

|  |  |
| --- | --- |
| Colegio Universitario **IES** *Siglo 21* | |
| EXAMEN FINAL | |
| **Materia: Introducción a la IA** | **Docente:** Ricardo Piña |
| **Modalidad: Presencial** | **Fecha: 22/07/2021** |

Reservado para el alumno

|  |  |
| --- | --- |
| **Alumno**: Sibello Gino | **Carrera:  INTELIGENCIA ARTIFICIAL Y CIENCIA DE DATOS** |
| **DNI**: 41 820 055 |

|  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Actividad** | **1** | **2** | **3** | **4** | **5** | **6** | **7** | **8** | **9** | **10** | **Total puntos** | NOTA |
| **Puntaje** |  |  |  |  |  |  |  |  |  |  |  |  |

* Modalidad de Evaluación:
  + Escrito e individual.
* Puntaje: Para aprobar el presente examen deberá haber obtenido **60 puntos** como mínimo.
  + - El puntaje se determinará a través de la siguiente escala:

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Puntaje** | **1-10** | **11-39** | **40-59** | **60-65** | **66-72** | **73-78** | **79-85** | **86-91** | **92-97** | **98-100** |
| **Nota** | **1** | **2** | **3** | **4** | **5** | **6** | **7** | **8** | **9** | **10** |

1. ¿Cuándo se dice que tenemos un problema de Regresión y cuándo uno de Clasificación? Dé un ejemplo inventado por Ud. de cada uno.

Se dice que tenemos un problema de regresión cuando los datos son Cuantitativos, ósea Números, un ejemplo seria determinar cuántas cantidades de un producto voy a vender, la respuesta sería X cantidades del mismo,

Tenemos un problema de clasificación cuando los datos son Cualitativos, ósea todo menos números, un ejemplo seria determinar si un individuo en particular va ser un potencial cliente, la respuesta aquí seria Si o No.

1. Ud tiene que resolver un problema de Clasificación con un Árbol de Decisión y cuenta con un DataSet. Cuál será la Accuracy (Exactitud) si entrena al modelo con todo el DataSet, lo evalúa sobre todo el DataSet y permite que el árbol **no tenga límite** en cuanto su profundidad (cantidad de niveles de nodos y hojas).

El Accuracy seria de 1 el cual en términos matemáticos seria perfecto y eso se debe a que estoy sufriendo de algo llamado overfiting ya que le estoy sometiendo a estudio los mismos datos con los que lo entrene y por ende se habría vuelto tan bueno clasificándolos que no cometería errores, el problema es que si le agrego nuevos datos para testearlo este seria muy malo con los nuevos datos ya que se habría ‘especializado’ únicamente en los datos anteriores y no seria para nada bueno generalizando nuevos casos.

1. ¿Qué se minimiza en el Método de Regresión Lineal?

Se minimiza la distancia entre los datos y nuestra recta o curva, en otras palabras

Cuando usamos regresión lineal siempre nuestros datos van a tener ruido o errores de

Medición y si lo vemos en una gráfica nos quedaría una serie de puntos no alineados

Perfectamente y para poder conseguir la recta o curva que mejor se acerque a todos los

Puntos debo minimizar lo máximo posible la distancia entre la recta y todos los puntos

1. ¿Cuándo tenemos que determinar los hiperparámetros de un modelo, por qué se suele utilizar el Validation Train y el Validation Test?

Se utiliza el val para así dejar que el test set sea únicamente datos puestos para testear y que nuestro modelo no se entre con ellos, así podemos determinar que tan bueno es nuestro modelo para clasificar de una manera general datos con los que nunca antes había tenido contacto, esta es una forma de evitar el overfiting

1. MODELIZACIÓN 1: En el archivo final\_data4.csv se le otorga los datos de las variables x e y, ambas numéricas. Genere en Orange3 el modelo de REGRESIÓN LINEAL que mejor pronostique los valores de y, puede utilizar hasta la potencia 5 de la variable x. Todas las divisiones del dataset deben ser 80/20.
   1. ¿Indique aquí cuál es la potencia elegida?

Potencia elegida = 4

* 1. Evalúe el modelo elegido e indique aquí el RMSE que esperaremos que obtenga al usarlo en el “mundo real”

RMSE = 0,127

* 1. Indique aquí los valores obtenidos para los **parámetros** del modelo elegido

*h*(*x*) = -0,0405 + 4.0283*x* + 0,1353*x*2 + 0,0728*x*3 -0,8538*x*4

* 1. Genere el modelo para producción
  2. Deberá entregar el archivo SuApellido\_SuNombre.ows.
  3. Guarde todos los archivos en un archivo comprimido y envíemelo por mail a ricardo.pina@gmail.com