Lineamientos para la evaluación del curso Inteligencia Artificial Supervisados

O. L. Quintero

26 de abril de 2023

Resumen

Este documento contiene los lineamientos para la presentación de la actividad evaluativa.

1. Introducción

El objetivo del curso es proveer elementos teóricos y conceptuales que le permitan a los estudiantes de Inteligencia Artificial de Ingenieria Matemàtica, enfrentar el problema de construir un modelo compacto (learning machine) que permita representar fenómenos del mundo real.

Consecuentemente, los principios de teoría de aprendizaje fueron adelantados en la clase. Se debe cuestionar y NO desviar la tarea de aprendizaje automático, es decir no "presuponer" la naturaleza del mismo. Debe explorarse la construcción de diversos modelos mediante la aplicación de los conceptos.

El curso pretende adelantar la correcta aplicación de conceptos que permiten construir el modelo, evaluarlo y juzgar con perspectiva científica su desempeño.

Esta actividad evaluativa consiste en entrenar varios modelos de aprendizaje supervisado

- 1. Arbol de decision
- 2. Maquina de soporte vectorial con tres kernels: lineal, polinomico y base radial
- 3. Modelos de regresion lineal y/o logistica

Los conceptos generales se pueden revisar directamente del libro "Machine Intelligence for decision making" (en borrador para uso de los estudiantes de este curso y bajo edicion por Springer), y de las diversas fuentes citadas en el libro con artículos científicos y otros libros mas especializados en cada tema.

Para comenzar a realizar su proceso de aprendizaje (me refiero a practicar en datos juguete antes de abordar el problema real), el estudiante puede usar conjuntos de datos sintéticos IRIS DATASET.

2. Contenidos

El estudiante debe seleccionar un problema relevante, y realizar este ejercicio de manera INDIVIDUAL.

El estudiante puede usar cualquier tipo de programa para realizar el ejercicio que se detallará en la siguiente sección.

Los estudiantes abordaran el entrenamiento de modelos de aprendizaje siguiendo los siguientes elementos:

- 1. Error real deseado
- 2. Error de entrenamiento deseado
- 3. Garantia probable de aprendizaje
- 4. Tamaño óptimo de la muestra
- 5. Dividir el conjunto de muestra en: entrenamiento, validación y prueba.
- Visualizar los datos con los algoritmos de visualización por defecto de la suite, extraer características si hace falta para aumentar la dimensionalidad
- 7. Visualizar los datos con el embebimiento BH tsne https://lvdmaaten.github.io/tsne/o con el UMAP que me gusta tanto
- 8. Entrenar el modelo con los datos en dimensiones originales
- 9. Entrenar el modelo con los datos en altas dimensiones (espacio original)
- 10. Entrenar el modelo con los datos en dimensiones reducidas (luego del embebimiento o con el autoencoder)
- 11. Realizar la comparación de los modelos en los espacios analizados

2.1. Entrenamiento de las maguinas de aprendizaje

Para su problema con m entradas, n salidas y N cantidad de instancias (patrones o tamaño de la muestra). Seleccionar las siguientes maquinas:

- 1. Regresion lineal o regresion logistica
- 2. Arbol de decision
- 3. Maquina de soporte vectorial con kernel lineal
- 4. Maquina de soporte vectorial con kernel polinomico
- 5. Maquina de soporte vectorial con kernel base radial

Plantear el problema usando las garntias para estas maquinas y su dimension VC. En este dibujito la tarea que puse en el 2022-2, pero pueden guiarse por el tablero de la clase que hicimos.

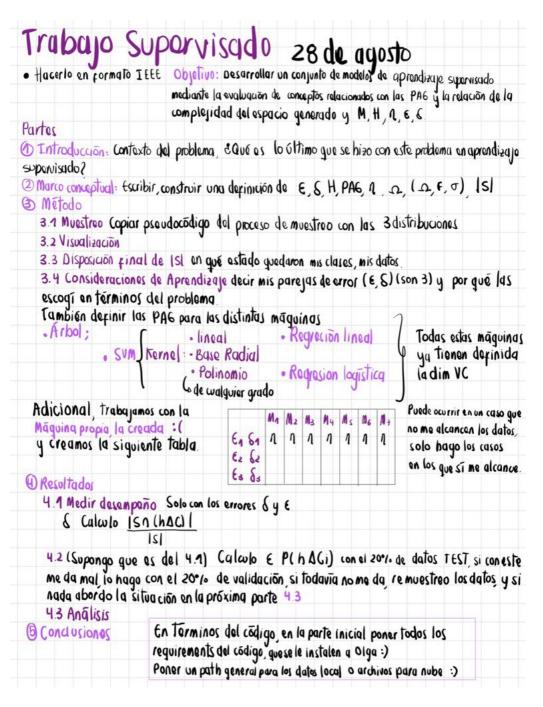


Figura 1: Tablerito

2.2. Entregables

Se debe entregar:

- Documento informe
- \blacksquare Codigos elaborados

Por favor asistir a asesoría con el profesor del curso en la plataforma teams y solicitar su cita con anticipación al correo oquinte1@eafit.edu.co.