

Taller

ceedcv
CENTRE ESPECÍFIC
D'EDUCACIÓ A DISTÀNCIA DE
LA COMUNITAT VALENCIANA

Introducción al aprendizaje automático



¿Que es el aprendizaje automático?

- Técnicas para que las computadoras puedan aprender, sin haber sido programadas explícitamente.
 - Estas técnicas detectan patrones en los datos y ajustan las acciones del programa en consecuencia.
-

¿Que vamos a aprender?

- Aprenderemos
 - Ejemplos de aprendizaje supervisado.
 - Ejemplos de aprendizaje no supervisado.
 - Usaremos Python 3 + Scikit-learn
-

Preparando el entorno

□ Python 3

■ Linux

- `sudo apt install python3 python3-pip`

■ Windows

- <https://www.anaconda.com/distribution/>

□ Scikit-learn

- `pip install -U scikit-learn`
 - `conda install scikit-learn`
-

Conceptos previos (1)

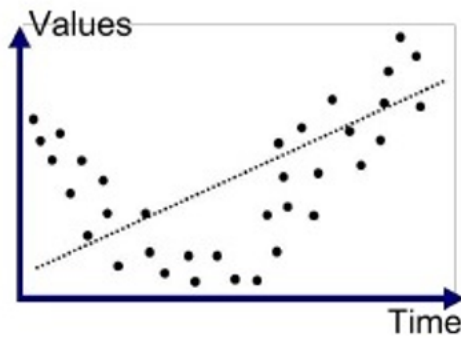
- **Corpus:** Conjunto de datos a aprender.
 - **Entrenamiento:** usando la técnica o tecnología que sea, se utilizan los datos del corpus para generar un modelo.
-

Conceptos previos (2)

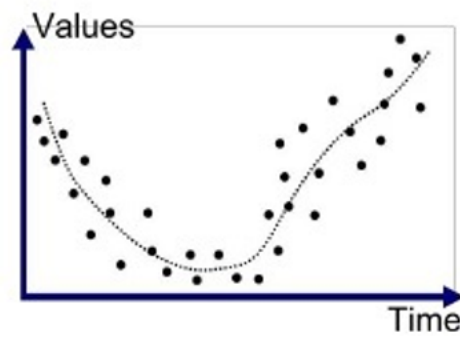
- **Modelo:** Estructura que representa el conocimiento aprendido (el resultado del entrenamiento). Por ejemplo una red neuronal. Depende de la técnica usada.
 - El modelo se utiliza para clasificar o reconocer datos nuevos o hacer una predicción
 - Por ejemplo, se entrena el modelo con un corpus de resultados de partidos para luego intentar predecir partidos nuevos.
-

Conceptos previos (3)

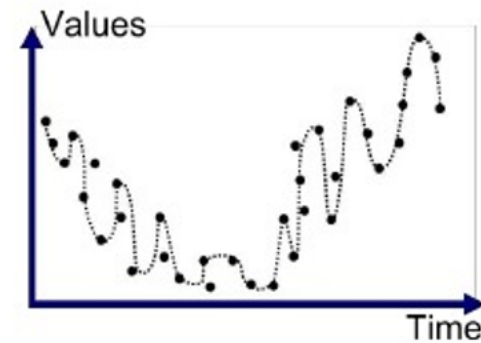
□ Entrenamiento



Underfitted



Good Fit/Robust



Overfitted

Aprendizaje supervisado (1)

- ¿Que es el aprendizaje supervisado?
 - Aprender a partir de datos de entrenamiento previos
 - Datos de entrenamiento formados
 - Vector de características
 - Etiqueta o resultado esperado
-

Aprendizaje supervisado (2)

- Problemas de regresión
 - A partir de una muestra se predice el valor de una o varias variables.
 - Problemas de clasificación
 - Se predice la clase de la muestra
-

Aprendizaje supervisado (3)

□ **Ejemplo Regresión 1:** Número de coches a una hora del día

- 08:00 → 1000
- 09:00 → 1500
- 10:00 → 700
- 14:00 → 170
- ...

□ **Ejemplo Regresión 2:** Horas estudiadas, nota obtenida

- 30 horas → 6
 - 50 horas → 10
 - 35 horas → 4
 - 20 horas → 5
 - 38 horas → 7.5
 - ...
-

Aprendizaje supervisado (4)

- **Ejemplo Clasificación 1:** Temperatura Frio/Calor
 - 25° → Calor
 - 22° → Calor
 - 10° → Frio
 - 0° → Frio
 - ...
 - **Ejemplo Clasificación 2:** Velocidad coche en relación a velocidad máxima de vía
 - 20, 50 → Lento
 - 50, 120 → Lento
 - 40, 50 → Adecuado
 - 50, 50 → Adecuado
 - 80, 60 → Peligro
 - 100, 120 → Adecuado
 - 140, 120 → Peligro
 - ...
-

Caso Estudio Regresión

- Caso de estudio Regresión
 - Algoritmo: Regresión lineal mínimos cuadrados
 - Corpus simple precio pisos en relación a metros cuadrados.
 - Ver "***RegresionPisos.py***"
 - Saber mas

<http://www.aprendemachinelearning.com/regresion-lineal-en-espanol-con-python/#more-5722>

Caso Estudio Clasificación (1)

- Caso de estudio Clasificación (1)
 - Juego que juega aleatoriamente a 3 en raya.
 - Guardamos todos los movimientos (de que precedía a como queda) y al finalizar la partida etiquetamos si ese movimiento llevó a la victoria o no.
 - Ver "***GeneraCorpus3EnRaya.py***"
-

Caso Estudio Clasificación (2)

- Caso de estudio Clasificación (2)
 - Entrenamos con el corpus generado y guardamos entrenamiento para ser usado en otros ficheros
 - Arbol de decision
 - Ver fichero **"EntrenaModeloTree.py"**
 - *Maquinas de vectores de soporte*
 - Ver fichero **"EntrenaModeloSVM.py"**
 - Usamos el entrenamiento para jugar y probar efectividad
 - Con score **"3EnRayaJuego.py"**
 - Sin score **"3EnRayaJuegoSinScore.py"**
-

Caso Estudio Clasificación (3)

- Caso de estudio Clasificación (3)
 - Uso de puntuación de la calidad de la predicción
 - Algoritmo usado: Árbol de decisión
 - Uso de clasificador y seleccionando aleatorio entre candidatos
 - Algoritmo usado: SVM
-

Caso Estudio Clasificación (4)

- Caso de estudio Clasificación (4)
 - Saber más
 - <http://www.aprendemachinelearning.com/principales-algoritmos-usados-en-machine-learning/>
-

Aprendizaje no supervisado (1)

- ¿Que es el aprendizaje no supervisado?
 - No hay entrenamiento previo.
 - Busca agrupaciones
 - Eso no implica que dichas agrupaciones tengan sentido o utilidad
 - El modelo se va ajustando según las observaciones.
-

Aprendizaje no supervisado (2)

- Caso de estudio (1)
 - Algoritmo K-means
 - Datos: número de goles y de asistencias en una temporada
 - Ajustamos el modelo en base a observaciones
 - Fichero "NoSupervisadoKMeans.py"
-

Aprendizaje no supervisado (3)

- Caso de estudio (2)
 - Observaciones
 - Grupos solo goleadores
 - Grupos solo asistentes
 - Grupos goleadores y asistentes
 - Grupos ni goleadores ni asistentes
-

Para saber más

□ Enlaces interesantes

- Mi curso favorito!!
- <https://es.coursera.org/learn/machine-learning>
- Otros enlaces interesantes
- <https://medium.com/datos-y-ciencia/introduccion-al-machine-learning-una-gu%C3%ADa-desde-cero-b696a2ead359>
- <https://relopezbriega.github.io/blog/2015/10/10/machine-learning-con-python/>
- <https://eu.udacity.com/course/intro-to-artificial-intelligence--cs271>
- <https://www.coursera.org/lecture/machine-learning-with-python/introduction-to-machine-learning-zSm8k>