

Analítica III - Caso de Estudio Recursos Humanos

Cesar Iván Ávila Díaz. cesar.avila@udea.edu.co

Verónica Andrea Morales González. veronica.moralesg@udea.edu.co

Repositorio GitHub: https://github.com/veronica1908/RH_Entregal

Resumen:

El análisis aborda la alta rotación de empleados en una empresa. Se propone una solución integral que incluye recopilación y análisis de datos, desarrollo de modelos predictivos, implementación de estrategias de retención personalizadas y seguimiento continuo. Se identificaron factores clave que influyen en los retiros, como edad y salario, y se diseñaron programas específicos para retener empleados. La solución busca reducir la rotación al 12% y promover un entorno laboral estable y productivo mediante un enfoque de mejora continua.

Palabras Clave: Análisis de datos, estrategias de retención, factores de riesgo, mejora continua, modelado predictivo, retención de personal, seguimiento y evaluación.

1. Descripción del caso problema:

Una empresa que tiene alrededor de 4000 empleados tiene una tasa de retiros de alrededor del 15% anual. Esto quiere decir que el 15% de los empleados en un año se retiran de la compañía por diferentes motivos. Los directivos de la empresa consideran que este nivel de retiros es muy alto, ya que trae altos costos para la empresa como: el costo de contratación que incluye todos los procesos desde el reclutamiento, la selección, la contratación y adicionalmente existe un costo relacionado con la capacitación, la curva de aprendizaje y las contrataciones no efectivas que son personas que se retiran de la empresa en el periodo de prueba. Adicionalmente, el retiro de las personas conlleva otras implicaciones negativas para la empresa que se presentan a continuación:

- Los proyectos en los que los empleados trabajan se atrasan y se pueden ver comprometidas las fechas establecidas y esto puede llegar a afectar satisfacción de clientes y usuarios.
- El trabajo del área de selección se aumenta y por lo tanto implica tener un área de mayor tamaño para el reclutamiento de las personas que se van.
- Los empleados que permanecen tienen que trabajar más para cubrir las labores de las personas que se retiran y para la capacitación de las nuevas que lleguen.

- El conocimiento que tenían las personas y su experiencia se pierde y, debido a esto, el tiempo para ejecutar procesos aumenta.

Los costos que implica la rotación varían mucho dependiendo de la persona y el cargo que tenía, sin embargo, se podría llegar a una estimación basada en criterios de expertos, consultar el siguiente artículo: <https://www.portafolio.co/economia/empleo/costos-de-la-alta-rotacion-de-personal-en-las-empresas-502333>.

Con base en la información anterior, se solicita al equipo de analítica proponer estrategias que permitan tomar acciones para reducir el porcentaje de retiros utilizando la información disponible que se tiene de los empleados. Asuma que la fecha actual es 1 de enero de 2017, y el plan y estrategia que proponga, son para reducir los retiros de los empleados durante el año 2017, para el cuál se espera que el porcentaje de retiros no super el 12% de los empleados.

2. Diseño de la solución para el caso de estudio

Para abordar el problema de retención de empleados y reducir el porcentaje de retiros al 12% durante el año 2017, se propone la siguiente solución:

2.1. Análisis de retención de empleados:

Recopilación de datos: Recolectar datos históricos de los empleados, incluyendo características

demográficas, histórico laboral, satisfacción laboral, desempeño, entre otros.

Análisis exploratorio de datos (EDA): Realizar un análisis exploratorio de los datos para identificar patrones y relaciones entre las variables.

2.2. Diseño y aplicación de modelos predictivos:

Desarrollo de modelos predictivos: Utilizar algoritmos de aprendizaje automático como Random Forest, Decision Trees, o Logistic Regression para desarrollar modelos predictivos. Estos modelos pueden predecir qué empleados tienen mayor probabilidad de retirarse en el futuro cercano.

Validación de modelos: Validar los modelos utilizando técnicas de validación cruzada y métricas de evaluación como precisión F1-score.

2.3. Implementación de estrategias de retención:

Identificación proactiva de empleados en riesgo: Utilizar los modelos predictivos para identificar proactivamente a los empleados en riesgo de retirarse.

Diseño de programas de retención: Diseñar programas de retención personalizados para los empleados identificados como en riesgo. Esto puede incluir oportunidades de desarrollo profesional, programas de bienestar laboral, incentivos económicos, entre otros.

Seguimiento y evaluación: Realizar un seguimiento continuo de la efectividad de las estrategias de retención implementadas y ajustarlas según sea necesario.

2.4. Evaluación de resultados:

Seguimiento del porcentaje de retiros: Monitorear el porcentaje de retiros de empleados a lo largo del año 2017 y compararlo con el objetivo del 12%.

2.5. Mejora continua y refinamiento del modelo:

Actualización de datos: Incorporar datos actualizados al modelo, de los años siguientes para

mejorar su precisión y capacidad predictiva. Esto puede incluir información sobre nuevos empleados contratados, cambios en las políticas de la empresa, eventos relevantes, entre otros.

Reentrenamiento del modelo: Reentrenar el modelo utilizando los datos actualizados de los años que van transcurriendo para capturar nuevas tendencias y patrones en los datos.

Validación del modelo mejorado: Validar el modelo actualizado utilizando técnicas de validación cruzada y métricas de evaluación para asegurar su precisión y fiabilidad.

Implementación de estrategias ajustadas: Ajustar las estrategias de retención en función de los nuevos *insights* obtenidos del modelo mejorado. Esto puede implicar la identificación de nuevos grupos de empleados en riesgo y la adaptación de las tácticas de retención para abordar sus necesidades específicas.

Seguimiento continuo: Mantener un seguimiento continuo del desempeño del modelo y de las estrategias de retención implementadas, y realizar ajustes adicionales según sea necesario.

Se muestra gráficamente la propuesta de solución:

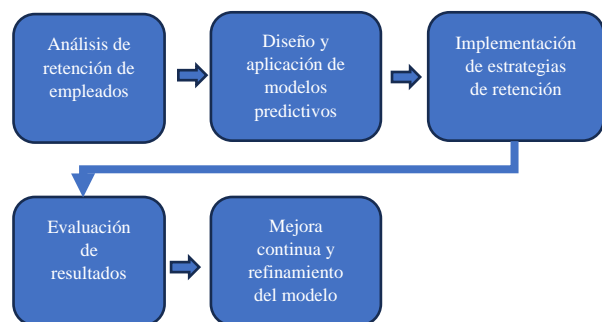


Ilustración 1. Solución teórica propuesta. Elaboración propia.

3. Desarrollo del caso de estudio

Se deben considerar la definición del objetivo y la estrategia general para abordar el problema:

El objetivo es desarrollar un modelo predictivo que identifique a tiempo a los empleados en riesgo de abandonar la empresa. Esto permitirá tomar acciones preventivas para retener a la mayoría de los empleados en riesgo. La solución propuesta

debe incluir un sistema de alerta temprana y recomendaciones personalizadas para retener a los empleados.

3.1. Análisis de retención de empleados

3.1.1. Recopilación de datos: Se recolectan de diferentes fuentes las siguientes bases de datos:

- **Data.dictionay.xlsx:** Descripción de los campos encontrados en las bases de datos.
- **employee_survey_data.csv:** Encuesta realizada a los empleados sobre satisfacción laboral a final de cada año, se tienen el histórico de dos encuestas, la realizada el 31-12-2015 y la que se realizó el 31-12-2016.
- **general_data.csv:** Información general de los empleados, dado que la información puede cambiar con el tiempo se extrajo la información con el último corte 31-12-2016 y un histórico con la información con corte 31-12-2015
- **manager_survey_data.csv:** Encuesta de desempeño de los empleados realizada por parte de los jefes, se tienen el histórico de dos encuestas, la realizada el 31-12-2015 y la que se realizó el 31-12-2016.
- **retirement_info.csv:** Información de retiro de los empleados que dejaron la empresa, se tiene la información de retiro de los empleados de 2015 y 2016, sin embargo, la información de 2015 no se pudo obtener completa por problemas en el almacenamiento de la información, por lo tanto, no están todos los empleados que se retiraron ese año.

3.1.2. Análisis exploratorio de datos (EDA):

Explorar los datos para comprender las características de los empleados que abandonan la empresa. Esto podría incluir análisis de estadísticas descriptivas, visualización de datos y la identificación de patrones preliminares.

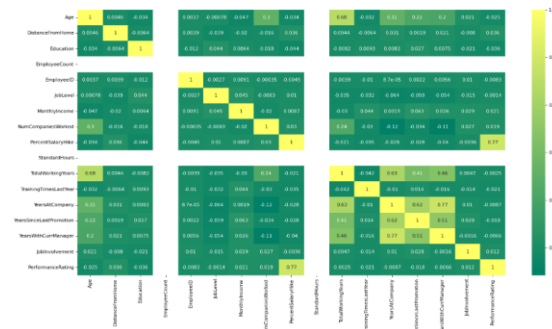


Ilustración 2. Matriz de correlación. Elaboración propia.

Se observa una correlación positiva significativa (igual o superior a 0,5) entre las variables:

- Correlación de 0.68 para edad y total de años trabajados.
- Correlación de 0.77 entre porcentaje de incremento del salario y clasificación de rendimiento.
- Correlación de 0.63 entre años trabajados y años en la compañía.
- Correlación de 0.62 entre años en la compañía y años desde la última promoción.
- Correlación de 0.77 entre años en la compañía y años con el gerente actual.
- Correlación de 0.51 entre años desde la última promoción y años con el gerente actual.

Además, se encontró la siguiente información relevante para esas variables:

- Los colaboradores que se retiran, en mayor proporción, tienen 29 y 31 años, mientras que los que permanecen, tienen entre 34 y 36 años.
- Hay mayor cantidad de retiros para los colaboradores que llevan en total un año de labores.
- Se tiene que hay mayores retiros para porcentajes de incremento de salario bajos.
- Se observa mayor cantidad de retiros para la clasificación de rendimiento igual a tres y menor para la clasificación cuatro.
- Las personas que llevaban un año en la compañía son las que mayor cantidad de retiros han tenido. Existe mayor cantidad de retiros en colaboradores que han sido

recientemente promovidos (menos de un año), lo que podría implicar una variable relacionada al motivo de retiro.

- Los colaboradores que recientemente han cambiado de gerente (en el último año), son los que mayor cantidad de retiros presentan, seguido de los que llevan 2 años.
- Las cantidades de colaboradores que permanecen en la compañía no son proporcionales a los retiros, por lo que esta variable también puede estar influyendo en el principal motivo de retiro y aportar a la solución.

3.2. Diseño y aplicación de modelos predictivos:

3.2.1. Desarrollo de modelos predictivos:

Se consideran los tres modelos planteados inicialmente; Random Forest, Decision Trees y Logistic Regression, se realiza una evaluación inicial de los modelos y se encuentra que tanto el decisión tree como Random Forest, tienen un rendimiento similar, siendo este de 0.999259, aunque es muy bueno, se realizará el afinamiento y validación con variables seleccionadas.

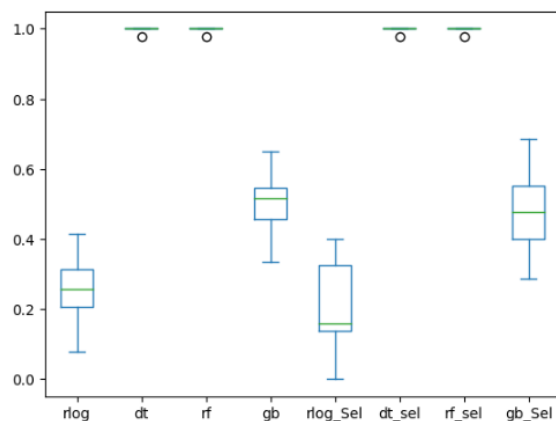


Ilustración 3. Selección de modelos. Elaboración propia.

```
rlog    0.255266
dt      0.999259
rf      0.999259
gb      0.496698
rlog_sel 0.200002
dt_sel  0.999259
rf_sel  0.999259
gb_sel  0.480975
dtype: float64
```

Ilustración 4. Rendimiento inicial de los modelos. Elaboración propia.

3.2.2. Validación de modelos: Se realiza el entrenamiento de los dos modelos seleccionados y se evalúa su rendimiento, obteniéndose mejores resultados con el modelo Random Forest, cuyo rendimiento final es de 0.99539 y puede verse representado en la matriz de confusión:

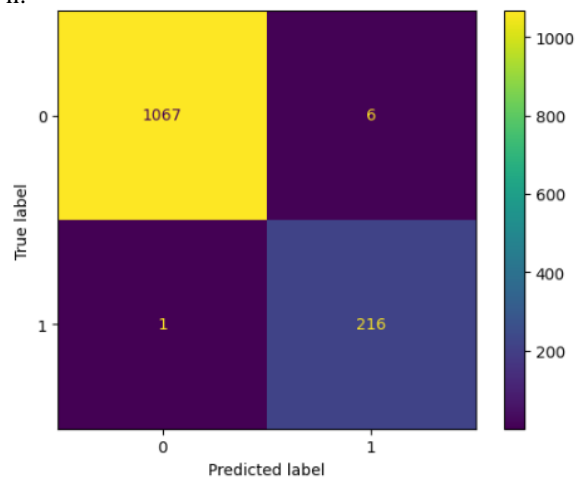


Ilustración 5. Matriz de confusión RF. Elaboración propia.

Esta matriz muestra lo siguiente:

Verdaderos positivos (TP): 1067 muestras fueron clasificadas correctamente como positivas.

Falsos negativos (FN): Aquí, hay 6 falsos negativos, lo que significa que el modelo clasificó erróneamente una muestra que debería haber sido positiva.

Falsos positivos (FP): El modelo clasificó erróneamente 1 muestra que deberían haber sido negativa.

Verdaderos negativos (TN): Aquí, 216 muestras fueron clasificadas correctamente como negativas (Verdaderos No Retiros).

3.3. Implementación de estrategias de retención:

3.3.1. Identificación proactiva de empleados en riesgo:

De acuerdo con la importancia de las variables o factores, se encontró que el de mayor importancia es la Edad (0,085668), luego está la variable Ingreso Mensual (0,072625), porcentaje de aumento salarial (0,061419), Viajes de negocios (0,053222), Nivel de trabajo (0,047964) y por último está el Nivel de opción sobre acciones (0,044076).



Ilustración 6. Importancia de las variables. Elaboración propia.

De acuerdo con esto y retomando el análisis exploratorio, los empleados en riesgo de renunciar a la compañía son:

Los que tienen edades entre 29 y 31 años, un ingreso mensual en promedio de 61564, porcentaje de aumento salarial inferior a 16% principalmente, raramente tienen viajes de negocios, su nivel de trabajo es inferior a 3 y su nivel de opción sobre acciones es inferior a 2.

3.3.2. Diseño de programas de retención:

- Desarrollo profesional para la franja de edad de 29 a 31 años: implementar programas de desarrollo profesional y oportunidades de crecimiento específicamente diseñados para empleados en este rango de edad.
- Revisión y ajuste de compensación: revisar y ajustar los paquetes de compensación para garantizar la competitividad salarial y ofrecer bonificaciones basadas en el desempeño.
- Aumentos Salariales por Desempeño: Proporcionar aumentos salariales significativos basados en el desempeño excepcional y el compromiso a largo plazo.
- Oportunidades de desarrollo a través de viajes de negocios: ofrecer oportunidades de desarrollo y exposición a través de viajes de negocios, rotaciones y asignaciones en diferentes ubicaciones.

- Programas de capacitación para mejorar el nivel de trabajo: implementar programas de capacitación para mejorar las habilidades y competencias de los empleados en niveles de trabajo inferiores a 3.
- Promoción de la participación en planes de acciones de la empresa: promover la participación en el plan de acciones de la empresa y ofrecer incentivos adicionales para fomentar la tenencia de acciones y el compromiso a largo plazo.

3.3.3. Seguimiento y evaluación: Para realizar un seguimiento y evaluación efectivos de las estrategias de retención, se propone los siguientes pasos: establecer indicadores clave de desempeño (kpis), definir metodologías de medición, establecer frecuencia de evaluación, recopilar datos y retroalimentación, analizar resultados e identificar tendencias, tomar acciones correctivas y de mejora, comunicar resultados y ajustes, iterar y mejorar continuamente.

3.4. Evaluación de resultados:

3.4.1. Seguimiento del porcentaje de retiros: Para realizar el seguimiento del porcentaje de retiros de empleados durante el año 2017, es fundamental establecer un objetivo claro de retención, en este caso, fijado en un 12%. Se debe implementar un sistema de monitoreo que permita recopilar y actualizar regularmente los datos sobre los retiros de empleados. Este sistema permitirá comparar periódicamente el porcentaje de retiros con el objetivo establecido, identificar cualquier desviación significativa y analizar las posibles causas. En caso de desviaciones, se deben tomar acciones correctivas para abordar las causas subyacentes y mejorar la retención de empleados. Es crucial registrar y documentar los resultados del seguimiento, comunicarlos a las partes interesadas pertinentes y, sobre esa base, iterar y ajustar continuamente las estrategias de retención para optimizar el rendimiento y acercarse al objetivo del 12% de retiros.

3.5. Mejora continua y refinamiento del modelo:

Para mejorar la precisión y capacidad predictiva del modelo de retención de empleados, se llevará a cabo una serie de pasos iterativos. En primer lugar, se actualizarán regularmente los datos del modelo, incorporando información actualizada de los años siguientes, como nuevos empleados contratados,

cambios en las políticas de la empresa y eventos relevantes. Luego, se reentrenará el modelo utilizando los datos actualizados, lo que permitirá capturar nuevas tendencias y patrones en los datos. Posteriormente, se procederá a validar el modelo mejorado mediante técnicas de validación cruzada y métricas de evaluación para asegurar su precisión y fiabilidad. Una vez validado, se ajustarán las estrategias de retención en función de los nuevos *insights* obtenidos del modelo actualizado, identificando y adaptando las tácticas para abordar las necesidades específicas de los diferentes grupos de empleados en riesgo. Finalmente, se mantendrá un seguimiento continuo del desempeño del modelo y de las estrategias implementadas, realizando ajustes adicionales según sea necesario para garantizar su efectividad a lo largo del tiempo.

4. Conclusiones

El análisis realizado proporciona una visión integral sobre los factores que influyen en la rotación de empleados en la empresa. A través de técnicas de análisis de datos y modelado predictivo, se identificaron patrones y tendencias significativas que permitieron desarrollar estrategias efectivas de retención de personal. Estas estrategias se basan en la detección proactiva de empleados en riesgo de renunciar, considerando criterios como la edad, el salario, el nivel de trabajo, los viajes de negocios y el nivel de opción sobre acciones. Además, se estableció un proceso de seguimiento y evaluación continuo para monitorear la efectividad de las estrategias implementadas y ajustarlas según sea necesario. En resumen, este trabajo proporciona una base sólida para mejorar la retención de empleados y reducir los costos asociados con la rotación de personal, lo que contribuirá a un entorno laboral más estable y productivo en la empresa.