

ST1613 GUÍA DEL ANTEPROYECTO

Trabajo Final: Plan de un proyecto de aprendizaje automático

NOTA 1: Revise bien los porcentajes de calificación y concentre su esfuerzo en lo que se considera más importante según la directriz dada.

NOTA 2: Para la entrega preliminar del anteproyecto (en la Clase 4), los únicos ítems obligatorios son "**pregunta de investigación y objetivos**" y "**datos para desarrollar el proyecto y análisis previo**". Para la entrega final, todos los ítems son obligatorios.

Debe producir un documento de 4 a 6 páginas (Arial 12) que contenga lo siguiente:

- **(20%) Pregunta de investigación y objetivos**

Describa la razón del proyecto, en ocasiones referida como la pregunta de investigación. Asegúrese que sea realista, alcanzable en el tiempo conferido y descrita en detalle para evitar ambigüedad. Presente el objetivo principal y específicos de su trabajo, los cuales formarán la base de las actividades que intenta lograr.

- **(15%) Metodología de investigación**

Resumen de la metodología a usar y los métodos de análisis propuestos. No se requiere un detalle completo de los métodos y técnicas, pero sí conocer un lineamiento de qué piensa probar y cuántas iteraciones de cada técnica, ya que esto se refleja en el plan detallado.

- **(10%) Revisión de la literatura, estado del arte y bibliografía**

Describa y referencie los trabajos similares o relacionados que han desarrollado los expertos en el tema que se está trabajando.

- **(25%) Datos para desarrollar el proyecto y análisis previo**

Es muy importante tener una intuición de antemano de si los datos disponibles al momento de comenzar brindan la información suficiente para lograr el objetivo. Por esto requiero ver que saben con qué datos cuentan, tanto en X (nombre de variable y descripción) como la y (en el caso supervisado). Especial foco en la variable dependiente (y) y la métrica a usar, ya que en mi experiencia el planteamiento de esta no suele alinearse con el objetivo del proyecto. Describan las fuentes de datos de ser posible, formato original de los mismos y tomen un tiempo para un análisis exploratorio básico (correlaciones o distribuciones, en especial entre las independientes y la dependiente) para saber si hay posibilidad de extraer $P(y|X)$. También describan las transformaciones a utilizar (en caso de variables no numéricas) y cualquier otro detalle a considerar con los datos elegidos.

- **(15%) Plan (diagrama Gantt o Pert)**

Usando un diagrama de preferencia (incluirlo en el documento de manera visible), detalle las tareas que requerirá para llevar a cabo los objetivos planteados. Enfóquese en demostrar la línea de tiempo relativa a cada aspecto del plan y las dependencias entre actividades. Para el tiempo trate de usar una escala de tiempo granular (días o semanas), para que pueda revisar su progreso hacia la entrega final. Propongan una escala realista, no es para impresionar al lector sino un plan para buscar seguirlo.

- **(10%) Implicaciones éticas**

La aplicación de la tecnología siempre tiene implicaciones éticas. Estas pueden ser genéricas y difíciles de definir, pero debe tomar un tiempo para revisarlas, ya sean que afecten a los humanos involucrados en el proceso (disminuir empleos, cambio de trabajo, entre otras), al medio ambiente o a la sociedad. Concéntrese en las implicaciones directas del proyecto en caso de ser logrado exitosamente o en caso poner en producción y se tengan fallas.

- **(5%) Aspectos legales y comerciales**

Realice una descripción corta del potencial comercial (dentro o fuera de la compañía) de su investigación. Adicionalmente revise si hay un marco o aspectos legales que deba considerar al momento de exponer los resultados o poner en producción el proyecto.