H. G., & Castro, A. M. A. (2018). Ausentismo y desempeño laboral en profesionales de enfermería de áreas críticas. *Revista Cuidarte*, 9(1), 1973-1987.

Montejo, A. P. (2001). Evaluación del desempeño laboral. *Gestión*, 2(9).

Palmar, R., & Valero, J. (2014). Competencias y desempeño laboral de los gerentes de los institutos autónomos dependientes de la Alcaldía del municipio Mara del estado Zulia. *Espacios públicos*, 159-188.

Paredes Mendo, J. A., & Urquia Rojas, D. B. (2020). Satisfacción laboral y el desempeño laboral en los trabajadores de la Municipalidad Distrital de Tabalosos, San Martín 2020.

Park, W. (2018). Artificial intelligence and Human resource management: New perspectives and challenges. School of Management-Kyung Hee University, 1-16.

Raschka, S., & VahidMirjalili. (2019). Python Machine Learning. Aprendizaje automático y aprendizaje profundo con Python, scikit-learn y TensorFlow. Bogotá: Marcombo.

Salazar Estrada, J. G., Guerrero Pupo, J. C., Machado Rodríguez, Y. B., & Cañedo Andalia, R. (2009). Clima y cultura organizacional: dos componentes esenciales en la productividad laboral. *Acimed*, 20(4), 67-75.

Todaro, R., Godoy, L., & Abramo, L. (2002). Desempeño laboral de hombres y mujeres: opinan los empresarios. *Cadernos pagu*, 197-236.