
Aplicación de Machine Learning para la estimación de la rotación de empleados en el sector BPO

Application of Machine Learning for estimating employee turnover in the BPO sector

Camilo Andrés Rey ^a
Camilo.rey@usantotomas.edu.co

Javier Sierra ^b
Javiersierra@usta.edu.co

Resumen

El presente estudio utiliza aplicaciones de técnicas estadísticas y Machine Learning para el estudio de la rotación de empleados Business Process Outsourcing, en términos de Recursos Humanos, el BPO (Business Process Outsourcing) implica externalizar ciertas funciones y procesos relacionados con la gestión del talento humano a una empresa especializada en este servicio. Esto puede incluir actividades como reclutamiento y selección de personal, administración de nóminas, gestión de beneficios, formación y desarrollo, gestión del desempeño, entre otros., a lo largo de este, se estudiarán las variables que afectan ese fenómeno y se crea un modelo de Machine Learning para la predicción de la rotación de empleados, cuyo objetivo es disminuir los costos asociados a la rotación de los mismos y los efectos negativos que se tengan a nivel organizacional.

Palabras clave: Rotación de empleados, BPO, Machine Learning, Regresión Logística, Random Forest.

Abstract

The present study uses applications of statistical techniques and Machine Learning to study employee turnover Business Process Outsourcing, in terms of Human Resources, BPO (Business Process Outsourcing) involves outsourcing certain functions and processes related to the management of human talent to a company specialized in this service. This may include activities such as recruitment and selection of personnel, payroll administration, benefits management, training and development, performance management, among others. Throughout this, the variables that affect this phenomenon will be studied and a model is created. of

^aEstudiante

^bProfesor

Machine Learning for the prediction of employee turnover, whose objective is to reduce the costs associated with employee turnover and the negative effects it has at the organizational level.

Keywords: Employee turnover, BPO, Machine Learning, Logistics Regression, Random Forest.

1. Justificación

La incorporación de sistemas predictivos en el proceso de contratación se presenta como un paso adelante estratégico para las empresas modernas. Al prever la posible rotación de personal desde las primeras fases del reclutamiento, las organizaciones pueden mejorar la gestión de sus recursos humanos, identificar individuos con un mayor potencial de permanencia y, por ende, disminuir los gastos relacionados con la rotación de personal. Esta capacidad de anticipación les permite implementar tácticas proactivas de retención, ofreciendo incentivos y un entorno laboral favorable que fomente el compromiso a largo plazo. Además, el enfoque predictivo facilita una planificación de la fuerza laboral más sólida, lo que permite a las empresas adaptarse rápidamente a los cambios del mercado y a las dinámicas internas. Por lo tanto, la inclusión de modelos predictivos desde la etapa de reclutamiento no solo mejora la eficacia operativa, sino que también establece un estándar de gestión del talento que impulsa el crecimiento sostenible y la competitividad en el panorama empresarial actual. En resumen, la capacidad de anticipación es fundamental para el éxito en la gestión del talento.

2. Objetivos

2.1. Objetivo General

El propósito principal es crear un modelo de predicción robusto para anticipar la rotación de personal en un contexto de reclutamiento dentro de una empresa de Externalización de Procesos Empresariales (BPO). Este modelo se desarrollará utilizando datos obtenidos durante el proceso de contratación, y se buscará comprender en profundidad los factores que influyen en la rotación del personal mediante un análisis detallado de variables recopiladas durante el reclutamiento, tanto cuantitativas como cualitativas. Se emplearán técnicas estadísticas como la regresión logística y el bosque aleatorio. Después, se compararon los resultados y el rendimiento de ambos modelos para evaluar su eficacia y precisión.

2.2. Objetivos específicos

- Identificar los principales determinantes de la rotación de personal en un Centro de Externalización de Procesos Empresariales (BPO): Este objeti-

vo implica examinar una amplia gama de variables, como la edad, nivel educativo, expectativa salarial previa, lugar de residencia, entre otras, para determinar cuáles tienen el mayor impacto en la decisión de los empleados de abandonar la empresa.

- Evaluar la eficacia y precisión del modelo predictivo desarrollado mediante pruebas y validación utilizando conjuntos de datos independientes y técnicas de validación cruzada, además de compararlo con estudios previamente realizados utilizando técnicas similares.
- Construir y entrenar modelos predictivos utilizando técnicas de aprendizaje automático y análisis estadístico para prever la rotación de empleados basándose en las variables recopiladas durante el proceso de reclutamiento.

3. Marco Teórico

En el dinámico mundo empresarial, caracterizado por una fuerte competencia y cambios constantes, las compañías dedicadas a la subcontratación de procesos de negocio enfrentan importantes obstáculos en la retención de su personal. La alta rotación de empleados en un BPO puede afectar negativamente la productividad, la calidad del servicio y la rentabilidad. Por consiguiente, resulta fundamental elaborar modelos predictivos eficaces que permitan comprender y anticipar la rotación de personal en este contexto (Fang et al., 2017).

Una estrategia habitual para crear modelos predictivos de rotación de personal en un BPO consiste en recopilar y examinar datos vinculados al proceso de contratación. Estos modelos a menudo emplean técnicas sofisticadas de análisis de datos, como el machine learning, para detectar patrones y relaciones que puedan prever la probabilidad de que un empleado renuncie a su puesto.

Además, es crucial considerar el impacto de factores externos, como la economía, el mercado laboral y las tendencias de la industria, en la rotación del personal. Los modelos predictivos deben incorporar estas variables para mejorar su precisión y utilidad en la toma de decisiones estratégicas.

Por otro lado, algunos estudios han profundizado en las técnicas de modelado predictivo utilizadas para prever la rotación de empleados. Por ejemplo, la regresión logística ha sido ampliamente empleada debido a su capacidad para modelar las relaciones entre variables predictoras y la probabilidad de rotación de empleados (Choi et al., 2016). Sin embargo, investigaciones más recientes han sugerido que técnicas de aprendizaje automático, como Random Forest, pueden ofrecer una mayor precisión en la predicción de la rotación de empleados en empresas de BPO (Smith et al., 2020).

A pesar de estos hallazgos, aún persiste una falta de consenso sobre cuál técnica de modelado predictivo es la más efectiva para predecir la rotación de empleados en empresas de BPO. Mientras algunos estudios respaldan el uso de técnicas tradi-

cionales como la regresión logística, otros sugieren que las técnicas de aprendizaje automático pueden proporcionar resultados más precisos y predictivos.

Para abordar esta disparidad y avanzar en la comprensión de la predicción de la rotación de empleados en empresas de BPO, se recomienda llevar a cabo un estudio comparativo que evalúe el rendimiento y la eficacia de diferentes técnicas de modelado predictivo. Este estudio podría comparar el desempeño de la regresión logística y Random Forest en la predicción de la rotación de empleados utilizando datos reales de empresas de BPO.

Según Ramírez (2016), el tener una tasa elevada de rotación de empleados en las empresas puede representar una gran preocupación, no solo debido al impacto en el ambiente laboral, sino también por las implicaciones económicas. Esto se debe a que una alta rotación implica costos considerables para la organización, los cuales resultan difíciles de recuperar.

Además de los aspectos financieros, la rotación de empleados también genera impactos sociales en toda la organización. Esto se debe a que provoca interrupciones en las estructuras sociales y en los canales de comunicación, lo que puede requerir un tiempo considerable para restaurar. En algunos casos, este fenómeno incluso puede resultar en la salida de más empleados recién contratados de la organización (Levin y Kleiner, 1992).

Los factores intervinientes para que una persona decida abandonar su actual empleo serían una mezcla de condiciones de sobreoferta del mercado de trabajo (pull) junto con características psicológicas y organizacionales que empujan (push) al individuo a dejar su trabajo actual de ahí la importancia que algunos autores han dado a las variables perceptivas y cognitivas del proceso (Bretones y González, 2009).

4. Resultado prácticos

4.1. Extracción de datos

Para la investigación se obtendrán los datos históricos de la empresa BPO center recopilados a lo largo de las contrataciones del primer trimestre del año 2024, así como también reportes extraídos de la plataforma de gestión humana donde se podrá contar con la información del personal activo y no activo, se tendrá acceso a reportes adicionales como las entrevistas realizadas a cada empleado y razones de terminación.

4.2. Datos disponibles

Los datos comprenden 1664 registros cuyas mediciones abarcan 19 variables, esta información contiene datos del personal activo/inactivo, dado que el objetivo del estudio es analizar factores que influyen en la rotación antes de 90 días del personal

no administrativo en BPO center, es de suma relevancia los datos con los que se va a realizar este análisis. Los 768(46.15 %) registros inactivos son importantes para enriquecer el análisis, dado que en esos se puede determinar, sin importar en qué momento se realice el corte de datos, el tiempo exacto que duraron en la empresa y en consecuencia si superan el umbral de los 90 días o no, por otra parte, los 884(53.125 %) registros activos no deberían ser tenidos en cuenta en su totalidad, puesto que, existen casos en los que los empleados a esa fecha de corte de datos, llevan menos de 90 días laborando y quedarían registrados en la variable respuesta como cero, en referencia a que duraron menos de 90 días antes de renunciar, caso contrario ocurre con los empleados activos que llevan más de 90 días en la empresa, dado que, independiente si renuncian o no en un futuro la variable respuesta no cambiaría, en consecuencia, de los empleados activos solo se tendrán en cuenta aquellos que lleven más de 90 días trabajados al corte de los datos, de lo contrario, utilizar los usuarios activos restantes, podría producir un sesgo significativo el los resultados obtenidos.

4.3. Procesamiento de datos

En función de la política de tratamientos de datos personales, la Constitución Política de Colombia, en su artículo 15, garantiza el derecho a la intimidad y al buen nombre, permitiendo a las personas conocer, actualizar y rectificar sus datos en entidades públicas y privadas. La Ley 1581 de 2012 regula la recolección y tratamiento de datos personales, y el Decreto 1074 de 2015 detalla su implementación. Estos instrumentos legales establecen las políticas de autorización, tratamiento, y derechos de los titulares de los datos, así como las responsabilidades de los encargados de su manejo. La Corte Constitucional ratificó la validez de esta legislación mediante la Sentencia C-748 de 2011. Estos marcos normativos buscan asegurar la protección y confidencialidad de la información personal, en el ejercicio normativo, con el fin de ejercer el cumplimiento de los mismo, se eliminó los datos personales asociaciones a los diferentes colaboradores del sector BPO seleccionado, dentro de los cuales se tienen la identificación y el correo electrónico del colaborador, datos sensibles cuya protección debe ser correcta.

Se analiza la integridad de los datos, eliminando 16(0.91 %) registros duplicados, no se registran datos faltantes en ninguna de las variables, aunque las variables Highest Level y Bpo experience registran valores de 0, lo cual se entenderá como “no registra”. Adicionalmente, se eliminaron las variables asociadas al tiempo, menos la fecha de la vinculación, que será utilizada como variable explicativa extrayendo el mes de esta para capturar comportamientos estacionales, por último, se eliminará la razón de la terminación del vínculo, porque no es algo que se pueda saber en el momento de la contratación, por ende, no dará valor agregado al análisis. En cuanto a la integridad de las variables, se realizó la depuración de datos orientada a la escritura de variables, donde se encontró que una misma característica dentro de la variable se escribía con y sin mayúsculas, paso seguido, se recalculó la variable objetivo indicadora de duración mayor a 90 días, dado que, cuando se replicó el cálculo se encontraron diferencias.

4.4. Análisis Exploratorio de Datos

En este apartado, se busca realizar un análisis de datos a nivel exploratorio e inferencial con el propósito de orientar los datos a la aplicación de metodologías de Machine Learning.

En un principio, se analiza la proporción de datos de la variable objetivo, que indica si el empleado duró más de 90 días como uno y cero en caso contrario, dado que, para la estimación de modelos de Machine Learning, es importante que la variable de interés esté balanceada en cuanto sus clases, para este caso, se tiene que un claro desbalance, puesto que las características marcadas como Si (uno) abarca la gran mayoría de los registros con 754(85.68 %) empleados, en contraparte, se tienen 126(14.32 %) que no duraron más de 90 días. Por ende, previo a la estimación de los modelos de Machine Learning se debe realizar un método de balanceo de clases.

Posterior a esto, se analiza la relación entre las variable de interés y las variables explicativas, encontrando resultados interesantes que ayudan a la preselección de variables que serán utilizadas en la parte de estimación de los modelos de Machine Learning, utilizando la prueba ANOVA para la variable Edad y la prueba chi cuadrado de independencia para estudiar posibles relaciones entre variables cualitativas, independientemente del método estadísticos a utilizar el sistemas de hipótesis a contrastar es:

$$\begin{cases} H_0 : \text{No hay asociación entre la variable respuesta y la variable explicativa } x_i \\ H_1 : \text{Hay asociación entre la variable respuesta y la variable explicativa } x_i \end{cases}$$

<i>Variable</i>	<i>Método</i>	<i>P - valor</i>	<i>Decisión</i>
Age	ANOVA	0.211	No hay relación
Program	Chi - Cuadrado	0	Relación
Marital Status	Chi - Cuadrado	0.2	No hay Relación
Gender	Chi - Cuadrado	0.015	Relación
Neighborhood	Chi - Cuadrado	0.5	No hay relación
English Level	Chi - Cuadrado	0.07(0.02)	No hay relación(relación)
Salary expectation	Chi - Cuadrado	0.2	No hay relación
Highest level	Chi - Cuadrado	0.8	No hay relación
BPO experience	Chi - Cuadrado	0.5(0.2)	No hay relación
ETA	Chi - Cuadrado	0.3(0.7)	No hay relación
Mes de incorporación	ANOVA	0	Relación

Tabla 1: Resultados pruebas de hipótesis relación entre variables.

Encontrando en un principio que las únicas variables que tienen relación con que un empleado dure más de 90 días o no en su cargo son el área a la que pertenece el empleado, género y mes de contratación , adicionalmente, entre paréntesis se anexaron los p - valores asociados a las pruebas en las que se crearon nuevas variables con menos clases y se volvió a calcular la prueba, encontrando la variable

nivel de inglés, con una relación con la variable objetivo, trabajando esta con solo dos clases A1 a B1 y B2 a C2, por ende, se trabajará esta variable de esta manera, resultado con 4 variables para ser utilizadas en los modelos.

Previo a la estimación del modelo, se realizó una técnica de sobre muestreo, con la finalidad de balancear las clases de la variable objetivo y reducir el sesgo que pueda presentar el modelo.

5. Estimación de modelos

En este apartado, se estimaron dos tipos de modelos, Regresión Logística y Random Forest, el primero, un método estadístico clásico, el cual por su estructura es fácil de interpretar, adicionalmente, tiende a presentar buenos resultados en comparación a modelos computacionales más avanzados, por otra parte, el Random Forest es un modelo computacionalmente muy completo, cuya estructura permite obtener resultados muy potentes en términos de la estimación de la variable objetivo. El objetivo de este apartado es comparar los resultados que se obtuvieron en cada uno de los desarrollos utilizando la metodología Cross Validation para garantizar los resultados.

5.1. Regresión Logística

Se realizó la estimación del modelo logístico calibrando únicamente el parámetro C , este controla la fuerza de la penalización aplicada a los coeficientes del modelo. Un valor bajo de C indica una fuerte regularización, lo que previene el sobreajuste pero puede llevar a una menor precisión en los datos de entrenamiento. Un valor alto de C permite un mayor ajuste a los datos de entrenamiento, pero puede aumentar el riesgo de sobreajuste. La selección de este parámetro es crucial para equilibrar la capacidad de generalización del modelo con la precisión en los datos de entrenamiento. Los resultados de las metricas para diferentes valores de C se observa en la siguiente tabla.

C	<i>exactitud</i>	<i>Precisión</i>
91	0.776973	0.716322
261	0.776973	0.708003
151	0.776143	0.710767
241	0.776143	0.708003
221	0.776143	0.709855

Tabla 2: Metricas Regresión Logistica.

Obteniendo para valores altos de C el mejor rendimiento en términos de exactitud y precisión, lo cual, nos permite concluir que el modelo estima bien general la

variable respuesta Y , además, de generar estimaciones aceptables para la deserción de empleados después de 90 días (Valor positivo en la variable explicativa $Y = 1$)

5.2. Random Forest

Se probaron diferentes valores para los parámetros $n_{estimators}$ y max_{depth} , el primero controla el número de árboles en el bosque, un valor mayor de $n_{estimators}$ generalmente produce un modelo más robusto que puede capturar mejor la relación entre las características y las etiquetas, pero también puede aumentar el costo computacional y el riesgo de sobreajuste, por su parte, el max_{depth} determina la profundidad máxima de cada árbol en el bosque, una profundidad mayor permite a los árboles capturar relaciones más complejas en los datos, pero también puede aumentar el riesgo de sobreajuste. Es importante ajustar max_{depth} para encontrar un equilibrio entre la capacidad de generalización del modelo y el riesgo de sobreajuste. Los resultados obtenidos se observan en la siguiente tabla.

$n_{estimators}$	max_{depth}	<i>Exactitud</i>	Precisión
100	20	0.791893	0.707876
100	30	0.791893	0.707874
100	40	0.791893	0.707875
100	50	0.791893	0.707876
400	50	0.789404	0.707876

Tabla 3: Metricas Random Forest.

Obteniendo, bajo la combinación de parámetros $n_{estimators} = 100$ y $max_{depth} = 20$ la mejor exactitud en el modelo y una precisión aceptable, además, cualquiera de los modelos con parámetro $n_{estimators} = 100$ pudo haber sido seleccionado, dado que no presenta diferencias abruptas independientemente del valor que toma el otro hiperparámetro max_{depth} .

5.3. Selección de modelo

Posterior a la estimación de los modelos *Regresión Logística* y *Random Forest*, se obtuvieron resultados muy similares bajo ambas metodología, sin embargo, como era de esperarse por el poder computacional que tiene el modelo *Random Forest* resultó brevemente superior en términos de rendimiento, dado que esa diferencia no es abrupta, se optara por la regresión logística dado que permite la explicación y interpretación de los resultados.

El modelo de regresión logística seleccionado se define como:

$$\ln \left(\frac{p}{1-p} \right) = -13.43 + 1.42x_{\text{Mes}} - 0.24x_{\text{Masculino}} - 0.74x_{\text{InglesB2aC2}} \\ - 0.05x_{\text{Programa}_2} + 0.46x_{\text{Programa}_3} + 0.73x_{\text{Programa}_4} + 0.09x_{\text{Programa}_5} \\ - 0.74x_{\text{Programa}_6} + 0.78x_{\text{Programa}_7} + 0.08x_{\text{Programa}_8} + 0.28x_{\text{Programa}_9} \\ + 0.003x_{\text{Programa}_{10}} - 0.09x_{\text{Programa}_{12}} - 0.02x_{\text{Programa}_{13}} + 0.16x_{\text{Programa}_{15}} \\ - 0.03x_{\text{Programa}_{16}} \quad (1)$$

En cuanto a la interpretación de los coeficientes, se interpretan los ODDS, que hacen referencia a la división de la probabilidad de que un empleado renuncie después de 90 días y la probabilidad de que el empleado renuncie antes de 90 días, en el caso del mes de contratación, manteniendo las demás variables constantes, cuando el mes de contratación aumenta en un mes, el ODDS aumenta en 4.15 veces, es decir, existe una relación positiva entre el mes de contratación y la renuncia después de 90 días por parte de un empleado. Por otra parte, en cuanto al nivel de inglés, tener un nivel alto de inglés tiene una asociación negativa con renunciar después de 90 días, lo cual supone que si una persona tiene un nivel de inglés medio a avanzado, es más probable que renuncie antes de 90 días, en cuanto al género del empleado, si este es hombre, es más probable que renuncie en tiempos menores a 90 días, por último, dentro de las variables que se consideraron en el modelo, el área a la que pertenece el empleado, aquellos que que hayan sido contratados en las áreas 6, 12, 2, 16, 13 tienen probabilidades mayores de desertar antes de los 90 días, siendo los del área 6 los que generan un menor aumento en el ODD, en contra parte, si el individuo fue contratado en las áreas 10, 8, 5, 15, 9, 3, 4 y 7, son áreas donde el coeficiente resultó positivo, asumiendo un aumento en la probabilidad de permanecer 90 días como mínimo en el puesto, siendo aquellos que trabajan en el área 7 aquellos que tienen un impacto mayor en el ODD.

6. Conclusiones

En el análisis de rotación de empleados en el sector BPO, se emplearon metodologías clásicas como la Regresión Logística y metodologías computacionalmente más potentes como Random Forest. Aunque Random Forest mostró un mejor rendimiento, se optó por el modelo de Regresión Logística debido a su capacidad de interpretación, lo que proporciona un valor agregado significativo sin una disminución considerable en el rendimiento.

El valor agregado del modelo es crucial para el negocio, ya que ofrece una herramienta adicional para los procesos de selección de personal. El análisis inicial permite conclusiones más robustas, indicando que variables tradicionalmente consideradas importantes para la estabilidad laboral, como el salario o el tiempo de desplazamiento al trabajo, no resultaron significativas en este estudio. Por ejemplo, se encontró que personas con un nivel de inglés más bajo tienden a rotar menos,

sugiriendo que la empresa debería considerar contratar personal menos calificado para reducir la rotación.

Para implementar estas conclusiones y mejorar los procesos de retención de empleados, es fundamental revisar y adaptar las políticas de selección, incorporando variables significativas como el nivel de inglés y ajustando los criterios para priorizar candidatos con menor tendencia a la rotación. Además, desarrollar estrategias de retención personalizadas, como programas de integración y mentorías, puede ser efectivo para empleados en áreas con alta rotación antes de los 90 días.

Es también esencial implementar un sistema de monitoreo continuo de las variables relevantes para la retención, utilizando modelos predictivos para anticipar posibles deserciones y tomar medidas preventivas. Mejorar el ambiente laboral y la estructura organizacional, revisando las dinámicas internas que contribuyen a la rotación, y ofreciendo programas de capacitación que desarrollen las habilidades de los empleados menos calificados, puede aumentar su lealtad y reducir la rotación.

En resumen, al considerar estos factores, la empresa puede minimizar la rotación de empleados, fomentar relaciones laborales duraderas, optimizar costos y mejorar el ambiente laboral.

No obstante, el estudio presenta ciertas limitaciones que deben considerarse. Una de las principales es el tamaño de la muestra, que no es muy grande. Esto implica que el modelo desarrollado podría mejorar significativamente si se entrenara con una cantidad mayor de datos. Con más datos, no solo se incrementarían la precisión y robustez del modelo, sino que también se podrían realizar pronósticos más precisos.

En este sentido, aunque el modelo actual proporciona pronósticos decentes, es recomendable explorar y proponer otros modelos que podrían ofrecer mejores estimaciones. Modelos avanzados o combinaciones de diferentes técnicas de machine learning podrían ser probados para mejorar la capacidad predictiva y la utilidad práctica del análisis.

En resumen, al considerar estos factores y las posibles mejoras, la empresa puede minimizar la rotación de empleados, fomentar relaciones laborales duraderas, optimizar costos y mejorar el ambiente laboral. A medida que se recojan más datos y se experimenten con nuevos modelos, se espera que la precisión de las predicciones y la efectividad de las estrategias de retención continúen mejorando.

Referencias

- A Gentle Introduction to Random Forests Using R. (2016, septiembre 20).
- Bretones, F. D., González, J. M. (2009). Absentismo y rotación laboral. *Psicología del trabajo*, 91-113.
- Cárdenas-Rojas, D. A., Martínez-Gómez, J. A. (2023). Medición y comparación del rendimiento de cuatro algoritmos de aprendizaje supervisado para formular

modelos predictivos sobre la rotación temprana de personal. *Revista de Investigación en Ingeniería y Ciencias Aplicadas*, 10(2), 23-35.

Carrillo-Cárdenas, J. D., Rodríguez-Gómez, D. A. (2020). Factores que influyen en la rotación de personal: Un análisis en empresas del sector servicios de la ciudad de Bogotá, Colombia. *TecnoLógicas*, 23(44), 167-188.

Centeno-González, M. A. (2017). La rotación de personal en las empresas del sector manufacturero en México: Un estudio empírico (Tesis de maestría). Universidad Nacional Autónoma de México, Ciudad de México, México.

Giraldo-Zuluaga, Juan Fernando. Análisis de la rotación de personal en las empresas del sector retail de la ciudad de Medellín, 2019-2021. Tesis de pregrado. Medellín: Universidad EAFIT, 2022.

Levin, J. M. y Kleiner, B. H. (1992). How to reduce organizational turnover and absenteeism. *Work Study*, 41(6):6-9.

Moirón-Sánchez, R. (2022). Análisis de la rotación de personal en empresas del sector servicios de la provincia de Granada (Trabajo fin de máster). Universidad de Granada, Granada, España.

Ramírez Peña, A. (2016, December 18). Impacto de la rotación de personal en las empresas. *Eltiempo.com*. Retrieved March 31, 2024, from <https://www.eltiempo.com/ec>

Revista SG. (2020, 14 de octubre). People Analytics: Prevención y predicción de la rotación del personal. SlideShare.

Ruiz, M. A. (2023, 14 de noviembre). Exploratory analysis of the Titanic dataset. RPubS.