Notas de defensa

**DIAP 3**

la ciencia de hacer que las computadoras actúen sin estar explícitamente programadas.

Fases:

Train: es la parte en la que nuestro modelo o algoritmo aprende

Test: es la parte en la que se evalua nuestro algoritmo

**DIAP 5**

Ventajas: al tener la información de las clases de cada dato podemos aprovecharla para desarrollar modelo muy precisios.

Desventajas: datos escasos y caros.

Ejemplo: nn, knn, svd, Regresiones

**DIAP 6**

Ejemplos: kmeans, árbol de decisión, random forest.

Campo más usado: Sistemas de recomendación.

Pros: los algoritmos suelen ser más “faciles”, it is easier work with them.

Desventajas: a años luz

**DIAP 7**

Modelo matemático usado para aproximar la relación de dependencia entre una variable dependiente Y, las variables independientes Xi

**DIAP 8**

Realizamos los cálculos del max/min para obtener las incógnitas de la hipótesis.

Esto se puede extrapolar a multivariable y a polinómica.

**DIAP 9**

esa es la solución que anula el gradiente y que el Hessiano es def pos, luego por convexidad la solución es el único mínimo global.

requiere O(N^3) operaciones por lo que se prefiere gradient descent, que requiere O(N^2) operaciones, y que hay otros algoritmos de aproximación incluso mejores que pueden llegar a O(N) operaciones y se usan dependiendo del problema.

**DIAP 10**

El intervalo que "descendemos" por el gradiente está condicionado por la variable learning rate.

Con un alto learning rate descenderemos más rápido, arriesgándonos a ignorar el mínimo de la función.

Con un learning rate muy bajo tendremos que recalcular el gradiente numerosas veces, probablemente sin hallar un resultado óptimo en un tiempo aceptable.

**DIAP 11**

Cuando nuestra nube de puntos no se ajusta a una línea recta, la regresión lineal no nos sirve para definir una hipótesis que se ajuste a nuestro modelo de datos.

Utilizamos el coeficiente de correlación lineal mide hasta qué punto es bueno el modelo lineal para nuestra nube de puntos. Cuanto más cercano sea este a 1 o -1, mejor será la aproximación, cuanto más próximo a 0 será peor.

En tal caso recurrimos a la regresión polinomial. Se utiliza cada modelo cuando el conjunto de datos tiende a ajustarse por una curva respectivamente polinomial.

Eso nos lleva al problema de determinar el número de parámetros del modelo. Necesitamos saber cómo aplicar regresión polinomial, en qué grado. ¿Cómo lo determinamos?: Spearman y Pearson.

**DIAP 12**

(Variance o sobreajuste). Se dice que nuestro modelo está sobreajustado o tiene problema de variance cuando nuestra hipótesis se ajusta demasiado a los datos de la muestra, siendo solo capaz de predecir correctamente dichos datos, y no nuevos. Es decir, cuando es incapaz de generalizar. También se conoce como overfitting.

(Bias o sobresimplificación). Se dice que nuestro modelo está sobresimplificado o tiene problema de bias cuando nuestra hipótesis no se ajusta correctamente a los datos. También se conoce como underfitting

**DIAP 13:** Se utiliza en problemas de clasificación binarea, donde la regresión lineal no sirve

**DIAP 14:** La función sigmoide tiene un dominio entre 0 y 1, lo que se interpreta como probabilidad

Para aplicar la regresión log requerimos que los errores muestrales sean normales y homocedásticos (misma varianza poblacional).

**DIAP 15**

Redes neuronales, qué son. Son un modelo (o estructura) inspirado en el comportamiento observado en el cerebro humano​. Consiste en un conjunto de unidades, llamadas neuronas artificiales, conectadas entre sí para transmitirse señales. La información de entrada atraviesa la red neuronal (donde se somete a diversas operaciones) produciendo unos valores de salida.

**DIAP 16**

Neurona, basada en neurona humana. Entradas -> cálculos -> salida. Unidad mínima de la red neuronal, realiza los cálculos e interacciona con otras neuronas (o no) para colaborar con el objetivo de la red.

**DIAP 17**

Entradas: datos del problema.

Pesos: valor de “relevancia que se le da a cada entrada”

Bias: actúa como sesgo ya que controla qué tan predispuesta está la neurona a disparar una determinada salida independiente de los pesos. Un sesgo alto hace que la neurona requiera una entrada más alta para generar una salida de alta.

Función de activación: Cálculo que se realiza con la salida de la neurona en función de la naturaleza del problema.

**DIAP 18**

Arquitectura: entrada, oculta/s y salida

es el algoritmo de aprendizaje más utilizado. Su objetivo es reducir la función de error, es decir, lo que falla la red. “La red actúa, se equivoca, tiene en cuenta si se ha pasado se ha quedado corta y para la siguiente vez, lo hará mejor”

**DIAP 19**

Red neuronal, como aprenden: error, forward y backward

Forward: la red avanza, computa los pesos con las entradas y bias y expone su respuesta. Los pesos aquí no cambian Lo que importa para que la red funcione bien son los pesos.

Backprop: se calcula el error, y se propaga hacia atrás, modificando levemente o no (mencionar lr) los valores definidos en el forward. Poner las 3 fórmulas de la memoria FALTA

**DIAP23**

Son las redes utilizadas en el campo de la visión artificial, están compuestas por distintas capas: convolución+relu,pooling y nn normal con sus capitas de entrada, oculta y salida.

Se utilizan porque trabajan con menos información, no hay un peso por entrada en cada neurona

**DIAP24**

Son la equivalencia a los pesos de las otras nn

Hablar de la operación de convolución (no es producto matricial)

**DIAP25**

Stride: el intervalo de pixeles que se arrastra el filtro

Pading: márgenes para hacer el filtrado equitativo si se desea

**DIAP26:**

Para condensar aún más la información.