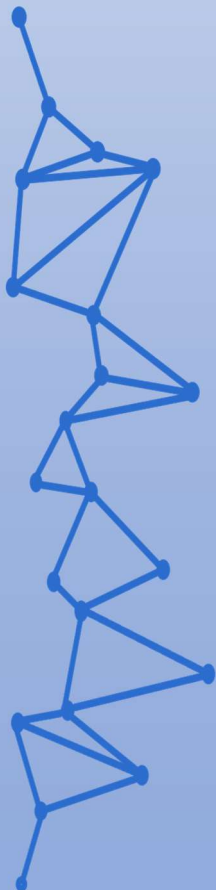




Curso de Especialización de Inteligencia Artificial y Big Data (IABD)



Sistemas de aprendizaje Automático

UD01. Justificación y antecedentes del Aprendizaje
Automático.
Resumen.

JUAN ANTONIO GARCIA MUELAS

Los **sistemas expertos**, marcaron la clara separación entre dos tipos de inteligencia artificial. La **inteligencia artificial fuerte o general**, similar a la mente de los seres humanos (el siguiente paso o punto al que dirigirse en un futuro), y la **inteligencia artificial débil o estrecha**, capaz de realizar con gran precisión una tarea concreta, dentro de un campo de especialización. Aquí es donde ha estado toda la evolución en el ámbito del aprendizaje automático profundo.

El campo de la Inteligencia Artificial no se centra en emular las capacidades o comportamientos humanos. La clave es la evolución y avance de la automatización.

El **concepto de inteligencia artificial débil** se refiere a agentes y modelos inteligentes que están especializados en un tipo de tarea, dentro de un campo de especialización concreto (juego ajedrez, reconocimiento imagen...)

Ventajas de la IA Débil:

- ✓ Actualmente es relativamente sencilla de conseguir y desarrollar.
- ✓ Se controla más fácilmente y no presenta comportamientos demasiado inesperados.
- ✓ Permite arquitecturas modulares, con las que combinar distintos talentos o capacidades según las necesidades del entorno en el que se despliega.

Inconvenientes de la IA Débil:

- ✓ Es un sistema "ciego", que no cuenta con todo el contexto de la situación, por lo que puede malinterpretar entradas inusuales.
- ✓ Es muy limitada; solo sirve para aplicaciones muy concretas.
- ✓ Suelen ser sistemas que no ofrecen una experiencia de usuario tan cercana o personalizada como la de un ser humano.

Ejemplos: Alexa, Siri, Google Assistant (asistentes conversacionales, ejemplo de IA estrecha). **IBM Watson** (sistema cognitivo que da información precisa en lenguaje natural, infiriendo conclusiones). **AlphaGo** (desarrollado por DeepMind como red neuronal). También están los coches autónomos o los algoritmos de recomendación.

Los algoritmos de recomendación simulan muy bien una característica humana, pero sin embargo son ejemplo de inteligencia artificial débil.

El **concepto de la Inteligencia Artificial Fuerte o General** parte de la posibilidad de que una máquina pueda abstraer hasta el nivel de interaccionar con seres humanos y con el resto de los objetos y seres del mundo físico en todo tipo de contexto y con objetivos de todo tipo. En definitiva, emular al ser humano.

Hay muchos contraargumentos, pues se entiende que la mente humana no es análoga a un ordenador y su hardware.

A partir de este punto, hay dos razones que hacen que tenga una mayor consideración la IA débil:

- ✓ Desde la vertiente lógica, se pueden estudiar los algoritmos y las funciones matemáticas con mejor claridad.
- ✓ Desde una perspectiva clásica de la filosofía continental, la debilidad intrínseca de la IA se podría argumentar a partir de la siguiente cuestión: nosotros los humanos fabricamos las herramientas y sistemas y, por tanto, incluye dentro nuestra propia fragilidad, la posibilidad del fallo.

Lo que si se está explorando en este momento es la construcción de modelos basados en redes neuronales capaces de ejecutar tareas en diferentes contextos.

La **conciencia artificial** es construir sistemas que puedan crear representaciones sobre si mismos. El crear IA fuerte no implica el surgimiento de conciencia artificial o emergente, pero es la posibilidad más pausable.

Gato de DeepMind es un modelo capaz, actualmente, de llevar a cabo varias tareas en contextos diferentes.

Es una inteligencia artificial generalista que puede llevar a cabo hasta 604 tareas diferentes.

La evolución de la inteligencia artificial se ha basado en tratar de ser capaz de actuar como los humanos y predecir el comportamiento más probable o deseable de un sistema en el planteamiento más actual. Para esto, las técnicas más prometedoras son las que utilizan modelos probabilísticos.

Las técnicas de aprendizaje automático miran el problema desde fuera y configuran sus propias reglas para emular el comportamiento del sistema. Consiguiendo una mayor tasa de acierto.

El Machine Learning (acuñado desde IBM en 1959) se basa en las estadísticas y el desarrollo de la informática.

Según Andrew Ng, profesor de inteligencia artificial en Stanford: El desarrollo del machine learning se ha basado en la disponibilidad de datos y la escala computacional.

Wasserman diferencia la estadística del machine learning en que la estadística modeliza problemas de baja dimensión mientras que el machine learning hace predicciones de alta dimensión.

Cada paso en un proyecto de Machine Learning requiere el uso de un método estadístico.

<https://machinelearningmastery.com/statistical-methods-in-an-applied-machine-learning-project/>

MODELOS BAYESIANOS.

El teorema de Bayes proporciona una manera de **calcular la probabilidad de una hipótesis basada en su probabilidad previa**, las probabilidades de observar diversos datos dada la hipótesis y los datos observados en sí maximizando la probabilidad de que un modelo se ajuste al conjunto de datos conocido como MAP (máximo a posteriori).

Es un sistema de **cálculo de probabilidades hecho al revés**. O sea, **conociendo las consecuencias** que se producen podemos calcular sus orígenes (siempre en porcentajes) y la probabilidad de que se hayan producido en uno u otro grupo.

Diagrama de la fórmula de Bayes con anotaciones explicativas:

$$P(A|B) = \frac{P(B|A) P(A)}{P(B)}$$

Las anotaciones son:

- Arriba a la izquierda: "THE PROBABILITY OF 'B' BEING TRUE GIVEN THAT 'A' IS TRUE" (con una flecha hacia abajo a $P(B|A)$)
- Arriba a la derecha: "THE PROBABILITY OF 'A' BEING TRUE" (con una flecha hacia abajo a $P(A)$)
- Abajo a la izquierda: "THE PROBABILITY OF 'A' BEING TRUE GIVEN THAT 'B' IS TRUE" (con una flecha hacia arriba a $P(A|B)$)
- Abajo a la derecha: "THE PROBABILITY OF 'B' BEING TRUE" (con una flecha hacia arriba a $P(B)$)

$$P(A|B) = \frac{P(B|A) P(A)}{P(B)}$$

B es el suceso que conocemos, A el conjunto de posibles causas, excluyentes entre sí, que pueden producirlo y, por tanto, $P(A|B)$ son las probabilidades a posteriori, $P(A)$ las probabilidades a priori y $P(B|A)$ la probabilidad de que se de B en cada hipótesis de A. Es un modo de variar nuestras expectativas

según la experiencia concreta que tenemos sobre algo.

Un algoritmo o modelo de aprendizaje automático es una forma específica de pensar sobre las relaciones estructuradas en los datos.

El Teorema de Bayes proporciona un modelo probabilístico para describir la relación entre los datos (D) y una hipótesis (H). Si lo representamos con una fórmula simplificada:

$$P(h | D) = P(D | h) * P(h) / P(D)$$

donde la probabilidad de una hipótesis h dados unos datos D, puede calcularse como la probabilidad de observar los datos dada la hipótesis, multiplicada por la probabilidad de que la hipótesis sea verdadera independientemente de los datos, y dividido todo por la probabilidad de observar esos datos independientemente de la hipótesis.

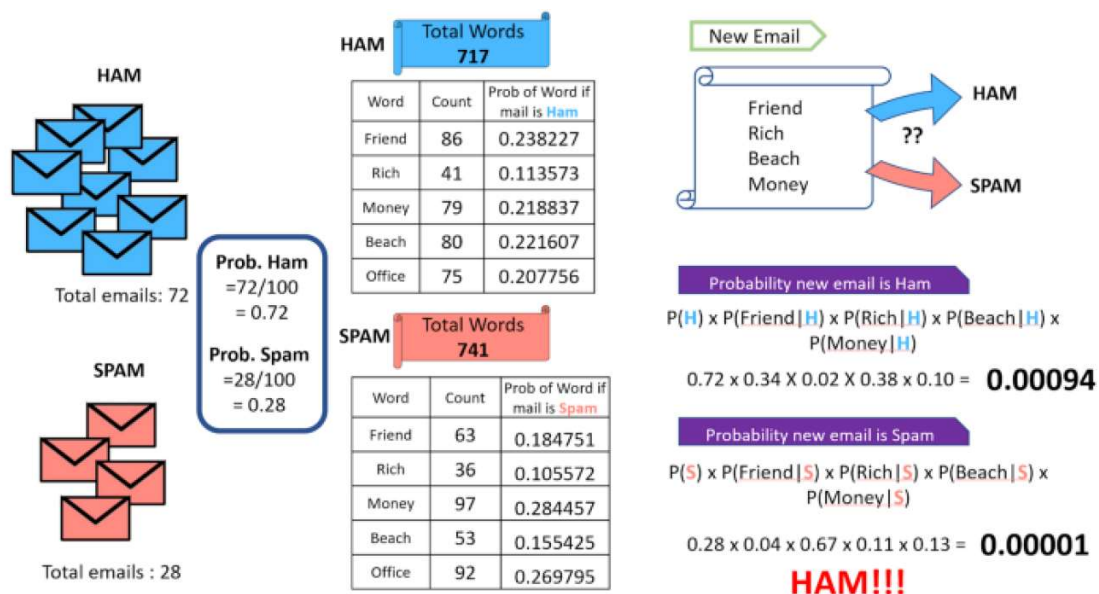
Los modelos de Naive Bayes son un tipo de algoritmos de aprendizaje automático basados en el teorema de Bayes. En ellos, se asume que las variables de entrada son independientes entre sí.

Sus principales **ventajas** son:

- ✓ Es una manera fácil y rápida de predecir clases, para problemas de clasificación binarios y multiclase.
- ✓ En los casos en que sea apropiada una presunción de independencia, el algoritmo se comporta mejor que otros modelos de clasificación, incluso con menos datos de entrenamiento.
- ✓ El desacoplamiento de las distribuciones de características condicionales de clase significa que cada distribución puede ser estimada independientemente como si tuviera una sola dimensión. Esto ayuda con problemas derivados de la dimensionalidad y mejora el rendimiento.

Y sus **desventajas** son:

- ✓ Aunque son unos clasificadores bastante buenos, los algoritmos Naive Bayes son conocidos por ser pobres estimadores. Por ello, no se deben tomar muy en serio las probabilidades que se obtienen.
- ✓ La presunción de independencia Naive muy probablemente no reflejará cómo son los datos en el mundo real.
- ✓ Cuando el conjunto de datos de prueba tiene una característica que no ha sido observada en el conjunto de entrenamiento, el modelo le asignará una probabilidad de cero y será inútil realizar predicciones.



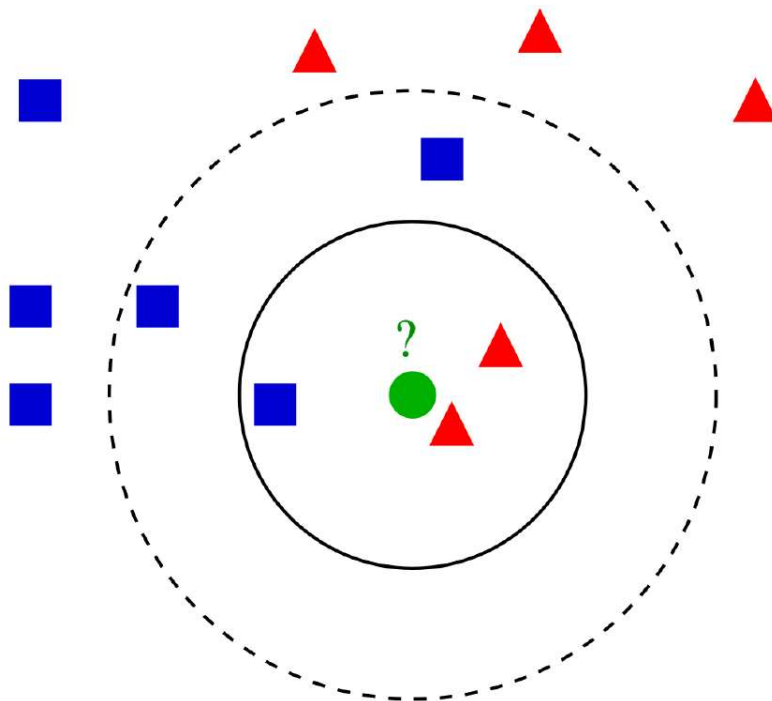
KNN

El método "**K Nearest Neighbors**" o los "**K vecinos más cercanos**" es un modelo de aprendizaje automático supervisado muy sencillo pero efectivo que se suele usar en problemas de clasificación.

Dada una distribución de casos situados en un mapa de coordenadas, clasificar un nuevo caso mirando los casos que tiene alrededor. Si la mayoría son de una clase A, el nuevo caso se clasificará como de esa misma clase A.

Este algoritmo no "aprende", no se entrena ni se fijan valores de parámetros internos. Simplemente se utiliza el conjunto de datos como "base de conocimiento" y se hacen las predicciones sobre dicha base.

Se fija el parámetro K a un cierto valor y se seleccionan los "k" elementos más cercanos (con menor distancia, según la función que se use). Por último, se realiza una "votación de mayoría" entre los k puntos: los de una clase/etiqueta que sean mayoría, decidirán su clasificación final.



Antti Ajanki ([CC BY-SA](#))

Es importante ajustar bien el valor de K y no elegir valores pares para evitar "empates".