

## Actividad | 2 | Software, Personal y

### Procesos

### Minería de Datos

---

## Ingeniería en Desarrollo de Software



academiglobal

TUTOR: Félix Acosta Hernández

ALUMNO: Marco Antonio García Sánchez

FECHA: 15/06/2024

# ÍNDICE

Portada	1
Índice	2
Introducción y descripción	3
Justificación	4
Desarrollo	5-9
Conclusión y referencias	10

# INTRODUCCIÓN

La minería de datos es un proceso fundamental en la era digital actual, que permite extraer información valiosa y patrones ocultos a partir de grandes volúmenes de datos. Su importancia radica en la capacidad de transformar datos crudos en conocimientos útiles que pueden guiar la toma de decisiones en diversos sectores, como negocios, salud, educación y finanzas. En el ámbito empresarial, la minería de datos permite identificar tendencias de mercado, mejorar la satisfacción del cliente y optimizar operaciones internas. En el campo de la salud, facilita la predicción de enfermedades y la personalización de tratamientos. Además, en la educación, puede mejorar la experiencia de aprendizaje mediante el análisis de comportamientos y resultados estudiantiles. En resumen, la minería de datos es esencial para aprovechar al máximo el potencial de los datos, proporcionando insights que impulsan la innovación, la eficiencia y la competitividad en un mundo cada vez más orientado a la información.

## DESCRIPCIÓN

Contextualización: Juan ha sido nombrado como el nuevo director del área de Inteligencia de Negocios y Minería de Datos Nacional. Por tanto, se le solicita un proyecto para diseñar e implementar en el área, como primer punto se pretende identificar los requisitos básicos como son el software necesario y el requisito del personal, posterior a esto en la actividad 3, Juan debe estimar los costos para la implementación de la propuesta del proyecto.

Actividad:

- Presentar una propuesta de proyecto
- Seleccionar 3 programas de Data Mining (considerar los que se mencionaron en el curso)
- Identificar los roles y/o perfiles requeridos para el desarrollo del proyecto.

Redactar una propuesta de software y motivo de selección, así como la propuesta del personal necesario y los requisitos para llevar a cabo el proyecto. Por último, describir ¿Cómo se realizaría la ejecución del proyecto para que sea exitoso al implementarlo en el área en el que Juan está a cargo? Escribir los pasos que se realizarían para el desarrollo del proyecto. Se pueden buscar ejemplos en Internet o en el material del curso.

## JUSTIFICACIÓN

La importancia de la minería de datos se justifica por su capacidad para convertir grandes volúmenes de datos en información procesable y valiosa. Esto, a su vez, puede traducirse en mayores ingresos y una ventaja competitiva más fuerte. En el ámbito de la salud, la minería de datos ayuda a predecir brotes de enfermedades, identificar factores de riesgo y personalizar tratamientos para los pacientes. Esto mejora significativamente la calidad de la atención médica y reduce los costos asociados con tratamientos innecesarios o ineficaces. En el sector educativo, el análisis de datos puede revelar patrones en el comportamiento y rendimiento de los estudiantes, permitiendo a los educadores adaptar métodos de enseñanza para satisfacer mejor las necesidades individuales de los estudiantes y mejorar los resultados educativos. En resumen, la minería de datos es esencial porque permite aprovechar al máximo el potencial de los datos, proporcionando insights que impulsan la innovación, mejoran la eficiencia operativa y fortalecen la competitividad en diversos sectores.

# PROPUESTA DE SOFTWARE

Hablando del tema central del proyecto, es importante aclarar que lo deseable si o sí es una propuesta de mejora, por lo cual, analizaremos una propuesta referente a las tecnologías de minerías de datos. A diferencia de como viene plasmado en la actividad como un recuadro, me gustaría hacerlo más a manera de puntos, esto debido a la investigación que estuve realizando, para explayarme de una manera más concisa pero no muy general, esta es mi propuesta:

## Propuesta de Software para el Área de Inteligencia de Negocios y Minería de Datos Nacional

El nombramiento de Juan como nuevo director del área de Inteligencia de Negocios y Minería de Datos Nacional marca un paso importante para el fortalecimiento del análisis de datos en la organización. La transformación de datos en información útil y estratégica es vital para la toma de decisiones informadas y para mantener una ventaja competitiva. A continuación, veremos los puntos importantes que son muy necesarios para que este proyecto funcione.

### Tecnologías de Minería de Datos Seleccionadas

#### 1. Machine Learning (Aprendizaje Automático):

- **Descripción:** Utiliza algoritmos y modelos matemáticos para permitir que las computadoras mejoren automáticamente a través de la experiencia.

#### 2. Big Data Analytics:

- **Descripción:** Consiste en el análisis de conjuntos de datos extremadamente grandes para descubrir patrones, tendencias y asociaciones, especialmente en relación con el comportamiento humano y las interacciones.

#### 3. Minería de Textos (Text Mining):

- **Descripción:** Proceso de derivar información de texto no estructurado a través del análisis de patrones, extracción de datos y aprendizaje automático.

### Justificación del Software Propuesto

El software propuesto es una solución integral de minería de datos que incluye herramientas avanzadas para el análisis y gestión de datos. Las principales razones para proponer este software incluyen:

# PROCESOS DE MINERÍA DE DATOS QUE REALIZARÁ EL SOFTWARE

Ahora que sabemos las tecnologías de minería de datos que vamos a utilizar para el proyecto, también es de vital importancia el conocer los procesos que tendrán o participarán dentro del área, que son las siguientes:

- **Análisis Predictivo:** Utiliza modelos de machine learning para prever tendencias futuras y comportamientos del mercado.
- **Segmentación de Clientes:** Agrupa a los clientes en segmentos basados en características y comportamientos similares, optimizando campañas de marketing y estrategias de venta.
- **Análisis de Sentimiento:** Evalúa las opiniones y percepciones de los clientes mediante el análisis de textos en redes sociales, encuestas y otros medios.
- **Detección de Anomalías y Fraudes:** Identifica patrones inusuales y potenciales fraudes en transacciones y actividades comerciales.
- **Análisis de Redes Sociales:** Estudia las interacciones y comportamientos en redes sociales para comprender mejor la influencia y el impacto de las estrategias de marketing.

## SOFTWARE DE DATA MINING

En la implementación de un sistema de minería de datos avanzado para la organización es importante contar con gestores que soporten grandes volúmenes de datos, utilizar múltiples gestores de bases de datos, como SQL Server, Oracle y PostgreSQL, puede ofrecer una serie de beneficios significativos.

### Beneficios de Utilizar SQL Server:

1. Integración con el Ecosistema de Microsoft
2. Facilidad de Uso y Gestión

### Beneficios de Utilizar Oracle

1. Rendimiento y Escalabilidad
2. Funcionalidades Avanzadas de Minería de Datos

### Beneficios de Utilizar PostgreSQL

1. Código Abierto y Flexibilidad
2. Soporte para Datos No Estructurados

## MEJOR GESTOR DE BASE DE DATOS PARA EL PROYECTO:

Para soportar el sistema de minería de datos propuesto, recomiendo el uso de PostgreSQL, en lo personal lo utilizo en mi trabajo de una manera muy sencilla, además de que la creación de funciones para modificar información y el volumen de datos que maneja la hace muy singular.

## PERFILES Y ROLES

Para la implementación y operación efectiva del sistema de minería de datos, se requiere un equipo con las siguientes competencias:

- **Científicos de Datos:** Expertos en análisis de datos, estadística y machine learning, encargados de desarrollar modelos predictivos y analizar grandes volúmenes de datos.
- **Ingenieros de Datos:** Profesionales en la construcción y mantenimiento de infraestructuras de datos, asegurando la integración y accesibilidad de los datos.
- **Analistas de Negocios:** Especialistas en traducir los insights derivados de los análisis de datos en estrategias y decisiones de negocio.
- **Desarrolladores de Software:** Responsables de la personalización e integración del software de minería de datos con los sistemas existentes.
- **Administradores de Bases de Datos:** Encargados de la gestión y mantenimiento de PostgreSQL, asegurando el rendimiento y seguridad del sistema de bases de datos.

## IMPORTANCIA DE ESTOS ROLES

Cada uno de estos roles es fundamental para el éxito del sistema de minería de datos debido a su especialización y contribución única al proyecto. Los científicos de datos y analistas de negocios aseguran que los datos se traduzcan en insights útiles para la toma de decisiones. Los ingenieros de datos y desarrolladores de software proporcionan la infraestructura y herramientas necesarias para el manejo y análisis eficiente de los datos. Los administradores de bases de datos garantizan la integridad, seguridad y disponibilidad de los datos. Juntos, forman un equipo unido que puede abordar los diversos desafíos del proyecto y asegurar su éxito a largo plazo.

## PERSONAL A CONTRATAR

- Científicos de Datos: 3
- Ingenieros de Datos: 3
- Analistas de Negocios: 2
- Desarrolladores de Software: 2
- Administradores de Bases de Datos: 2

Este equipo garantiza que todos los aspectos del proyecto, desde la infraestructura de datos hasta la traducción de insights en estrategias de negocio, sean manejados con la máxima eficiencia y efectividad.

# PROCESO DEL PROYECTO

Como todo proyecto, es necesario tener claro y de una manera organizada las actividades a realizar dentro de cada etapa, por lo cual son las siguientes:

## 1. Fase de Planificación y Recolección de Requisitos

- Entrevistas y Encuestas: Realizar entrevistas con usuarios clave y encuestas para recopilar requisitos detallados.
- Evaluación de la Infraestructura Actual: Analizar la infraestructura existente para identificar brechas y necesidades.
- Documentación de Requisitos: Crear un documento formal de requisitos que sirva como guía para las fases posteriores.

## 2. Fase de Selección y Adquisición de Software y Hardware

- Pruebas de Concepto: Realizar pruebas de concepto (PoC) para evaluar el software en un entorno controlado.
- Negociación y Contratación: Negociar contratos y adquirir el software y hardware seleccionados.
- Proceso de Contratación: Anunciar y contratar personal con las competencias requeridas.

## 3. Fase de Diseño de la Arquitectura del Sistema

- Reuniones de Diseño: Realizar talleres de diseño con los equipos técnicos y de negocio.
- Creación de Diagramas: Desarrollar diagramas de arquitectura que describan la estructura del sistema.
- Revisión de Diseño: Validar el diseño con todas las partes interesadas para asegurar alineación.

## 4. Fase de Desarrollo e Integración

- Configuración de Bases de Datos: Instalar y configurar SQL Server, Oracle y PostgreSQL.
- Desarrollo de Modelos: Utilizar herramientas de data mining para desarrollar y entrenar modelos predictivos.
- Integración con Sistemas Existentes: Asegurar que el nuevo sistema se integre sin problemas con las soluciones actuales de la organización.



## **5. Fase de Pruebas y Validación**

- **Planificación de Pruebas:** Crear planes de prueba detallados para todas las partes del sistema.
- **Ejecución de Pruebas:** Realizar pruebas funcionales, de integración, de rendimiento y de seguridad.
- **Ajustes Basados en Pruebas:** Hacer ajustes y optimizaciones según los resultados de las pruebas.

## **6. Fase de Capacitación y Implementación**

- **Capacitación del Personal:** Organizar sesiones de capacitación para todos los usuarios y administradores del sistema.
- **Documentación de Procesos:** Documentar todos los procesos y procedimientos del sistema.
- **Lanzamiento en Producción:** Implementar el sistema en el entorno de producción y realizar pruebas finales.

## **7. Fase de Monitoreo y Mejora Continua**

- **Monitoreo del Sistema:** Implementar herramientas de monitoreo para vigilar el rendimiento y la seguridad del sistema.
- **Recolección de Feedback:** Establecer canales de comunicación para recibir feedback continuo de los usuarios.
- **Ciclo de Mejora Continua:** Analizar el feedback y las métricas de rendimiento para identificar áreas de mejora y realizar actualizaciones periódicas.

Este plan de ejecución proporciona una hoja de ruta detallada para la mejora del área de Inteligencia de Negocios y Minería de Datos Nacional. Siguiendo estas fases y actividades, la organización puede asegurar una implementación exitosa que maximice el valor del análisis de datos y apoye la toma de decisiones estratégicas basadas en datos.

## CONCLUSIÓN

Realizar proyectos ficticios de tecnologías de minería de datos es una práctica sumamente valiosa. Estos proyectos permiten a estudiantes y profesionales desarrollar habilidades cruciales en un entorno controlado y sin los riesgos asociados con datos reales. Además, fomentan la creatividad y la innovación al permitir la experimentación con diferentes métodos y técnicas sin las limitaciones del mundo real. A través de estos ejercicios, los individuos pueden profundizar su comprensión de algoritmos, herramientas y procesos de minería de datos, preparándose mejor para abordar desafíos complejos en el futuro. En última instancia, estos proyectos ficticios sirven como un campo de entrenamiento para perfeccionar competencias y fomentar una mentalidad analítica, esencial en la era de la información. Por lo tanto, invertir tiempo y recursos en ellos no solo enriquece el conocimiento teórico, sino que también potencia la capacidad práctica para aplicar la minería de datos de manera efectiva en situaciones reales, contribuyendo al avance tecnológico y al desarrollo profesional.

## REFERENCIAS

Ortega, C. (2024, 9 marzo). *Técnicas de minería de datos: Qué es e importancia*. QuestionPro.

<https://www.questionpro.com/blog/es/tecnicas-de-mineria-de-datos/>

Ali, A. (2023, 9 junio). Top 10 data mining techniques | Astera. Astera.

<https://www.astera.com/es/type/blog/top-10-data-mining-techniques/>

Daniel. (2023, 30 octubre). *Text Mining o minería de textos: definición, técnicas, casos de uso*. Formación En Ciencia de Datos | DataScientest.com.

<https://datascientest.com/es/text-mining-o-mineria-de-textos-definicion-tecnicas-casos-de-uso>

De Ceupe, B. (2019, 5 septiembre). Ceupe. Ceupe.

<https://www.ceupe.com/blog/proceso-del-data-mining.html>

## CODIGO GITHUB

<https://github.com/Marcogarcia3103/Codigo>