Título del Proyecto

Seminario de Analítica y Ciencia de Datos

Resumen Descriptivo del Proyecto

Este proyecto se distingue por su enfoque estratégico y preciso en la aplicación de técnicas avanzadas de análisis de datos, integrando herramientas de aprendizaje automático y análisis predictivo, que busca no solo identificar productos con alto potencial de ventas, si no tambien comprender las dinámicas del nicho de mercado según su categoría. A través de un análisis exhaustivo de grandes conjuntos de datos de productos, se pretende descubrir patrones y tendencias significativas que proporcionen una visión precisa del mercado para un producto. Este enfoque se fundamenta en la necesidad de abordar los desafíos actuales en las ventas en línea, donde una proporción considerable de vendedores no logra el éxito debido a una investigación de mercado insuficiente, ya que, en plataformas como Amazon, entre un 30% y un 40% de las cuentas no logran éxito, esto según el blog “https://medium.com/@libradafigg42/que-porcentaje-de-vendedores-de-amazon-fracasan-amzalgo-com-2dccd18826fc"

Basado en el Data Product Canvas1 resuelve:

1. Problema.

a. ¿Cuál es mi problema alrededor de mi proyecto?

R:/ El problema alrededor del proyecto radica en la falta de éxito al momento de seleccionar un producto con un alto potencial de ventas. Existe una falta de claridad en la identificación de la categoría y del producto específico para un nicho de negocio determinado.

b. ¿Por qué se manifiesta el problema?

R:/ El problema se manifiesta debido a una investigación de mercado insuficiente el cual me impide seleccionar un producto con alto potencial de venta y comprender las dinámicas del mercado en el cual se desea incursionar. Esta falta de comprensión del mercado y de sus demandas contribuye a que un alto porcentaje de proyectos

Fracasen al no lograr esta selección.

c. ¿De quién es el problema?

R:/ El problema afecta principalmente a los emprendedores o empresarios que buscan tener éxito en sus negocios ya que enfrentan dificultades para alcanzar objetivos.

2. Datos.

a. Describa la fuente de datos (Sea más específico que en documento pasado)

R:/ Para este proyecto nos enfocaremos en la información que nos provee Amazon, dicha información se extraerá con una herramienta de suscripción llamada Hellium10, la cual es una plataforma diseñada específicamente para extraer información de Amazon la cual incluyen datos detallados sobre palabras clave relevantes para productos específicos, métricas de rendimiento de productos, datos de ventas históricas, tendencias de búsqueda y análisis de competidores.

b. ¿Cuál es la calidad que posee dicha fuente?

R:/ La calidad de la información proporcionada por Helium10 suele ser muy alta y confiable, ya que la plataforma utiliza datos extraídos del Web Services de Amazon y otras fuentes de datos externas.

c. ¿Es accesible y disponible?

R:/ La información es accesible y disponible luego de pagar una suscripción ya sea mensual o anual

Informacion de la herramienta en: https://www.helium10.com/

3. Hipótesis.

a. ¿Qué es lo que pretende probar?

R:/ Con el proyecto se pretende mejorar la capacidad de los vendedores para identificar productos con alto potencial de ventas y comprender mejor las dinámicas del mercado, ofreciendo insights precisos y detallados sobre qué porcentaje de éxito tendría el producto en el mercado, en la categoría seleccionada

b. ¿Cuáles pueden ser las respuestas esperadas a lo que pretende probar?

R:/

- Elegir e identificación productos con alto potencial de ventas

- Comprender el mercado y las dinámicas del nicho seleccionado

-Rentabilidad y ROI mas alto

c. ¿Qué acciones debo tomar para cada respuesta anterior?

R:/

- Analizar tendencias de mercado y datos históricos de ventas para identificar productos con alta demanda y bajo nivel de competencia.

-Realizar análisis de mercado para comprender las necesidades y preferencias de los consumidores en el nicho seleccionado.

-Realizar un seguimiento regular del rendimiento financiero y realizar análisis comparativos para identificar áreas de mejora

4. Solución.

a. Tipo. (¿Qué tipo de algoritmo de ML es? ¿Qué técnica es?

R:/ uno de los algoritmos de ML que podemos usar es de tipo “Aprendizaje Supervisado”, usando la técnica de “Regresión lineal múltiple”, todo esto bajo el lenguaje de programación Python

b. ¿Cómo debe ser implementado?

R:/Ya que usaremos regresión lineal multiple, seleccionaremos la variable independiente que seria el potencial de ventas de los productos, y las variables independientes serian diversas caracteristicas, precio, # de competidores, ventas, calificacion, entre otras). Luego dividiremos los datos separando el conjunto de entrenamiento con el de pruebas para luego realizar la evaluación del modelo utilizando métricas como el error cuadratico medio (MSE), el coeficiente de determinación (R2), entro otros indicadores de precisión.

c. ¿Qué resultados se esperan?

R:/ obtendremos los coeficientes de regresión que nos ayuda a relacionar la variable independiente con el potencial de venta y un valor intercepto que será el valor esperado del potencial de ventas. Adicional podemos realizar el entrenamiento del modelo podremos realizar predicciones sobre el numero de ventas, que se puedan generar

5. KPI’s

a. ¿Cómo espera evaluar el modelo?

R:/ para evaluar el modelo se realizara una validación cruzada k-fold, el cual me evaluara la robustez y generalización del modelo, esto para verificar si el modelo tiene la capacidad de realizar predicciones precisas.

b. ¿Qué métricas debe usar?

R:/ para evaluar el modelo con la validación cruzada k-fold, podremos usar algunas de las siguientes métricas: Error cuadrático medio, raíz del error cuadrático medio, error absoluto medio, coeficiente de determinación, coeficiente de correlación.

6. Actores.

a. ¿Quién es su cliente?

R:/ Serán empresas, emprendedores enfocados en importación y venta de productos ya sea en línea o de manera presencial y que desene mejorar la capacidad de identificación de productos potenciales en ventas, comprendiendo mejor la dinámica del mercado.

b. ¿Quiénes son los interesados (Stakeholders)?

R:/ Equipos comerciales, empresas o emprendedores enfocados en importación y ventas en plataformas de e-commerce, como Amazon, Mercado Libre, Walmart.

c. ¿Quién usará la solución?

R:/ La solución la podrán usar gerentes de producto, equipos de marketing, vendedores en línea, analistas de datos

d. ¿A quién impactará?

R:/ Impactara a toda persona que este interesada en lanzar un producto con posibilidades altas de éxito, ayudando a mejorar con la utilidad y el ROI para las compañías.

1

https://medium.com/@leandroscarvalho/data-product-canvas-a-practical-framework-for-building-highperformance-data-products-7a1717f79f0

Rúbricas

R01: Completitud y consistencia: el estudiante realiza una entrega del momento evaluativo con todos los elementos desarrollados, manteniendo una relación coherente entre los elementos solicitados. El estudiante usa una consistencia argumentativa que incluye el uso de referencias cuando hay mérito de autoría entre otros.

R02: Pertinencia: el proyecto del estudiante está dentro del ámbito de la analítica y la

ciencia de datos y es evidente el desarrollo de cada uno de los elementos del Data Product Canvas

R03: Solución: el estudiante es capaz de responder a las preguntas del Data Product

Canvas con la consistencia necesaria para relacionarla con el proyecto que desarrolla. Hay una relación evidente entre ellos. Las herramientas de MAchine Learning son una solución viable dentro de la descripción de la necesidad y objetivos del proyecto.

R04: Datos: los datos sugeridos son pertinentes al proyecto y al alcance delimitado en el

momento evaluativo anterior.