

| **1. Informe final Proyecto APT** |
| --- |
| El objetivo de este informe es que describas los aspectos más relevantes de tu Proyecto APT. Es importante que fundamentes las decisiones que tuviste que tomar a lo largo del proceso.  A continuación, encontrarás distintos campos que deberás completar con la información solicitada, los que dan cuenta del resumen de tu proyecto APT y sus principales resultados. |

| Nombre del proyecto | Análisis Predictivo de Comportamientos de Usuarios en Productos Copec |
| --- | --- |
| Área (s) de desempeño(s) | Se analizarán datos de cupones electrónicos y tarjetas de combustible para comprender el comportamiento de los clientes y mejorar su satisfacción, utilizando la metodología CRISP-DM. Se aplicarán modelos predictivos en Python, como Árboles de Decisión y Random Forest, para clasificar a los clientes y predecir su comportamiento futuro. La visualización de datos con herramientas como matplotlib y seaborn facilitará la interpretación de resultados. Además, se usarán herramientas colaborativas como Google Colab para optimizar el análisis. Los insights obtenidos permitirán a Copec diseñar estrategias personalizadas y mejorar su gestión de recursos, adaptándose a cambios en el mercado. |
| Competencias | 1.Análisis de Datos: Se aplicará la metodología CRISP-DM para analizar los datos de productos digitales de Copec, identificando patrones y tendencias con modelos predictivos en Python, como Árboles de Decisión y Random Forest.  2.Desarrollo de Modelos Predictivos: Se utilizarán técnicas de machine learning, como Regresión Lineal y Random Forest, implementadas con scikit-learn, para predecir el comportamiento futuro de los clientes.  3.Visualización de Datos: Se generarán gráficos con matplotlib y seaborn para representar claramente los resultados, facilitando su comprensión para los stakeholders.  4.Uso de Herramientas Tecnológicas: Se utilizará R y RStudio para colaborar en la ejecución del código, junto con RShiny y sus librerías para el análisis de datos.  5.Optimización de Estrategias y Recursos: Los insights del análisis predictivo permitirán a Skip diseñar estrategias de marketing más eficientes y optimizar la gestión de recursos, mejorando la adaptación a cambios en el consumo. |

| **Contenidos del informe final** | |
| --- | --- |
| 1. Relevancia del proyecto APT | * *Señala qué problema buscó solucionar tu proyecto y su relevancia para el contexto de la profesión. Algunas preguntas que pueden ayudarte a responder este apartado son:*   **Relevancia del tema para el campo laboral:**  El tema del Análisis Predictivo de Tendenciases altamente relevante para el campo laboral de mi carrera como analista de datos y consultor en ciencias de la computación, ya que se enfoca en el uso de herramientas de machine learning y análisis de datos para la toma de decisiones estratégicas. Este tipo de análisis es crucial para empresas que buscan optimizar sus recursos, mejorar la experiencia del cliente y anticipar comportamientos, lo que se alinea perfectamente con el perfil de un analista de datos o ingeniero informático. Las habilidades y conocimientos necesarios para implementar modelos predictivos son altamente demandados en la industria actual, ya que permiten a las organizaciones obtener insights accionables basados en datos, lo que mejora su competitividad.  **Ubicación de la situación a abordar:**  El proyecto se sitúa en Chile, específicamente en la empresa Skip, que tiene operaciones a nivel nacional. Skip es una de las principales empresas en el mercado de combustibles y productos energéticos, con una vasta red de clientes distribuidos en todo el país. La situación abordada se enfoca en el uso de datos generados por sus productos digitales, como cupones electrónicos, TCT y TAE.  Las principales características del lugar incluyen una amplia base de clientes, tanto individuales como empresariales, y una operación a nivel nacional con presencia en regiones urbanas y rurales. Esto hace que el análisis de datos sea clave para gestionar eficientemente los recursos y mejorar la relación con los clientes.  **Impacto en los afectados:**  La situación impacta a diferentes grupos dentro del mercado de Skip:   * **Clientes empresariales** que usan la TAE para gestionar el abastecimiento de sus flotas de vehículos. * **Consumidores individuales** que usan la TCT o cupones electrónicos para obtener descuentos y promociones. * **Segmentos de clientes** como nuevos, perdidos, decrecientes, crecientes y reactivados que serán categorizados por el modelo predictivo.   El análisis permite a Skip personalizar sus estrategias de marketing y servicio al cliente, mejorando la relación y satisfacción de estos usuarios.  **Aporte de valor del Proyecto APT:**  El aporte de valor de este proyecto para Skip es real, ya que el modelo predictivo desarrollado ofrece la capacidad de anticipar el comportamiento de los clientes, lo que permite:   * Optimizar las campañas de marketing al identificar qué segmentos requieren más atención o están en riesgo de perderse. * Mejorar la asignación de recursos mediante la personalización de ofertas y comunicaciones para cada tipo de cliente. * Aumentar la retención y lealtadde los clientes mediante estrategias de reactivación y prevención de pérdida. * Maximizar el valor de vida del cliente , lo que puede tener un impacto positivo en los ingresos y la posición competitiva de Skip. |
| 2. Objetivos | * *Objetivo general y específicos.*   *¿Cuál es el objetivo general de tu Proyecto APT? ¿Cuáles son los objetivos específicos de tu Proyecto APT?*  **Objetivo general del Proyecto APT:**  El objetivo general del Proyecto APT Análisis Predictivo de Tendencias es desarrollar un modelo predictivo robusto que permita identificar y prever con precisión las tendencias de comportamiento de los clientes en el uso de productos de Skip, como cupones electrónicos, TCT (Tarjeta de Combustible) y TAE (Tarjeta de Abastecimiento Empresarial). El propósito es que Copec pueda utilizar este modelo para personalizar estrategias de marketing, optimizar la asignación de recursos y mejorar la experiencia del cliente, anticipando la pérdida de clientes y maximizando el valor de vida del cliente (CLV).  **Objetivos específicos del Proyecto APT:**   1. **Categorizar a los clientes** en segmentos clave como nuevos, perdidos, decrecientes, crecientes y reactivados, mediante el análisis detallado de los datos históricos de consumo. 2. **Entender las dinámicas subyacentes** que impulsan los comportamientos de los clientes, identificando los factores que influyen en su lealtad, abandono o crecimiento. 3. **Proporcionar insights accionables** para Skip, permitiendo la personalización de estrategias de marketing y comunicación en tiempo real, con el fin de mejorar la satisfacción y lealtad del cliente. 4. **Optimizar la asignación de recursos** de Skip al anticipar cambios en el comportamiento de los clientes, permitiendo una distribución más eficiente de las ofertas y promociones. 5. **Desarrollar estrategias de reactivación** efectivas, basadas en la identificación temprana de patrones de pérdida de clientes, para maximizar la retención y reducir la tasa de abandono. 6. **Monitorear y adaptar en tiempo real** la oferta y las comunicaciones de Copec para mejorar la experiencia del cliente y fortalecer su posición competitiva en el mercado. |
| 3. Metodología | * *Metodología utilizada y su pertinencia para cumplir objetivos.*   *¿Qué metodología utilizaste para desarrollar tu Proyecto APT?*  .**Metodología utilizada:**  Para desarrollar el Proyecto APT Análisis Predictivo de Tendencias, utilicé la metodología CRISP-DM, ampliamente reconocida en el campo de la ciencia de datos y análisis predictivo. Esta metodología me permitió estructurar el proceso en fases claramente definidas, facilitando la comprensión del problema, el manejo adecuado de los datos y el desarrollo del modelo predictivo.  **Fases y procedimientos del proyecto:**   1. **Comprensión del negocio**: En esta fase, profundicé en el contexto del negocio de Skip, identificando los objetivos clave, como la necesidad de prever el comportamiento de los clientes en productos como TCT y TAE, y categorizar a los clientes en segmentos estratégicos. Esto me permitió definir los objetivos específicos del modelo predictivo y alinearlo con las necesidades de Copec para optimizar la personalización de estrategias de marketing y la asignación de recursos. 2. **Comprensión de los datos**: Aquí realicé un Análisis Exploratorio de Datos para examinar las características de los datos proporcionados, como los patrones de consumo, frecuencia de uso de los productos y características demográficas. Identifiqué valores atípicos, datos faltantes y patrones relevantes para alimentar el modelo. 3. **Preparación de los datos**: Esta fase fue crucial para la calidad del modelo. Implicó la limpieza de los dato**s**, como la eliminación de valores atípicos y la imputación de datos faltantes, así como la transformación de variables categóricas a numéricas. Además, se realizaron ajustes en las variables de entrada para mejorar la precisión del modelo predictivo. Este paso también incluyó la creación de nuevas características a partir de los datos históricos que aportaron información clave al modelo. 4. **Modelado**: En esta fase, implementé distintos algoritmos de machine learning (como regresión logística, árboles de decisión y random forest) para crear el modelo predictivo que clasificaría a los clientes en segmentos y predeciría cambios en sus comportamientos. Realicé un proceso de entrenamiento y validación cruzada para garantizar que el modelo fuese robusto y generalizable. 5. **Evaluación**: Evalué el desempeño del modelo utilizando métricas como R2, RMSE y accuracy (cuando fue pertinente), para comprobar que cumplía con los objetivos de precisión y fiabilidad. También se generaron informes de evaluación para analizar el ajuste del modelo y realizar posibles mejoras. 6. **Despliegue**: Implementé un dashboard interactivo que permite a Skip visualizar los segmentos de clientes y sus comportamientos. El modelo está diseñado para ser escalable y adaptable, y el dashboard facilita la toma de decisiones basadas en los insights del modelo predictivo.   **Fundamentación de la metodología:**  Elegí CRISP-DM porque es una metodología flexible y adaptada a proyectos de minería de datos y machine learning, lo cual se alinea perfectamente con el desarrollo del Proyecto APT. Esta metodología fue la más pertinente para cumplir los objetivos del proyecto por las siguientes razones:   1. **Estructura clara y eficiente**: CRISP-DM proporciona una guía detallada para cada fase, asegurando que se realicen todos los pasos necesarios, desde la comprensión del negocio hasta el despliegue del modelo, de manera ordenada. 2. **Adaptabilidad**: Debido a la naturaleza iterativa de CRISP-DM, puede ajustar las fases del proyecto según los resultados obtenidos en cada etapa. Por ejemplo, tras evaluar el modelo, fue posible regresar a la fase de preparación de datos para hacer ajustes y mejorar el rendimiento. 3. **Orientación a resultados de negocio**: Esta metodología no solo se enfoca en los aspectos técnicos, sino también en garantizar que el modelo predictivo ofrezca valor real al negocio, algo fundamental para cumplir con los objetivos planteados de mejorar la retención de clientes y optimizar las estrategias de marketing en Skip. |
| 4. Desarrollo | * *Descripción de las etapas o actividades del Proyecto APT.*   *¿Cuáles fueron las etapas o actividades que desarrollaste en tu Proyecto APT?*  En el desarrollo del Proyecto APT Análisis Predictivo de Tendencias, seguí una serie de etapas clave, alineadas con la metodología CRISP-DM:   1. **Comprensión del negocio**: Esta etapa consistió en entender a fondo los objetivos del proyecto dentro del contexto de Skip, enfocados en la identificación de patrones de comportamiento de los clientes a través de productos como TCT y TAE. También se definieron los objetivos específicos y se delinearon las expectativas del modelo predictivo. 2. **Comprensión de los datos**: Aquí analicé los datos proporcionados por Skip, identificando las variables más relevantes, como las transacciones históricas, frecuencia de uso y características de los clientes. También detecté la presencia de datos faltantes y patrones anómalos que debían abordarse. 3. **Preparación de los datos**: En esta etapa, realicé la limpieza y transformación de los datos. Esto incluyó la imputación de valores faltantes mediante técnicas como KNNImputer, y la conversión de variables categóricas a numéricas para alimentar los algoritmos de machine learning. Adicionalmente, generé nuevas características a partir de los datos históricos para optimizar el rendimiento del modelo. 4. **Modelado**: Implementé varios algoritmos de machine learning, como regresión logística, árboles de decisión y random forest, para predecir los comportamientos de los clientes. Cada modelo fue evaluado mediante técnicas de validación cruzada para asegurar su precisión y capacidad de generalización. 5. **Evaluación**: Realicé un análisis exhaustivo de los resultados del modelo mediante métricas como R2, RMSE y accuracy. Esta fase me permitió identificar oportunidades de mejora en el modelo y realizar ajustes en los parámetros. 6. **Despliegue**: Se desarrolló un dashboard interactivo para que Skip pueda visualizar y analizar los segmentos de clientes en tiempo real, permitiendo tomar decisiones estratégicas basadas en los resultados del análisis predictivo.   **Dificultades y facilitadores en el desarrollo del Proyecto APT:**  **Facilitadores:**   1. **Acceso a datos históricos completos**: Tener acceso a los datos históricos de consumo de Skip fue clave para la fase de modelado y permitió construir un modelo predictivo robusto. 2. **Uso de herramientas de análisis avanzadas**: La implementación de RShin, R y Rshiny facilitó el análisis, procesamiento y visualización de los datos, agilizando las etapas de modelado y evaluación. 3. **Soporte metodológico de CRISP-DM**: La estructura clara de la metodología CRISP-DM proporcionó un marco que guió el desarrollo del proyecto de forma ordenada y coherente.   **Dificultades:**   1. **Datos faltantes y desbalanceados**: Enfrenté problemas con datos incompletos y un desbalance en los segmentos de clientes. Algunos segmentos (por ejemplo, clientes perdidos) eran mucho más pequeños que otros, lo que afectaba la precisión del modelo. 2. **Complejidad en la integración de nuevas características**: A pesar de los avances en la preparación de los datos, la creación de nuevas características fue un desafío. Hubo dificultades para generar variables que realmente mejoran el rendimiento del modelo. 3. **Evaluación del modelo**: La validación del modelo con múltiples métricas y la necesidad de ajustes reiterados en los hiperparámetros complicaron la optimización del modelo predictivo.   **Ajustes realizados:**  **Ajustes ante las dificultades:**   1. **Imputación de valores faltantes**: Para abordar los datos incompletos, utilicé el método de imputación, lo que mejoró significativamente la calidad de los datos y permitió un mejor ajuste del modelo. 2. **Manejo del desbalance de datos**: Implementé técnicas de resampling y ajuste de pesos en los algoritmos de machine learning para corregir el desbalance en los datos. Esto ayudó a que el modelo tratará de manera más equitativa a todos los segmentos de clientes, mejorando la precisión en la predicción de clientes menos representados. 3. **Optimización de hiperparámetros**: Para mejorar el rendimiento del modelo, realicé ajustes finos en los hiperparámetros de los algoritmos utilizando Grid Search CV, lo que permitió obtener un mejor balance entre precisión y generalización. 4. **Refinamiento de las características**: Durante la fase de modelado, ajusté algunas de las nuevas características generadas para asegurar que aportarán valor al modelo. Esto implicó volver a la fase de preparación de datos para mejorar la ingeniería de características. |
| 5. Evidencias | * *Adjunta evidencias que permitan dar cuenta del desarrollo del Proyecto APT y sus resultados finales.*   *¿Qué evidencias pueden servir para que los demás puedan visualizar y entender las distintas etapas de tu Proyecto APT y el resultado final?* |
| 6. Intereses y proyecciones profesionales | * *Reflexión sobre el aporte del Proyecto APT en el desarrollo de los intereses profesionales.*   El Proyecto APT fue una experiencia invaluable que me permitió explorar a fondo el mundo del análisis predictivo y su aplicación en un entorno empresarial real como Copec. Este proyecto me sirvió para confirmar mi interés en el análisis de datos y en el desarrollo de modelos predictivos como herramientas clave para la toma de decisiones. Además, me permitió profundizar en la optimización de procesos empresariales, un área que siempre me ha interesado y que en este proyecto puede abordar desde una perspectiva de data science.  Uno de los aspectos más reveladores fue la capacidad de traducir datos en insights accionables que permiten personalizar estrategias y optimizar recursos. Esto reforzó mi deseo de continuar desarrollando habilidades relacionadas con la inteligencia de negocios y el machine learning. Aunque ya tenía claro mi enfoque hacia el análisis de datos, el proyecto me permitió entender mejor la importancia de la visualización de datos y cómo comunicar los resultados a niveles gerenciales, lo cual será crucial para mi crecimiento profesional.  Al final del proyecto, mis intereses profesionales han evolucionado hacia un enfoque más interdisciplinario, donde puedo integrar machine learning, optimización de procesos, y gestión de proyectos para resolver problemas complejos en un contexto empresarial. Aunque mi objetivo inicial sigue siendo el mismo, ahora veo más claramente el potencial de ampliar mi experiencia en áreas como la arquitectura de software y el desarrollo de modelos de datos, que son fundamentales para la creación de soluciones integrales.  **Proyecciones laborales a partir de Proyecto APT:**  A partir del Proyecto APT, me gustaría seguir profundizando en varias áreas clave:   1. **Machine Learning y Data Science**: Este proyecto ha reafirmado mi interés por el análisis predictivo. Quiero seguir desarrollando mis habilidades en modelos avanzados de machine learning, especialmente en áreas como la clasificación y predicción de comportamiento de clientes, para abordar desafíos más complejos. 2. **Inteligencia de Negocios (BI)**: Me gustaría seguir explorando las plataformas de Business Intelligence, La capacidad de transformar grandes volúmenes de datos en información útil para la toma de decisiones estratégicas me parece un área muy prometedora. 3. **Optimización de Procesos y Gestión de Proyectos**: La combinación de data science con la gestión de proyectos me permitió ver el valor de un enfoque estructurado y sistemático en la implementación de soluciones. Me veo trabajando en proyectos que incluyan tanto la optimización técnica de procesos como la coordinación de equipos multidisciplinarios para lograr objetivos de negocio.   **Proyección laboral**:  Después de haber completado el Proyecto APT, me veo orientando mi carrera hacia posiciones más senior en análisis de datos o como consultor de inteligencia de negocios. La experiencia en proyectos como este me ha dado una visión más clara de cómo quiero integrar mis conocimientos de machine learning, gestión de proyectos y análisis predictivo para contribuir al éxito empresarial. Eventualmente, me gustaría asumir un rol de liderazgo en proyectos de ciencia de datos, donde pueda no solo aplicar técnicas avanzadas, sino también mentorizar a otros profesionales en el área.  Este proyecto ha sido un paso fundamental en mi desarrollo, y estoy entusiasmado por seguir creciendo en el campo del análisis de datos, explorando nuevas metodologías y tecnologías, y aplicándolas para resolver problemas reales en el mundo laboral. |