



## **Actividad 2 - Software, Personal y Procesos**

### **Minería de datos**

### **Ingeniería en Desarrollo de Software**

**Tutor: Félix Acosta Hernández**

**Alumno: Jusi Ismael Linares Gutiérrez**

**Fecha: 01/10/2023**

## Índice

Introducción .....	3
Descripción .....	3
Justificación .....	3
Software de Data Mining.....	4
Perfiles y roles .....	5
¿Qué roles o perfiles escogiste para el desarrollo del proyecto? .....	5
¿Por qué son fundamentales?.....	5
Con base en los roles seleccionados, ¿cuánto personal se va a contratar? .....	5
Proceso del proyecto .....	5
Conclusión .....	6
Referencias.....	6

## Introducción

En esta nueva actividad de la materia minería y análisis de datos, Juan se ha convertido en el director de Inteligencia de Negocios y Minería de Datos Nacional, lo que implica un emocionante proyecto por delante. El objetivo principal es diseñar e implementar un enfoque sólido en minería de datos, algo fundamental en el mundo empresarial actual.

Para llevar a cabo este proyecto, primero debemos abordar algunos aspectos básicos. Esto incluye elegir el software adecuado, una herramienta esencial para explorar y extraer información valiosa de grandes conjuntos de datos. También es crucial definir los roles y perfiles necesarios para el equipo, ya que cada miembro desempeñará un papel importante en el éxito del proyecto. En las próximas etapas, profundizaremos en los detalles de la propuesta del proyecto y seleccionaremos las herramientas de minería de datos más adecuadas. Además, analizaremos los perfiles requeridos, asegurándonos de que tengamos un equipo competente y comprometido para llevar a cabo este desafío.

## Descripción

Juan ha asumido la responsabilidad de liderar el área de Inteligencia de Negocios y Minería de Datos Nacional. Su primera tarea es diseñar e implementar un proyecto esencial para esta área. El primer paso consiste en identificar los requisitos clave para el proyecto, que incluyen la selección del software necesario y la definición de los roles y perfiles del personal involucrado.

Para llevar a cabo este proyecto de manera efectiva, es esencial elegir cuidadosamente el software de Data Mining adecuado. Esto implica seleccionar tres programas de Data Mining que se alineen con los objetivos y necesidades del área.

Además de la elección de software, Juan también debe identificar los roles y perfiles necesarios para el desarrollo exitoso del proyecto. Esto implica definir claramente las responsabilidades y habilidades requeridas para cada miembro del equipo, asegurando así que todos estén alineados con los objetivos del proyecto.

## Justificación

La elección de implementar Data Mining en nuestro proyecto se justifica por su capacidad para desentrañar información crucial de grandes conjuntos de datos. Esto es fundamental en el contexto actual de negocios, donde la toma de decisiones basada en datos es esencial para mantenerse competitivo.

Data Mining nos permite identificar patrones, tendencias y relaciones ocultas en nuestros datos, lo que puede conducir a una comprensión más profunda de nuestro mercado y clientes. Esto, a su vez, facilita la formulación de estrategias más efectivas y la optimización de nuestros procesos. Además, contar con perfiles especializados en análisis de datos y minería de datos es esencial para llevar a cabo con éxito este proyecto. Estos expertos desempeñarán un papel fundamental en la extracción de información valiosa de nuestros datos y en la toma de decisiones informadas. La implementación de Data Mining y la contratación de personal calificado son pasos esenciales para mejorar nuestra capacidad de análisis y toma de decisiones.

## Software de Data Mining

Tecnología de Minería de datos	¿Por qué propondrías ese software?	¿Qué procesos de minería de datos puede realizar el software?
<b>Orange</b>	Porque es de código abierto y gratuito por lo que no se necesitara invertir para utilizarlo	repesaje como preprocesador, volver a pesar un conjunto de datos, posprocesamiento de igualdad de probabilidades, desestabilización adversaria, sesgo del conjunto de datos,
<b>KNIME</b>	Es altamente visual, lo que evita tener que saber programar	Manipulación de filas, columnas, etc. Histogramas, Creación de modelos estadísticos y de minería de datos, Scoring, Creación de informes a la medida.
<b>SAS</b>	Por su capacidad para trabajar con múltiples formatos de datos.	Modelos generados con diferentes métodos, Muestran resultados analíticos en diagramas, Se genera automáticamente código de valoración para todas las etapas de la generación del modelo, Opción de implementación en la nube.

### ¿Cuál es el mejor gestor de base de datos para este proyecto?

Microsoft SQL Server es un sistema de gestión de base de datos relacional desarrollado como un producto de software con la función principal de almacenar y recuperar datos según lo solicitado por otras aplicaciones.

## Perfiles y roles

¿Qué roles o perfiles escogiste para el desarrollo del proyecto?

Analista de negocios, analista de datos, gerente de datos y análisis, ingeniero de datos, administrador de BD, estadístico y arquitecto de datos.

¿Por qué son fundamentales?

Estos roles trabajan en conjunto para garantizar que una empresa pueda aprovechar al máximo sus activos de datos. Desde la comprensión de las necesidades comerciales hasta la implementación de soluciones técnicas y análisis profundos.

Con base en los roles seleccionados, ¿cuánto personal se va a contratar?

1 analista de negocios, 1 analista de datos, 1 gerente de datos y análisis, 1 ingeniero de datos, 1 administrador de BD, 1 estadístico y 1 arquitecto de datos.

## Proceso del proyecto

El proyecto será una herramienta de predicción de ventas, se desarrollara un software que en función de datos recolectados se ponga a predecir lo que se va a vender. Esto con el fin de optimizar inventarios e implementar estrategias de ventas.

Pasos para desarrollar el proyecto

1.- Definición de objetivos y alcance

Identificar lo que se desea alcanzar con el proyecto y lo que debe cumplirse para que el proyecto sea considerado un éxito.

2.- Recopilar datos

Recolectar datos históricos de ventas, incluyendo información sobre productos, regiones, temporadas y otros factores relevantes.

3.- Selección de modelos de aprendizaje automático

Evaluar y seleccionar algoritmos de aprendizaje automático adecuados para el problema de predicción de ventas.

4.- Desarrollo del proyecto

Diseñar e implementar la interfaz de usuario del software, que permitirá a los usuarios (nuestros empleados) cargar datos y obtener predicciones para trabajar con esa información.

5.- Pruebas

Realizar pruebas exhaustivas para garantizar la precisión y confiabilidad de las predicciones.

6.- Despliegue y capacitación

Implementar el software en el área de Juan y empezar a capacitar a los usuarios para hacer uso del proyecto

7.- Monitoreo y mantenimiento continuo

Monitoreo de los resultados para confirmar que aun funciona con precisión, además de actualizarlo periódicamente para que no quede obsoleto ante otras tecnologías

8.- Documentación

Documentar todo el proceso de desarrollo del proyecto para futuras referencias.

## Conclusión

La elección de programas de Data Mining adecuados es crucial para que el proyecto tenga éxito, ya que estos programas proporcionan las herramientas necesarias para encontrar patrones y tendencias ocultas en los datos. Esto puede llevar a obtener información valiosa que beneficie a la empresa.

Además, es esencial identificar los roles y perfiles necesarios para el proyecto. Cada uno de estos roles cumple una función importante en la ejecución de las tareas, desde analizar datos hasta tomar decisiones basadas en los análisis realizados.

Estos proyectos en el campo de la Inteligencia de Negocios y Minería de Datos tienen un gran impacto en mejorar la eficiencia de las operaciones de la empresa y en tomar decisiones más informadas. Esto se traduce en un valor real tanto en el trabajo como en la vida cotidiana. En un mundo donde comprender y usar datos es fundamental, proyectos como este son esenciales para aprovechar al máximo su potencial y obtener beneficios significativos.

## Referencias

Bioinformatics Laboratory, & University of Ljubljana. (n.d.). *Orange Data Mining - Blogs*.

Orangedatamining.com. Retrieved October 1, 2023, from

<https://orangedatamining.com/blog/>

Wikipedia contributors. (n.d.). *KNIME*. Wikipedia, The Free Encyclopedia.

<https://es.wikipedia.org/w/index.php?title=KNIME&oldid=152381520>

*Data mining software, model development and deployment, SAS Enterprise Miner*. (2019,

January 15). Sas.com. [https://www.sas.com/es\\_mx/software/enterprise-miner.html](https://www.sas.com/es_mx/software/enterprise-miner.html)

Nextech. (2021, February 3). ¿Qué es SQL Server? y ¿Para qué sirve? *Nextech*.

<https://nextech.pe/que-es-sql-server-y-para-que-sirve/>

*Tipos de roles en Analytics (Business Intelligence, Big Data)*. (2019, June 21). TodoBI - Business

Intelligence, Big Data, ML y AI. <https://todobi.com/tipos-de-roles-en-analytics-business/>

## GitHub

[https://github.com/JusiLinGu/Practicas\\_UMI](https://github.com/JusiLinGu/Practicas_UMI)