}

1. **PARTE I**

| **1. Antecedentes Personales** |
| --- |
| A continuación, se presenta una tabla en la que debes completar la información solicitada. |

| Nombre estudiante | * **Benjamin Garcia** * **Felipe Galleguillos** * **Denisse Ponce** * **Macarena Vergara** |
| --- | --- |
| Rut | * **20782913-7** * **19406021-1** * **18189059-2** * **20538318-2** |
| Carrera | **Ingeniería en Informática** |
| Sede | **Plaza Vespucio** |

| **2. Descripción Proyecto APT** |
| --- |
| En la descripción debes señalar brevemente el nombre de tu proyecto APT y las competencias del perfil de egreso que vas a poner en práctica. Si en tu carrera están definidas las áreas de desempeño, también menciona a qué áreas de desempeño está vinculado el proyecto. |

| Nombre del proyecto | *CiberBI* |
| --- | --- |
| Área (s) de desempeño(s) | *Análisis de datos, desarrollo de sistemas de información, gestión de proyectos de TI.* |
| Competencias | * *Aplicar técnicas de ciencia de datos y Machine Learning.* * *Diseñar soluciones de software para la toma de decisiones.* * *Utilizar herramientas de visualización y análisis de datos (Power BI, SQL, Python).* |

| **3. Fundamentación Proyecto APT** |
| --- |
| A continuación, se presentan distintos campos que debes completar con la información solicitada. Esta sección busca que describas en detalle tu proyecto y justifiques su relevancia y pertinencia. |

| Relevancia del proyecto APT | *El proyecto busca apoyar la toma de decisiones en retail mediante predicción de ventas y análisis de datos. Aporta al campo laboral de la informática aplicada a negocios, ya que las empresas requieren sistemas de apoyo predictivo y visualización efectiva.* |
| --- | --- |
| Descripción del Proyecto APT | *Desarrollo de un sistema de análisis de datos y modelos predictivos con herramientas de Machine Learning y series de tiempo, evaluados con métricas de error, complementado con un dashboard en Power BI.* |
| Pertinencia del proyecto con el perfil de egreso | *Aplica competencias de análisis de datos, modelado predictivo y desarrollo de dashboards, alineado al perfil de egresado de Ingeniería en Informática.* |
| Relación con los intereses profesionales | *Se vincula con mi interés en ciencia de datos y BI, aportando experiencia práctica para mi desarrollo en esa área.* |
| Factibilidad de desarrollo del Proyecto APT | *El proyecto es factible de desarrollar dentro del semestre, considerando que la duración del mismo permite avanzar de manera progresiva en las distintas etapas: recolección de datos, análisis exploratorio, modelado, validación y visualización de resultados. Las horas asignadas a la asignatura resultan suficientes si se mantiene una adecuada planificación semanal y se avanza de forma iterativa.*  *En cuanto a materiales, se requieren principalmente recursos computacionales de nivel medio (un equipo personal con capacidad de ejecutar Python, Power BI y entornos de datos), además de acceso a datasets de ventas disponibles públicamente en plataformas como Kaggle, lo cual asegura que no existan restricciones críticas de insumos.*  *Existen factores externos que facilitan su desarrollo, tales como la disponibilidad de bibliotecas y frameworks open-source (scikit-learn, Prophet, pandas), plataformas de versionado como GitHub y herramientas de visualización accesibles como Power BI. Asimismo, el proyecto está alineado con tendencias actuales en ciencia de datos y analítica de negocios, lo que respalda su pertinencia académica y profesional.*  *En conclusión, el proyecto se considera viable en tiempo, recursos y contexto, con un nivel de complejidad adecuado al perfil de egreso de la carrera, siempre que se mantenga una gestión adecuada del cronograma y se priorice el alcance definido.* |

1. **PARTE II**

| **4. Objetivos** |
| --- |
| En este apartado debes definir objetivos generales y específicos del Proyecto APT. Es importante aclarar que los objetivos se deben plantear en forma clara, concisa y sin dar mayores explicaciones, es decir, deben entenderse por sí solos. Se sugiere redactarlos utilizando un verbo en infinitivo, pues ello obliga a precisar acciones concretas. |

| Objetivo general | *Desarrollar un modelo de predicción de ventas en empresas del sector retail, utilizando técnicas de análisis de datos, algoritmos de Machine Learning y visualización interactiva, con el propósito de anticipar la demanda, detectar patrones de consumo y apoyar la toma de decisiones estratégicas en la gestión comercial.* |
| --- | --- |
| Objetivos específicos | * *Recolectar y organizar un conjunto de datos históricos de ventas de retail, considerando variables relevantes como fechas, productos, promociones y localización de las tiendas.* * *Realizar un análisis exploratorio de datos (EDA) para identificar tendencias, patrones de consumo, estacionalidades y posibles anomalías en las series de ventas.* * *Implementar modelos predictivos basados en técnicas de Machine Learning (ej. Random Forest, XGBoost) y en modelos de series de tiempo (ej. ARIMA, Prophet) para estimar ventas futuras.* * *Evaluar el rendimiento de los modelos aplicados, utilizando métricas estadísticas de error como MAE, RMSE y MAPE, con el fin de determinar cuál ofrece mejores resultados.* * *Diseñar y construir un dashboard interactivo en Power BI que integre las predicciones, métricas clave y visualización de patrones de consumo, para facilitar la interpretación por parte de usuarios no técnicos.* * *Documentar de manera detallada la metodología, los resultados y las conclusiones del proyecto, elaborando un informe técnico y una presentación académica que respalde el trabajo realizado.* |

| **5. Metodología** |
| --- |
| En el siguiente apartado deberás describir la metodología, propia de tu disciplina, que utilizarás para resolver el proyecto APT antes descrito, incluyendo las etapas y métodos de trabajo. |

| Descripción de la Metodología |
| --- |
| *El proyecto se desarrollará bajo una* ***metodología ágil tipo Scrum adaptada al contexto académico****, ya que permite organizar el trabajo en ciclos cortos, generar avances incrementales y facilitar la coordinación entre los cuatro integrantes del equipo. Se trabajará en* ***sprints de dos semanas****, con reuniones de planificación al inicio de cada sprint para definir actividades y responsables, reuniones breves de seguimiento semanal para revisar avances, y una retrospectiva al finalizar cada sprint que permitirá identificar mejoras para el siguiente ciclo. Para la gestión de tareas se utilizarán tableros digitales (Trello o GitHub Projects), lo que permitirá visibilizar el estado de las actividades y asegurar el cumplimiento de plazos.*  *Cada integrante asumirá un rol principal, sin dejar de colaborar en las demás etapas:*   * ***Benjamín García (Data Engineering):*** *responsable de la recolección, limpieza y preparación de los datos.* * ***Felipe Galleguillos (Analista de Datos):*** *encargado del análisis exploratorio y la visualización inicial de patrones de consumo.* * ***Denisse Ponce (Científica de Datos):*** *responsable del desarrollo y validación de los modelos predictivos.* * ***Macarena Vergara (Desarrolladora BI):*** *encargada del diseño del dashboard interactivo en Power BI.*   *El* ***Scrum Master*** *será rotativo entre los integrantes, fomentando liderazgo compartido y responsabilidad en la gestión del proyecto. El rol de* ***Product Owner*** *será asumido de manera colectiva por el equipo, alineando las decisiones con las directrices del profesor guía.*  *En paralelo, desde la perspectiva técnica, el proyecto se estructurará en* ***etapas de análisis de datos y desarrollo iterativo****, propias de la informática y la ciencia de datos:*   1. ***Recolección y preparación de datos***     * *Obtener dataset histórico de ventas en retail (ejemplo: Walmart Store Sales).*    * *Realizar limpieza de datos: eliminación de nulos, duplicados y outliers.*    * *Estandarización de formatos de fechas, variables categóricas y numéricas.* 2. ***Análisis exploratorio de datos (EDA)***     * *Aplicar técnicas estadísticas y visualizaciones para identificar patrones, estacionalidades y tendencias.*    * *Generar gráficas de distribución, correlación y series de tiempo.* 3. ***Selección e implementación de modelos predictivos***     * *Evaluar modelos basados en series de tiempo (ARIMA, Prophet).*    * *Probar modelos de Machine Learning (Regresión, Random Forest, XGBoost).*    * *Seleccionar el modelo más adecuado en base a métricas (RMSE, MAE, MAPE).* 4. ***Evaluación y validación del modelo***     * *Comparar resultados de diferentes enfoques.*    * *Validar el modelo con un conjunto de datos de prueba.*    * *Documentar fortalezas y limitaciones del modelo escogido.* 5. ***Visualización de resultados***     * *Diseñar un dashboard interactivo en Power BI que muestre:*       + *Predicciones de ventas.*      + *Tendencias históricas.*      + *Métricas clave de rendimiento.*    * *Asegurar que la visualización sea clara y útil para la toma de decisiones estratégicas.* 6. ***Documentación y entrega final***     * *Redactar un informe técnico con metodología, experimentos y resultados.*    * *Publicar el código y dataset procesado en un repositorio de GitHub.*    * *Preparar presentación final para la defensa del proyecto.*   *La elección de esta metodología se fundamenta en que* ***favorece la entrega continua de evidencias, la adaptación a cambios en los requerimientos y la organización eficiente del tiempo****, aspectos clave en un proyecto de titulación con duración semestral.* |

| **6. Evidencias** |
| --- |
| A continuación, describe qué evidencias serán evaluadas en el informe de avance y en el informe final de tu proyecto APT. Estas evidencias deben ser acordadas con tu docente. Se entenderá por evidencia los productos que se desarrollen durante el proyecto y cuyo propósito sea visibilizar o documentar cómo se ha implementado el trabajo. |

| **Tipo de evidencia**  **(avance o final)** | **Nombre de la evidencia** | **Descripción** | **Justificación** |
| --- | --- | --- | --- |
| *Avance* | *Dataset limpio y documentado* | *Conjunto de datos históricos de ventas retail procesados (limpieza de nulos, outliers, formatos estandarizados) y descripción técnica del proceso* | *Asegurar la calidad y confiabilidad de los datos base del proyecto.* |
| *Avance* | *Resultados de EDA (Análisis Exploratorio)* | *Informe con gráficas, tendencias, estacionalidades y correlaciones detectadas en el dataset.* | *Permite comprender los patrones de consumo y fundamenta la selección de modelos.* |
| *Avance* | *Primer modelo predictivo* | *Implementación inicial de modelos de predicción y reporte de métricas de error preliminares.* | *Entrega un punto de comparación para evaluar mejoras en modelos más complejos.* |
| *Final* | *Modelo predictivo validado* | *Modelos seleccionados (series de tiempo y ML) evaluados con métricas (MAE, RMSE, MAPE) y documentación de resultados.* | *Constituye el núcleo técnico del proyecto y permite demostrar capacidad predictiva.* |
| *Final* | *Dashboard en Power BI* | *Tablero interactivo con predicciones, métricas de error y visualización de patrones de consumo.* | *Facilita la interpretación de los resultados por parte de usuarios no técnicos y apoya la toma de decisiones.* |
| *Final* | *Informe técnico y repositorio GitHub* | *Documento con metodología, resultados y conclusiones, junto con un repositorio con código y dataset procesado.* | *Evidencia la trazabilidad, reproducibilidad y formaliza los resultados del proyecto.* |

| **7. Plan de Trabajo** |
| --- |
| En la siguiente tabla define la planificación de tu Proyecto APT de acuerdo a lo requerido. |

| **Plan de Trabajo Proyecto APT** | | | | | | |
| --- | --- | --- | --- | --- | --- | --- |
| Competencia o unidades de competencias | Nombre de Actividades/Tareas | Descripción Actividades/Tareas | Recursos | Duración de la actividad | Responsable | Observaciones |
| Nombra las competencias o unidades de competencias que se relacionan con las diferentes actividades requeridas para el desarrollo de la actividad. | Señale el nombre de la tarea o actividad. | Describe la tarea o actividad. | Nombra los recursos necesarios para llevar a cabo las actividades definidas. | Escribe la duración de actividades o tareas. | Escribe el nombre del integrante del equipo responsable de la actividad y tareas asociadas. | Escribe las dificultades o facilitadores que se podrían presentar durante la ejecución de cada una de las actividades propuestas para llevar a cabo el plan de trabajo. |
| Ciencia de datos y gestión de información | Recolección y preparación de datos | Obtención del dataset histórico, limpieza de nulos y outliers, estandarización de variables. | Python, Pandas, SQL, Kaggle | 2 semanas | Benjamín García | Posible dificultad: calidad y volumen de datos; facilitador: dataset abierto disponible en Kaggle. |
| Análisis de datos e interpretación | Análisis exploratorio de datos (EDA) | Identificación de patrones, estacionalidades y tendencias. Visualización inicial de series temporales. | Python (Matplotlib, Seaborn), Jupyter | 2 semanas | Felipe Galleguillos | Dificultad: identificar patrones complejos; facilitador: bibliotecas estadísticas y gráficas. |
| Desarrollo de modelos predictivos | Implementación de modelos | Desarrollo de modelos de series de tiempo (ARIMA, Prophet) y Machine Learning (Random Forest, XGBoost). | Python (Scikit-learn, Statsmodels, Prophet) | 4 semanas | Denisse Ponce | Dificultad: ajuste de hiper parámetros; facilitador: conocimientos previos de ML. |
| Validación y comparación de resultados | Evaluación de modelos | Aplicación de métricas (MAE, RMSE, MAPE) y validación con backtesting. | Python, librerías métricas | 2 semanas | Denisse Ponce y Felipe Galleguillos | Dificultad: sobreajuste; facilitador: pruebas con diferentes enfoques. |
| Visualización de resultados | Desarrollo del dashboard en Power BI | Diseño del dashboard con predicciones, métricas y patrones de consumo. | Power BI, SQL | 2 semanas | Macarena Vergara | Dificultad: integración de datos; facilitador: experiencia previa en BI. |
| Documentación y comunicación | Informe final, GitHub y presentación | Redacción del informe técnico, carga de código en GitHub y preparación de presentación final. | Word, GitHub, PowerPoint | 2 semanas | Todo el equipo | Dificultad: tiempo limitado al final del semestre; facilitador: trabajo colaborativo y división de tareas. |

| **8. Carta Gantt** |
| --- |
| Busca un formato de Carta Gantt que te acomode y organiza en este las actividades planificadas en el punto anterior considerando el periodo asignado para el desarrollo de tu Proyecto APT. Debes mantener la temporalidad del periodo académico en el desarrollo de las tres fases que contempla la Asignatura de Portafolio de Título. |

| **Actividad** | **Fase 1** | | | | **Fase 2** | | | | | | | | | | | | **Fase 3** | | | |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **S 1** | **S 2** | **S 3** | **S 4** | **S 5** | **S 6** | **S 7** | **S 8** | **S 9** | **S 10** | **S 11** | **S 12** | **S 13** | **S 14** | **S 15** | **S 16** | | **S 17** | **S 18** |
| **Definición de alcances de proyecto y presentación de propuesta** |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  | |  |  |
| **Recopilación de datos y análisis de recursos** |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  | |  |  |
| **Análisis exploratorio de datos** |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  | |  |  |
| **Preparación de ambientes y recursos informáticos** |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  | |  |  |
| **Selección de modelos de predicción** |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  | |  |  |
| **Diseño módulo de usuario** |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  | |  |  |
| **Programación backend** |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  | |  |  |
| **Limpieza base de datos** |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  | |  |  |
| **Construcción modelos de predicción** |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  | |  |  |
| **Entrenamiento modelos de predicción** |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  | |  |  |
| **Ajustes de FrontEnd** |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  | |  |  |
| **Presentación propuesta** |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  | |  |  |
| **Testing de módulos** |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  | |  |  |
| **Ajustes de errores y optimización** |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  | |  |  |
| **Presentación Final** |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  | |  |  |