

Analítica en Recursos Humanos

Descripción de la situación

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:

1. 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.
2. 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.
3. 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 llegan.
4. 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.

Las bases de datos entregadas son las siguiente:

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.

Guía de trabajo

El trabajo que se entrega a final del módulo debe contener:

1. Un proyecto en GitHub con el código generado para la solución del problema, organizado, documentado y dónde se refleje la **contribución de todos los integrantes del grupo**.
2. Un informe con el diseño de la solución propuesta, los análisis realizados, las conclusiones y recomendaciones del trabajo (El informe debe tener máximo 5 páginas letra tamaño 12).

Se propone que los siguientes elementos sean analizados en el informe:

- a. Diseño de solución propuesto
- b. Limpieza y transformación
- c. Análisis exploratorio
- d. Selección de algoritmos y técnicas de modelado
- e. Selección de variables
- f. Comparación y selección de técnicas
- g. Afinamiento de hiperparámetros
- h. Evaluación y análisis del modelo
- i. Despliegue del modelo

Nota 1: Dependiendo de la solución que el equipo proponga, que es libre, puede que se omitan pasos o se incluyan adicionales, en este caso, se debe analizar por qué no se incluyó el paso o por qué se adicionaron.



Facultad de Ingeniería
Departamento de Ingeniería Industrial
Aplicaciones en analítica
Analítica en recursos humanos

Criterios de evaluación serán los siguientes:

1. Creatividad en la propuesta de solución 10%
2. Pertinencia en la propuesta de solución 5%
3. Estructura (funciones, varios archivos, uso de sql), documentación y Justificación del código 30%
4. Análisis de resultados 35%.
5. Redacción adecuada del informe (incluye ortografía) 20%

Presentación de avances:

- Cada grupo presentará avances en **una sola sesión por modulo**.
- Los avances presentados son de temas diferentes en cada sesión de la siguiente manera:

Presentación de Avance 1: Diseño de la solución, preprocesamiento, limpieza y transformación, selección de algoritmos.

Presentación de Avance 2: Diseño de la solución, preprocesamiento y selección de variables.

- Cada día de avances se seleccionará aleatoriamente los grupos que presentarán. Todos los grupos que no hayan presentado hasta la fecha deben estar preparados para exponer.
- La presentación es de máximo 10 minutos.
- Se puede presentar en notebook o archivo de códigos.
- La presentación de avances también es para resolver dudas en el avance del proyecto.

Los criterios de evaluación son:

- Claridad y orden en la presentación (presentar lento y escoger bien los contenidos a mostrar)
- Grado de avance adecuado
- Compresión adecuada del problema
- Propuestas innovadoras en los desarrollos
- No tener errores conceptuales (mal uso o mala interpretación de conceptos o herramientas)