**Nombre:** **Armando** **Apellidos: Olivares**

Se esperan respuestas breves que en ningún caso deberían sobrepasar media cara

1. Aunque tenemos dos features muy correlacionadas, decidimos construir un modelo de regresión lineal. ¿Qué peligro corremos? ¿Qué podemos hacer para mitigar este peligro?

**Existe el riesgo de colinealidad con lo cual podemos caer en el diseño de un modelo mal ajustado, para mitigar ese peligro debemos seleccionar una de las 2 features en cuestión.**

1. Estoy construyendo un modelo de regresión logística y me preocupa que esté haciendo overfit. ¿Cómo podría verlo en la función de coste?

**Nos damos cuenta que nuestro modelo está haciendo overfit cuando la función de coste se acerca a cero en los datos de entrenamiento, mientras se va incrementando en los datos de validación.**

1. Vamos a aplicar un modelo de 3 vecinos próximos en un problema de clasificación en 3 clases. Enumerar las posibles votaciones (p.e., 3-0-0 indica 3 votos para la clase 1 y 0 para el resto) e indicar los posibles valores que tomarán las probabilidades a posteriori de cada clase.

|  |  |  |  |
| --- | --- | --- | --- |
| **Posibles Valores** | **Probabilidad Clase1** | **Probabilidad Clase1** | **Probabilidad Clase1** |
| 0-0-3 | 0 | 0 | 1 |
| 0-3-0 | 0 | 1 | 0 |
| 3-0-0 | 1 | 0 | 0 |
| 0-2-1 | 0 | 0.66 | 0.33 |
| 2-0-1 | 0.66 | 0 | 0.33 |
| 2-1-0 | 0.66 | 0.33 | 0 |
| 1-1-1 | 0.33 | 0.33 | 0.33 |
| 0-1-2 | 0 | 0.33 | 0.66 |
| 1-2-0 | 0.33 | 0.66 | 0 |
| 1-0-2 | 0.33 | 0 | 0.66 |

1. ¿Qué es la maldición de la dimensionalidad? ¿Qué modelos de los estudiados en el curso se van a ver más o menos afectados por ella?

**The curse of dimensionality se presenta cuando tenemos un dataset con muchos features lo cual es un problema ya que esto significa un aumento en el espacio de búsqueda en la función de optimización, en este sentido al incrementarse las dimensiones la cantidad de datos necesarios para que el estimador funcione correctamente debe aumentarse exponencialmente.**

**De los modelos estudiados el que se verá más afectado son las redes neuronales, mientras las SVM debido a la regularización se verán menos afectadas.**

1. Tengo un problema de regresión con muchas features por lo que creo que un modelo SVR **lineal** me puede dar buenos resultados. ¿Qué hiperparámetros debo optimizar? Me da la impresión de que el modelo va a tener la pega de ser homogéneo. ¿Qué puedo hacer para mitigar esto y tener finalmente un buen modelo?

**Debemos optimizar el C, y el epsilon para lograr una mejor regularización y un modelo menos complejo, y con respecto a la homogeneidad del modelo debemos utilizar el algoritmo de sequential minimal optimization (SMO), y en sklearn se debe activar la opción de dual= True.**

1. ¿Qué cantidad es lineal en un modelo de regresión logística? ¿Qué cantidad se obtiene como salida del modelo cuando se aplica a un patrón x? ¿Cómo podría usar dicha cantidad para asignar x a una de las dos clases de un tal problema?

**La parte lineal es el logaritmo de los odds:  log(P(1|x)/P(0|x)) = wx +b,**

**Para predecir podemos utilizar el resultado de los odds P(1|x)/P(0|x), ya que el threshold que maximiza el accuracy es 0.5.**

1. ¿Qué es un problema de clasificación desequilibrado? Si voy a optimizar los hiperparámetros del mismo por validación cruzada, ¿qué estrategia de construcción de folds debería seguir?

**Un problema de clasificación desequilibrado se nos presenta cuando en nuestros datos hay una presencia muy desproporcionada de una de las clases a predecir.**

**Para optimizar los hiperparametros en un problema desequlibrado, debería seguir una estrategia de selección de folds estratificado es decir los folds estarán seleccionados para tener la misma cantidad de muestras de cada clase, por otro lado se puede hacer la selección de hiperparámetros mediante optimización bayesiana.**

1. Tengo un problema de clasificación muy desequilibrado y he aprendido suficiente Python como para poder programar mi propia función de scoring para usarla en GridSearchCV. ¿Qué combinación de accuracy o recall o precisión podría usar?

**Para un problema muy desequilibrado se debería usar o programar una función que calcule la media armónica entre la precisión y el recall, esta función será nuestra función objetivo.**

1. ¿Cuántos pesos y sesgos va a tener un PMC de regresión sobre patrones de dimensión 40 y una red de dos capas ocultas (100, 10)? ¿Qué dimensión tendrá el gradiente de la función de error? ¿Y el Hessiano?

**Pesos = 40\*100+ 100\*10 + 10\*1**

**Sesgos = 100 + 10+1**

**Gradiente = 40\*100+ 100\*10 + 10\*1**

**Hessiano = (101\*40)^2 + (101\*10)^2 +(11\*1)^2**

1. Si en un problema dado hacer el fit de un MLP de regresión tarda 10 segundos, ¿qué tiempo de ejecución debo esperar si mantengo la arquitectura MLP pero triplico la muestra? ¿Qué tiempo debería esperar si se tratara de un modelo SVR?

**Para un MLP tardara aproximadamente 30 segundos y para el SVR tardara 90 segundos.**