

¡Bienvenidos al reto!

Primera Parte
Análisis de Datos

01 Introducción

¡Bienvenidos al desafío que combina innovación, estrategia y talento! En TechNova Solutions, reconocemos que el éxito no solo depende de nuestras avanzadas soluciones tecnológicas, sino también del equipo humano que impulsa nuestra visión. Hoy, ustedes tienen la oportunidad de transformar un problema en una ventaja competitiva. Analicen, propongan y diseñen soluciones basadas en datos para abordar uno de los retos más importantes de las organizaciones modernas: la retención de talento. Este es su momento de brillar como consultores, aportar ideas disruptivas y marcar la diferencia en una empresa líder. ¡El futuro del talento en TechNova está en sus manos!

TechNova Solutions es una empresa líder en el desarrollo de soluciones tecnológicas avanzadas, con más de 60 años de experiencia en el mercado. Se dedica a la implementación de herramientas de inteligencia artificial, automatización y análisis de datos para transformar digitalmente empresas en sectores como la salud, energía y comercio minorista. Actualmente está posicionada en varios países de la región y tiene perspectivas de crecimiento muy importantes.

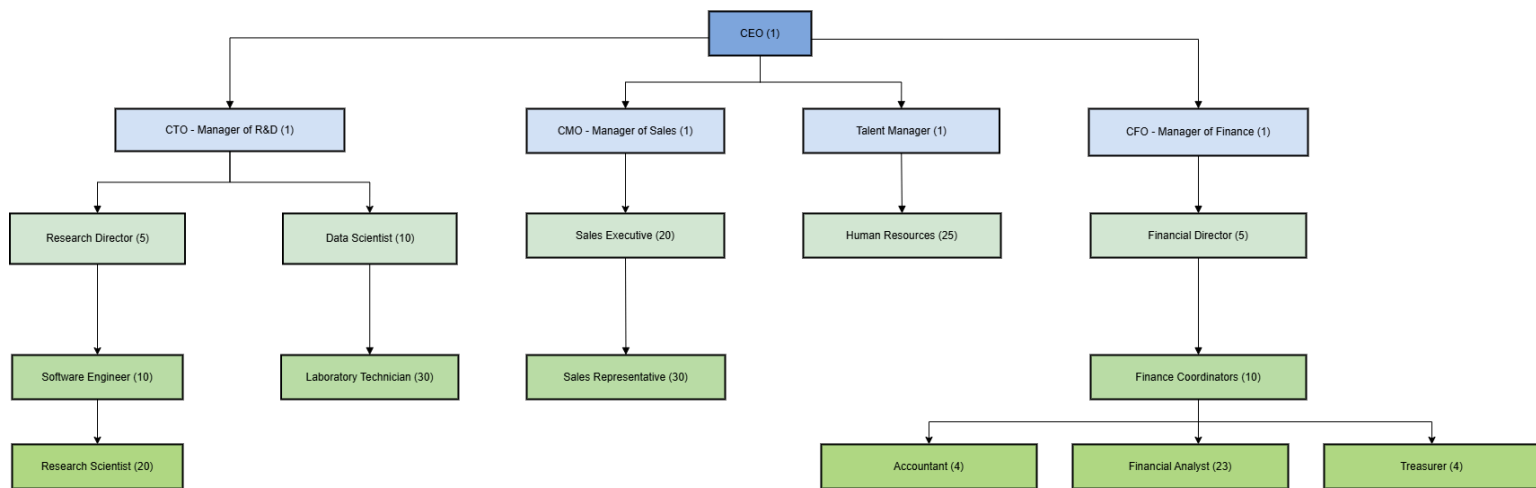
Misión: Impulsar un futuro más eficiente y conectado mediante soluciones tecnológicas disruptivas.

Visión: Ser reconocidos como el principal socio tecnológico para empresas globales que buscan liderar en sus industrias a través de la innovación.

Valores Corporativos:

- **Innovación constante:** Enfocados en la creación de soluciones que estén un paso adelante.
- **Colaboración interdisciplinaria:** Equipos diversos trabajan juntos para resolver desafíos complejos.
- **Orientación al cliente:** Priorizan las necesidades y objetivos de sus socios comerciales.
- **Compromiso con la sostenibilidad:** Buscan soluciones tecnológicas que beneficien a las personas y al planeta.

Actualmente, TechNova Solutions cuenta con un equipo de **200 empleados y un CEO**, distribuidos estratégicamente para mantener la excelencia en el servicio y el liderazgo en el mercado. A continuación se detalla el organigrama de la compañía:



[Organigrama TechNova Solutions](#)

02

Problemática de la Empresa

A pesar de su éxito, TechNova enfrenta un desafío importante: **la alta rotación de empleados**. Esto no solo afecta el clima laboral, sino que tiene implicaciones económicas significativas, incluyendo:

- Costos de reclutamiento y capacitación de nuevos empleados.
- Reducción en la productividad debido a la pérdida de conocimiento empresarial.
- Impacto negativo en la satisfacción y motivación del equipo restante.

Con base en esta problemática, ustedes como consultores tendrán el objetivo de analizar y comprender las causas de la rotación del personal, cuantificar su impacto y proponer soluciones basadas en datos, y en su conocimiento en la gestión del talento humano.

Para lograr esto, la empresa les proporcionará dos bases de datos, las cuales son las siguientes:

1. Data_empleados_historico. Base de datos con más de mil empleados que han estado en la compañía.
2. Data_empleados_actuales_2024: Base de datos con los 200 empleados actuales.

Las bases de datos tienen los siguientes datos:

- **edad:** Edad del empleado (en años).
- **abandono:** Variable dummy (Yes o No) que indica si el empleado abandonó la empresa. La base de datos histórica muestra algunas personas con “Yes” y otras con “No”, es importante recalcar que de esta base de datos ya todos los empleados abandonaron la compañía, el “Yes” se refiere a que el empleado renunció voluntariamente y el “No” se refiere a que se fue por otros motivos diferentes al abandono voluntario.
- **viajes:** Frecuencia de viajes laborales del empleado.
- **departamento:** Departamento al que pertenece el empleado dentro de la empresa.
- **distancia_casa:** Distancia, en kilómetros, que el empleado recorre desde su casa hasta la empresa.
- **educación:** Nivel educativo alcanzado por el empleado.

- **carrera:** Carrera profesional cursada por el empleado.
- **empleados:** Número de empleados a cargo que tiene el empleado (si aplica).
- **id:** Identificación única asignada a cada empleado.
- **satisfacción entorno:** Nivel de satisfacción del empleado con su entorno laboral.
- **sexo:** Género del empleado.
- **implicación:** Nivel de implicación o compromiso del empleado en los procesos de la empresa.
- **nivel_laboral:** Nivel jerárquico dentro de la empresa, donde 1 es bajo y 5 alto.
- **puesto:** Cargo o posición ocupada por el empleado.
- **satisfacción trabajo:** Nivel de satisfacción del empleado respecto a su trabajo actual.
- **estado_civil:** Estado civil del empleado.
- **salario_mes:** Salario mensual percibido por el empleado (en dólares).
- **num_empresas_anteriores:** Número de empresas en las que el empleado trabajó antes de unirse a la actual.
- **mayor_edad:** Indica si el empleado es mayor de edad.
- **horas_extra:** Determina si el empleado ha realizado horas extras **incremento_salario_porc:** Porcentaje de incremento salarial recibido por el empleado en el último año.
- **evaluación:** Puntuación de la última evaluación de desempeño del empleado.
- **Satisfacción compañeros:** Nivel de satisfacción del empleado con sus compañeros de trabajo.
- **horas_quincena:** Número de horas trabajadas por el empleado durante cada quincena.
- **nivel_acciones:** Cantidad de acciones de la compañía que posee el empleado (si aplica)
- **años_experiencia:** Total de años de experiencia laboral del empleado.
- **num_formaciones_ult_año:** Número de capacitaciones recibidas por el empleado en el último año.
- **conciliacion:** Variable que se elimina (no será utilizada en el análisis).
- **años_compania:** Número de años que el empleado lleva trabajando en la compañía actual.
- **años_en_puesto:** Número de años que el empleado lleva en su puesto actual.
- **años_desde_ult_promocion:** Años transcurridos desde el último ascenso del empleado.
- **años_con_manager_actual:** Años que el empleado ha trabajado bajo la supervisión de su jefe actual.

Con base en estos datos, ustedes como consultores en gestión del talento humano, tendrán que desarrollar las siguientes fases del reto.



Fase 1: Análisis Inicial de Datos

Objetivo: Obtener un entendimiento profundo de los datos proporcionados y usarlos para diagnosticar problemas clave relacionados con la rotación de empleados.

Actividades:

1. Exploración y limpieza de datos:

- Revisar las variables disponibles en ambas bases de datos.
- Identificar datos faltantes, inconsistencias y realizar la limpieza necesaria, teniendo en cuenta que hay algunas columnas con muchos datos faltantes, que recomendamos

eliminar, pero hay otras columnas con pocos valores faltantes que quedan a criterio de cada consultor qué hacer con ellas (eliminarlas o imputarlas).

2. **Análisis exploratorio:** Responder las siguientes preguntas clave:

- ¿Cuántos abandonan?
- ¿Qué tipos de personas dejan la empresa?
- ¿Qué nivel de escolaridad tienen los que abandonan?
- ¿La distancia a casa influye?
- ¿El estado civil influye?
- ¿Los que hacen horas extras abandonan más?
- ¿Qué puestos abandonan más?
- ¿Cuánto ganan los que abandonan en comparación de los que no?
- ¿Qué correlaciones hay entre el abandono y las demás variables?
- Buscar en internet datos que permitan cuantificar cuánto le cuesta a una empresa tener que reemplazar a un empleado.
- ¿Cuánto cuesta que los empleados no estén motivados?
- ¿Cuánto se puede ahorrar solucionando el problema?
- Entre otras que consideren importantes para entender el problema.

Estos análisis deben ir acompañados de gráficas para mejorar el entendimiento, además les serán útiles en las partes finales del reto.

04

Fase 2: Modelo Predictivo de Abandono

Objetivo: Desarrollar un modelo predictivo que identifique a los empleados con mayor probabilidad de abandonar la empresa.

Actividades:

1. **Preparación de datos:**

- Dividir la base histórica en **entrenamiento** y **prueba**.
- Seleccionar y transformar variables relevantes, esto incluye: Transformar algunas variables a variables dummy.
- Hacer una transformación a la variable “Abandono” para que sea numérica. (Ojo, no se debe volver una variable dummy, solo hay que cambiar el “Yes” y el “No” por 0 y 1)

2. **Selección del modelo:**

- Probar diferentes algoritmos de machine learning, como:
 - Regresión logística.
 - Árboles de decisión (Random Forest, Decision Tree).
 - Redes neuronales.
 - SVM (Máquinas de Soporte Vectorial).
- Elegir el modelo con mejor desempeño basado en métricas como accuracy, recall, F1-score, entre otros.

3. **Análisis de resultados:**

- Identificar las variables más influyentes en el abandono.

- Interpretar los resultados del modelo para ofrecer recomendaciones accionables.

4. Aplicación del modelo:

- Usar el modelo entrenado para predecir la probabilidad de abandono de los empleados actuales, teniendo en cuenta que a los datos de los empleados actuales le tendrán que hacer transformaciones muy similares a los que le hicieron a la base de datos histórica.

Al finalizar este reto, esperamos recibir un análisis integral que incluya la exploración y limpieza de los datos, hallazgos clave acompañados de gráficas, y respuestas fundamentadas en los planteamientos sobre la rotación de empleados. Además, deberán entregar un modelo predictivo funcional, detallando los métodos utilizados, las métricas de desempeño y las variables más influyentes, junto con su aplicación a los datos actuales. Finalmente, les sugerimos que piensen en recomendaciones basadas en datos que reduzcan la rotación y alineen la gestión del talento humano con los objetivos estratégicos de TechNova Solutions. Para ello, pueden apoyarse en herramientas como Python, Excel, Power BI u otras que consideren pertinentes para desarrollar su análisis y presentar sus propuestas de manera clara y efectiva.

05

Recomendaciones

- ☐ Si leen los datos en python con pandas, haganlo de la siguiente manera:

```
df= pd.read_csv('Data_empleados_historico.csv',index_col='id', na_values='#N/D')
```

- ☐ Cuando vayan a aplicar modelos predictivos, hagan una copia del dataframe original, así podrán hacer manipulaciones en la copia sin afectar el original.
- ☐ La limpieza de datos si se hace en los datos originales.

Contáctanos en caso de cualquier duda o inquietud

Juan Esteban Rendón

jurendon@unal.edu.co

Evelin Florez

evflorezm@unal.edu.co

Andrea Henao

anhenao@unal.edu.co

Nahomy Morales

kemoraless@unal.edu.co

Sara Restrepo

sarestrepoo@unal.edu.co