

**Materia:** Tópicos Selectos en Inteligencia Artificial  
**Semestre:** Segundo 2022  
**Docente:** Ariel Hugo Alba Ríos  
**Fecha:** 04-10-2022

### Descripción

Uno de los tópicos de mayor tendencia y crecimiento en el campo de la investigación y económico es el de la Inteligencia Artificial (IA).

Precisamente, la Inteligencia Artificial es una de las disciplinas científicas que hace uso de la información habiente para ayudarnos a través de sus procesos / modelos o algoritmos a aumentar la productividad, reducir costos y a obtener más valor de los datos.

El siguiente documento describe los objetivos, reglas, reglas, procesos y deadlines del proyecto final para la materia Tópicos Selectos en Inteligencia Artificial.

Con el proyecto se espera aplicar el aprendizaje basado en problemas, para lo cual, se les entregará un dataset del mundo real y se aplica diferentes técnicas y algoritmos para poder dar conclusiones enriquecedoras respecto a los datasets.

### Objetivo

El objetivo es que los estudiantes apliquen todo lo aprendido en la materia (procesos / métodos / algoritmos) para poder extraer resultados de un dataset que ellos mismos elegirán.

### Reglas

Los estudiantes podrán formar grupos compuestos de 2 personas como máximo. Los estudiantes también tendrán la posibilidad de presentar el proyecto de forma individual.

El ponderaje será el mismo ya sea para grupos formados por 2 personas o individuales.

Sin embargo, algunas procesos variarán de acuerdo al tipo de grupo que se forme (Ver Proceso); además se exigirá un agregado a los grupos formados por 2 personas, en cuanto a pulcritud de código y documentación.

Se exigirá pulcritud en el documento final, además de legibilidad en cuanto al código. Por código legible se entiende que el mismo deberá estar autodocumentado o comentado. No debe existir códigos comentados, a no ser que se esté explicando el motivo del comentario. Esto puede influir en la nota final del grupo / estudiante en el proyecto. En caso de que el código

Se puede utilizar cualquier biblioteca o herramienta externa escrita en Python, como también se puede utilizar los snippets de código hechos en clase.

En base a la promoción de la HONESTIDAD, el plagio en el proyecto será sancionado con la nota mínima para el proyecto (0 puntos). Por plagio se entiende si el código presentado por un estudiante es igual o similar (definido por el docente) al de otro estudiante.

El estudiante deberá explicar el código en la defensa final, en caso que el docente lo requiera.

## Proceso

**1.** Escoger al menos 2 datasets del archivo DatasetProyecto.ods en UPB Virtual. Uno será el principal, y el otro será utilizado como back-up en caso que las preguntas no sean lo suficiente para el proyecto.

**1.1** El grupo o el estudiante puede escoger otros dataset de internet, sin embargo, estos debe estar aprobados por el docente. Para la aprobación, es necesario presentar las opciones hasta el primer checkpoint (06-10-2022), para que en caso de rechazo, se tenga tiempo para escoger otros dataset o seleccionar algunos de la URL del punto 1.

**2.** Especificar 5 preguntas por dataset, siempre pensado que estas se puedan resolver con métodos / algoritmos llevados en clase (En caso de que la pregunta sea compleja, se alertará al estudiante en el segundo checkpoint).

**2.2** (Con el Docente) Escoger las preguntas a responder y el dataset final a utilizar.

**3.** Preprocesar el dataset. Limpiar los datos, normalizar, discretizar y graficar las distribuciones de los features. Sacar conclusiones de los gráficos

**4.** Preparar modelos.

Para los grupos compuestos por 2 personas, 3 modelos serán obligatorios, 2 modelos básicos y una red neuronal (puede ser cualquier combinación que incluyan las 2 clases).

Para los estudiantes individual 2 modelos serán obligatorios (1 caso base y una red neuronal).

Cada modelo deberá incluir: "feature engineering" acompañado de gráficos.

**5.** Sacar conclusiones que respondan las preguntas de **(2.2)** y plasmar todo en el documento final a presentar.

## Checkpoints / Deadlines / Nota del Segundo Parcial

**Primer Checkpoint (1 y 1.1)** **06-10-2022** **15%**

### Entregables:

- Lista de alumnos que componen el grupo
- Datasets escogidos

**Segundo Checkpoint (2 y 2.1)** **12-10-2022** **35%**

### Entregables:

- Lista de preguntas a realizar con los 2 datasets escogidos
- Modelos a utilizar para responder las preguntas (Ideas iniciales)

**Defensa Parcial (3)** **19-10-2022** **50%**

### Entregables:

- Documento del Proyecto (1.5 hojas max) en el que se describe:
  - Dataset a utilizar
  - Objetivos (Responder a las preguntas)
  - Preprocesamiento del dataset (Ayudado por gráficos si es necesario)
- Archivos .ipynb o .py que corroboran el preprocesamiento
  - Será necesaria una documentación extra en caso de que se requiera pasos para ejecutar el código
- Pequeña defensa donde el estudiante deberá explicar el documento en 10 min máximo

## Checkpoints / Deadlines / Nota del Final

**Primer Checkpoint (4)** **26-10-2022** **30%**

### Entregables:

- Mostrar avance en cuanto al modelaje y optimización. Hablar sobre los modelos escogidos e hiperparámetros a optimizar

**Defensa Parcial (5)** **01-10-2022** **70%**

### Entregables:

- Documento del Proyecto (5 hojas max) en el que se describe:
  - Dataset a utilizar
  - Objetivos (Responder a las preguntas)
  - Preprocesamiento del dataset
  - Gráficos de los features
  - Modelos a utilizar y el porqué se los escogieron
  - Feature Engineering (Hyperparameter tuning)
  - Reducción de dimensionalidad (si es aplicado, en caso contrario colocar las razones porqué no se lo utilizó)
  - Resultados gráficos de las optimizaciones (Importante!)
  - Conclusiones
- Archivos .ipynb o .py que corroboran el modelaje (entrenamiento y optimización)
  - Será necesaria una documentación extra en caso de que se requiera pasos para ejecutar el código

- Defensa del Proyecto, describir en máximo 20 min todo el proceso del proyecto.