

Actividad 2 – Definición de software, personal y procesos

Minería y análisis de datos

Ingeniería en Desarrollo de Software

Tutor:

Félix Acosta Hernández

Alumno:

Alejandro Abarca Gerónimo

Fecha:

06 de febrero de 2024

Indice

Introducción	3
Descripción	3
Justificación	4
Software de Data Mining	4
Perfiles y roles.....	5
Proceso del proyecto	6
Conclusion	7
Referencias	7

Introducción

En el desarrollo de esta actividad se nos solicita crear un proyecto y realizar una investigación a detalle de diferentes softwares que permitan realizar Data Mining. Analizaremos 3 propuestas de software así como también se consideraran los perfiles y roles del personal que formara parte de equipo de trabajo.

El análisis del proyecto consta de diferentes etapas:

1. **Fase de comprensión del negocio:** encargada de la comprensión del negocio o problema.
2. **Fase de comprensión de datos:** se encarga de la comprensión de los datos.
3. **Fase de preparación de los datos:** se procede a su preparación para adaptarlos a las técnicas de Data Mining.
4. **Fase de modelado:** selección de las técnicas de modelado más apropiadas para el proyecto de Data Mining.
5. **Fase de evaluación:** se realiza la evaluación del modelo, teniendo en cuenta el cumplimiento de los criterios de éxito del problema.
6. **Fase de implementación:** finalizado el modelo y validado, se transforma el conocimiento obtenido en acciones dentro del proceso de negocio.

Descripción

Para la elaboración de la actividad se nos plantea lo siguiente:

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**.

Para llevarlo a cabo realizaremos una investigación de 3 diferentes softwares comparando y eligiendo uno para que en la actividad 3 podamos realizar un presupuesto en cuanto a los costes que este proyecto generaría para la empresa.

Lo que se entregará en la actividad:

1. Presentar una propuesta de proyecto.
2. Seleccionar 3 programas de Data Mining.
3. Identificar los roles y/o perfiles requeridos para el desarrollo del proyecto.

Justificación

El emplear este tipo de análisis y recaudación de información nos ayuda a tener más control y conocimiento sobre las cosas que se están trabajando, por ejemplo la mayoría de los datos comerciales tienen varias dimensiones: varias categorías en las que se desglosan los datos para su presentación, seguimiento o análisis.

El emplear Data Mining ayuda a detectar patrones de conducta en los clientes y predecir su comportamiento. Su objetivo es generar nuevas oportunidades de mercado. Hoy en día, la búsqueda, el análisis y la gestión de datos representan un mercado de grandes oportunidades laborales. El profesional de data mining trabaja las bases de datos para evaluar la información y descartar la que no sea útil o confiable.

Software de Data Mining

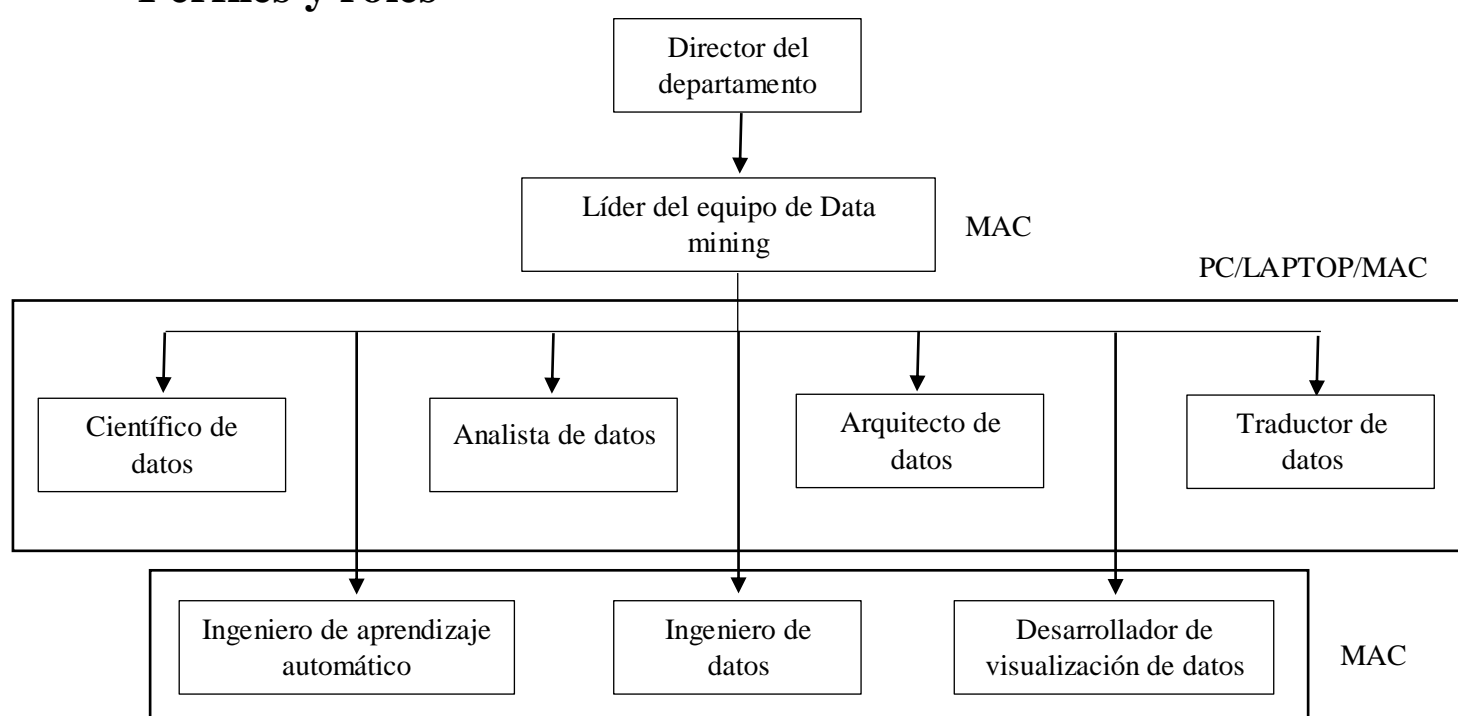
Tecnología de minería de datos	¿Por qué propondrías ese software?	¿Qué procesos de minería de datos puede realizar el software?
SQL Developer de Oracle	Es un software que cuenta con seguridad integrada y protege los entornos en la nube, permite lograr un alto rendimiento en todas las cargas de trabajo, impulsa la sostenibilidad con informática en la nube que funciona con energía	<ul style="list-style-type: none"> ✓ Análisis e inteligencia empresarial. ✓ Big data y data lake. ✓ Servicios de bases de datos. ✓ Nube distribuida. ✓ Multicloud.

	renovable, tiene presencia global ya que tiene una nube distribuida causando así experiencias de éxito entre los principales consumidores.	✓ Aprendizaje automático (ML) e IA.
Microsoft Power BI	En este software es posible convertir los datos en objetos visuales con herramientas avanzadas de análisis de datos, funcionalidades de inteligencia artificial y una herramienta de creación de informes, además permite reunir información desde cualquier origen y agregarlos al centro de datos OneLake.	✓ Crear cuadros de mando para la toma de decisiones. ✓ Importar datos. ✓ Creación de gráficas, y cálculos. ✓ Inteligencia artificial la cual es encargada de realizar las operaciones.
AWS	Es una plataforma innovadora diseñada con informática en la nube, la cual tiene la nube más flexible y con mayor seguridad, además es una de las más grandes comunidades entre clientes y socios, con millones de personas activas a nivel global.	✓ Almacenamiento en la nube. ✓ Redes. ✓ Bases de datos. ✓ Inteligencia artificial. ✓ Servicios móviles. ✓ Seguridad informática.

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

Utilizaría la plataforma de Amazon Web Services.

Perfiles y roles



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

R: Elegí 8 perfiles, de acuerdo a las funciones que realiza cada uno dentro del proyecto estos serían suficientes para lograr el desarrollo, los cuales son líder del equipo, científico de datos, analista de datos, arquitecto de datos, traductor de datos, ingeniero de aprendizaje automático, ingeniero de datos y desarrollador de visualización de datos.

2. ¿Por qué son fundamentales?

R: Por que cada uno cumple una función específica dentro del organigrama y del desarrollo del proyecto, es decir cada uno tiene sus tareas asignadas para así lograr el objetivo y llegar al resultado esperado.

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

R: Son 9 personal en total incluyendo a Juan como director del proyecto.

Proceso del proyecto

El software que se eligió para llevar a cabo el desarrollo del proyecto fue Amazon Web Services ya que esta plataforma es una de las que tienen mayor seguridad en la nube y es apta para trabajar desde cualquier parte del mundo ya que tiene una red distribuida a nivel global que permite realizar trabajos desde la nube, permitiendo administrar bases de datos, almacenamiento, inteligencia artificial, seguridad informática, etc.

El personal que se ha contemplado cumple con los requisitos para la formación y desarrollo del proyecto.

El proceso a realizar para la selección de la herramienta de Data Mining se realizará mediante el estudio de mercados, que ofrece cada una de las plataformas y comparando en relación precio/calidad para así aprovechar al máximo cada una de las funcionalidades de la plataforma a elegir, se realizarán diferentes presupuestos para así comprobar el gasto a la empresa una vez que se haya seleccionado.

Posterior a esto se lanzaran vacantes en diferentes plataformas de empleo para entrevistar a diferentes postulantes y contratar al más capacitado y apto para el puesto, se tienen 8 puestos para los cuales se pide un grado de estudios y conocimiento.

Una vez seleccionado el personal se comenzara con la recaudación de información y el análisis de esta para el desarrollo del proyecto.

Conclusion

Durante el desarrollo de la actividad hemos conocido algunas herramientas dedicadas al Data Mining como lo son SQL Developer Oracle, Microsoft Power BI y Amazon Web Services, en las cuales durante todo el desarrollo de la actividad nos enfocamos, realizando investigaciones acerca de los servicios que cada uno de estos ofrecen. Nos centramos en AWS para el desarrollo y ejemplificación del proyecto, ya que esta fue la herramienta de Data Mining que se eligió para trabajar, solicitando presupuestos para comprar costos en base a las funcionalidades con las que cuentan los 3 softwares mencionados anteriormente.

La finalidad del desarrollo de esta actividad es introducirnos al mundo de los datos y su análisis para obtener datos relevantes, la minería de datos es una rama de la tecnología que mediante la inteligencia artificial nos ayuda a realizar operaciones con mayor rapidez sin tanto esfuerzo, como bien se dice, se busca realizar algo que automatice la mayor cantidad y parte de procesos.

Referencias

1. Basado en la tesis: "Metodología para la definición de requisitos en proyectos de Data Mining (ER-DM)" de José Alberto Gallardo Arancibia.
2. SQL Developer. (s. f.). Oracle México. <https://www.oracle.com/mx/database/sqldeveloper/>
3. Power BI: Visualización de datos / Microsoft Power Platform. (s. f.). <https://www.microsoft.com/es-mx/power-platform/products/power-bi/>
4. ¿Iquest;Cómo me aseguro de no se me vuelve a cobrar cuando utilizo el nivel gratuito de AWS? (s. f.-b). [Vídeo]. Amazon Web Services, Inc. https://aws.amazon.com/es/free/?all-free-tier.sort-by=item.additionalFields.SortRank&all-free-tier.sort-order=asc&awsf.Free%20Tier%20Types=*all&awsf.Free%20Tier%20Categories=*a