

Aplicación de Analítica a un Caso Estudio en Recursos Humanos

Autores: Laura Herrera, Carlos Vásquez y Yuddy Giraldo

Problema de negocio

Una empresa con 4000 empleados experimenta una tasa de retiros anual del 15%, lo que genera altos costos de contratación y capacitación. Los proyectos pueden retrasarse, afectando la satisfacción del cliente, y el área de selección debe aumentar para cubrir las vacantes. Los empleados restantes enfrentan una mayor carga laboral y pérdida de conocimiento, lo que prolonga los procesos.

Diseño de la solución.

1. Procesos impactados y descripción

Proceso de contratación y capacitación: El proceso de contratación incluye la identificación y selección de candidatos mediante la revisión de aplicaciones y entrevistas, seguido de la toma de decisiones sobre la contratación. Posteriormente, se lleva a cabo la orientación y capacitación del nuevo personal.

Retención de empleados: La retención de empleados busca conservar el talento más valioso mediante estrategias que reduzcan la rotación, evitando la pérdida de conocimiento y los costos asociados a la contratación y capacitación de nuevos empleados, y manteniendo el compromiso de los trabajadores con la empresa.

Tabla 1. Problemas asociados.

Problema de negocio	Problemas analíticos
<ul style="list-style-type: none">• Alta tasa de retiros de empleados. Impacto en proyectos y satisfacción del cliente.• Aumento del trabajo en el área de selección.• Sobrecarga de trabajo para los empleados restantes.• Pérdida de conocimiento y experiencia.	<ul style="list-style-type: none">• Identificación de las variables influyentes en la alta tasa de retiros.• Evaluación de la efectividad de las estrategias de retención.• Predicción de los retiros del personal.

2. Variables

Para predecir	Para entrenamiento
$y_{pred} = \text{análisis de cada empleado } 2017(t + 1)$	$y_{real} = \text{desempeño de cada empleado } 2016(t + 1)$
$X = \text{información empleados } 2016(t)$	$X = \text{información empleados } 2015(t)$

Tabla 2. Variables.

3. Diagrama del diseño de la solución

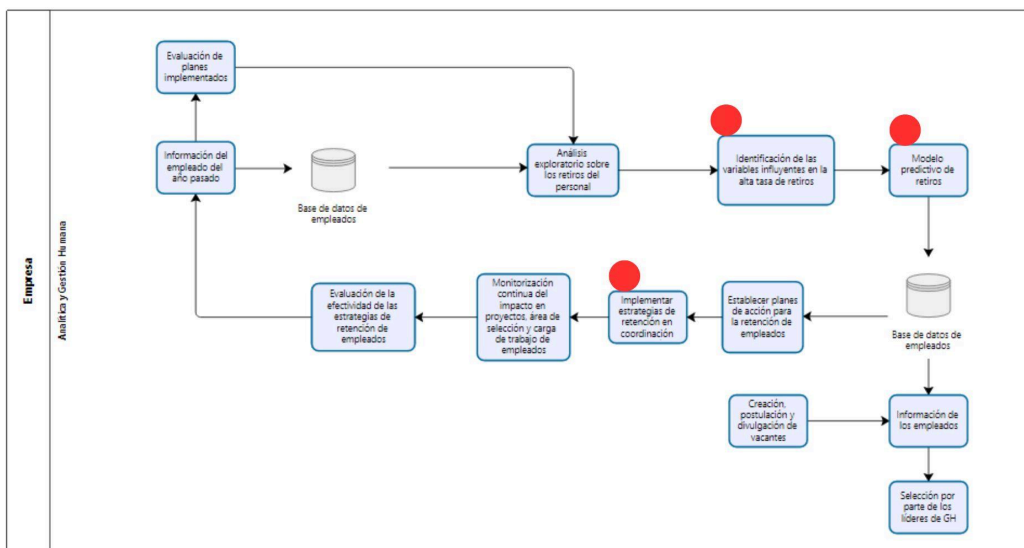


Figura 1. Diseño de la solución.

4. Zoom a la solución analítica

Se desarrolla un modelo de clasificación para predecir la renuncia de empleados en el próximo año, cuyos resultados se utilizarán junto con recursos humanos para tomar medidas preventivas. Se identifican las variables clave relacionadas con las renuncias para implementar acciones preventivas y reducir la tasa de rotación. El modelo se actualiza anualmente para mantener su precisión y relevancia, esto se puede observar en la figura 1, los puntos rojos es donde el campo de la analítica actúa.

5. Selección de algoritmos

Para abordar la predicción del retiro de empleados, se emplean algoritmos de clasificación con interpretabilidad clave. Se seleccionan la Regresión Logística para evaluar la probabilidad de retiro, los Árboles de Decisión para identificar factores relevantes, Random Forest para mejorar la precisión mediante combinación de árboles, y Gradient Boosting para ajustarse a errores previos y mejorar gradualmente la precisión de las predicciones.

Selección de variables

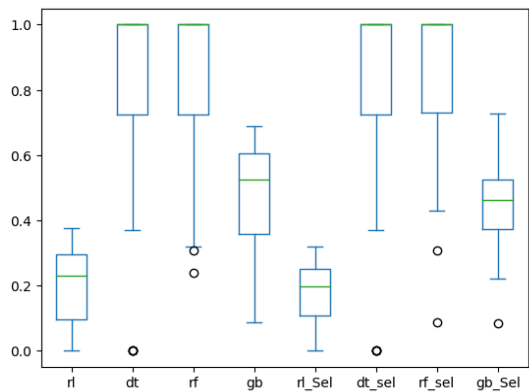
Tabla 3. Variables importantes.

Variable	Peso
MonthlyIncome	0.148964
Age	0.124493
YearsAtCompany	0.109916
TotalWorkingYears	0.096863
DistanceFromHome	0.079593
PercentSalaryHike	0.076699
YearsSinceLastPromotion	0.070497
EnvironmentSatisfaction	0.063077
YearsWithCurrManager	0.060775
NumCompaniesWorked	0.056513

JobSatisfaction	0.055827
MaritalStatus_Single	0.034241
BusinessTravel_Travel_Frequently	0.022540

Para mejorar la retención de empleados, se pueden implementar diversas estrategias basadas en algunas de las variables más importantes del modelo. Para el "MonthlyIncome", se sugiere realizar revisiones periódicas de sueldos para asegurar que los empleados están siendo compensados de manera justa y competitiva en el mercado laboral, y ofrecer incentivos basados en el desempeño. Respecto a la "Age", programas de desarrollo profesional y oportunidades de mentoría diseñados para empleados de diferentes grupos de edad para fomentar el crecimiento continuo. Para "YearsAtCompany", se pueden establecer programas de reconocimiento y oportunidades de ascenso. Para "TotalWorkingYears", brindar opciones flexibles de trabajo y horarios para ayudar a los empleados a equilibrar su vida laboral y personal. "DistanceFromHome" ofrecer incentivos de transporte o programas de reembolso para ayudar a cubrir los costos de viaje para empleados que tienen que desplazarse largas distancias para llegar al trabajo. Para el "PercentSalaryHikel" implementar un sistema transparente de revisión de salarios que permita a los empleados entender cómo se determinan sus aumentos salariales. Los "YearsSinceLastPromotion", hacer programas de desarrollo profesional para mejorar el compromiso y crecimiento profesional de los empleados.

Evaluación y análisis del modelo



rl	0.212993
dt	0.790158
rf	0.842176
gb	0.467928
rl_sel	0.190291
dt_sel	0.790158
rf_sel	0.848142
gb_sel	0.445383

Media de F1-score para cada uno de los modelos

Figura 2. Gráfica de los modelos.

Figura 3. Porcentaje de explicación de los modelos.

Para afrontar el desafío de predicción de retiros de empleados en un contexto de desbalance de clases, se opta por modelos como Random Forest y Decision Tree, que ofrecen un equilibrio entre precisión y exhaustividad, evidenciado por altos puntajes de F1 score. Se elige un modelo con variables seleccionadas para mantener un rendimiento comparable, reduciendo así la complejidad y el costo computacional sin comprometer la capacidad predictiva. Este enfoque eficiente nos permite identificar las variables más influyentes en las tasas de retiro y pronosticar con precisión las salidas de los empleados.

- **Prueba de sobre ajuste**

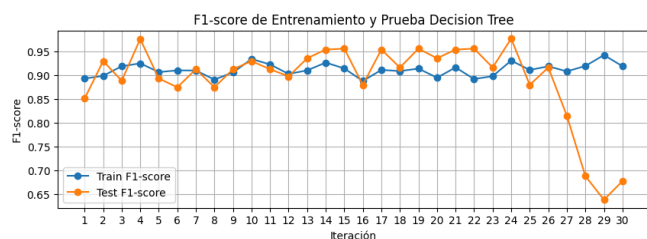


Figura 4. Prueba de sobre ajuste 1.

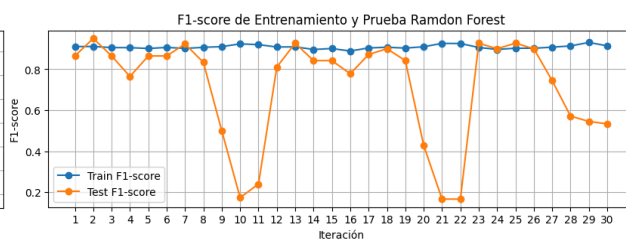


Figura 5. Prueba de sobre ajuste 2.

En el Decision Tree el desempeño de los datos de prueba y de test tienen comportamientos similares, resaltando que varios puntos el desempeño mejora en las pruebas y algunos son significativamente peores. En el Random Forest el desempeño en los datos de entrenamiento es consistente y más alto que en los datos de evaluación, sin embargo el desempeño de los datos de prueba tiene picos donde es muy bajo. Esto sugiere que hay un grado de sobreajuste en el modelo. En conclusión se elige el modelo Decision Tree por que tiene menor sobre ajuste y mayor desempeño que el modelo Ramdon Forest

- **Medidas de desempeño modelo Decision Tree**

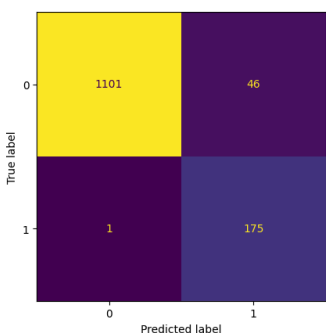


Figura 6. Medida de desempeño DcT.

Accuracy test: 0.9644746787603931
 Precision: 0.9990925589836661
 Recuperacion (Sensibilidad): 0.959895379250218
 F1-score: 0.8816120906801007
 Especificidad: 0.9943181818181818

Figura 7. Resultados de la medida de desempeño.

Se destaca una alta precisión y exhaustividad en el análisis, donde se prioriza minimizar los falsos negativos, este error podría resultar en costos significativos para la empresa, ya que no se tomarían medidas preventivas para retener a ese empleado antes de que decida dejar la empresa. Esto conduciría a la necesidad de contratar y capacitar a un nuevo empleado para reemplazar al que se fue, lo que implica altos

costos adicionales. Por otro lado, aunque predecir incorrectamente que un empleado se va cuando en realidad no lo hace (falso positivo) también puede generar costos, estos podrían ser menores en comparación con los costos asociados con la rotación real del personal.

El modelo presenta un rendimiento sólido en la clasificación con una precisión de prueba del 96.45%. La alta precisión (99.91%) indica una gran proporción de verdaderos positivos entre las predicciones de renunciaciones. Además, la alta sensibilidad (95.99%) sugiere una efectiva identificación de casos positivos reales. El puntaje F1 (88.16%) refleja un equilibrio adecuado entre precisión y sensibilidad, especialmente útil en clasificaciones binarias desequilibradas. La especificidad (99.43%) muestra una capacidad para evitar falsas alarmas en la predicción de casos negativos.

6. Despliegue del modelo

Se obtuvo una tasa de retiros del 17.36% equivalente a 766 empleados para el año actual, lo cual está aproximadamente un 5% por encima de lo deseado por la empresa, por este motivo es importante implementar estrategias de retención para poder cumplir con el objetivo de retiros menor al 12%.

Tabla 4. Pesos de las variables de la predicción.

EmployeeID	2	peso	Aporte al pronóstico
PercentSalaryHike	2,129306	0,076699	0,16
MaritalStatus_Single	1,458650	0,034241	0,05
BusinessTravel_Travel_Frequently	2,075297	0,022540	0,05

Basándonos en los datos de un trabajador y en el peso de cada variable en el pronóstico de renuncia del empleado, podemos identificar las siguientes estrategias para aumentar la retención de este empleado:

Incremento Salarial: Implementar revisiones salariales periódicas y transparentes basadas en el desempeño y la inflación del mercado laboral. Ofrecer bonificaciones por mérito y aumentos salariales significativos para reconocer y recompensar el buen desempeño.

Estado Civil: Proporcionar beneficios adicionales que se adapten a las necesidades de los empleados solteros, como proporcionar incentivos adicionales, programas de desarrollo profesional adaptados metas individuales, bonificaciones por logros alcanzados, con el fin de fortalecer su compromiso y conexión con la empresa

Viajes Frecuentes: Compensar adecuadamente a los empleados que viajan con frecuencia con incentivos adicionales, como días libres compensatorios, bonificaciones por viajes y opciones flexibles de trabajo que les permitan equilibrar mejor su vida laboral y personal.