}

**Guía1. Definición Proyecto APT**

**Asignatura Capstone**

1. **PARTE I**

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

|  |  |  |  |
| --- | --- | --- | --- |
| Nombre estudiante | **Franco Martín Orsi Arroyuelo** | **María de los Ángeles Plaza de los Reyes Sanhueza** | **Anthony Jesús Silva Galindez** |
| Rut | **20.535.481-6** | **13.272.261-7** | **25.620.376-6** |
| Carrera | **Ingeniería en Informática** | | |
| Sede | **Antonio Varas** | | |

|  |
| --- |
| **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 | ***Diagnóstico temprano de Cáncer de Pulmón*** |
| Área (s) de desempeño(s) | * *Análisis y Evaluación de soluciones informáticas.* * *Desarrollo de software.* |
| Competencias | * *Desarrollar una solución de software utilizando técnicas que permitan sistematizar el proceso de desarrollo y mantenimiento, asegurando el logro de los objetivos.* * *Construir modelos de datos para soportar los requerimientos de la organización de acuerdo con un diseño definido y escalable en el tiempo.* * *Realizar pruebas de certificación tanto de los productos como de los procesos utilizando buenas prácticas definidas por la industria.* |

|  |
| --- |
| **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 cáncer de pulmón es una de las principales causas de muerte a nivel mundial, representando una amenaza significativa para la salud pública. En Chile, el cáncer es la segunda causa de muerte más común en cánceres, lo que resalta la importancia de un diagnóstico temprano para mejorar las tasas de supervivencia. Este proyecto se centra en la predicción temprana del cáncer de pulmón utilizando algoritmos de machine learning, lo que permitiría un diagnóstico precoz y, por lo tanto, un tratamiento más efectivo.*  *Elegimos este tema porque consideramos que el uso de herramientas tecnológicas avanzadas, como el machine learning, puede ser un gran aporte para el campo de la salud. La capacidad de predecir el riesgo de cáncer de pulmón puede beneficiar no solo a los profesionales de la salud, sino también a los pacientes, permitiendo tomar decisiones informadas y mejorar su calidad de vida.*  *Este proyecto tendrá un impacto en instituciones de salud, médicos y pacientes, especialmente aquellos en áreas urbanas con acceso a tecnología. Además, es una propuesta que contribuye a la innovación en el diagnóstico médico y puede integrarse en sistemas hospitalarios para asistir a los profesionales en la toma de decisiones.*  *Aporte de valor: Este sistema permitirá a los profesionales de la salud identificar a pacientes en riesgo de desarrollar cáncer de pulmón y así tomar decisiones proactivas sobre pruebas diagnósticas y tratamientos. Además, la reducción de tiempo y recursos para diagnosticar cáncer beneficiará a los sistemas de salud, mejorando la eficiencia y reduciendo costos.* |
| Descripción del Proyecto APT | *Nuestro proyecto se enfoca en la detección temprana del cáncer de pulmón mediante un sistema predictivo utilizando algoritmos de machine learning. El objetivo principal es desarrollar una herramienta que permita predecir si un paciente tiene un alto riesgo de desarrollar cáncer de pulmón basándonos en sus datos médicos y condiciones físicas. Este sistema analizará un conjunto de datos extraídos de fuentes como Kaggle e ICCR, que incluyen imágenes de radiografías y datos histopatológicos. Igualmente, se buscará fuentes de datos que correlacionen hábitos de vida y/o antecedentes familiares con la ocurrencia de cáncer pulmonar.*  *Para abordar la problemática, utilizaremos un enfoque basado en machine learning, explorando diferentes algoritmos predictivos y evaluando su rendimiento para seleccionar el más adecuado.*  *Dentro de los algoritmos, se considera para regresión y clasificación: árboles de decisión, regresiones lineales y logísticas y random forest; para estudio de imágenes: redes neuronales, KNN. Y para su evaluación, usar accuracy, precisión, recall, F1 Score, curva ROC y AUC para clasificación; error absoluto medio, error cuadrático medio, raíz del error cuadrático medio y coeficiente de determinación para regresión. Igualmente, se evaluará si es necesario utilizar validación cruzada para estos algoritmos.*  *No obstante, de acuerdo con los resultados obtenidos, se podrían incluir otros algoritmos como métodos de evaluación.*  *Asimismo, el proyecto generará informes en Power BI (o similares) para visualización de los resultados, asegurando que el sistema sea funcional y usable por profesionales de la salud no especializados en informática. Además, consideraremos preguntas clave como: ¿Qué factores de riesgo contribuyen al cáncer de pulmón? ¿Qué grupos de pacientes están más expuestos?*  *El proyecto sigue la metodología CRISP-DM para asegurar una correcta estructuración en las etapas de análisis, desarrollo, y prueba del modelo.*  *Para asegurar que la plataforma sea intuitiva y que la visualización de los resultados sea clara para el usuario final, se incluirán pruebas de usabilidad con profesionales de la salud.*  *Finalmente, se estudiará la factibilidad de la implementación de herramientas para el monitoreo continuo y mejora del modelo, como MLFlow, Prometheus + Grafana y/o Evidently AI, de acuerdo con la disponibilidad de tiempo y recursos.* |
| Pertinencia del proyecto con el perfil de egreso | *La pertinencia del proyecto con el perfil de egreso de nuestra carrera se refleja en la integración de las competencias clave que hemos desarrollado a lo largo de nuestra formación. Nuestro proyecto de diagnóstico temprano de cáncer de pulmón está directamente relacionado con el área de desarrollo de software, análisis de datos y machine learning, que son competencias fundamentales en el perfil de egreso.*  *El proyecto exige el uso de técnicas avanzadas de machine learning para predecir el cáncer en pacientes, lo que está alineado con nuestra competencia en el desarrollo y mantenimiento de software escalable, así como en la construcción de modelos de datos. Además, la capacidad de analizar y evaluar soluciones informáticas para problemas reales, como es el caso del diagnóstico médico, es una competencia central de nuestro perfil profesional.*  *Para abordar la problemática del diagnóstico temprano, necesitamos emplear competencias específicas como el diseño de modelos de datos sólidos y escalables, el desarrollo de algoritmos predictivos eficientes y la implementación de buenas prácticas en el desarrollo de software. Estas competencias nos permiten asegurar que el sistema no solo sea funcional, sino también confiable y fácil de usar, incluso para profesionales de la salud que no tienen conocimientos técnicos profundos en informática.*  *En resumen, el Proyecto APT no solo está alineado con el perfil de egreso, sino que también pone a prueba las habilidades técnicas y analíticas necesarias para enfrentar desafíos complejos en el campo de la informática aplicada a la salud.* |
| Relación con los intereses profesionales | *El Proyecto APT que estamos desarrollando está estrechamente relacionado con nuestros intereses profesionales en varias áreas clave, incluyendo el análisis de datos, machine learning y desarrollo de software orientado a la salud. Estos son campos que hemos priorizado en nuestra formación y que juegan un papel crucial en el diagnóstico temprano de enfermedades graves como el cáncer.*  *Este proyecto nos permitirá aplicar habilidades en la creación de soluciones predictivas basadas en datos, utilizando técnicas avanzadas de machine learning para contribuir a la mejora del sistema de salud, lo que refleja directamente nuestros intereses en utilizar la tecnología para solucionar problemas reales en el área de la salud. Además, trabajar en este tipo de proyecto contribuirá de manera significativa a nuestro desarrollo profesional, preparándonos para futuros desafíos en áreas críticas como la investigación y el desarrollo de sistemas inteligentes de apoyo a la toma de decisiones en el sector médico.*  *Este proyecto nos permitirá desarrollar competencias en el análisis y modelado de datos, que son esenciales para crear soluciones escalables y efectivas en el contexto médico, consolidando nuestros intereses en el uso de tecnologías emergentes para mejorar la calidad de vida y el cuidado de la salud.* |
| Factibilidad de desarrollo del Proyecto APT | *El desarrollo del Proyecto APT es totalmente factible dentro del marco de la asignatura y el semestre académico. El tiempo disponible es suficiente, considerando la duración del semestre y las horas asignadas a la asignatura, que nos permitirán avanzar de manera organizada utilizando la metodología Scrum, que promueve la planificación iterativa e incremental. Esta metodología es especialmente útil para dividir el trabajo en Sprints y cumplir con los plazos establecidos de manera eficiente.*  *En cuanto a los materiales, contamos con las herramientas necesarias para el desarrollo del sistema, incluyendo entornos de programación como Python, bibliotecas de machine learning, plataformas de almacenamiento de datos como Google Cloud, y acceso a datasets públicos relacionados con el cáncer de pulmón. Estas herramientas son suficientes para llevar a cabo tanto la parte de desarrollo de software como la implementación de los algoritmos de machine learning.*  *Factores externos que facilitan el desarrollo incluyen el acceso a datasets confiables y ya preprocesados, lo que nos permite enfocarnos directamente en el análisis y modelado de datos sin necesidad de realizar extensas tareas de limpieza de datos. Además, la amplia disponibilidad de recursos de aprendizaje en línea y documentación relacionada con las tecnologías que utilizaremos también favorece el avance del proyecto.*  *En cuanto a los posibles factores que podrían dificultar el desarrollo, el principal desafío es el manejo de grandes volúmenes de datos y la necesidad de ajustar adecuadamente los modelos de machine learning para obtener predicciones precisas. Sin embargo, para mitigar estos riesgos, implementaremos mecanismos de control de calidad durante el proceso de desarrollo, como pruebas iterativas y validación cruzada de modelos, que permitirán ajustar el rendimiento a tiempo.* |

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 sistema de diagnóstico temprano de cáncer de pulmón utilizando algoritmos de machine learning, que permita identificar la probabilidad de que un paciente presente cáncer basado en sus características físicas y condiciones ambientales.* |
| Objetivos específicos | * *Investigar y evaluar, al menos, tres algoritmos predictivos de machine learning para el diagnóstico temprano de cáncer de pulmón.* * *Seleccionar el algoritmo con el mejor rendimiento para predecir la probabilidad de cáncer en pacientes.* * *Implementar un sistema de software que permita la carga de datos, el procesamiento de los mismos y la generación de predicciones.* * *Desarrollar un dashboard interactivo en Power BI, o similar, para visualizar los resultados predictivos de manera clara y concisa.* * *Asegurar que el sistema cumpla con los atributos de calidad funcionalidad, usabilidad y confiabilidad para su uso por personal de salud.* |

|  |
| --- |
| **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 |
| *Para la organización y desarrollo del presente proyecto durante la Fase 1, el líder de equipo, Franco Orsi, ha tenido la tarea de organizar las actividades y capacitar a los demás integrantes sobre las herramientas de comunicación y desarrollo que serán utilizadas en el transcurso del proyecto.*  *En cuanto a la gestión del proyecto durante la Fase 2, se ha decidido utilizar la metodología ágil scrum, para lo cual, en principio, se designa a María de los Ángeles Plaza de los Reyes como Product Owner y Scrum Master, pudiendo asumir este(os) rol(es) otro integrante del equipo en un futuro si fuese necesario.*  *Los tres integrantes tendrán el rol de Development Team.*  ***Funciones, tareas y responsabilidades de cada rol en el equipo\****  ***Product Owner***  *Funciones:*   * *Maximizar el valor del producto.* * *Gestionar el Product Backlog.*   *Tareas:*   * *Definir los elementos del Product Backlog y priorizarlos según el valor de negocio.* * *Refinar y desglosar las historias de usuario junto con el equipo.* * *Definir claramente los requisitos y aceptar o rechazar el trabajo completado por el equipo.* * *Asegurarse de que el equipo comprenda los elementos del Product Backlog en el nivel necesario.*   *Responsabilidades:*   * *Mantener una visión clara y compartida del producto que se va a construir.* * *Alinear las necesidades del cliente y el negocio con las capacidades del equipo de desarrollo.* * *Garantizar que el trabajo realizado por el equipo se enfoque en lo que genera mayor valor para el negocio.*   ***Scrum Master***  *Funciones:*   * *Facilitar los procesos Scrum.* * *Mejorar la efectividad del equipo.*   *Tareas:*   * *Dirigir las ceremonias Scrum (Reuniones diarias, Planning, Sprint Review, Retrospectiva).* * *Ayudar al equipo a eliminar impedimentos que puedan afectar el progreso.* * *Asegurarse de que todos sigan los principios de Scrum.* * *Fomentar la autoorganización y la mejora continua del equipo.*   *Responsabilidades:*   * *Asegurar la correcta implementación de Scrum dentro del equipo.* * *Guiar al equipo en la adopción de principios ágiles y de Scrum.* * *Mantener el foco en los objetivos del Sprint.* * *Fomentar un ambiente de colaboración y comunicación abierta.*   ***Development Team***  *Funciones:*   * *Desarrollar y entregar incrementos del producto.* * *Auto-organizarse para cumplir los objetivos del Sprint.*   *Tareas:*   * *Diseñar, desarrollar, probar y entregar el trabajo comprometido para el Sprint.* * *Colaborar para dividir el trabajo y asignar las tareas necesarias.* * *Participar activamente en la planificación del Sprint y estimación de historias de usuario.* * *Garantizar la calidad del producto a través de revisiones y pruebas.* * *Desarrollar de manera incremental el producto para que sea funcional al final de cada Sprint.*   *Responsabilidades:*   * *Garantizar que el trabajo entregado al final de cada Sprint sea de alta calidad y cumpla con los criterios de aceptación.* * *Colaborar y coordinar el trabajo dentro del equipo de desarrollo sin depender de factores externos.* * *Auto-organizarse para encontrar la mejor manera de cumplir con los objetivos del Sprint.* * *Compartir la responsabilidad de la entrega del producto con el resto del equipo.*   *En cuanto al desarrollo de los algoritmos de machine learning se utilizará la metodología CRISP-DM.*  *\*Dado que puede haber cambios en la asignación de los roles de Product Owner y Scrum Master y que todos los integrantes cumplen el rol de Development Team, no se asigna directamente funciones, tareas o responsabilidades a cada integrante del equipo de forma directa en este documento, sin que serán asignadas mediante las tarjetas del tablero de Jira y documentadas en el informe de avance.* |

|  |
| --- |
| **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** | **Historias de Usuario** | *Listado de historias de usuario relevantes para el desarrollo del proyecto con su respectiva clasificación en las épicas correspondientes* | *Proporcionar la visión que tiene el equipo respecto a los desarrollos que se pretenden hacer para satisfacer las necesidades del usuario.* |
| **Avance** | **Roadmap del Producto** | *Plan estratégico de alto nivel que describe la visión del producto a lo largo del tiempo* | *Proporcionar una visión a largo plazo del producto y comunicar hitos clave y prioridades del proyecto.* |
| **Avance** | **Product Backlog** | *Lista priorizada y detallada de los requisitos del producto, contiene historias de usuario, tareas técnicas, épicas, bugs, y criterios de aceptación.* | *Establecer qué se decide hacer para construir el producto y cómo evoluciona con el tiempo.* |
| **Avance** | **Definition of Done (DoD)** | *Lista de criterios que un incremento debe cumplir para ser considerado completo y listo para entrega.* | *Reportar el aseguramiento de la calidad del producto y alineamiento de las expectativas de qué significa que una tarea esté "terminada" para el equipo.* |
| **Avance** | **Sprint Backlog** | *Tarjetas asignadas por Sprint, con las tareas asignadas a cada integrante.* | *Hacer trazabilidad a nivel individual y grupal del aporte al proyecto.* |
| **Avance** | **Acta de la Reunión de Retrospectiva del Sprint** | *Documentación de los aspectos discutidos en la reunión de retrospectiva del Sprint.* | *Recoger aprendizajes clave, identificación de áreas de mejora, y acciones acordadas en cada Sprint* |
| **Avance** | **Acta de la Reunión de Revisión del Sprint (Sprint Review)** | *Resumen de la reunión de revisión del Sprint, donde el equipo presenta el trabajo completado.* | *Hacer trazabilidad de la eficiencia que se ha tenido en cada Sprint.* |
| **Final** | **Interfaz funcional de proyecto** | *Interfaz que muestra el despliegue del (los) modelo(s) elegidos para cumplir con el objetivo del proyecto.* | *Evidenciar el término del desarrollo del proyecto.* |

|  |
| --- |
| **7. Plan de Trabajo** |
| En la siguiente tabla define la planificación de tu Proyecto APT de acuerdo con 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[[1]](#footnote-1) | Observaciones |
| * *Gestión de Proyectos Informáticos* * *Trabajo en Equipo y Comunicación* * *Ética Profesional* | *Organización del equipo* | *Organización del equipo de trabajo en cuanto al proyecto a realizar, sus objetivos, alineación de vías de comunicación, capacitación sobre las tecnologías a ocupar e implementación del repositorio.* | *Github, whatsapp, Google drive* | *4 semanas* | *Franco* | *Un facilitador para esta tarea es el hecho de que Franco y Ángeles han trabajado juntos toda la carrera, lo que plantea la posible dificultad para integrar a Anthony a una dinámica ya establecida.* |
| * *Gestión de Proyectos Informáticos* * *Ética Profesional* | *Entrega de la documentación individual de la fase 1* | *Completar y subir al repositorio los documentos 1.1, 1.2 y 1.3* | *Github, AVA* | *4 semanas* | *Franco*  *Ángeles*  *Anthony* | *La guía del profesor se presenta como un facilitador para llevar a cabo estas tareas. Como dificultad se presenta la falta de comprensión de algunos ítems de los documentos.* |
| * *Gestión de Proyectos Informáticos* * *Trabajo en Equipo y Comunicación* * *Ética Profesional* | *Entrega de la documentación grupal de la fase 1* | *Completar y subir al repositorio los documentos 1.4, 1.5 y presentación del proyecto* | *Github, AVA* | *1 semana* | *Franco*  *Ángeles* | *La guía del profesor se presenta como un facilitador para llevar a cabo estas tareas. Como dificultad se presenta la falta de comprensión de algunos ítems de los documentos y la falta de comunicación fluida del equipo de trabajo.* |
| * *Gestión de Proyectos Informáticos* * *Análisis y Diseño de Sistemas* * *Trabajo en Equipo y Comunicación* * *Ética Profesional* | *Sprint 0* | *Elaboración de Historias de Usuario, Roadmap del Producto, Product Backlog y DoD* | *Office, github, whatsapp, Google drive, Jira* | *1 semana* | *Franco*  *Ángeles*  *Anthony* | *La dinámica de trabajo ya establecida entre Franco y Ángeles facilitarían el desarrollo de esta tarea, podría dificultarse si no se alinean con Anthony.* |
| * *Gestión de Proyectos Informáticos* * *Análisis y Diseño de Sistemas* * *Trabajo en Equipo y Comunicación* * *Dominio de Tecnologías de la Información y Comunicación (TIC)* * *Adaptabilidad y Aprendizaje Continuo* * *Desarrollo y Mantenimiento de Software* * *Ética Profesional* | *Sprint 1-8* | *Desarrollo de 8 Sprints* | *Office, github, whatsapp, Google drive, Jira, VSCode, Anaconda, Colab, Datasets de Kaggle* | *9 semanas* | *Franco*  *Ángeles*  *Anthony* | *La dinámica de trabajo ya establecida entre Franco y Ángeles facilitarían el desarrollo de esta tarea, podría dificultarse si no se alinean con Anthony, además, podrían presentarse dificultades a nivel técnico por falta de conocimientos en alguna área de desarrollo o por insuficiencia de hardware que rinda en modelos muy exigentes.* |
| * *Trabajo en Equipo y Comunicación* | *Preparación de presentación de proyecto APT* | *Confección del material audiovisual e informe(s) correspondiente(s) a la presentación del proyecto* | *Office, documentación del proyecto* | *2 semanas* | *Franco*  *Ángeles*  *Anthony* | *La experiencia del equipo en elaboración de material audiovisual es un facilitador.* |
| * *Trabajo en Equipo y Comunicación* | *Presentación de proyecto APT* | *Presentación del proyecto frente a la comisión* | *Office, documentación del proyecto* | *20 minutos (2 semanas\*)* | *Franco*  *Ángeles*  *Anthony* | *La experiencia del equipo presentando frente a público es un facilitador. Una dificultad sería el nerviosismo frente a la comisión.* |

|  |
| --- |
| **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 |
| *Organización del equipo* |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| *Entrega de la documentación individual de la fase 1* |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| *Entrega de la documentación grupal de la fase 1* |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| *Sprint 0* |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| *Sprint 1-8* |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| *Preparación de presentación de proyecto APT* |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| *Presentación de proyecto APT* |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |

1. En caso de que el Proyecto APT sea grupal, en esta columna deben indicar el nombre de los responsables de cada tarea o actividad. Esto posteriormente permitirá diferenciar la evaluación por cada integrante. [↑](#footnote-ref-1)