

|  |  |
| --- | --- |
| Colegio Universitario **IES** *Siglo 21* | |
| IEFI – PARTE PRÁCTICA | |
| **Materia: Introducción a la IA** | **Docente:** Ricardo Piña |
| **Modalidad: Presencial Semipresencial Distancia** | **Fecha:** |

Reservado para el alumno

|  |  |
| --- | --- |
| **Alumno**: Lautaro Santos Da Silveira | **Carrera:  INTELIGENCIA ARTIFICIAL Y CIENCIA DE DATOS** |
| **DNI**: 43879787 |

|  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **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. MODELIZACIÓN 1: El archivo credit\_japan.csv contiene datos **reales** brindados por una empresa japonesa que se dedica a evaluar candidatos para otorgar tarjetas de crédito. ¡Por este motivo han ocultado los nombres de todas las variables (y los clientes)!
   1. Sólo sabemos que la variable target es la denominada A16 y que asume los valores “+” y “-“. Nuestro objetivo será pronosticar el valor de dicha variable A16, por lo cual será un problema de Clasificación.
   2. Encuentre el árbol de decisiones que mejor resuelva nuestro problema considerando la métrica Accuracy, utilice como hiperparámetro la profundidad del árbol. Para dividir DataSets utilice la proporción 80/20.
   3. Guarde el Archivo con el nombre credit\_japan.ows
   4. A continuación escriba un resumen de lo obtenido:

d) La profundidad del árbol que usaremos será de 1, si bien la profundidad 3 tiene el mismo valor de AUC, este no generalizaría los futuros casos tan bien como el de la profundidad 1. La variable que usaremos para generar el árbol será A9

1. MODELIZACIÓN 2: En el archivo data\_regresion.csv se le otorga los datos de las variables x e y, ambas numéricas, por lo cual será un problema de Regresión. Genere en Orange3 el modelo de Regresión Lineal (polinómica) que mejor pronostique los valores de y. Para evaluar utilizaremos RMSE.
   1. Para el modelo de Regresión Lineal analice hasta los polinomios de grado 6 generando las potencias necesarias.
   2. Para dividir DataSets utilice la proporción 80/20.
   3. En base al RMSE seleccione el modelo (y el hiperparámetro) que esperamos generalice mejor.
   4. Calcule el RMSE que esperamos tenga el modelo seleccionado al utilizarlo en las observaciones no conocidas.
   5. Deje también listo el modelo para producción.
   6. Guarde el archivo con el nombre regresion.ows.
   7. A continuación, escriba un resumen de lo obtenido: indicando claramente modelo elegido, valor del hiperparámetro (potencia) y RMSE obtenido.
   8. Finalmente comprima en una carpeta todos los archivos necesarios para corrección de tal forma que el nombre del archivo comprimido sea Su\_Apellido\_Su\_Nombre y envíemelo por la Mensajería del Aula Virtual.

G) el modelo utilizado fue de regresión lineal con un polinomio de 3er grado, donde la ecuación de la función será: y=w0 + w1\*x + w2\*x^2+ w3\*x^3, reemplazando los valores de los w por sus originales quedaría: **y = 10.04 + (-4.90) \* x + (-0.003) \* x^2 + 0.099 \* x^3**. Nuestro RMSE será de 2.03