**Información general del proyecto**

**Objetivo del proyecto**: Verificar aplicación de modelos de IA en perfilamiento de estudiantes de cursos de programación tras investigación.

**Entregables del proyecto**:

* Dataset con códigos fuente y notas para conjunto de estudiantes.
* Vectores de Código generados por modelos pre entrenados.
* Predictor que correlacione la vectorización con cuartil de rendimiento académico de los estudiantes.

**Plazo del proyecto**: El avance de Proyecto debe estar listo al finalizar junio 2024 (final de 1er semestre)

**Tareas**:

**Marzo / Abril**:

* Recopilar notas y códigos fuente de un conjunto representativo de estudiantes de programación.
* Realizar el preprocesamiento necesario
* Investigar modelos pre entrenados para la vectorización de código.
* Generar vectorizaciones de código utilizando los modelos seleccionados.
* Desarrollar método para comparar las vectorizaciones generadas.
* Analizar y evaluar las diferencias y similitudes entre las vectorizaciones de código obtenidas.

**Mayo**:

* Diseñar modelo predictivo que utilice las vectorizaciones de código.
* Explorar algoritmos de aprendizaje automático para construcción del predictor.
* Entrenar y validar predictor utilizando el conjunto de datos construido.
* Evaluar rendimiento del predictor mediante métricas adecuadas (precisión y el coeficiente de determinación).

**Junio**:

* Analizar relación entre vectorizaciones de código y rendimiento académico de estudiantes.
* Interpretar resultados del predictor para comprender cómo las características del código influyen en el rendimiento estudiantil.
* Identificar patrones y tendencias significativas que puedan guiar futuras investigaciones o intervenciones educativas.

**Recursos para el Proyecto**:

Lenguajes de programación: Python

Bibliotecas: SKLearn, TensorFlow, PyTorch

Modelos Pre-Entrenados: CodeBERTa (por ahora)

Plataformas de Desarrollo \*?

Herramientas de Gestión \*?

**Notas**:

* ¿Establecer Cronograma?
* Documentar descripciones de avance (README o archivos de texto)
* Actualizar repositorio al completar tareas y/o obteniendo resultados
* Comunicar resultados
* Actualizar Planificación a medida que se necesite durante el desarrollo