

Introducción a la Inteligencia Artificial

Índice

| | |
|--|----|
| La inteligencia artificial..... | 2 |
| Introducción General..... | 2 |
| Introducción al módulo..... | 3 |
| Evolución..... | 4 |
| Orígenes..... | 4 |
| Desarrollo..... | 5 |
| Presente y futuro..... | 6 |
| Aplicaciones..... | 8 |
| Aplicaciones en logística, transporte y agricultura..... | 8 |
| Aplicaciones en Salud, Biotecnología, Ventas al detalle y Educación..... | 9 |
| Aplicaciones de entretenimiento y la industria aeroespacial..... | 10 |
| Aplicaciones en Servicios Financieros, Manufactura, Cadena de Suministro, Socioeconomía y Mercadotecnia..... | 12 |
| Modelos..... | 13 |
| Aprendizaje Automático..... | 13 |
| Aprendizaje profundo (deep learning)..... | 16 |
| Robótica..... | 18 |
| Sistemas Expertos..... | 22 |
| Componentes de un Sistema Experto..... | 24 |
| Procesamiento de lenguaje natural (PLN)..... | 25 |
| Introducción al procesamiento de lenguaje natural..... | 25 |
| Introducción a la sección..... | 25 |
| Introducción al procesamiento de lenguaje natural..... | 26 |
| Técnicas de procesamiento de lenguaje natural..... | 27 |
| Análisis del lenguaje..... | 27 |
| Técnicas y modelos de análisis del lenguaje..... | 29 |
| Ejemplos de modelo probabilístico..... | 31 |
| El aprendizaje automático en el procesamiento del lenguaje natural (Intro SVN)..... | 35 |
| Aplicaciones de las técnicas de procesamiento de lenguaje natural..... | 39 |
| Aplicaciones de la IA..... | 40 |
| Introducción al análisis de sentimiento en Twitter..... | 40 |
| Instalación de Rapid Miner..... | 42 |
| Tutorial de RapidMiner..... | 42 |
| Árboles de clasificación..... | 42 |
| Introducción a los árboles de clasificación..... | 42 |
| Introducción a los árboles de decisión – clasificación..... | 43 |
| Los árboles de decisión – clasificación..... | 44 |
| La técnica de árboles de clasificación..... | 45 |
| Introducción a conceptos de árboles de decisión y clasificación..... | 45 |
| Estructura de árboles de clasificación..... | 47 |
| Criterios para la construcción de un árbol..... | 48 |

| | |
|--|----|
| Entrenamiento y prueba de los árboles..... | 49 |
| Construcción de modelos de árboles de clasificación..... | 49 |
| Evaluación del modelo (Matriz de confusión)..... | 50 |
| Métricas de la Matriz de Confusión..... | 51 |
| Aplicaciones de árboles de decisión..... | 52 |
| Clasificador Bayesiano..... | 52 |
| Introducción al clasificador bayesiano..... | 52 |
| Introducción al módulo..... | 52 |
| Introducción a Naive Bayes..... | 53 |
| El algoritmo de Naive Bayes..... | 53 |
| Técnica del clasificador bayesiano..... | 54 |
| Características del algoritmo Naive Bayes..... | 54 |
| Teorema de Bayes..... | 55 |
| Tablas de frecuencias y probabilidades..... | 55 |
| Probabilidad posterior de la clase SI..... | 57 |
| Probabilidad posterior de la clase NO..... | 58 |
| Resultado de la predicción..... | 59 |

La inteligencia artificial

Introducción General

¿Será posible construir una máquina o robot tan inteligente como el ser humano? El cerebro humano es tan fascinante como complejo, por lo que, parte de su funcionamiento sigue siendo un misterio para los investigadores.

Los seres humanos somos inteligentes y eso nos hace diferentes a los demás, a las máquinas, a los robots y a las cosas.

La Inteligencia Artificial permite crear dispositivos o algoritmos que sean capaces de comportarse como un ser humano, con inteligencia propia, lo cual implica que deben tener las capacidades para percibir el entorno, realizar tareas y tomar decisiones en forma exitosa.

La Inteligencia artificial está tomando cada vez más importancia en nuestra vida diaria y laboral, ya que, permite la comunicación con nuestros dispositivos como si ellos también fueran personas.

Ejemplos:

- Nuestro teléfono es un asistente personal gracias a que utiliza procesamiento de lenguaje natural
- las redes sociales nos sugieren el etiquetado de fotos porque tienen técnicas de reconocimiento de imagen
- cuando compramos por internet nos ofrecen recomendaciones de productos basadas en otras compras que ya hemos realizado anteriormente
- cuando nos trasladamos a un lugar podemos recibir información optimizada del tráfico y navegación en tiempo real.
- en nuestras tareas laborales, mediante el algoritmo de inteligencia artificial, minería de datos y big data, podemos predecir la permanencia o ausencia de nuestros clientes

- analizar nuestra cuenta de twitter para conocer la satisfacción de los clientes con nuestra empresa y nuestros productos.

Estamos en la era de la Inteligencia Artificial, rodeados de máquinas, robots, algoritmos y dispositivos inteligentes. Usamos un teléfono inteligente y accedemos a las redes sociales que usan algoritmos de inteligencia artificial.

Por esto, formar parte de los desarrolladores de estas tecnologías, representa una enorme oportunidad.

En este curso aprenderás los fundamentos teóricos de la inteligencia artificial con sus principales ramas como el **aprendizaje automático, el aprendizaje profundo, la robótica, el procesamiento del lenguaje y lo sistemas expertos.**

Además, aplicarás los conceptos en la creación de modelos como el de **procesamiento de lenguaje natural, árboles de clasificación y el clasificador bayesiano.**

Adicionalmente, aprenderás a utilizar la plataforma de ciencia de datos llamada RapidMiner.

Introducción al módulo

Vamos a comenzar esta sección de Introducción a la Inteligencia Artificial.

¿Recuerdas las preguntas centrales?: **¿Será posible construir una máquina o robot tan inteligente como el ser humano? ¿Existen actualmente algoritmos inteligentes que sean capaces de tomar decisiones por las personas?**

A lo largo de esta sección conocerás más sobre esto. Aprenderás los fundamentos teóricos de la inteligencia artificial, conocerás sus orígenes y evolución, así como las distintas aplicaciones.

Además analizarás sus principales ramas o modelos, como el aprendizaje automático, aprendizaje profundo, robótica, procesamiento del lenguaje y sistemas expertos.

La inteligencia artificial se encuentra en un auge, gracias al enorme desarrollo tecnológico que permite mayor rapidez en el procesamiento y mayor capacidad de almacenamiento, por esto, actualmente podemos contar con una enorme cantidad de datos, que se le ha llamado el **Big Data**, gracias a esto, la inteligencia artificial ha podido desarrollarse con mayor rapidez.

Otra gran tecnología que ha impulsado la inteligencia artificial es la **minería de datos**, que está formada por el conjunto de técnicas y algoritmos que permiten el análisis del Big Data, estos algoritmos son de inteligencia artificial.

En esta sección podrás conocer las aplicaciones en diversas áreas como la Logística, transporte, agricultura, salud, ventas al detalle y educación, entre otras.

Además, conocerás los modelos o ramas de inteligencia artificial, como el **aprendizaje automático, aprendizaje profundo, robótica y sistemas expertos**

Evolución

Orígenes

Los orígenes de la inteligencia artificial pueden ser descritos desde diferentes áreas o nociones, que se han observado de ella a lo largo de todo su desarrollo durante estos años.

Se puede considerar como el inicio formal en el año 1950, porque desde entonces se consideró la intención de dotar de recursos a las máquinas, para que fueran capaces de resolver problemas sin el apoyo del ser humano, esto es, de manera autónoma.

Aunque la noción de máquina está presente desde muchos años atrás, fue **Alan Turing** a finales de la II Guerra Mundial, en el año 1950, quien sentó las primeras bases de la inteligencia artificial, en particular estudió en forma científica a las máquinas inteligentes.

Desde los años 30's, Alan Turing había estado trabajando en la teoría de computabilidad y creó una de las primeras computadoras electromecánicas.

Años después, junto con Alan Turing, diversos investigadores le confieren existencia formal al área de exploración y estudio de estas materias. Entre ellos podemos mencionar a los siguientes: McCarthy, Minsky, Edmonds, Papert y Carbonell, entre otros.

Línea de tiempo de la evolución de la IA en sus orígenes:

1. **1937** Turing plantea la teoría de la computabilidad.
2. **1940** Turing crea uno de los primeros computadores electromecánicos
3. **1943:** McCulloch y Pitts publican el primer trabajo sobre inteligencia artificial basándose en la fisiología básica, funcionamiento de las neuronas, la lógica proposicional (Russell y Whitehead) y la teoría de la computación (Turing).
4. **1950:** Turing publica su artículo Computing Machinery and Intelligence, donde introduce la prueba de Turing, el aprendizaje autónomo, los algoritmos genéricos y el aprendizaje por refuerzo. Shanon estudia la capacidad simbólica de las computadoras.
5. **1951:** Minsky y Edmonds construyen el primer computador a partir de una red neuronal: el SNARC.
6. **1955:** Simon, Newell y Shaw desarrollan el primer lenguaje orientado a la resolución de problemas: GPS.
7. **1956:** Se organizó un taller en Dartmouth, a través del cual se fundó formalmente el campo de inteligencia artificial, adoptando dicho nombre.
8. **1957:** Se crea el Primer lenguaje de programación: FORTRAN.
9. **1958:** McCarthy definió el lenguaje de programación LISP (que se utiliza hasta la actualidad), también publicó el artículo "Programs with common sense" donde se describe el primer sistema completo de inteligencia artificial.

Newell, Shaw y Simon diseñan el primer programa inteligente basado en procesamiento de información.

10. **1961:** Samuel diseña el 1º programa que juega damas, el cual es capaz de aprender de sus errores.
11. **1962:** McCarthy y Raphael construyen el primer robot móvil: SHAKEY.
12. **1963:** Minsky creó el laboratorio de Inteligencia Artificial en Stanford.
Slagle crea el programa SAINT capaz de resolver problemas de integración de cálculo en forma cerrada.
13. **1964:** Bobrow muestra que las computadoras pueden entender el lenguaje natural, resolviendo problemas de álgebra.
14. **1966:** Se crea el programa ELIZA para el estudio de la comunicación hombre-máquina.
15. **1968:** Minsky y Papert publican “Perceptrones”, texto que muestra los límites de las redes neuronales.
16. **1970:** Carbonell desarrolla SCHOLAR, programa interactivo para la enseñanza.

Desarrollo

Después de las aportaciones de Alan Turing, McCarthy y Minsky, entre otros, a partir del año 1971, la inteligencia artificial se comenzó a desarrollar en el mundo, gracias a diversas aportaciones muy importantes de investigadores de diversos países.

En los años 70's, se crea el primer Robot que comprende el inglés y el primero que tienen visión y se construye el primer vehículo autónomo. En los años 80's, se construye el primer sistema experto y la Inteligencia Artificial se convierte en ciencia.

Después de una etapa difícil en su desarrollo, a partir de los años 90's se logran importantes avances, como el prototipo del robot NAO, los sistemas que detectan emociones e interactúan con niños autistas y el chatbot Eugene.

Esta época de los años 90's se conoce como la edad de oro de la Inteligencia Artificial, ya que las empresas comienzan a invertir en tecnología creando el llamado mundo digital, en el que se mejoró la capacidad de procesamiento y análisis de enormes cantidades de datos.

Una etapa cúlmine de la Inteligencia Artificial se presentó cuando IBM lanzó la computadora Deep Blue, que fue capaz de ganar a Gari Kasparov, campeón del mundo de ajedrez. Este acontecimiento, puso a la Inteligencia Artificial como protagonista del mundo tecnológico.

Línea de tiempo de desarrollo de IA:

1. **1971:** Winograd logra que un robot comprenda frases en inglés para ejecutar una acción. Su programa de comprensión del lenguaje se llama SHRDLU.
2. **1973:** Se construye el 1º robot que utiliza la visión para ubicar y ensamblar bloques.
3. **1974:** 1º sistema experto.
4. **1977:** Schank y Abelson diseñan un programa para la comprensión del lenguaje natural.
5. **1979:** Moravec construye el 1º vehículo autónomo controlado por computador.

6. **1981:** Comienza en Japón el proyecto Quinta generación de computadoras.
7. **1982:** McDermott crea el primer sistema experto comercial: R1. El programa se utiliza para el pedido de nuevos sistemas informáticos. Gray crea la 1ª computadora con procesamientos en paralelo
8. **1986:** Rumelhart y McClelland publican las aplicaciones de los algoritmos de aprendizaje en informática y psicología, estableciendo el conexionismo basado en redes neuronales.
9. **1987:** La inteligencia artificial se convierte en ciencia.
10. **1988:** Pearl publica el texto "*Probabilistic Reasoning in Intelligence System*", que incorpora las probabilidades y la teoría de decisiones en la inteligencia artificial.
11. **1988-1993:** Invierno de la inteligencia artificial, ya que no se cumplen las expectativas y disminuye el financiamiento. Esto producto de la capacidad limitada de las computadoras, lo que imposibilita la solución de los problemas complejos.
12. **1995:** Russell y Norvig retoman el trabajo en agentes totales, sistemas que buscan soluciones a información de entrada continua.
13. **1997:** Computadora Deep Blue vence al ajedrecista Gasparov.
14. **2000:** Se construye el robot ASIMO, que se moviliza en dos pies, da la mano y puede contestar preguntas simples.
15. **2002:** Puchol crea el sistema de reconocimiento digital de huellas digitales.
16. **2005:** Desarrollan el prototipo de NAO, robot programable y autónomo
17. **2009:** Desarrollan sistemas que detectan emociones e interactúan con niños autistas.
18. **2014:** El bot de chat Eugene logra superar el test de Turing.

Presente y futuro

Después de la época de oro de la Inteligencia Artificial, en los últimos años, la inteligencia artificial ha sido potenciada por diversos factores, tales como la transformación digital que representa una nueva transformación de las empresas hacia el desarrollo tecnológico con la finalidad de lograr sus ventajas competitivas y mantenerse como líderes en los mercados cambiantes y altamente competitivos.

Esta transformación digital está generando una enorme cantidad de datos que no es posible analizarlos con las herramientas y algoritmos que tradicionalmente se utilizan con las bases de datos.

Por eso, la **inteligencia artificial**, el **big data** y la **minería de datos**, han permitido el desarrollo de algoritmos sofisticados de inteligencia artificial, que están permitiendo a las empresas tomar decisiones basadas en datos y a los ciudadanos les está ofreciendo la comodidad e inmediatez de la información para sus actividades diarias.

Entre las aplicaciones más representativas e impactantes que actualmente se utilizan, se encuentran los asistentes por voz Siri, Cortana, Alexa, Bixby y muchos otros. Todos ellos utilizan técnicas de procesamiento de lenguaje natural.

Google Search, Google Maps y Netflix utilizan algoritmos de inteligencia artificial. También, en las redes sociales como Facebook, Pinterest y Fotos de Google, los algoritmos de inteligencia artificial permiten el reconocimiento facial, esto es, la visión por computadora.

En las empresas, la robótica, los sistemas expertos y el aprendizaje de máquina, están revolucionando los procesos, permitiendo a las empresas la innovación en sus negocios.

Se están desarrollando herramientas que son plataformas de ciencia de datos que permiten la aplicación de los algoritmos de inteligencia artificial. Algunas plataformas son el RapidMiner, Anaconda, Python.

Asimismo Watson, es la plataforma de IBM que integra el aprendizaje automático para automatizar el ciclo de vida de la IA.

Esquema de áreas de la IA:



Figura 1: Áreas de la inteligencia artificial

Según Hanio Morillo, experta en Inteligencia Artificial de Google, el mundo será IA First, es decir, estará liderado y mediado por la inteligencia artificial, asegura que actualmente ya lo es para Google. Menciona que los gigantes como Google, Microsoft, Facebook, Amazon, Oracle están comprometidos con la inteligencia artificial, porque esta tecnología está cada vez más presente en el día a día de todos, en lo que hacemos, en lo que elegimos; y en las decisiones de grandes empresas, de salud pública, de seguridad o de finanzas.

“La inteligencia artificial está evolucionando, estamos en un momento histórico muy importante: por primera vez, tenemos un gran poder computacional y una grandísima cantidad de datos. Cuando ambos se combinan, el resultado es asombroso porque podemos utilizar esos datos de una forma beneficiosa para el usuario.” Asentó Morillo.

- En el futuro se vislumbra que se continuará con las investigaciones para el desarrollo de una **inteligencia artificial general**, más que especializada, es decir, una inteligencia más apegada al ser humano, lo cual es muy difícil de lograr.
- Asimismo, se seguirán desarrollando aplicaciones relacionadas con la web, los videojuegos y los robots autónomos (vehículos autónomos, robots sociales, robots para la exploración de planetas, etc.).
- También las aplicaciones para el medioambiente y el ahorro energético serán importantes, así como para la economía, la sociología y el arte.
- Se desarrollará con mayor profundidad el **massive data-driven AI**, es decir, la inteligencia artificial con datos masivos, esto es, la explotación de cantidades masivas de datos y para procesarlos con hardware cada vez más rápido con el fin de descubrir relaciones entre ellos, detectar patrones y realizar inferencias y aprendizaje mediante modelos probabilísticos.
- El auto autónomo también está enriquecido con nuevas investigaciones, prototipos y pruebas. Se han realizado diversos experimentos y se han registrado algunos accidentes fatales. Pero las investigaciones continúan y los expertos aseguran que será una tecnología que estará al alcance de nuestras manos en un futuro cercano.

Aplicaciones

Aplicaciones en logística, transporte y agricultura

Introducción

Aplicaciones en logística transporte y agricultura la inteligencia artificial es tan potente que sus aplicaciones son variadas y extensas por mencionar algunos ejemplos los sistemas de inyección de gasolina en nuestros automóviles están diseñados utilizando **algoritmos genéticos** así como las turbinas de los aviones a reacción.

Algoritmo genético: Método adaptativo que pueden usarse para resolver problemas de búsqueda y optimización.

Además en el metro de hong kong durante la noche los ingenieros llevan a cabo trabajos de mantenimiento que han sido programados con un software de planeación la detección automática de transportación fraudulentas de tarjetas de crédito, se realiza mediante algoritmos de aprendizaje automático.

También el enrutamiento de las llamadas de teléfonos móviles se basa en una inteligencia artificial

La detección de hábitos de consumo se basa en el análisis automático de grandes cantidades de datos mediante algoritmos de aprendizaje automático.

Asimismo, la aplicación de la inteligencia artificial está presente en los sistemas de traducción automatizada mediante la cual se utilizan algoritmos para traducir texto o audio de un idioma a otro, además de sustituir simplemente una palabra por otra puede incorporar técnicas estadísticas que aumentan la probabilidad de identificar correctamente las frases expresiones idiomáticas nombres propios y otras anomalías.

Algunas aplicaciones en la logística transporte y agricultura:**Logística y transporte:**

- En ciertas ciudades como mountain view de california es común ver autos automáticos recorriendo en las calles cada día es muy probable que en los próximos años sea normal encontrar estos vehículos circulando en todo el mundo además los trenes autobuses y otros medios de transporte también pueden ser conducidos de forma automática.

Existen 5 niveles de autonomía los autos autónomos más avanzados actualmente están en el nivel 3 lo que significa que el auto puede manejarse al volante y los pedales asimismo como estar atento al entorno pero depende del complemento de la supervisión humana. A partir de 2020 se cree que se logrará al nivel 4 en el que solo en algunas ocasiones específicas requerirá intervención humana

- Grandes empresas de comercio electrónico están realizando pruebas para entregar sus productos en estos medios inclusive mediante drones sin embargo el camino es complejo han existido fallas. Google en el 2016 reportó 400 fallas en sus pruebas que obligaron a desconectar el piloto automático. Nissan reportó también 400 y Mercedes Benz más de mil veces en solamente unos 3.000 kilómetros unas 500 por fallo técnico y el resto por decisión del conductor que expresó no sentirse cómodo en decenas de estos casos

Se evitaron accidentes gracias a la intervención del conductor humano

Por último en los autos completamente autónomos otra dificultad no técnica vendrá por aspectos relacionados con la responsabilidad legal en caso de accidentes **¿Quien será el responsable?**; la respuesta no es evidente y quizás será necesario equipar los coches con cajas negras similares a las de los aviones con el fin de poder definir responsabilidades.

También será imprescindible reflexionar sobre aspectos éticos por ejemplo en el software de un auto completamente autónomo se tendrá que haber previsto que hacer ante alternativas como por ejemplo salvar la vida de los pasajeros o de otras personas

Agricultura

En esta área la inteligencia artificial se utiliza con la finalidad de simplificar y acelerar la toma de decisiones de un ingeniero agrónomo quien debe saber cuándo es el mejor momento para sembrar y cosechar, además qué insumos debe utilizar.

Ya existen plataformas específicas en el mercado para la industria de la agricultura que analizan el tipo de suelo semillas y clima para seguir la mejor mezcla para lograr el éxito de la cosecha

Aplicaciones en Salud, Biotecnología, Ventas al detalle y Educación

En estas áreas, la inteligencia artificial ayuda a médicos y pacientes a tener un diagnóstico más rápido y preciso.

Salud y Biotecnología

Por ejemplo, en la detección del cáncer mediante muestras de sangre en los pacientes. El material recolectado tiene una enorme cantidad de datos que necesitan ser analizados para encontrar ciertos patrones.

La técnica ayuda a identificar los factores genéticos que podrán conducir el desarrollo o no de una enfermedad.

Una de las técnicas de inteligencia artificial más usadas en las aplicaciones médicas son los **sistemas expertos**. Por ejemplo, el sistema **Athena** ayuda a la toma de decisiones de los médicos a la hora de gestionar pacientes con problemas de hipertensión. Procesa los datos clínicos de cada paciente y, gracias a su base de conocimientos sobre hipertensión, produce una serie de recomendaciones sobre cómo gestionar mejor la atención clínica personalizada.

Otro ejemplo, es el de los científicos de la universidad de Carnegie Mellon en colaboración con cuatro hospitales de Chicago. han desarrollado un sistema capaz de predecir infartos con cuatro horas de antelación en enfermos, mejorando en más de tres horas los tiempos de predicción para los cardiólogos.

Este software fue entrenado con datos de 133 mil pacientes incorporando 72 parámetros presentes en la historia clínica de los enfermos, incluyendo signos vitales, edad, glucemia y recuento de plaquetas, entre otros.

Retail o Ventas al detalle

La inteligencia artificial ayuda a realizar predicciones de ventas, de comportamientos de clientes, de la oferta de productos y además permite elegir el producto adecuado para recomendar sea un cliente en particular, es decir, ofrece un servicio personalizado a los clientes.

Empresas como Amazon utilizan algoritmos para identificar si un libro tendrá o no éxito, inclusive antes de su lanzamiento.

Otra excelente aplicación para el sector minorista, es la optimización del inventario, donde la inteligencia artificial ayuda a las empresas a prevenir ingresos y determinar cuánto de un insumo debe adquirirse.

Educación

La inteligencia artificial permite la creación de diferentes tipos de servicios, por ejemplo, permite identificar si un estudiante está a punto de darse de baja o reprobar para entonces sugerir nuevos cursos para un estudiante, nuevas alternativas de pago o incluso crear ofertas personalizadas para optimizar el aprendizaje y fomentar la educación.

La inteligencia artificial permite conocer en forma detallada a los maestros y estudiantes, mediante **técnicas de segmentación**.

En resumen, permite que se ofrezca una educación personalizada con rutas de aprendizaje de acuerdo al perfil del estudiante que se van desarrollando en forma adaptativa.

Aplicaciones de entretenimiento y la industria aeroespacial

Entretenimiento, videojuegos

Existen videojuegos que utilizan técnicas de inteligencia artificial tales como el ajedrez las damas y los juegos de combate o estrategia

Hay videojuegos en las que el jugador compite con agentes inteligentes los llamados non player characters (NPCs) que son personajes animados presentes en el juego que el jugador no controla.

Los NPCs tienen que desplazarse de un lugar a otro evitando obstáculos y tienen que tomar decisiones acerca de cuál de sus posibles acciones es la más adecuada en función del contexto.

Entre las técnicas de inteligencia artificial que usan estos videojuegos tenemos:

- la búsqueda heurística
- la planificación
- las redes neuronales
- los sistemas deductivos basado en reglas
- y el aprendizaje de tácticas e incluso
- estrategias para adaptarse a las habilidades de cada jugador

Watson es un programa desarrollado por IBM que ganó el concurso de Jeopardy que consiste en dar respuestas a preguntas de cultura general. Watson ganó a brillantes contrincantes humanos en forma clara con las mismas reglas de juego de los jugadores humanos

Asistentes personales virtuales

En los próximos años los bancos tendrán empleados digitales ayudándonos a realizar las operaciones y además para responder a nuestras preguntas agilizando así la atención al público.

Por ejemplo **Siri** que fue desarrollada para los iPhones de Apple, se comunica mediante el habla escucha las preguntas del usuario y responde, recomienda y realiza acciones mediante los servicios web por ejemplo:

- encontrar un restaurante cercano o comprar las entradas de cine.
- ofrece una interacción conversacional con otras aplicaciones como los recordatorios, el estado del tiempo, la bolsa, la mensajería, el correo electrónico, el calendario, los contactos las notas, la música, el reloj, el navegador web y los mapas.

Siri también puede devolver llamadas, leer los mensajes del buzón de voz etcétera y dispone de soporte en varios idiomas.

Los datos que son utilizados por Siri son muy extensos; siri utiliza **conexiones de alta velocidad para reconocimiento de voz e integra a otros proveedores de información consultando fuentes como Bing, Wikipedia y Twitter sobre preguntas concretas puede incluso conectarse a:**

- The New York Times
- Planeadores de viaje: mediante apoyo virtual también será posible organizar mejores viajes de vacaciones por ejemplo planeando y programando los detalles tales como vuelos hoteles y traslados

Industria Aeroespacial

Los robots viajan al espacio desde el año 1997 en ese entonces la nasa envió al robot **Sojourner** de 65 centímetros de largo 48 de ancho y 30 de alto con un peso de 10 kilogramos estuvo 83 días en marte.

Después se enviaron a los robots Spirit, Opportunity y Curiosity.

Aplicaciones en Servicios Financieros, Manufactura, Cadena de Suministro, Socioeconomía y Mercadotecnia.

Se revisan las simulaciones que se utilizan para apoyar el crecimiento urbano, los sistemas de recomendaciones y los análisis de transacciones fraudulentas, entre otros.

Servicios Financieros

En esta área los algoritmos de la inteligencia artificial ayuda a las instituciones financieras a identificar los riesgos que un cliente puede representar y hasta predecir patrones del mercado y sus consecuencias así como recomendaciones de operación.

Los análisis de transacciones probablemente fraudulentas son un área en la que la inteligencia artificial permite tener éxito.

Manufactura y cadena de suministros.

Las técnicas y algoritmos de la inteligencia artificial ayuda en el **análisis de los productos** y las piezas que requieren mantenimiento, para **predecir los problemas que se podrían presentar en el futuro, esto ayuda a las empresas de manufactura a decidir cuánto comprar o producir así como a predecir impactos y riesgos de proveedores.**

O sea:

- **Predecir impactos**
- **Riesgos de proveedores**

Socioeconomía

En estas áreas también se usa cada vez más, modelos de inteligencia artificial, principalmente de **simulación basada en agentes.**

Un ejemplo, es la simulación de interacciones entre grandes cantidades de agentes y la predicción de **los efectos que causaría introducir elementos nuevos en un determinado sistema** como por ejemplo:

- Los efectos que tendría sobre la movilidad urbana
- La construcción de un parque o una zona peatonal
- Los efectos sobre la economía y la ecología de la construcción de una autopista o un aeropuerto. sistema como por ejemplo los efectos que tendría sobre la movilidad urbana, la construcción de un parque o una zona peatonal o los efectos sobre la economía y la ecología de la construcción de una autopista o un aeropuerto, de esta forma.

Las decisiones sobre dichas actuaciones se pueden tomar con muchas más y mejor

De esta forma las decisiones sobre dichas actuaciones se pueden tomar con muchas más y mejor información disponible

Chatbot

Otro ejemplo de los robots de conversación que son programas de computadora que usa un conjunto de reglas para conducir una conversación basada en el habla o en el texto con un humano a través de una interfaz de chat en línea.

Los robots de conversación **son alimentados por la inteligencia artificial y usan la técnica del aprendizaje de máquina para detectar e imitar la conversación humana**, se desarrollan comúnmente para proporcionar contenido específico o servicio automatizado o utilidad a los usuarios.

También, específicamente el mercadotecnia, **los sistemas de recomendación.**

Sistemas de recomendación: son algoritmos filtrados de información basada en inteligencia artificial que puede predecir automáticamente las preferencias del usuario y las respuestas a las consultas basadas en el comportamiento pasado, la relación de un usuario con otros usuarios.

La simulación entre los elementos comparados y el contexto.

Ejemplos:

- Sistemas de frecuentemente comprado de amazon
- el algoritmo de cine match de netflix
- Además estos algoritmos también se utilizan en las redes sociales como facebook y linkedin para encontrar conexiones entre personas y datos e identificar objetivos para las campañas de marketing.

Modelos

Aprendizaje Automático

Conceptos:

- **Aprendizaje automático:** es una rama de la inteligencia artificial que también es conocido como aprendizaje de máquinas (machine learning en inglés). Se enfoca en crear técnicas y algoritmos para que las computadoras aprendan.

También es el proceso por el cual las computadoras desarrollan el reconocimiento de patrones o la capacidad de aprender continuamente y hacer predicciones basadas en datos, tras lo cual realizan ajustes sin haber sido programados específicamente para ello.

Como forma de inteligencia artificial, el aprendizaje automático automatiza el proceso de creación de modelos analíticos y permite que las máquinas se adapten a nuevas situaciones de manera independiente.

- **El aprendizaje de máquinas:** consiste en la implementación de procedimientos que le permiten a las máquinas mejorar su rendimiento a través de la búsqueda de patrones que permitan predecir futuras conductas.

Como forma de inteligencia artificial, el aprendizaje automático automatiza el proceso de creación de modelos analíticos y permite que las máquinas se adapten a nuevas situaciones de manera independiente.

Se aplica en análisis financieros, reconocimiento de lenguaje, los videojuegos, detección de fraudes, en medicina y en algoritmos de búsquedas de datos, entre otras.

Algunos ejemplos de estos tipos de algoritmos son los árboles de clasificación, segmentación k.means, reglas de asociación y clasificador bayesiano entre otros.

Por ejemplo, las noticias de Facebook, utilizan un algoritmo con aprendizaje automático para personalizar las noticias que se muestran a cada usuario. Si tu detienes frecuentemente tu desplazamiento para leer o darle “like” a las publicaciones de un amigo en particular, las noticias empezarán a mostrarte más actividad de ese amigo. En caso de que ya no te detengas a leer, darle “like” o comentar a tu amigo, las noticias se ajustarán en consecuencia.

Aplicaciones

Se aplica en análisis financieros, reconocimiento de lenguaje, los videojuegos, detección de fraudes, en medicina y en algoritmos de búsquedas de datos, entre otras.

Algunos ejemplos de estos tipos de algoritmos son los árboles de clasificación, segmentación k.means, reglas de asociación y clasificador bayesiano entre otros.

Por ejemplo, las noticias de Facebook, utilizan un algoritmo con aprendizaje automático para personalizar las noticias que se muestran a cada usuario.

Si tu detienes frecuentemente tu desplazamiento para leer o darle “like” a las publicaciones de un amigo en particular, las noticias empezarán a mostrarte más actividad de ese amigo. En caso de que ya no te detengas a leer, darle “like” o comentar a tu amigo, las noticias se ajustarán en consecuencia.

Historia

Cuando surgió este concepto durante los años sesentas y setentas, su enfoque era principalmente hacia el uso de las redes neuronales artificiales que servían para aprender a clasificar datos. No se le daba tanta importancia en el ámbito de la investigación en inteligencia artificial.

Sin embargo a partir de los años 80 se empezaron a desarrollar nuevos métodos de aprendizaje que lo convirtieron en una de las áreas más activas de la inteligencia artificial.

Método de aprendizaje

El método de aprendizaje más conocido a partir de esta década es el **aprendizaje inductivo de árboles de decisión**.

Fundamento

La mayoría de **los métodos de aprendizaje automático infieren hipótesis a partir de datos**.

Dichas inferencias son inductivas contrariamente a las inferencias deductivas típicas de la lógica matemática.

El razonamiento inductivo va de lo particular a lo general.

El contraste es que las **inferencias inductiva son consecuencia lógica de un conjunto de premisas** y por lo tanto son conclusiones falseables.

Inductivas vs Deductivas

Mientras que las inferencias inductivas son solamente **hipótesis susceptibles de ser falseadas según se vayan conociendo datos adicionales** es decir, **las premisas apoyan la conclusión pero no la garantizan.**

Un ejemplo típico es el siguiente:

- Si un conjunto de datos sobre pájaros solo contiene información sobre pájaros blancos entonces la hipótesis a la que llegaría un algoritmo de aprendizaje inductivo es que todos los pájaros son blancos.
- Si más adelante se proporcionan datos adicionales que contienen información sobre pájaros negros entonces la hipótesis anterior queda falseada.

Tipos de algoritmos:

- **aprendizaje supervisado:** en los algoritmos de aprendizaje supervisado se dispone de un conjunto de datos de entrenamiento que permiten aprender para generar el modelo.

En el caso de un algoritmo. aprenda a clasificar, el conjunto de datos de entrenamiento, consiste en datos de los que se conoce a qué clase pertenecen y esta información que le dan estos datos ya clasificados permiten aprender para generar un modelo

- **ejemplo el árbol de clasificación:** es a partir del cual poder clasificar nuevos datos cuya clase es inicialmente desconocida. En el diseño del modelo es necesario seleccionar o etiquetar la variable o dato que se desea fijar o predecir, los datos son separados en la etapa de entrenamiento y la de prueba.

La primera etapa es para crear el modelo de acuerdo con el comportamiento de los datos y la segunda es para probar el modelo construido y para evaluarlo

- **Aprendizaje no supervisado:** en este tipo de algoritmos **no se tiene la información sobre la clasificación de los datos** por lo tanto en este caso el sistema tiene que ser capaz de reconocer patrones para poder clasificar las nuevas entradas es decir, no se dispone de un conjunto de datos de entrenamiento que permitan aprender para generar el modelo.

En el diseño del modelo no se seleccionan o etiquetas variables y los datos no son separados en la etapa de entrenamiento y la de prueba.

- **Aprendizaje por refuerzo:** los algoritmos de aprendizaje por refuerzo se pueden considerar que se sitúan entre los supervisados y los no supervisados.

En el aprendizaje por refuerzo, un agente **aprende mediante la interacción con su entorno en particular a partir de las consecuencias de acciones que se seleccionan** ya sea según

su experiencia previa que es la explotación de lo ya aprendido o según la selección aleatoria de las acciones posibles en cada situación mediante un proceso de prueba y error.

El agente recibe un valor numérico llamado refuerzo que puede ser positivo o negativo, éste codifica el éxito o fracaso de las acciones **el objetivo de la gente es seleccionar aquellas acciones que maximizan el refuerzo acumulado.**

No es aprendizaje supervisado porque no se proporcionan datos correctos de pares estado acción pero tampoco es completamente no supervisado pues el agente recibe un valor de refuerzo que guía su aprendizaje.

Un ejemplo típico es el de un ratón que debe aprender a recorrer un laberinto en la salida del laberinto la recompensa podría hacer por ejemplo un trozo de queso la posición inicial.

Dentro del laberinto sería el estado inicial y la salida el estado final dentro del laberinto también tiene que haber situaciones fácilmente identificables que corresponden a situaciones en las que se deben tomar decisiones que en este caso son acciones para realizar un avance en el laberinto.

Por ejemplo las esquinas los cruces de los pasillos y los extremos de pasillos sin salidas serían estados de este problema en cada estado hay o más acciones posibles.

Si el agente llega a un cruce hay cuatro acciones posibles pero en el extremo de un pasillo solamente cabe una sola acción que es ir hacia atrás.

El agente no conoce el mapa del laberinto así que debe aprender dicho camino por eso debe recorrer el laberinto en forma aleatoria y aprender mediante la prueba y error.

El aprendizaje por refuerzo consiste precisamente en aprender aquellas acciones que conduzcan al camino más corto para atravesar el laberinto.

Aprendizaje profundo (deep learning)

Definición

El aprendizaje profundo también conocido como deep learning, en inglés. Es una familia de métodos que aprenden representaciones de datos con múltiples niveles de abstracción. .

También es conocido como redes neuronales profundas porque tratan de emular el enfoque de aprendizaje que los seres humanos utilizan para obtener ciertos tipos de conocimiento en su forma más simple,

El aprendizaje profundo puede considerarse como una forma de automatizar el análisis predictivo estos algoritmos.

Redes neuronales profundas

Aplican en cascada un conjunto de transformaciones no lineales de forma que cada nivel recibe como entrada la salida del nivel anterior, aprendiendo así representaciones cada vez más complejas por ejemplo, para aprender a reconocer sillas basándose en imágenes de muchas sillas como datos de entrenamiento.

1. **El primer nivel de un sistema de aprendizaje profundo:** recibirá como entrada una matriz de contenidos de valores de píxeles de una imagen y lo que aprendes son características de bajo nivel como por ejemplo líneas, rectas en distintas orientaciones estas líneas constituyen los datos de entrada del segundo nivel.
2. **Segundo nivel,** a su vez aprende una representación de características un poco más abstracta como por ejemplo una serie de líneas formando contornos que delimitan regiones de la imagen a su vez estos contornos alimentan al tercer nivel.
3. **Tercer nivel:** aprende otra representación de características todavía más abstracta como por ejemplo, partes de una silla.
4. **Cuarto nivel:** es el último nivel y es en el que el sistema reconocerá la silla completa.

Otro ejemplo:

Imagina a un niño que está aprendiendo la palabra perro o el concepto perro. El niño aprende lo que es y lo que no es un perro señalando objetos y diciendo la palabra perro el niño señala y el padre evalúa si es un perro, no es un perro mientras el niño continúa apuntando a los objetos se vuelve más consciente de las características que poseen todos los perros.

Lo que el niño hace sin saberlo es aclarar una abstracción compleja en este concepto el concepto de perro construyendo una jerarquía en la que cada nivel de abstracción se crea con el conocimiento que se obtuvo de la capa precedente.

El aspecto clave del aprendizaje profundo es que estas características cada vez más abstractas.

No están diseñadas manualmente por el programador del sistema sino que son automáticamente extraídas mediante el aprendizaje, es decir, los programas informáticos que utilizan el aprendizaje profundo en cada algoritmo en la jerarquía aplica una transformación en su entrada y utiliza lo que aprende para crear un modelo estadístico como salida.

Las iteraciones continúan hasta que la salida ha alcanzado un nivel de precisión aceptable el número de capas de procesamiento a través de los cuales los datos deben pasar es lo que inspiró la etiqueta de profundidad "deep".

Redes neuronales profundas 2

Uno de los métodos de aprendizaje profundo más conocidos son las redes neuronales profundas, en este caso el algoritmo básico de aprendizaje que se usa en cada uno de los niveles es el mismo que en las redes neuronales clásicas que contienen únicamente **dos niveles**.

Este algoritmo propuesto por Hinton y Williams en 1986 se llama **Backpropagation y consiste en llevar a cabo un descenso por gradiente esto es por variaciones actualizando los pesos de las conexiones entre neuronas de forma que se minimice el error de clasificación de acuerdo a los datos de entrenamiento.**

Aunque las redes neuronales son múltiples no es un concepto nuevo; hasta ahora no se había podido usar debido a que requiere una enorme cantidad de datos de entrenamiento para que sean capaces de aprender además del uso del software y hardware sofisticado gracias a la disponibilidad de cientos de miles incluso millones de datos de entrenamiento.

Ejemplo las imágenes en internet de toda clase de objetos animales etcétera y gracias al acceso a computadoras de altas prestaciones ha sido aplicar el método del aprendizaje profundo con éxito para reconocer patrones.

Actualmente se están desarrollando muchas aplicaciones del aprendizaje profundo, entre ellas el diagnóstico médico basado en imágenes y el reconocimiento visual de objetos en robótica los algoritmos de aprendizaje profundo al ser parte del aprendizaje automático incluyen aplicaciones de modelización de datos y reconocimiento de patrones además también pueden utilizar aprendizaje supervisado o no supervisado.

En este caso es mejor el no supervisado como vemos en este ejemplo, en un algoritmo supervisado construido para aprender el concepto de perro inicialmente el algoritmo sería provisto de datos de entrenamiento que son un conjunto de imágenes para las cuales un ser humano ha etiquetado cada imagen como perro o como no perro.

El algoritmo **utiliza la información que recibe de los datos de entrenamiento** para crear un conjunto de características para el perro y así construir un modelo predictivo en este caso el modelo que la computadora crea por primera vez podría predecir que cualquier cosa en una imagen que contenga cuatro patas y una cola, debería estar etiquetada como, perro.

Robótica

Conceptos

- **Robótica:** es la ciencia que estudia los fundamentos de las máquinas capaces de imitar tareas humanas, hasta llegar a recrear inteligencia. Es una rama de la inteligencia artificial. Un robot es un dispositivo manipulador controlado por computadora. La robótica es la ciencia que estudia los robots
- **Robot:** Cualquier agente autónomo que reside en la propia máquina y además son inteligentes ya que tienen la capacidad de tomar decisiones con base en la información de sus sensores.

Partes de un robot

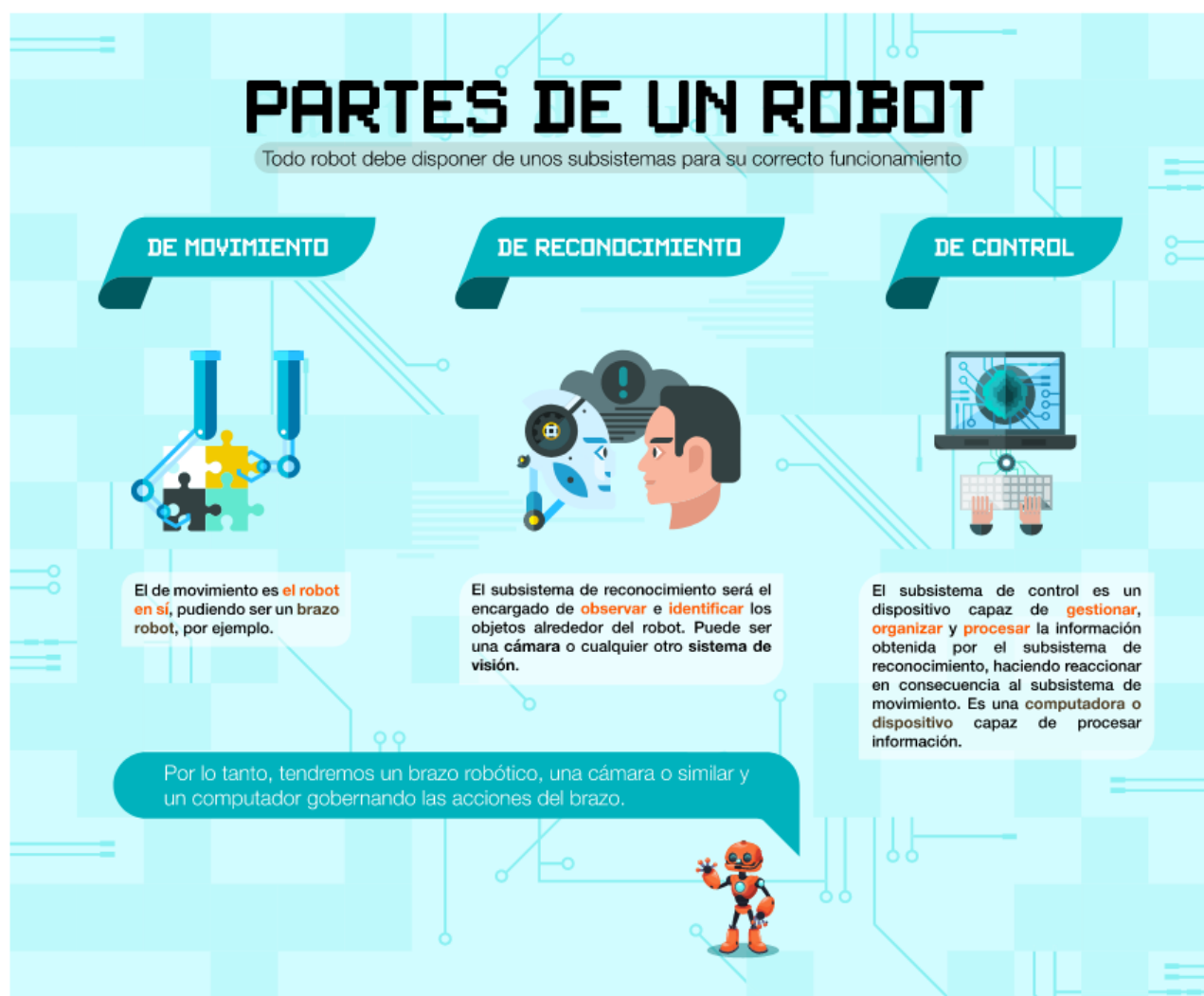


Figura 2: Partes de un robot

Estamos en los inicios de la cuarta revolución industrial. El mundo interconectado que nos proporcionan internet, la inteligencia artificial y la robótica, ya están modificando nuestra rutina y lo hará aún más en un futuro próximo.

La investigación en estos campos avanza a gran velocidad, deslumbrándonos con descubrimientos que, hasta ahora, formaban parte de la ciencia ficción.

Sin embargo, la sociedad, que es el campo real de experimentación de esta revolución, y que finalmente tendrá que vivir con los robots y la inteligencia artificial, ¿está preparada para un cambio tan drástico?

Mucha de esta tecnología del futuro ya existe; sin embargo, son muchos los que aseguran que su implantación masiva tardará. Incluso puede que más de lo que tardaron en aceptarse los inventos durante la Revolución Industrial.

Por el contrario, otros la ven inminente. **Es un trabajo de mentalización y de educación.**

Los científicos nos presentan unas máquinas que se parecen a nosotros físicamente y que son capaces de mejorar nuestra eficiencia. Pueden analizar con más lógica y rapidez que nosotros, trabajar más horas incluso sin dormir.

Áreas de la robótica:

- **Robótica Industrial.** Es uno de los campos de más auge de la robótica. El robot industrial suele ser el típico brazo que suelda mientras ensambla partes de un producto en la cadena de montaje de una fábrica.

La industria que más utiliza robots de este tipo es la automovilística.

Un **robot industrial** es cualquier máquina que posea brazos articulados móviles que permita cierta manipulación. Los más comunes son los brazos robot.

- **Robótica de servicio:** Según la Federación Internacional de Robótica, un robot de servicio opera de manera automática o semiautomática para realizar servicios útiles al bienestar de los seres humanos o a su equipamiento, excluyendo las operaciones de fabricación.

Pueden utilizarse para cuestiones militares, aspirar en la limpieza, para el mantenimiento de línea de alta tensión o para rescates submarinos, entre otros usos.

Dentro de los robots de servicio también se pueden incluir:

- **los robots domésticos:** son aquellos que realizan este tipo de trabajos, como barrer, aspirar, limpiar, etc.

Un robot doméstico tiene que reunir una serie de características: Debe realizar los trabajos sin intervención humana. Debe realizar sus trabajos con total autonomía.

Por eso cuentan con sensores que le permiten esquivar obstáculos, reconocer la superficie que va a aspirar y encontrar la estación de recarga cuando sus baterías están a punto de agotarse. Sólo requiere la intervención humana para introducir el programa deseado.

- **robots de ocio:** Entre estos, tenemos los robots mascota, cada vez más frecuentes en los hogares, desde un Furby hasta mascotas más sofisticadas y robotizadas, como la mascota Pleo. Estos robots mascota intentan imitar a las fases humanas de crecimiento emocional y de aprendizaje.

A medida que la mascota va creciendo, gracias a los cuidados del niño o la niña, el robot se transforma hasta llegar a la etapa adulta.

También se han puesto al servicio de niños con problemas de autismo para apoyar sus terapias.

También, se están desarrollando robots humanoides para la asistencia a personas con discapacidad, ancianos o incluso para el cuidado de niños pequeños.

Como ves, la robótica se abre paso y a la vez se interrelaciona con otras disciplinas, como la psicología, sociología, con el fin de poder crear robots asistentes que aprendan a reconocer y gestionar expresiones faciales o comportamientos humanos para interactuar con las personas en la vida cotidiana.

- **robots de educación:** Los centros educativos están empezando a introducir la robótica como medio de aprendizaje, ya sea para explicar otras materias en sus clases o para

introducir a los alumnos en la mecánica, robótica, mecatrónica, la programación o la electrónica.

También, existen robots para enseñar a los niños idiomas, principalmente la pronunciación y el vocabulario.

Desarrollo de proyectos de Robótica

Arduino es una de las plataformas más utilizadas para llevar a cabo proyectos tecnológicos de robótica. Esto, debido a su versatilidad, a la cantidad de sensores que puede incorporar, a su facilidad de uso y a los precios bajos.

Se emplea un lenguaje de programación especial, que se basa en la sintaxis de otros lenguajes, como C y C++.

Aún así, este lenguaje que utiliza Arduino, posee sus características propias, orientado a una fácil programación de los sensores y dispositivos externos.

El futuro del trabajo

Existen estudios que han analizado la cantidad de trabajos que desaparecerán con la llegada de los robots y la inteligencia artificial.

En su informe "El Futuro del empleo 2018", el Foro Económico Mundial pronostica que se perderán empleos. En la actualidad, debido a los avances tecnológicos, el 29% de las tareas laborales son realizadas por una máquina.

Este porcentaje significa la pérdida de 75 millones de empleos para 2025. Sin embargo, se señala que se crearán nuevos roles o funciones que necesitarán de un mayor grado de especialización.

Así, se abrirán oportunidades que permitirán el surgimiento de más de 133 millones nuevos roles de trabajo.

Robots destacados en 2018

- **Pepper:** Es un robot semi-humanoide fabricado por SoftBank Robotics. Fue presentado en 2014 y en 2015 ya se vendía en Japón.

Es el primer robot humanoide social del mundo. Puede reconocer rostros y emociones humanas básicas.

Actualmente, el robot está presente en empresas y escuelas. Más de 2.000 compañías ya lo adoptaron como asistente para recibir, informar y guiar a los visitantes que reciben.

Pepper se ha convertido en un recurso instrumental para las empresas en los Estados Unidos. Ha mejorado su capacidad para aprender, evolucionar y adoptar nuevas características. Continúa comprometiendo e informando a los clientes, capacitando a los empleados, optimizando las operaciones y facilitando el aprendizaje. Según palabras de sus diseñadores.

El futuro para Pepper es conectarse a las herramientas de una empresa (sistemas de gestión de inventario, por ejemplo). Así, recibirá información y análisis en tiempo real, que irá procesando

- **Atlas:** Es el robot desarrollado por la compañía de Estados Unidos Boston Dynamics, se dio a conocer en 2013. Mide 1.50 metros y pesa 75 kilogramos, fue diseñado para tareas de búsqueda y rescate. Puede correr a gran velocidad, saltar obstáculos y hasta hacer parkour (es una actividad física conocida como "acrobacia urbana")
- **Insight:** Es el robot geofísico que lanzó NASA y llegó a Marte el 26 de noviembre de 2018. Tiene sensores que miden el suelo y la temperatura para analizar la evolución geológica del planeta.

La tecnología de InSight permite entender mejor los planetas similares (Venus, Mercurio y etc). En resumen estudia la estructura interna del planeta.

Sistemas Expertos

En la presente unidad, se muestra una introducción a los sistemas expertos. Se describen sus antecedentes, conceptos, ventajas y componentes, asimismo, se describen algunos ejemplos. Además, se describen las problemáticas de su implementación.

Conceptos

Sistema experto: es una máquina que piensa y razona como un experto lo haría en una cierta especialidad o campo. Un sistema experto puede definirse como un **sistema informático de hardware y software** que simula a los expertos humanos en un área de especialización.

El sistema experto es capaz de procesar y memorizar información, aprender y razonar en situaciones deterministas e inciertas, comunicarse con los humanos y otros sistemas expertos, tomar decisiones apropiadas y explicar por qué se han tomado tales decisiones.

Por ejemplo, un sistema experto en diagnóstico médico, requiere como datos los síntomas del paciente, los resultados de análisis clínicos y otros hechos relevantes, y, utilizando estos, busca en una base de datos la información necesaria para poder identificar la correspondiente enfermedad.

Antecedentes

Las investigaciones the new world y simon sobre la forma en que un ser humano resuelve problemas, dieron como resultado un modelo basado en lo que llamaron **sistema de producción**.

Éste consistía en representar conocimientos sobre cómo resolver un problema mediante reglas de producción.

Estas reglas eran del tipo sí entonces donde la parte si contenía alguna premisa que debía cumplirse con el fin de poder aplicar la parte entonces qué consistía en una acción.

Las reglas de producción se almacenaron en lo que llamaron **memoria a largo plazo** que procesaría datos sobre tareas concretas a resolver que estaban contenidas en otra **memoria a corto plazo** de forma que cuando algún dato cumpliera la condición expresada en la premisa de alguna regla ésta se activaría y se llevaría a cabo la acción indicada en dicha regla.

Esta idea de representar conocimientos mediante reglas del tipo sí entonces dio lugar a los sistemas expertos una de las técnicas más importantes que ha dado la inteligencia artificial.

Ventajas

- Personal con poca experiencia puede resolver problemas que requieren un conocimiento de expertos. Ayudan en los casos en los que hay pocos expertos humanos.
- ayudan en los casos en que la complejidad del problema impide al experto humano resolverlo. Los sistemas expertos pueden responder a preguntas y resolver problemas mucho más rápido que un experto humano, además, ayudan en los casos en que la complejidad del problema impide al experto humano resolverlo.
- El número de personas con acceso al conocimiento aumenta con el uso de sistemas expertos. El conocimiento de varios expertos humanos puede combinarse, lo que da lugar a sistemas expertos más fiables, ya que se obtiene un sistema experto que combina la sabiduría colectiva de varios expertos humanos en un lugar.
- Los sistemas expertos pueden ser utilizados para realizar operaciones monótonas aburridas e incómodas para los humanos.

Ejemplos

Heuristic Dendral

El primer sistema experto propiamente dicho fue heurística desarrollado durante 10 años desde principios de los años 70 en la nueva universidad de Stanford en el ámbito de la química orgánica, era capaz de hipotetizar la estructura topológica de un compuesto químico, es decir la disposición espacial de sus átomos para interpretar su espectrograma de masas usando conocimientos de químicos expertos en analizar la información proporcionada por el espectrómetro. Éstos conocimientos se representaron mediante reglas si/entonces

MYCIN

Otro ejemplo fue un proyecto que se desarrolló en la universidad de Stanford en el que los investigadores “your life” **Buchanan y Cohen** desarrollaron el famoso sistema experto llamado MYCIN también utilizaba reglas del tipo sí/entonces para representar el conocimiento de médicos expertos en enfermedades infecciosas.

La parte si de una regla típica de MYCIN describía posibles síntomas y la parte entonces expresaba una posible causa de dichos síntomas.

Algo de historia

A mediados de los años 80s había centenares de sistemas expertos desarrollados en todo el mundo en distintos dominios de aplicación.

Los espectaculares éxitos alcanzados con los sistemas expertos cambiaron significativamente las prioridades de investigación en la inteligencia artificial ya que en lugar de focalizar esfuerzos en desarrollar sistemas generales de resolución de problemas la mayoría de los investigadores se dedicaron a resolver problemáticas muy específicas.

Hoy día entre muchos ejemplos cotidianos tenemos a los cajeros automáticos y los semáforos como sistemas expertos.

Algunas áreas en las que más se utilizan son las siguientes: negocios, industria, medicina, ingeniería y ciencias aplicadas.

Existen algunos problemas en su implementación:

- El costo de los procesos de adquisición de conocimientos, el desarrollo de un sistema experto es generalmente caro, pero el mantenimiento y el costo marginal de su uso repetido es relativamente bajo
- la ganancia en términos monetarios tiempo y precisión resultantes del uso de los sistemas expertos son muy altas y la amortización es muy rápida sin embargo antes de desarrollar o adquirir un sistema experto debe realizarse un análisis de factibilidad y de costo-beneficio otro
- fragilidad ya que únicamente pueden resolver situaciones específicas que estén representadas con reglas sí/entonces

Por ejemplo, si un sistema experto en medicina no contiene el conocimiento de que únicamente las mujeres pueden tener embarazos dicho sistema podría preguntar por los resultados de una prueba prenatal para pacientes varones.

Esta fragilidad está relacionada con la ausencia de conocimientos de sentido común en los sistemas expertos.

Componentes de un Sistema Experto

1. **Expertos humanos.** Se requiere la colaboración de uno o varios expertos humanos especialistas en el tema de estudio y los ingenieros del conocimiento.
2. **Base de Conocimiento.** Los especialistas son responsables de suministrar a los ingenieros del conocimiento una base de conocimiento ordenada y estructurada, y un conjunto de relaciones bien definidas y explicadas.
3. **Subsistema de Adquisición de Conocimiento.** Controla el flujo del nuevo conocimiento que fluye del experto humano a la base de datos.
4. **Control de la Coherencia.** Controla la consistencia de la base de datos y evita que unidades de conocimiento inconsistentes entren en la misma.
5. **El Motor de Inferencia.** Es el corazón de todo sistema experto. Obtiene las conclusiones aplicando el conocimiento a los datos.
6. **Subsistema de adquisición de conocimiento.** Se utiliza para obtener el conocimiento necesario y continuar con el proceso de inferencia hasta que se hayan sacado conclusiones.
7. **Interfase de Usuario.** Es el enlace entre el sistema experto y el usuario.
8. **Subsistema de ejecución de órdenes.** Es el componente que permite al sistema experto iniciar acciones.
9. **Subsistema de Explicación.** Da explicaciones de las conclusiones obtenidas o de las acciones iniciadas por el sistema experto.

10. **Subsistema de Aprendizaje.** Lleva a cabo el aprendizaje estructural al aprender reglas, distribuciones de probabilidad, etc. y también el aprendizaje paramétrico al estimar los parámetros necesarios para construir la base de conocimiento. Obtiene experiencia a partir de los datos disponibles.

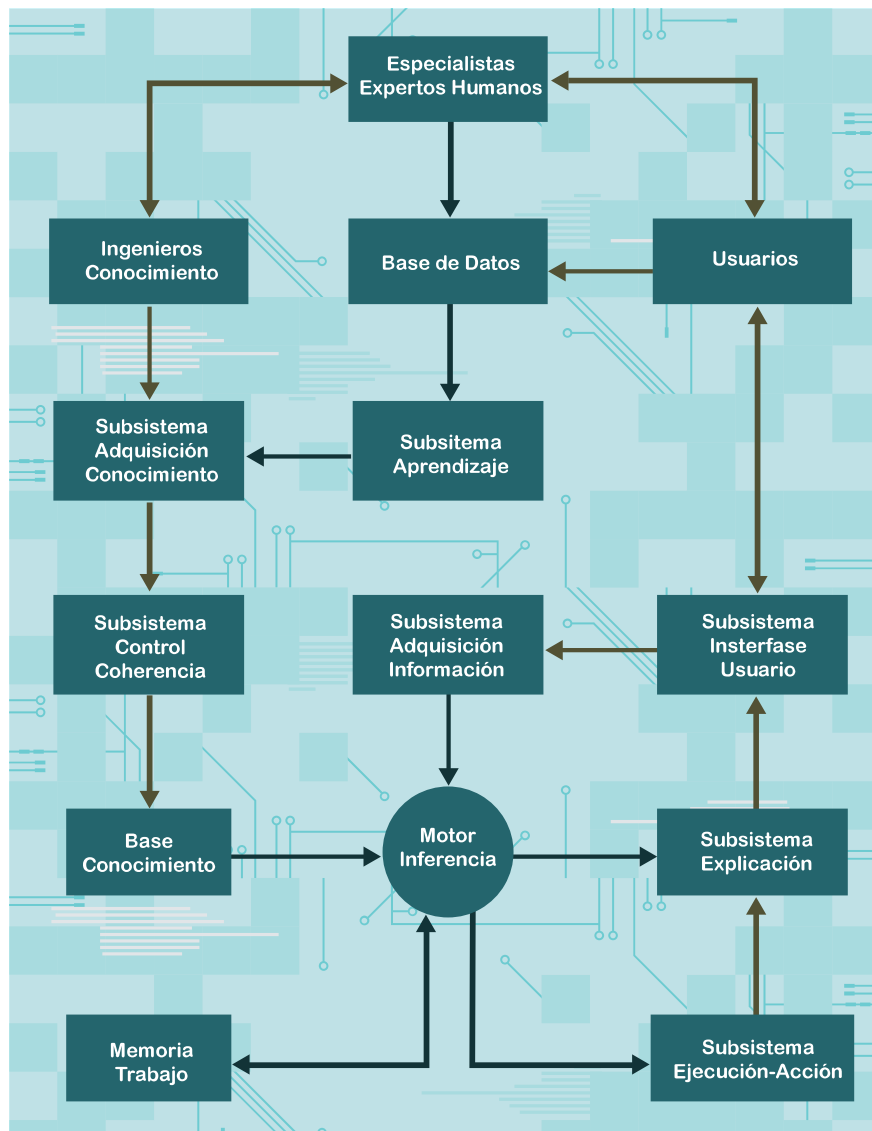


Figura 3: Modelo de Sistema Experto

Procesamiento de lenguaje natural (PLN)

Introducción al procesamiento de lenguaje natural

Introducción a la sección

Esta sección trata sobre el procesamiento del lenguaje natural, también conocido como PLN, que se enfoca en analizar las interacciones entre las computadoras y el lenguaje humano.

En estos contenidos se muestra una introducción al procesamiento del lenguaje natural. Se describen los conceptos relacionados con el PLN, los tipos y modelos de procesamiento y las problemáticas relacionadas con el análisis del lenguaje en sus formas hablada y escrita.

También se examinan las diferentes técnicas de análisis, en particular el aprendizaje automático. Se muestran las diversas aplicaciones, mostrando un esquema de cómo se realiza un análisis de sentimiento en twitter y se realizan ejercicios prácticos mediante herramientas en internet.

Además se realiza una práctica paso a paso con el software RapidMiner de análisis de sentimiento en twitter.

El procesamiento del lenguaje natural hablado o escrito, es una importante área de desarrollo de la inteligencia artificial. Es esencial comprender que todavía no se logra una exactitud total en la traducción y comprensión de textos y audios, debido a los problemas de ambigüedad y estructura del lenguaje.

Introducción al procesamiento de lenguaje natural

- **PLN:** El procesamiento del lenguaje natural (PLN), consiste en las técnicas de representación del lenguaje humano. El PLN estudia las interacciones entre las computadoras y el lenguaje humano.

Es una **disciplina** que se ocupa de la formulación e investigación de mecanismos computacionales para la comunicación entre personas y máquinas mediante el uso de Lenguajes Naturales.

- **Lenguaje Natural** es el que se utiliza en la comunicación humana, ya sean escritos, hablados o mediante signos.
- **Linguística computacional:** La ciencia que estudia el PLN. En su etapa actual de desarrollo, es principalmente una rama de las tecnologías de aprendizaje automático que son parte de la inteligencia artificial.

Esta habilidad de la máquina para procesar la información comunicada, no simplemente se refiere a las letras o los sonidos del lenguaje, sino, todo el mensaje como tal. Es decir, una contestadora telefónica, una impresora o un procesador de textos, **no** son dispositivos de inteligencia artificial que utilicen algoritmos de procesamiento del lenguaje natural.

Un programa de traducción automática de textos o audios, sí lo es.

El PLN es un campo de las ciencias de la computación, inteligencia artificial y lingüística.

Tipos de PLN:

- **tipo simbólico** que se caracteriza por sistemas que almacenan la fonología/fonética, morfología, sintaxis, semántica, pragmática y discurso, a través de representaciones.
- **tipo estadístico**, que se caracteriza por la aplicación de técnicas matemáticas para inferir conocimiento.

Estos procesamiento de lenguajes son fundamentales para la traducción de textos, donde actualmente la tasa de error alcanza el 20-30%. Esto se produce por los problemas de ambigüedad

que sólo tienen solución con el conocimiento del entorno donde se utiliza el lenguaje. Es decir, el procesamiento del lenguaje natural es complicado y no se tiene aún la confiabilidad al 100%, aunque se ha investigado mucho sobre este tema.

El conocer el contexto en donde se emitió un mensaje de texto que se quiere analizar, ayuda mucho en la exactitud del análisis.

¿Cuántas veces hemos malinterpretado un mensaje de WhatsApp? Nos imaginamos que se quiso decir algo, pero en realidad era otra cosa la que se pretende expresar, o el mensaje no tenía la intención que nosotros asumimos que tenía.

Más aún, cuántas veces no comprendemos al cien por ciento una frase que nos dice una persona cara a cara, no sabemos si es sarcasmo, si está vacilando o si está hablando en serio.

Ahora, imagina la dificultad que esto representa para los programadores de algoritmos de análisis de procesamiento de texto o del habla.

Además, el procesamiento del lenguaje natural hablado también es fundamental para la inteligencia artificial.

Esto representa diversos problemas como los siguientes:

- **Procesamiento de habla.** Para lograr una comunicación hombre - robot eficiente, las máquinas deben entender el lenguaje hablado.
- **El diálogo.** No debe ser un monólogo. El problema es que en los diálogos se usan mucho las oraciones incompletas o recortadas a una sola palabra, lo cual dificulta su análisis.
- **Generación del lenguaje.** El robot debe componer, estructurar y ordenar las oraciones.
- **Relacionar las palabras con las acciones,** objetos y circunstancias en la conversación. Si le decimos al robot, “barre”, él debe tomar una escoba y barrer, lo cual implica comprensión y movimientos .

El procesamiento del lenguaje natural hablado o escrito, representa una enorme área de la inteligencia artificial.

Se han realizado muchas investigaciones y se tiene un gran avance al respecto, sin embargo, en ambos casos, todavía no se logra una exactitud total, debido a los problemas de ambigüedad y estructura del lenguaje.

Técnicas de procesamiento de lenguaje natural

Análisis del lenguaje

La estructura del lenguaje se realiza a 4 niveles:

1. **Análisis Morfológico:** Análisis de las palabras para extraer raíces, unidades léxicas. Se identifican los componentes en su mínima expresión. Por ejemplo, en la frase “El niño corre” debe extraer que “El” es un artículo, “niño” es sustantivo y “corre” es un verbo.

Las **unidades léxicas** es la expresión mínima de un lenguaje, entonces la expresión mínima del lenguaje que son pues identificadores o sea las palabras las letras que forman que forman palabras estas palabras es una expresión mínima que en un lenguaje se puede llamar un identificador se puede llamar. Por ejemplo aquí le estamos poniendo artículo sustantivo verbo que son identificadores que se van dando como resultado todo el análisis que tú hagas de esta frase.

Éste análisis morfológico que es lo que hace pues **recorre cada uno de las letras o de los elementos o caracteres que tú tienes en el mensaje** y entonces llega y dice esta es una letra la letra e y la guarda en la memoria ésta es la letra l y la guarda en la memoria; entonces en la memoria ya tiene la e y la l tiene el viene la siguiente y dice es un espacio en blanco. Al momento que dice que es un espacio en blanco agarra lo que tiene en la memoria que era la letra e y la letra ele y la guarda en una tabla y le asigna un identificador que en éste caso es un artículo.

Entonces cada vez que se encuentra un espacio en blanco va tomando y va almacenando lo que él tiene en la memoria en la siguiente área el recorrido es decir esta es una letra n la guarda la y la guarda a la palabra niño y la guarda y la o la guarda y luego se encuentra un espacio en blanco al encontrarse el espacio en blanco viene por lo que guardo que fue la palabra niño, a la unidad léxica niño y lo mete la tabla de otra palabra.

Entonces vean cómo va separando el niño corre ahora es una tabla de ser una frase de texto sin estructura vean como ahora se convierte en una estructura que es una tabla que tiene dos columnas en la primera columna va a tener la unidad léxica que extrajo de ese texto y en la segunda el tipo de unidad léxica que nosotros tenemos aquí entonces en ese consiste el análisis morfológico.

Pasa de ser un texto sin estructura a una tabla de 2 columnas donde en la primer columna va la unidad léxica que encontró y en la segunda el tipo de unidad léxica asignado.

2. **Análisis sintáctico:** Análisis de la estructura sintáctica de la frase mediante una gramática, es decir, el orden de los elementos identificados en el punto 1. Ej: sujeto="El" (artículo) + "niño" (sustantivo) + "corre" (verbo).

El análisis sintáctico va a verificar que éstos estén en orden porque aquí primero se encontró el luego el niño y luego corre, entonces el orden fue artículo sustantivo y verbo.

Bien, el análisis sintáctico revisa que estén en orden correcto pero quién le dice el orden el orden se lo dice la gramática en la gramática es donde está guardado el orden.

Una gramática es algo como esto que dice el sujeto va a ser igual al artículo más un sustantivo más un verbo entonces esto podríamos nosotros interpretarlo como una regla gramatical en donde para formar un artículo tiene que venir primero para formar un sujeto tiene que venir primero un artículo después venir un sustantivo y después venir un verbo entonces con esto se refiere al orden gramatical.

Entonces dada esta tabla dice es un artículo pregunta lo que sigue es un sustantivo porque aquí sigue un sustantivo si es correcto y después sigue un verbo sí es correcto entonces si se

da que está en la tabla un artículo más un sustantivo más un verbo en ese orden entonces esto es un sujeto entonces “el niño corre ahora” es un sujeto que es lo que tenemos aquí.

3. **Análisis semántico:** Se analiza el significado de la frase

4. **Análisis pragmático:** Es el análisis profundo de los significados

Las etapas 3 y 4 que es el análisis semántico y el análisis pragmático que revisan el significado de todos estos elementos que tenemos como resultado de 1 y 2, entonces hay cuestiones muy propias del lenguaje en particular que se esté realizando en donde estas dos etapas pues les dan el sentido a esta frase que yo estoy analizando entonces eso es lo que hace el análisis semántico y el análisis pragmático

Técnicas y modelos de análisis del lenguaje

Hay 2 tipos de técnicas y modelos de análisis de lenguaje:

1. **Técnicas lingüísticas formales:** una técnica lingüística formal se refiere a que se tiene un conjunto de reglas desarrolladas basadas en una estructura y sobre ese conjunto de reglas se realiza en las fases de análisis del lenguaje
2. **Técnicas probabilísticas** Se basan en el estudio de un conjunto de textos de referencia a los que se les conoce como un corpus que es un conjunto en este caso un conjunto de textos de referencia que tienen características de tipo probabilístico asociadas a las distintas fases del análisis del lenguaje.

Aquí en estas técnicas probabilísticas yo no voy a tener una de reglas de entrada o de inicio sino que primero voy a tener un conjunto de textos en donde ya se están utilizando todos los textos y los voy a estar analizando y probabilísticamente

Entonces lo que va a identificar es la forma en que se van organizando todos esos textos de referencia y ahí están implícitas todas las reglas gramaticales.

Modelos para el procesamiento del lenguaje natural:

Derivado de estos dos tipos de técnicas tengo los modelos que son los modelos para el procesamiento del lenguaje lógico y el probabilístico.

1. **Lógicos (Gramática)** el modelo lógico que se basa precisamente en las gramáticas o sea es utilizando la técnica lingüística formal. Es decir la gramática es ese conjunto de reglas estructurales
2. **Probabilísticos (basados en conjuntos de textos):** Hay los modelos probabilísticos que están basados en un conjunto de textos que son estas técnicas probabilísticas que toman como base un conjunto de textos de referencia es decir un corpus entonces los dos modelos que yo tengo para realizar este

Análisis mediante una Lingüística Formal

Lo compone (o es) una serie de reglas gramaticales que yo tengo aquí entonces esto que tenemos aquí.

Imaginemos que es una estructura gramatical es decir es una gramática y esta gramática se entendería de la siguiente forma por ejemplo yo voy a decir una oración se compone de una Fnominal que es una frase nominal, más una frase verbal, más un complemento

Entonces esto es una regla. Si yo identifico que tengo una frase nominal y en seguida una frase verbal y en seguida un complemento entonces yo estoy hablando de una oración.

Ahora bien, esta frase nominal de que se compone aquí está la regla es decir la frase nominal se va a componer de un artículo más un sustantivo.

Y ésta es otra regla gramatical la frase verbal que tenemos aquí de que se compone de un verbo el artículo que tenemos aquí de qué se compone es una cadena de caracteres. El sustantivo que lo tenemos es una cadena de caracteres y el verbo es una cadena de caracteres entonces esta es la gramática.

Y si la viéramos en forma de un árbol es decir la oración que tenemos aquí se va a formar de una frase nominal más una frase verbal si se fijan, aquí no tenemos un complemento pero es válido que yo pueda tener la parte de la frase nominal más la frase verbal.

Ahora la frase nominal de que se compone dice la regla de un artículo más un sustantivo entonces yo tengo aquí un artículo y un sustantivo y este artículo de qué se compone el artículo se compone de una cadena de caracteres aquí está el y el el sustantivo de que se compone de una cadena de caracteres.

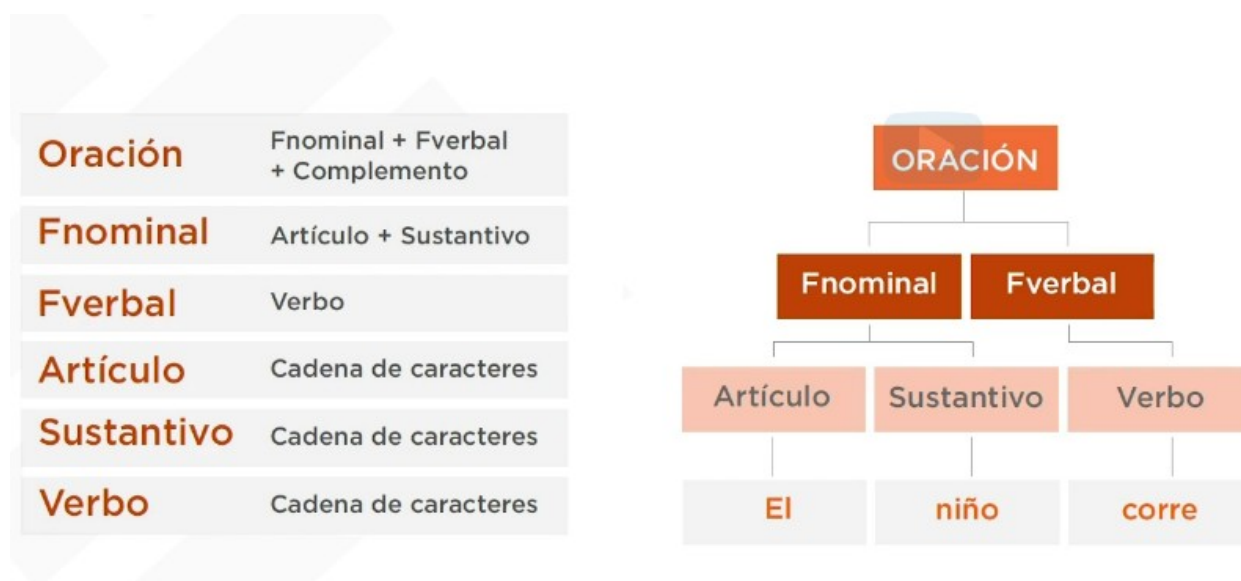


Figura 4: Árbol de reglas gramaticales por lingüística

La frase verbal de que se compone de un verbo y el verbo de que se compone de una cadena de caracteres la “c”, la “o”, la “r” dos veces y la “e”

Análisis por métodos probabilísticos

Éstos modelos probabilísticos funcionan a la inversa o sea yo no parto de que tengo una gramática y que debo analizar mi frase para ver si hay encajan ahí.

Lo que sucede es que las personas o lingüistas recopilan un corpus o un conjunto de ejemplos de datos, es decir van a tener como muchos documentos para ver cómo es que se están utilizando todos los textos y qué estructura tienen todos los textos.

Procedimiento

1. **Tener este corpus** o sea este conjunto de ejemplos ahora
2. **Cálculo de la frecuencia con la que las unidades lingüísticas aparecen en un contexto determinado** es decir ya que tengo este conjunto de documentos yo veo cómo están las diferentes estructuras pero también considerando el contexto dice.

Por ejemplo frecuencia en la que aparece la palabra pasaporte en la colección de ejemplos y datos es decir yo tengo un conjunto de documentos y veo que en todos esos documentos la palabra pasaporte aparece mil veces

3. **Predecir cuál será la siguiente unidad en un contexto dado** es decir si yo ya dije que la palabra pasaporte aparece una mil veces por ejemplo y yo puedo seguir con ese ejemplo y decir la frecuencia en que la palabra pasaporte aparece seguida ya sea antes o después de la palabra vuelo

Entonces si yo en esta tabla que genere en la etapa 2 me dice que pasaporte aparece junto a la palabra vuelo en el 98% de ocasiones quiere decir que cuando yo esté analizando un nuevo documento allí el 98 por ciento de probabilidad de que cuando aparezca la palabra pasaporte va a aparecer junto a la palabra vuelo

Esta forma de analizar en forma probabilística esto es el paradigma del aprendizaje automático precisamente dice los algoritmos infieren las posibles respuestas a partir de los datos observados anteriormente en el corpus, es decir en el conjunto de datos de entrada entonces esto es la forma en que se analiza mediante el modelo probabilístico

Ejemplos de modelo probabilístico

Ejemplo 1

Frase “Juan usa un gato para reparar su coche”

Esta parte del lenguaje que sería un texto que en este caso es una pregunta que dice Juan usa un gato para reparar su coche entonces esto se tiene que analizar para identificar vamos a suponer qué tipo de gatos el que utiliza Juan para identificar, si es este gato que es un animal o este gato que es una herramienta cuál es el gato que se refiere en esta frase.

La computadora tiene que hacer un análisis porque una persona pues el sentido común le diría que el gato que está utilizando es el de la herramienta pero la computadora tiene que aprender eso.

Entonces que tiene que hacer pues tiene que utilizar un modelo y una técnica de análisis del lenguaje para saber cuál es el gato que está utilizando Juan.

¿Cómo lo realiza esto? bueno pues tenemos textos de entrenamiento entonces, al ser probabilístico recordemos que va a necesitar muchos textos donde se utilicen la palabra gato para ir viendo cual se

asemeja más a este tipo de frase para poder identificar cuál de los dos gatos es el que debe funcionar.

Aquí entonces tiene por ejemplo un conjunto de textos y empieza a sacar texto como “Pedro usa un martillo para ...”, “Anna usa un desarmador para...”, “El obrero usa una grúa para ...”; y por otro lado tiene “El gato comió ratón”, “El perro come la carne”, “El hámster come avena”.

¿Qué gato tiene Juan?

Juan usa un gato para reparar su coche

¿Qué gato?

El diagrama muestra un gato y una grúa como opciones. A continuación, se detallan los textos de entrenamiento y el diccionario monolingüe utilizados para la clasificación.

| Textos de Entrenamiento | Diccionario monolingüe |
|---|--|
| <p>Pedro usa un martillo para Ana usa un desarmador para el obrero usa una grúa para alguien usa éstos para algo</p> <p>el gato come ratones el perro come la carne el hámster come avena éstos comen algo</p> <p>El gato de Juan ha de ser más parecido a un martillo, un desarmador o una grúa que a un perro o un hámster.</p> | <p>Martillo: una herramienta que... Desarmador: una herramienta que... Grúa: una herramienta que... Gato 1: un animal doméstico peludo. Gato 2: una herramienta que...</p> <p>De las dos acepciones de gato, la segunda es la que más se parece a martillo, desarmador o grúa. ¡Ya sabemos cuál gato!</p> |

Figura 5: Ejemplo 1 Modelo probabilístico, usa un diccionario y usa textos precargados

Así tiene un conjunto de este es el corpus todos estos son el corpus con la palabra gato y algunas se refieren al gato animal y la otras al gato herramienta.

Entonces el gato de Juan ha de ser más parecido a un martillo a un desarmador o a una grúa que a un perro o un hámster, entonces tiene todo este conjunto de datos esa es la primera etapa que decíamos ya que tiene todo este conjunto de datos todo el corpus empieza a ver todas las frases cómo se están formando, cómo se van formando cada una de las frases y con qué se tiene que buscar texto y sea cual es la que más se asemeja a este contexto de esta frase.

Entonces tienen que analizar pues cuál de los dos gatos es el que más se asemeja y tiene que encontrar que el más parecido es el gato que esa herramienta entonces ya formando un diccionario dice martillo es una herramienta desarmador es una herramienta grúa una herramienta gato uno animal doméstico gato 2 es una herramienta.

Entonces pues al parecer aquí lo que más se asemeja al contexto que él tiene que es esta frase es que es una herramienta entonces dice de las dos acepciones de gato la segunda es la que más se parece a un martillo a un desarmador o una grúa entonces el gato que corresponde a esta frase es el gato de herramienta.

Entonces así es como trabaja el aprendizaje automático

Ejemplo 2: traductor de Google

Según el director Peter Norvig comenta que ellos usan el aprendizaje automático a partir de ejemplos entonces utiliza un modelo probabilístico

Dice en el inicio del proyecto los programadores intentaron enseñarle el sistema reglas gramaticales y de vocabularios: dice esto es un sustantivo esto es un verbo y este es el orden es decir lo primero que intentaron en google fue trabajar con el modelo de reglas gramaticales o sea basado en una gramática usa técnicas lingüísticas formales y se dieron cuenta que no era tan sencillo trabajar con estas reglas gramaticales porque el lenguaje es tan fluido y extenso que los programadores no pudieron con las millones de palabras y los billones de combinaciones que se pueden dar entre ellas y sobre cómo cambian con el tiempo.

Es decir se tiene una regla gramatical pero el lenguaje es un lenguaje vivo un lenguaje activo y el lenguaje se va actualizando y hay nuevas palabras, entonces ya no se encaja en la misma gramática y eso hace mucho más difícil. Una frase puede decir “El niño corre” o puede decir “Corre el niño” y ambas formas está correcto de decir; entonces se dan muchas probabilidades que yo puedo tener y es difícil encajar se a una sola gramática.

Entonces el proyecto del traductor de google primero trato de hacerlo con reglas gramaticales dice este enfoque no era el adecuado porque tenía dos problemas:

1. Las reglas del idioma son absolutas o sea la gramática es absoluta pero el lenguaje está muy lejos de ser absoluto y está lleno de excepciones entonces esto es lo que les decía este hay nuevas palabras.
2. El lenguaje demás es complejo lleno de excepciones también es cambiante con nuevas palabras que se agregan todo el tiempo y reglas que cambian para dar paso al lenguaje día a día el lenguaje es un ente vivo entonces se está actualizando constantemente entonces la regla el gramatical que tenemos ya no funciona.

Entonces lo que usaron un nuevo enfoque basado en datos, el primer reto fue lograr tener suficientes ejemplos es decir ellos dejaron el proyecto de la regla y emprendieron el proyecto a través de los métodos probabilísticos y lo primero que buscaron es ese corpus, o sea, un conjunto de ejemplos.

Para tener ese conjunto de ejemplos dice se programaron algoritmos para buscar en la web millones de ejemplos de traducciones en el mundo real y que analizara todos esos datos para encontrar los patrones acerca del modo en que las frases completas son traducidas a estas frases.

Es decir buscaron todas las traducciones que había en internet y cómo se estaban realizando se utilizó el aprendizaje automático para buscar patrones que se repitieran.

Por ejemplo “Una frase tal en inglés aparece que se traduce en esta otra frase al español pero sólo si está cerca de esta otra palabra”. Entonces ellos fueron buscando de acuerdo a los diferentes contextos así ellos pudieron elaborar este traductor de Google.

Trabaja con base en probabilidades y que no es cien por ciento exacto pero tiene es una muy buena herramienta.

Sin embargo pues tiene un porcentaje de exactitud entonces así es como trabaja porque no lo pudieron implementar con base en las gramáticas.

También está el problema de la segmentación de los textos dice yo tengo este problema de terminar las palabras de un texto en el que no hay espacios en blanco entonces imagínense esta frase que yo tengo aquí está este tipo de problemas dice se aplica en la traducción del japonés y el chino en dónde pueden haber varias palabras sin tener espacios en blanco.

Entonces lo que se quiere es identificar las palabras que yo tengo aquí es el objetivo es obtener la frase original con un cierto grado de confianza que es fácil leer frases sin espacios en blanco entonces dice este proceso se puede llevar a cabo mediante un modelo n-gram.

n-Gram trabaja una secuencia de palabras: $p_1 p_2 \dots p_n$ calculando sus probabilidades de aparición en la frase final dado un conjunto de palabras como corpus, es decir se trabaja con un diccionario de palabras y va analizando la frase.

La probabilidad de cada palabra se calcula como la frecuencia relativa de aparición de dicha palabra en el corpus.

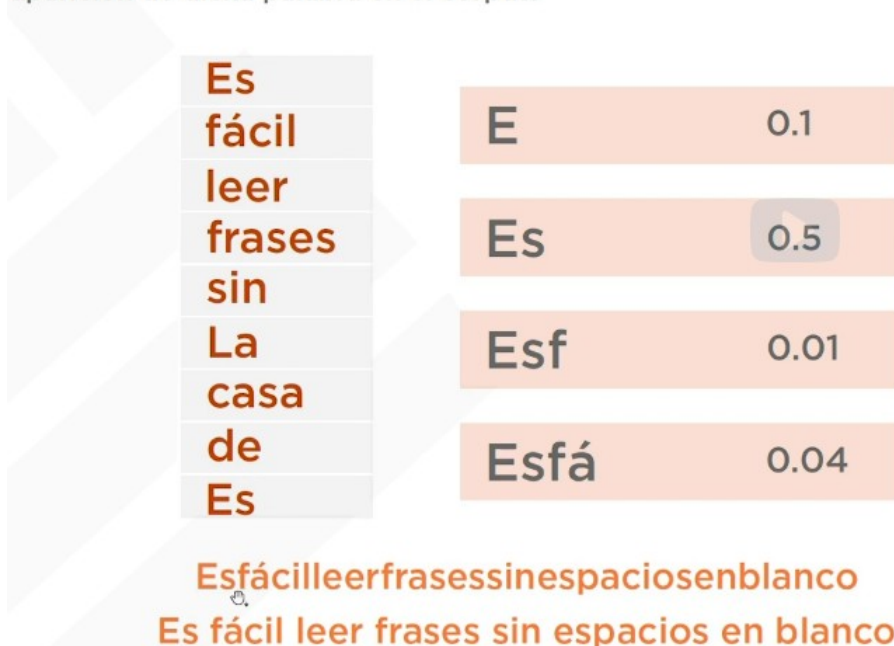


Figura 6: Traductor: mide las frecuencias de aparición de letras y frases

Entonces voy sacando y voy separando palabras yo tengo la e y la que aparece tantas veces aquí luego leo la s y digo ese aparece tantas veces aquí luego leo la f y digo que ese f aparece tantas veces aquí.

Luego leo la siguiente que es la y luego que eso aparece tantas veces aquí y entonces yo voy formando palabras o estructuras.

Por ejemplo aquí es tiene una alta probabilidad de aparición entonces yo empiezo a leer y digo en ese entonces ahora consulto con toda mi lista de palabras que yo fui identificando y veo que es tiene una alta aparición o sea tiene una alta probabilidad entonces la saco como una palabra dada su probabilidad alta como una palabra que es dentro de esta frase y entonces ya tengo la e y la s y

ahora que sigue sigue la f y así voy a seguir buscando y voy a buscar una palabra fácil y voy a buscar la que tiene un alto porcentaje de probabilidad de aparición porque en todo este corpus aparece con una alta probabilidad.

Entonces ya tengo otra palabra y así voy a ir analizando dada esta probabilidad las diferentes palabras para lograr separar esta frase que no tiene espacios.

El aprendizaje automático en el procesamiento del lenguaje natural (Intro SVN)

Dando un recordatorio acerca del aprendizaje automático; bueno el aprendizaje automático es importante señalar que la máquina se programa para que extraiga patrones de comportamiento, es decir, tú le das un conjunto de elementos y va a extraer los patrones nos va a decir cómo tienden a comportarse eso a partir de las entradas recibidas es decir de todos los datos que se le da y con base a esta información con los patrones que se van identificando realizar la evaluación de las nuevas entradas.

Entonces también tenemos que recordar que hay dos tipos de aprendizaje el aprendizaje supervisado y el no supervisado.

- En el aprendizaje **no supervisado** se realiza el procesamiento con base únicamente en las entradas.
- En el aprendizaje **supervisado** se cuenta con un corpus o un conjunto de elementos de entrada que ya están clasificados con sus fases de entrenamiento y de prueba.

Bien esto está relacionado con todos qué es el aprendizaje automático y ya hemos hablado que el aprendizaje supervisado va a recibir un conjunto de elementos para estar entrenando entrenando genera el modelo y lo prueba entonces así es cómo funciona el aprendizaje automático

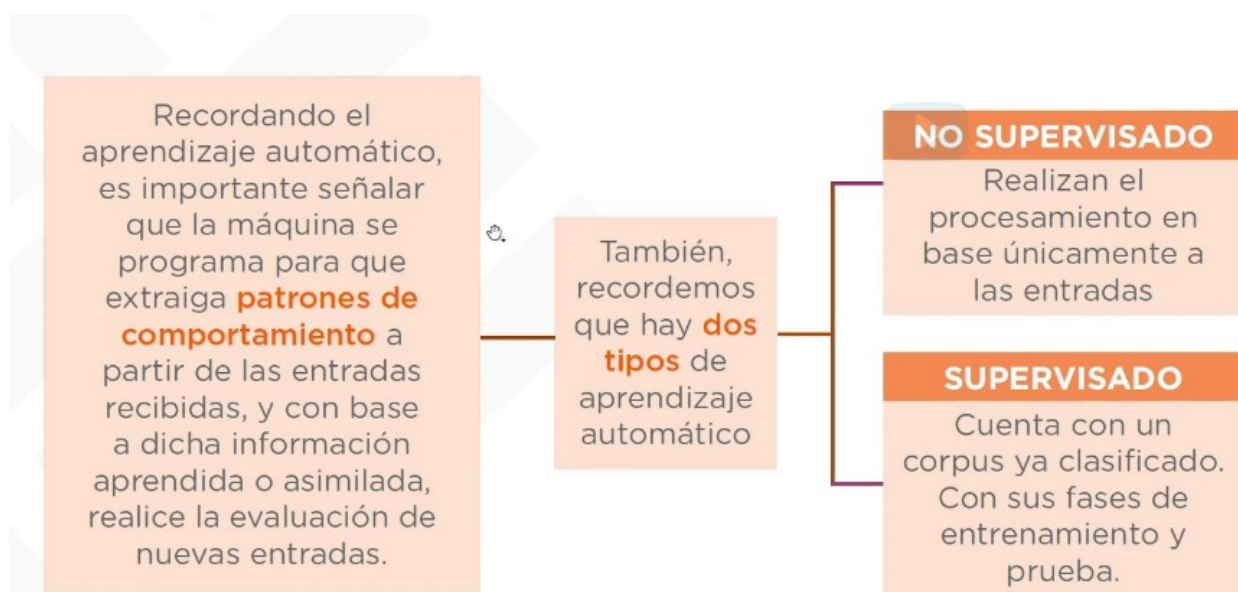


Figura 7: Aprendizaje automático

Aquí tenemos precisamente un esquema de este aprendizaje automático aquí tenemos las dos etapas

- a) **el training o el entrenamiento.** Entonces yo tengo entradas tengo varios elementos en la etapa de entrenamiento voy a estar dándole entradas entradas entradas entonces introduzco un elemento que va a extraer una característica la va a dar de alta en todas las características de este nuevo elemento porque hay otros elementos que ya pasaron por aquí y que tienen diferentes características y todas estas características van a ir alimentando mi modelo de máquina de aprendizaje.

Entonces todo este algoritmo va a estar trabajando y se le llamamos que estos datos ya están es decir todas estas entradas ya son parte del modelo porque se ha entrenado con el entonces se van interpretando viene otro elemento que es también de entrenamiento se extraen las características de este nuevo elemento se agregan o se confirman algunas características que ya había entrado al modelo al algoritmo aquí está trabajando y trabajando

Todo esto va a ser entrenamiento que es estar revisando todos los elementos de entrada para ver qué características tienen y eso lo hace el algoritmo.

Entonces éste algoritmo que está trabajando aquí va a ir formando el modelo de clasificador es decir yo voy a suponer que tengo un árbol de decisión un árbol de clasificación aquí voy a estar entrenando con una serie del corpus que son los datos de entrada que va a estar extrayendo características el algoritmo está trabajando está etiquetando es decir está trabajando con todas estas entradas va a generar el modelo.

- b) hay que aquí se le llama **predicción** pero viene siendo la prueba del modelo la aplicación del modelo para probarlo.

Aquí ya se generó el modelo, éste modelo se generó base en todas estas entradas con todos los documentos que fueron parte del corpus es decir con todo el entrenamiento que se le dio se genera este modelo y ahora sí ya que tengo yo este modelo lo voy a aplicar con un dato nuevo es decir un dato que no fue parte de la etapa de entrenamiento es decir no está etiquetado entonces en la predicción yo voy a tener una nueva entrada voy a extraer una característica tengo otras características que ya están asociadas y las voy a meter al modelo entonces el modelo lo va a caracterizar o le va a decir de qué tipo es por qué por qué ya conoce muchos otros ejemplos que vio acá en la etapa de entrenamiento.

Finalmente lo vuelve a clasificar lo vuelvo a etiquetar para que sea que un elemento de entrada más a este nuevo algoritmo entonces así es como trabajan en términos generales estos algoritmos.

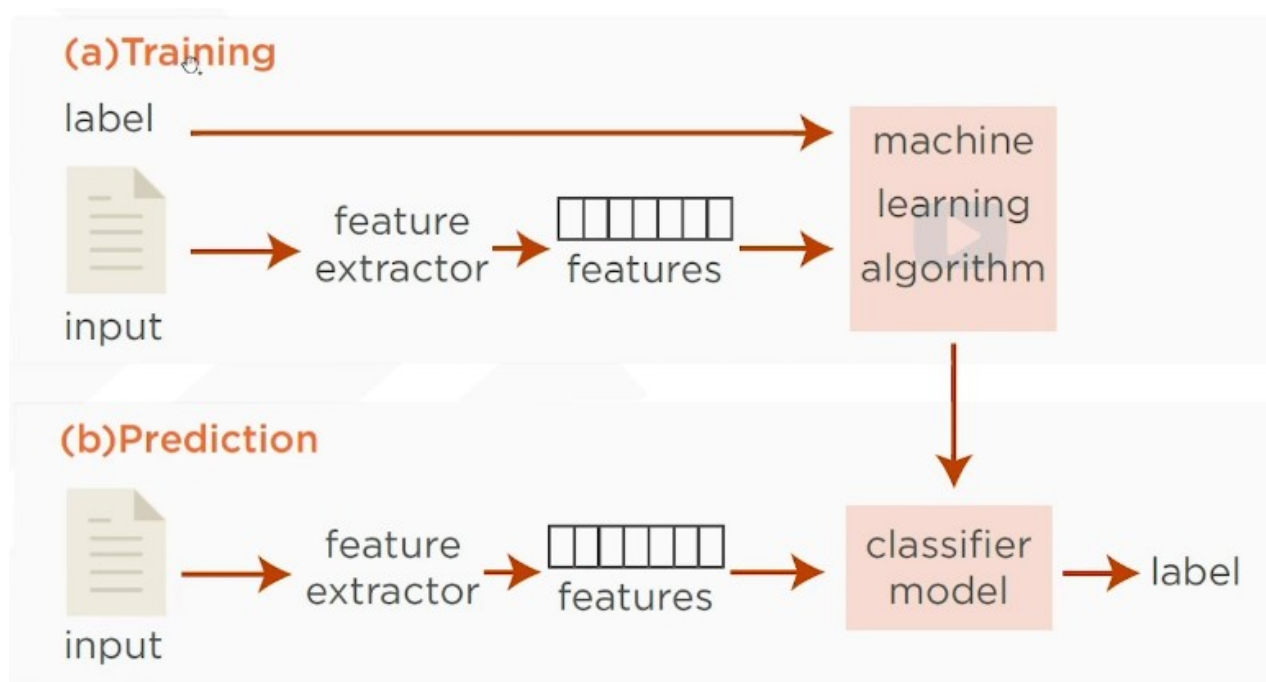


Figura 8: Funcionamiento de machine learning

Ejemplo: Problema de detección de Spam

Por ejemplo el problema de detección de dice las principales técnicas que se utilizan son las máquinas de vectores de soporte (supported vector machines SVM), Naive Bayes y clasificadores de máxima entropía

- Se utiliza la categoría gramatical de las palabras, la presencia y la frecuencia de algunos términos y su composición semántica.
- Sin embargo, la mayoría de éstos métodos van acompañados de un diccionario que entrega información a priori de los términos para obtener las polaridades respectivas. En algunos casos, éstos diccionarios son realizados por personas y en otros, se ocupa un sistema automático.

Support Vector Machine

- Tiene aplicaciones en el procesamiento del lenguaje natural, el habla, el reconocimiento de imágenes y visión artificial. Además son utilizadas en la detección y clasificación de sentimientos
- Se basa en **métodos kernel**, los cuales toman datos y los ponen en un espacio de características apropiados
- Se basa principalmente en vectores donde, usando aprendizaje computacional, logra tomar decisiones de límite entre 2 categorías separándolas lo más posible.

Yo tengo de este plano que hablaba en este espacio aquí tengo vectores que son estas líneas que yo tengo aquí por eso se le llama máquinas de vectores de soporte.

Entonces yo tengo estos puntos (Figura 9) y estos triángulos aquí imaginémonos que esto es un conjunto de documentos con unas ciertas características y que yo los he graficado en el espacio y tengo todos estos aquí y acá tengo otro conjunto de características que me salieron de los documentos.

Entonces tengo agrupados estos puntos porque es una clase de características que yo tengo aquí y tengo agrupados estos otros puntos porque es otra clase de características

Entonces yo hago un vector que me va a representar todos estos puntos y hago otro vector que me va a representar todos estos puntos y este vector representa esta clase y este vector representa a esta otra clase.

Entonces entre las dos clases una distancia y yo voy a generar un vector en este plano que es el que yo tengo aquí que me dé el **margen máximo** o sea la distancia mayor entre cada uno de ellos que es lo que yo tengo aquí y con este vector ahora en este plano que se está formando aquí con las 2 clases: la clase la clase de triángulos y la clase de círculos negros

Yo voy a tener un nuevo dato o sea un nuevo documento con una nueva característica y el algoritmo lo que va a decir es que si está en una clase o en otra

A eso se refiere esta explicación de máquinas de vectores de soporte que tenemos aquí entonces así es cómo funcionan estas máquinas

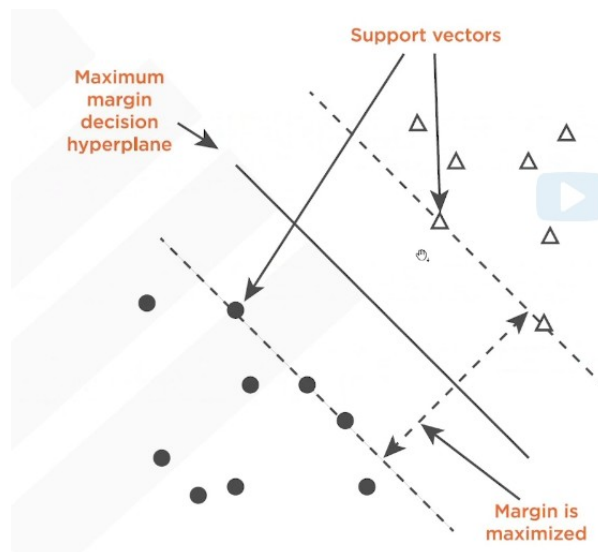


Figura 9: Support Vector Machine

Aplicaciones de las técnicas de procesamiento de lenguaje natural

Aplicaciones de la IA

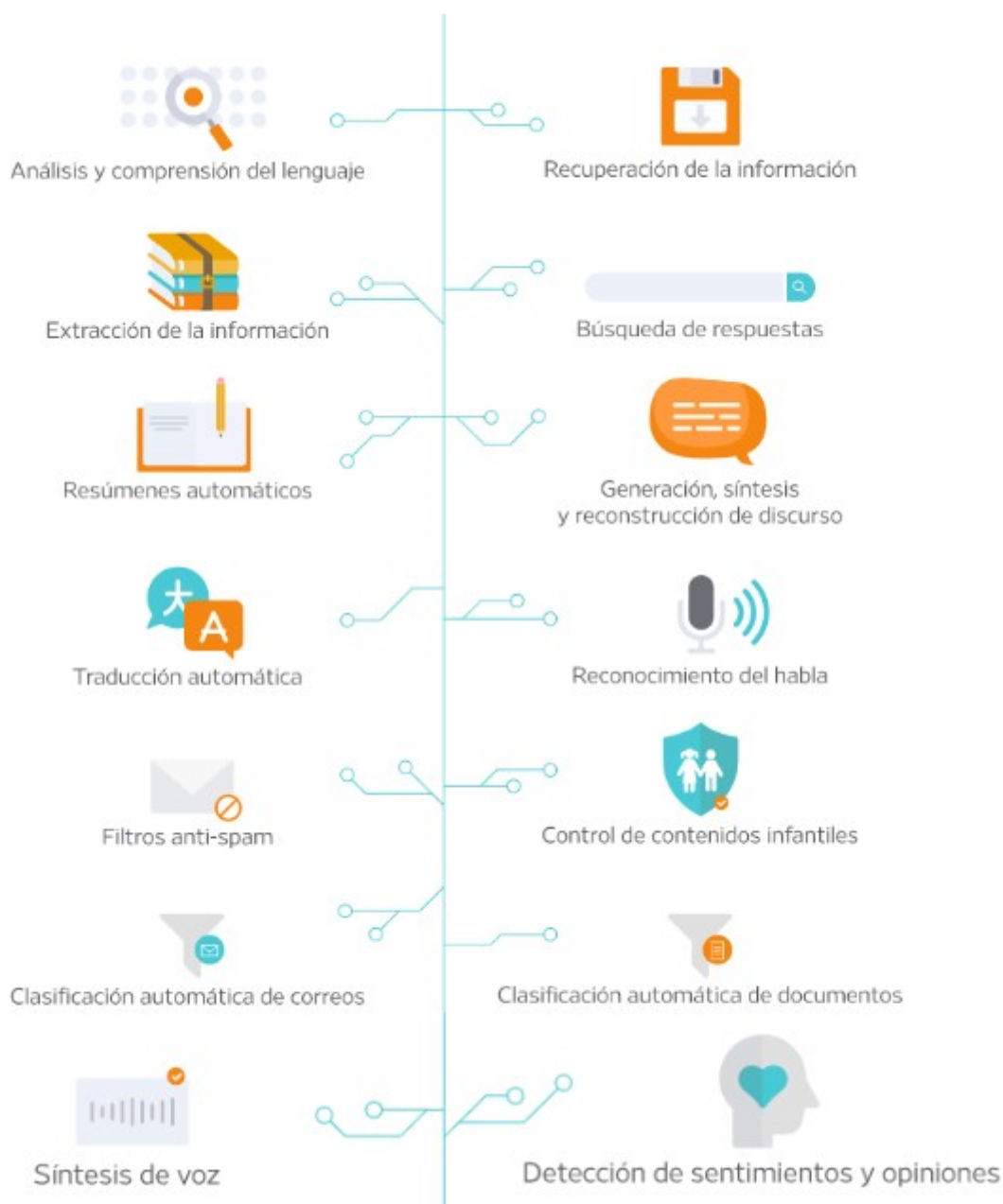


Figura 10: Aplicaciones de la Inteligencia Artificial

Introducción al análisis de sentimiento en Twitter

En la presente sección, se muestra una introducción al procesamiento del lenguaje natural utilizado para realizar un análisis de sentimiento en Twitter.

Se describe esta red social y se muestran los detalles que se tienen que considerar al realizar el proceso de análisis.

El análisis consiste en determinar la Polaridad de un mensaje, que puede ser con sentimiento positivo, negativo o neutro. Además, se puede identificar la intensidad y emoción del mensaje.

Twitter es una de las plataformas sociales más importantes en el mundo, en la cual sus más de 326 millones de usuarios pueden escribir sobre cualquier tema en sólo 280 caracteres de extensión del mensaje.

Los usuarios describen en forma simple, sus vidas, opiniones o discusiones; y por otra parte, las empresas lanzan campañas publicitarias, describen sus productos y anuncian sus eventos.

Para identificar las opiniones en Twitter, es necesario realizar un análisis de sentimientos, técnica que utiliza procesamiento de lenguaje natural, análisis de texto y herramientas computacionales para clasificar comentarios subjetivos de diferentes usuarios, ya sean sentimientos propiamente u opiniones sobre diversos temas.

Los métodos usados para este tipo de análisis tienen cerca de 15 años de aplicación, los cuales se han usado para clasificar correos, reseñas de clientes o publicaciones digitales, entre muchos otros

Algunas cuestiones que debes considerar para realizar un análisis de sentimiento en twitter, son determinar si existe opinión en el tweet o no y reconocer las abreviaciones y modismos típicos como las tildes.

El análisis consiste en determinar la polaridad de una oración pudiendo tener palabras positivas y negativas en la misma frase (Ej. “Me alegro que se haya terminado, pésimo el espectáculo”, “La película no fue nada buena”).

Lo que se mide son 3 cosas:

- **Polaridad:** indica si el mensaje tiene un sentimiento positivo, negativo o neutro.
- **Intensidad:** proporciona un valor numérico en relación con la intensidad del sentimiento.
- **Emoción:** clasifica el texto según los distintos tipos de emociones, como puede ser la alegría, la tristeza o la ira.

El algoritmo para realizar el análisis de sentimiento, debe contar con la fase de aprendizaje, que consiste en tener un corpus o conjunto de mensajes clasificados para generar el modelo.

Por esto, se divide este corpus en dos conjuntos: uno más grande corresponde al conjunto de entrenamiento y el otro se llama conjunto de prueba.

El primero se dedica propiamente al aprendizaje del algoritmo, mientras que el segundo sirve para calcular la precisión del algoritmo ya entrenado.

Algunas acciones son necesarias en la etapa de preprocesamiento de los textos:

- Eliminar las URL del mensaje
- tokenizar o extraer las palabras del tweet
- pasar a minúsculas
- quitar las palabras vacías, que no aportan información de sentimiento como los artículos y las preposiciones.

Por otro lado las abreviaciones, las negaciones (como “No tengo frío”) y los emoticones que tienen los textos de los mensajes, se les debe dar un tratamiento especial.

También reciben un análisis específico los intensificadores y reductores. Estos términos son los que hacen reducir o aumentar el sentimiento total del mensaje a analizar.

Estos, no siempre son fáciles de identificar y pueden variar dependiendo del contexto.

Por ejemplo, palabras como “muy”, “más”, “bastante” intensifican el mensaje. Y “poco”, “menos”, “casi”, lo reducen.

Asimismo, las palabras en mayúsculas y signos de exclamación también enfatizan los mensajes. Como ves, se tiene que considerar muchos detalles. Recuerda las dificultades al procesar el lenguaje, en este caso, existe la dificultad del análisis, ya que identificar la polaridad o la emoción no siempre resulta exacto al 100%.

Instalación de Rapid Miner

Follow the video: <https://learning.edx.org/course/course-v1:AnahuacX+IEST-CP1-2x+3T2019/block-v1:AnahuacX+IEST-CP1-2x+3T2019+type@sequential+block@a0d0528ac6494d18a45896062c14f99c/block-v1:AnahuacX+IEST-CP1-2x+3T2019+type@vertical+block@63fcf1083b864df0aecbe192e312d47f>

Tutorial de RapidMiner

Follow the video: <https://learning.edx.org/course/course-v1:AnahuacX+IEST-CP1-2x+3T2019/block-v1:AnahuacX+IEST-CP1-2x+3T2019+type@sequential+block@a0d0528ac6494d18a45896062c14f99c/block-v1:AnahuacX+IEST-CP1-2x+3T2019+type@vertical+block@c193cfc1d0e3488386eea6600a861010>

Árboles de clasificación

Introducción a los árboles de clasificación

Vamos a iniciar la sección de árboles de clasificación.

Los árboles de decisión, representan uno de los modelos de minería de datos más comunes y estudiados son fáciles de construir y de interpretar.

Los árboles de clasificación son un tipo de árboles de decisión, que clasifican los datos del conjunto de entrada tomando como base una variable clasificadora categórica, es decir que toma un conjunto finito de valores esto es, por ejemplo, que puede valer solamente si o no.

Las aplicaciones de árboles de clasificación son muy variadas, por ejemplo, se utilizan en las predicciones de siniestros en abandono de clientes en logro de metas de producción o bajas de alumnos en el sector educativo.

Los temas que revisaremos en esta sección son los siguientes:

- Introducción a los árboles de clasificación
- La técnica de los árboles de clasificación

- Aplicaciones de los árboles de decisión
- Ejercicio práctico de clasificación.

Introducción a los árboles de decisión – clasificación

- La modelación predictiva permite descubrir patrones de comportamientos de clientes, productos, servicios y competidores para generar estrategias que permitan tomar decisiones basadas en los conocimientos generados.
- La predicción de estos patrones se logra mediante la aplicación de algoritmos y técnicas como los árboles de decisión.
- La modelación predictiva se puede implementar mediante la minería de datos con el apoyo de este software especializado.

El árbol de decisión es el modelo de clasificación más utilizado debido a su enorme potencial para la predicción y a su simplicidad y facilidad de entendimiento, mediante estos árboles se pueden construir modelos que se interpretan analizando las diferentes variables que están consideradas en este modelo en las diversas ramificaciones del árbol.

En esta unidad se analizará el funcionamiento de estos árboles de decisión y se utilizará el software especializado RapidMiner para realizar una práctica del diseño de un modelo predictivo basado en un árbol de decisión para proponer el diseño de un esquema de mejora del proceso que apoye a las decisiones estratégicas.

Los temas que se desarrollarán en esta unidad son los siguientes: Introducción a los árboles de decisión y Aplicación de los árboles de decisión con el software RapidMiner.

Los objetivos específicos son:

- Comprender el algoritmo de árboles de decisión para identificar las oportunidades de su aplicación
- Utilizar el software RapidMiner para diseñar modelos predictivos que mejoren los procesos
- Realizar una práctica con el software RapidMiner con la técnica de árboles de decisión para diseñar estrategias que permitan lograr una ventaja competitiva.

Recuerda la siguiente frase de Lee Lacocca: "Incluso una decisión correcta es incorrecta cuando se toma demasiado tarde."

En este mundo de alta competencia las empresas deben tomar decisiones en forma constante, rápida y eficaz.

La predicción, como los árboles de decisión le permiten a las empresas tomar decisiones basadas en datos sobre predicciones del futuro lo cual los pondrá a la vanguardia en el mercado.

El software especializado en minería de datos, como el RapidMiner, les permitirá disminuir drásticamente el tiempo de generación de los modelos por lo que podrán tener predicciones en muy corto tiempo.

Los árboles de decisión – clasificación

El árbol de decisión se conoce como CART (*Clasificación And Regression Tree*) es un tipo de aprendizaje de máquina se utiliza para predecir valores.

Hay 2 tipos de árboles de decisión:

- **Árboles de clasificación:** Clasifican los datos del conjunto de entrada en función de una variable clasificadora categórica, es decir, que toma un conjunto finito de valores.
- **Árboles de regresión:** Igual que el anterior solo que la variable de clasificación es continua hablaríamos.



Figura 11: Ejemplo de árbol de decisión, se ve que cosas caen en las sentencias si/entonces

Los árboles de decisión son uno de los modelos más utilizados en la práctica su interpretación y evaluación también es sencilla, pueden combinar variables numéricas y categóricas en el mismo modelo.

Su implementación práctica se reduce a una serie de reglas que pueden ser fácilmente escritas como un conjunto de sentencias del tipo si/entonces.

Los árboles de decisión son modelos de minería de datos, y en su funcionamiento subdividen los datos para generar grupos separados de datos de forma que **todos los que pertenezcan a un mismo grupo sean de la misma clase**, la cual es utilizada como de dicho grupo o región.

Si una región contiene datos de diferentes clases es a su vez subdivididas en subgrupos o subregiones más pequeñas siguiendo el mismo criterio hasta separar todo el espacio de entrada en regiones y subregiones separadas que solamente contienen elementos de una misma clase.

Usos:

Los árboles de decisión se utilizan en sectores como el bancario y las compañías aseguradoras para tomar decisiones respecto sobre la concesión de créditos o el cálculo de las pólizas dado que permiten determinar qué características relativas a los usuarios son las de mayor o menor riesgo, siendo posible segmentar a los usuarios en función de dichas características.

Es importante considerar que cuando se realiza una práctica de creación de un modelo de predicción con árboles de decisión se deben llevar a cabo diferentes etapas en forma ordenada como:

- el diseño del modelo la evaluación
- la implementación

Las herramientas especializadas para llevar a cabo estas etapas son muy útiles ya que ahorran una gran cantidad de tiempo

La técnica de árboles de clasificación

Introducción a conceptos de árboles de decisión y clasificación

Términos

- **Árboles de decisión:** decisión que dice que es uno de los modelos de minería de datos más comunes y que más se han estudiado porque tiene una alta capacidad para explicarse

La forma en que nosotros podemos interpretar este árbol es sencilla entonces es la forma en que funciona un árbol de decisión nos permite realizar una fácil interpretación de este modelo que se ha generado y gracias a eso pues es muy utilizado este modelo de árboles de decisión ahora bien hablamos de árboles de decisión.

- **Árboles de clasificación:** son árboles de decisión donde se clasifican los datos del conjunto de entrada en función de una variable categórica es decir una variable que toma un conjunto finito de valores.

Entonces un tipo de árbol de decisión es el árbol de clasificación y la característica principal de un árbol de clasificación es que actúa como un árbol de decisión nada más que la variable se va a clasificar, la variable que se va a predecir es una variable de tipo categórica que puede valer un conjunto de valores.

Vamos a suponer que “sobrevivió” es una variable categórica las respuestas a sobrevivió a los posibles valores de sobrevivió es sí o no.

Entonces es un conjunto finito de valores sí o no entonces y es de tipo categórica por el tipo de dato entonces estos son los árboles de decisión en particular árboles de clasificación

- **Árboles de regresión:** Consisten en un árbol de decisión en que la variable clasificadora es continua es decir la variable que yo voy a predecir en este modelo no es categórica sino que es una variable continua, es decir puede tomar valores 1, 1.1, 1.2, 1.3, 1.4, 5, 10, 15.

Esa es la variable clasificadora, entonces tu puedes predecir, por ejemplo, las ventas o el número de fallas, entonces aquí pues son variables de tipo continuas

Entonces los árboles de decisión son un modelo de minería de datos que intenta **subdividir el espacio de datos de entrada** para generar regiones distintas de forma que **todos los elementos que pertenecen a una misma región sean de la misma clase**, la cual es utilizada como representante de dicha clase.

Entonces yo tengo un conjunto de datos aquí de entrada a este modelo y trata de dividir, dice regiones distintas entonces esta entrada de datos trata de separarla. Hay ciertas variables que te dicen si las separas o si no las separas.

Entonces lo que trata el árbol es de ir separando, entonces aquí tengo un gran conjunto de datos que dentro de ese conjunto de datos tienen diferentes características entonces, lo empiezo a separar por ejemplo en dos o en tres elementos, también es otra variable cuántas ramas, cuántas separaciones haces aquí, a cuántas ramas le pones a este árbol.

Entonces este dato de entrada tú lo vas a separar en diferentes subconjuntos de datos que le estamos llamando generar regiones disjuntas, de tal manera que todos los elementos de esta rama pertenezcan a una misma clase

Por ejemplo si yo tengo aquí una entrada de datos, puedo separar en hombres y en mujeres por ejemplo entonces todos se van a venir para una rama todos los que tienen la característica de que son hombres y para otra rama todos los que tienen la característica de que son mujeres

Entonces básicamente eso es lo que contiene este árbol de decisión dice si una región contiene datos de diferentes clases es subdividida en regiones más pequeñas entonces por ejemplo vamos a suponer aquí tengo los datos de entrada.

Dijimos aquí voy a poner los hombres aquí las mujeres si aquí dentro de los hombres todavía hay otras características a lo mejor de la edad en donde la edad me puede decir que que es muy variable o sea unos adultos mayores hay unos niños por ejemplo entonces yo puedo subdividir entonces voy a particionar y así **voy a ir particionando y particionando mientras haya variables diferentes** claro que hay otro dato que me va a decir en qué momento o cuando me conviene ya dejar de particionar este conjunto de entrada.

O sea:

- Si una región contiene datos de diferentes clases, es **subdividida en regiones más pequeñas** siguiendo un criterio
- **Se sigue hasta particionar todo el espacio de entrada en regiones disjuntas que contengan elementos de una misma clase**
- **Un árbol de decisión es una secuencia de condiciones que son interrogadas con respecto a los datos de entrada tomando una decisión parcial que lleva hacia una rama u otra repitiendo este proceso hasta llegar a una hoja donde se toma una decisión final** entonces esto es en conclusión yo tengo un conjunto de datos el cual voy a tomar decisiones.
- **Profundidad máxima de un árbol de decisión:** Ahora dice la profundidad máxima de un árbol de decisión es el máximo número de condiciones que es necesario resolver, es entonces cuántas preguntas voy a hacer aquí para hacer una partición esa es la profundidad

que me va a dar este árbol. O sea, es el máximo número de condiciones que es necesario resolver para llegar a una hoja.

Estructura de árboles de clasificación

Las estructuras son 3:

1. **Nodos hoja o terminales:** Representan regiones etiquetadas o clasificadas de acuerdo a una clase
2. **Nodos internos o splits:** Representan condiciones que permiten decidir a qué subregión va cada elemento que llega a dicho nodo, es decir, la partición realizadas
3. **Nodo raíz:** Es el nodo superior. Cuando se presenta un nuevo dato a un árbol, se empieza por el nodo raíz, que contiene una condición por la cual determina por qué rama del árbol debe descender (o sea, a qué subregión pertenece el dato) hasta alcanzar la siguiente condición o bien a una hoja terminal.

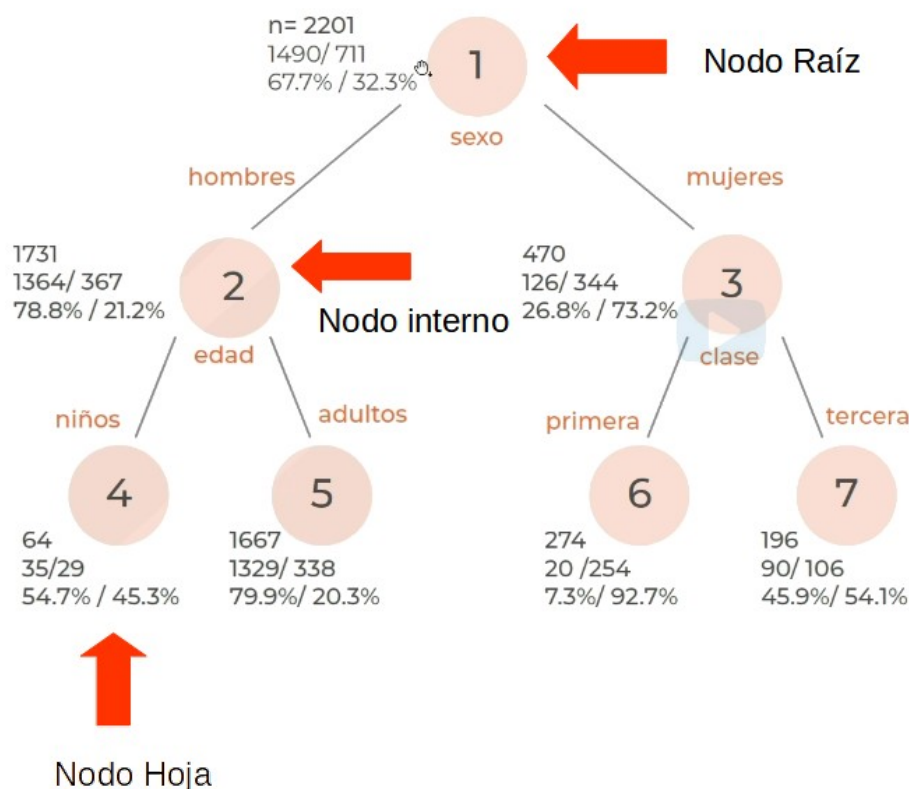


Figura 12: Ejemplo de árbol de clasificación según supervivencia de embarcados en el titanic. donde se indican los tipos de nodos.

En la figura 12 el nodo 2 es un sub-árbol que yo pudiera considerar que el nodo 2 es el nodo raíz de los nodos 4 y 5 pero como tal todo esto es una rama entonces todo esto es es un nodo intermedio de todo el árbol general.

Si quisiera analizar nada más esta parte del árbol entonces tiene la misma estructura del nodo general tiene un nodo raíz y tiene otros nodos y tiene una variable que lo divide que por ejemplo

edad y para acá tengo niños y para acá tengo adultos, entonces este es el nodo raíz de todo el árbol. luego dice

Tengo nodos internos o splits representan condiciones que permiten decidir a qué supresión va a cada elemento que llega a dicho nodo entonces yo dije que este era mi nodo raíz y estos son los nodos que yo voy a irme encontrando en el camino.

En las nodos 4, 5, 6 y 7 como ya no hay subdivisiones a esto se les llama nodos hoja.

Como lo tenemos en la figura 12 de supervivencia de los pasajeros del Titanic, entonces supongamos que en total eran 2021 pasajeros entonces estos son los datos. 1490 no sobrevivieron y 711 si sobrevivieron entonces no sobrevivieron, que es el 67.7%, y los que si sobrevivieron el 32.3%.

Entonces de éstos yo hago una primera partición de acuerdo al sexo al nodo 2 son hombres y al 3 los que son mujeres.

Supongamos que hombres fueron en total 1731 ahora, bien de éstos no sobrevivieron 1364 y si sobrevivieron 367 y hacia el otro lado yo tengo visión de mujeres y en total fueron 470 mujeres no sobrevivieron 126 y si sobrevivieron 344.

Así yo voy partición ando acá en las mujeres partición o por clase y en los hombres puedo particionar por edad entonces de niños cuántos niños fueron 64 niños aquí están los que no sobrevivieron y los que si sobrevivieron.

Acuérdense que la variable categórica aquí o sea la que estoy prediciendo es la variable de sobrevivió o no sobrevivió.

Acá tengo a los adultos que fueron mil 667 en total no sobrevivieron mil 329 y si sobrevivieron 338.

En las mujeres tengo una división por clase o sea en qué nivel estaban viajando en primera o en tercera clase entonces yo tengo aquí que 196 se fuera una tercera clase y 274 se fueron la primera clase y de estas de las 274 no sobrevivieron 20 pero si sobrevivieron 254 que es el 92.7 %

Entonces si observamos aquí en estos nuevos terminales, quienes tienen el mayor porcentaje de supervivencia aquí tenemos el 45.3 el 20.3% que son este hombres adultos sobrevivieron 20.3 hombres niños.

Por otra parte sobrevivieron 45.3% de mujeres en clase tercera sobrevivieron 54.1% mujeres en clase primera sobrevivieron 92.7% entonces las que más probabilidades tienen de sobrevivir son las que son mujeres y que viajan en clase de primera.

Entonces así es como se va analizando este árbol de decisión en particular árbol de clasificación.

Criterios para la construcción de un árbol

- **Cp o criterio de parada:** Determina en qué momento se deja de seguir seleccionando nodos a ser subdivididos

- **Cs o criterio de selección:** Determina qué nodo es seleccionado para ser particionado en 2 o más subnodos.
Entonces yo voy a dividir y voy a decir este nodo voy a dividir pero tengo que definir bajo que el criterio lo voy a dividir.
- **Cc o criterio de clasificación:** Determina qué clase se asigna a un nodo hoja. Normalmente se trata de la clase que minimiza el error de clasificación
- **Cd o criterio de partición:** Determina cómo se particiona un nodo en 2 o más subnodos. Normalmente los árboles de decisión son binarios ($P=2$ siendo P las partes), indicando que cada región se subdivide en 2 regiones disjuntas.

Entrenamiento y prueba de los árboles

Construcción de modelos de árboles de clasificación

Lo primero que hay que construir es el **modelo de árbol**. El objetivo es asegurar que los modelos construidos a partir de los datos disponibles funcionen correctamente para nuevos datos que haya que procesar en un futuro, es decir, asegurar que el modelo es válido y capaz de ser usado en la etapa de producción

Para esto hay que:

1. **Preparar los datos:** Quiere decir que si tengo las tablas de datos pero no siempre son adecuadas para generar inmediatamente el modelo sino que tienen que tener algún tratamiento.

Por ejemplo quitar los valores nulos, los vacíos, los blancos seleccionar las tablas las columnas o variables que yo quiero que entren al modelo, eliminar las que no necesito. Entonces hay una serie de operaciones.

2. **Entrenar los datos:** Se que llevar a cabo una operación de entrenamiento que consiste en ir aplicando el modelo con varios conjuntos de datos para ir ajustando este modelo.

Voy a generar un modelo con un conjunto de datos entonces esos datos se les llama datos de prueba.

Una vez que ya tengo el modelo ya digo este es mi modelo lo voy a aplicar a nuevos datos que tienen que ser datos diferentes a los datos que yo utilicé.

Entonces normalmente en esta etapa de entrenamiento yo toda la tabla la dividida en dos partes que son datos de prueba y datos nuevos. Entonces con los datos de prueba hago el entrenamiento para generar el modelo y con esos datos nuevos aplico ese modelo.

Se debe evitar que los datos de prueba sea dependiente los datos utilizados durante su entrenamiento evitando el problema conocido como **sobre-entrenamiento**.

Este problema consiste en que el modelo sólo responda estrictamente a las propiedades del conjunto de datos de entrenamiento y que sea incapaz de aplicarse con niveles de acierto adecuados a otros nuevos conjuntos de datos que pueden aparecer en un futuro

3. **Probar el modelo:** Para probarlo pues le voy a aplicar el modelo nuevos datos diferentes a los que yo usé en el entrenamiento entonces pruebo el modelo
4. **Evaluar el modelo:** Se evalúa con una matriz de confusión donde hay que ver qué tanta exactitud tiene mi modelo
5. **Aplicar el modelo**

Proceso de creación y validación del modelo

Para aprendizaje supervisado es necesario contar con un conjunto de datos de entrenamiento y de prueba.



Figura 13: Proceso de generación y validación de un modelo

Entrenamiento y prueba

- Se utiliza el conjunto de datos de entrenamiento para crear un modelo supervisado
- El conjunto de datos de prueba se utiliza para **medir la precisión** alcanzada por el modelo
- El modelo se forma mediante la repetición iterativa de entrenamiento y verificación hasta conseguir niveles de precisión y de capacidad de predicción aceptables
- Los conjuntos de datos de entrenamiento y de prueba suelen ser **extracciones aleatorias** del juego de datos inicial
- En función de la cantidad de datos disponibles, existen diferentes técnicas para la generación de los conjuntos de entrenamiento y prueba:
 - 70% de los datos para entrenamiento y 30% para prueba
 - Dividir la tabla de datos en K conjuntos, entrenar con K-1 conjuntos y usar el restante para prueba
 - Todos los datos menos 1 para entrenar y un solo dato para probar.

Evaluación del modelo (Matriz de confusión)

Acá se debe:

- Cuantificar el grado de “bonanza” de la solución encontrada

- Lo anterior permite comparar entre los distintos métodos sobre los mismos conjuntos de datos
- La evaluación del modelo de clasificación se calcula comparando las predicciones generadas por el modelo en un conjunto de datos D con las etiquetas de clases verdaderas de las instancias de éste conjunto de datos

Matrices de confusión

Se usa para evaluar modelos y es una tabla con una visión gráfica de los errores cometidos por el modelo de clasificación. También es encontrada en la literatura como tabla de contingencia o matriz de errores.

VERDADERO POSITIVO (TRUE POSITIVE , TP)

Número de clasificaciones correctas en la clase positiva (P).

VERDADERO NEGATIVO (TRUE NEGATIVE , TN)


Número de clasificaciones correctas en la clase negativa (N).

FALSO NEGATIVO (FALSE NEGATIVE , FN)

Número de clasificaciones incorrectas de clase positiva clasificada como negativa.

FALSO POSITIVO (FALSE POSITIVE , FP)

Número de clasificaciones incorrectas de clase negativa clasificada como positiva



| | | Clase predicha | |
|-----------------|---|----------------|----|
| | | P | N |
| Clase verdadera | P | TP | FN |
| | N | FP | TN |

Figura 14: Matriz de confusión

Métricas de la Matriz de Confusión

Las métricas proporcionan información general sobre el número de instancias incorrectamente clasificadas.

Estas son:

- **Error de Clasificación (ERR):** Es la suma de predicciones incorrectas sobre el número total de predicciones

$$ERR = \frac{FP + FN}{FP + FN + TP + TN}$$

Figura 15: Error de predicción

- **Exatitud:** Es el número de predicciones correctas sobre el número total de predicciones



$$ACC = \frac{TP + TN}{FP + FN + TP + TN} = 1 - ERR$$

Figura 16: Cálculo de exactitud

Aplicaciones de árboles de decisión

- Medición de desempeño académico en alumnos
- Evaluación de riesgos en créditos
- Asignación de seguros
- Lanzamiento de nuevos productos
- Predicción de siniestros
- Predecir bajas de alumnos
- Predecir abandono de clientes
- Predecir logro de metas de producción
- Análisis de riesgos en proyectos
- Decisiones sobre demandas en juicios
- Planificación de producción
- Proyectos de inversión
- Diagnósticos médicos
- Predicciones meteorológicas
- Controles de calidad

Clasificador Bayesiano

Introducción al clasificador bayesiano

Introducción al módulo

El clasificador bayesiano se basa en el algoritmo Naive Bayes, el cual se basa en la estadística bayesiana, es decir, se asume que las variables predictoras son independientes entre sí

Suele utilizarse en:

- Análisis de sentimiento
- Detección de spam
- Sistemas de recomendación

- **Clasificación de texto**

Permite construir modelos de predicción de forma fácil por, justamente, la condición de independencia de sus variables, y su exactitud es alta.

Introducción a Naive Bayes

Se basan en clasificación estadística que se hace bajo el **Teorema de Bayes** y se asume que las variables predictoras son independientes entre sí, o sea, que una característica presente en un conjunto de datos no se relaciona con la presencia de otra característica.

- Se utiliza para crear modelos predictivos dada su sencillez y gran facilidad de uso e interpretación.
- Los modelos que genera tienen altos niveles de exactitud gracias a que su método es simple.

Tiene como fundamento la Estadística Bayesiana la cual difiere de la estadística tradicional en que, entre otras cosas, la estadística tradicional se basa en un razonamiento inductivo y sus parámetros son constantes desconocidas, mientras que en la Estadística Bayesiana, el razonamiento es deductivo y sus variables son aleatorias.

Los métodos de tipo Naive Bayes, estiman los parámetros, los cuales expresan la probabilidad condicionada de cada clase dadas las propiedades de una muestra que se expresa en forma de atributos. A partir de esto, los parámetros pueden ser combinados para formar las clases que maximizan sus probabilidades a nuevas muestras.

El objetivo de estos métodos, es maximizar las probabilidades de nuevas muestras, a partir de ejemplos del conjunto de entrenamiento y por lo tanto, la probabilidad condicional es un elemento central del algoritmo de Naive Base.

Es decir, el algoritmo expresa una forma de **calcular la probabilidad posterior** de que ocurra un cierto evento, **dadas las probabilidades de que hayan ocurrido eventos anteriores**.

El algoritmo de Naive Bayes

Se basa en probabilidades condicionales y lo que busca es maximizar el modelo, es decir, otorga más importancia a aquellos eventos que sean más relevantes en el conjunto de datos.

Para construir el modelo hay que:

1. **Preparar los datos**
2. **Construir el modelo**
3. **Evaluación del modelo:** en esta fase se evalúa si el modelo armado en base a una base de datos que se le dió para nuevos datos que haya que analizar. O sea, cerciorarse de que el modelo es capaz de ser usado en producción.

O sea, los pasos de modelado, en este caso el algoritmo de aprendizaje es Naive Bayes:

1. Se divide el set de datos en training y test set
2. Con el training set se pasa por el algoritmo y se genera el modelo.

3. Se testea el modelo generado y se calcula su **precisión** contabilizando cuantas veces predijo en forma correcta y se obtiene un **porcentaje de exactitud** usando matriz de confusión



Figura 17: Pasos de modelado, en éste caso el algoritmo de aprendizaje es Naive Bayes. Primero se divide el set de datos en training y test set, luego con el training set se pasa por el algoritmo y se genera el modelo. Por último, se testea el modelo generado y se calcula su precisión. Se debe evitar que el modelo sea dependiente de los datos utilizados durante su entrenamiento evitando el problema de **sobre-entrenamiento**.

El **sobreentrenamiento** es el hecho de que el modelo sólo responda bien al conjunto de datos de entrenamiento y que no sea capaz de predecir bien en el conjunto de datos que se prueba después.

Técnica del clasificador bayesiano

Características del algoritmo Naive Bayes

- Es una técnica de clasificación basada en el Teorema de Bayes
- Se fundamenta en el supuesto de **independencia entre los predictores**. O sea, supone que la presencia de una característica particular en una clase no está relacionada con la presencia de ninguna otra característica.

Por ejemplo, si yo tengo una variable A, la cual está impactada o explicada por las variables B, C y D. Entonces el algoritmo se basa en que B, C y D son independientes entre sí siendo éstos los predictores y la variable que se quiere predecir es A.

Siendo la presencia de una característica particular la clase A, no está relacionada con las otras características o sea, la presencia de B no depende de la presencia de C ni de D

Se llama Naive Bayes porque es muy difícil que en la vida real, la condición de independencia se cumpla, es un algoritmo ingenuo.

Pero por otra parte, ésta independencia es lo que garantiza es que los cálculos sean sencillos, y aún siendo así de ingenuo, su eficacia es muy alta.

Ejemplo:

Una fruta puede considerarse que es una manzana si es roja, redonda y de aproximadamente de 3 pulgadas de diámetro.

Incluso si éstas características dependen unas de otras o de la existencia de otras características, todas éstas propiedades **contribuyen independientemente a la probabilidad de que esa fruta sea una “manzana”** o sea, no hay una correlación entre que sea roja, redonda y de aproximadamente de 3 pulgadas de diámetro y por eso se lo conoce como “ingenua”.

El modelo Naive Bayes es fácil de construir y particularmente útil para conjuntos de datos muy grandes; y junto con su simplicidad, se sabe que Naive Bayes supera incluso a los métodos de clasificación altamente sofisticados.

Teorema de Bayes

Proporciona una forma de calcular la probabilidad posterior.

Si:

- $P(c|X)$ es la probabilidad que ocurra c dado que pasó X
- $P(c)$ es la probabilidad de que pasa c (la clase)
- $P(X|c)$ es la probabilidad de que pasó x dado que pasó c
- $P(X)$ es la probabilidad del predictor

La relación entre ellas es: (Figura 18)

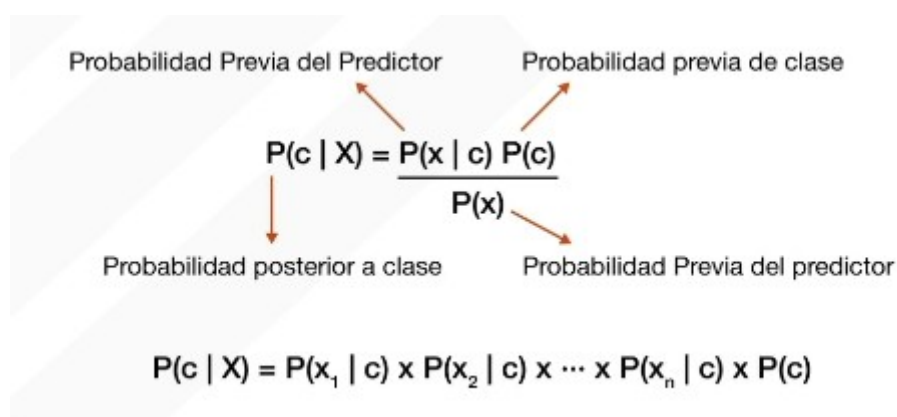


Figura 18: Lo que se quiere calcular es $P(c|X)$ dado que en el predictor ya tengo calculadas $P(c)$, y $P(X|c)$

Tablas de frecuencias y probabilidades

Pasos del algoritmo Naive Bayes:

1. Tener la tabla de datos
2. Convertir el conjunto de datos en una tabla de frecuencias. Es decir, cuanto se repiten cada uno de esos datos
3. Crear una tabla de probabilidades calculando las correspondientes probabilidades de que ocurran los diversos eventos
4. Calcular la probabilidad posterior de cada clase con la ecuación de Naive Bayes

5. **Mostrar el resultado.** La clase con la probabilidad posterior más alta es el resultado de la predicción

Ejemplo:

Si yo quiero predecir para un tipo de estado del tiempo, si se juega o no un partido, aplicando los pasos.

1. **Tabla de datos:** acá se carga la información previa que tengo

| Clima | Juegan |
|----------|--------|
| Soleado | No |
| Nublado | Si |
| Lluvioso | Si |
| Soleado | Si |
| Soleado | Si |
| Nublado | Si |
| Lluvioso | No |
| Lluvioso | No |
| Soleado | Si |
| Lluvioso | Si |
| Soleado | No |
| Nublado | Si |
| Nublado | Si |
| Lluvioso | No |

Figura 19: Tabla ejemplo

2. **Convertir los datos a una tabla de frecuencias:** En el ejemplo se tiene el tipo de clima y se ve en la tabla que “Jugaron” puede tener solo 2 valores, sí y no. Entonces se llena la tabla de frecuencias siendo la primera columna el estado del tiempo, la segunda es cuantas veces sí se jugó en un estado de tiempo particular y en la tercera, cuantas veces no se jugó.



| Clima | Juegan |
|----------|--------|
| Soleado | No |
| Nublado | Si |
| Lluvioso | Si |
| Soleado | Si |
| Soleado | Si |
| Nublado | Si |
| Lluvioso | No |
| Lluvioso | No |
| Soleado | Si |
| Lluvioso | Si |
| Soleado | No |
| Nublado | Si |
| Nublado | Si |
| Lluvioso | No |

| Clima | No | Si |
|----------|----|----|
| Nublado | | 4 |
| Lluvioso | 3 | 2 |
| Soleado | 2 | 3 |
| Total | 5 | 9 |

Figura 20: Construcción de tabla de frecuencias en base a los datos iniciales

3. Se crea una **tabla de probabilidades** calculando las probabilidades correspondientes de que ocurran los diversos eventos

Calculamos los diversos eventos.

| Clima | Juegan |
|----------|--------|
| Soleado | No |
| Nublado | Si |
| Lluvioso | Si |
| Soleado | Si |
| Soleado | Si |
| Nublado | Si |
| Lluvioso | No |
| Lluvioso | No |
| Soleado | Si |
| Lluvioso | Si |
| Soleado | No |
| Nublado | Si |
| Nublado | Si |
| Lluvioso | No |

| Clima | No | Si |
|----------|----|----|
| Nublado | | 4 |
| Lluvioso | 3 | 2 |
| Soleado | 2 | 3 |
| Total | 5 | 9 |

| Clima | No | Si | | |
|----------|---------|---------|---------|-----|
| Nublado | | 4 | $=4/14$ | 0.2 |
| Lluvioso | 3 | 2 | $=5/14$ | 0.3 |
| Soleado | 2 | 3 | $=5/14$ | 0.3 |
| Total | 5 | 9 | | |
| | $=5/14$ | $=9/14$ | | |
| | 0.36 | 0.64 | | |

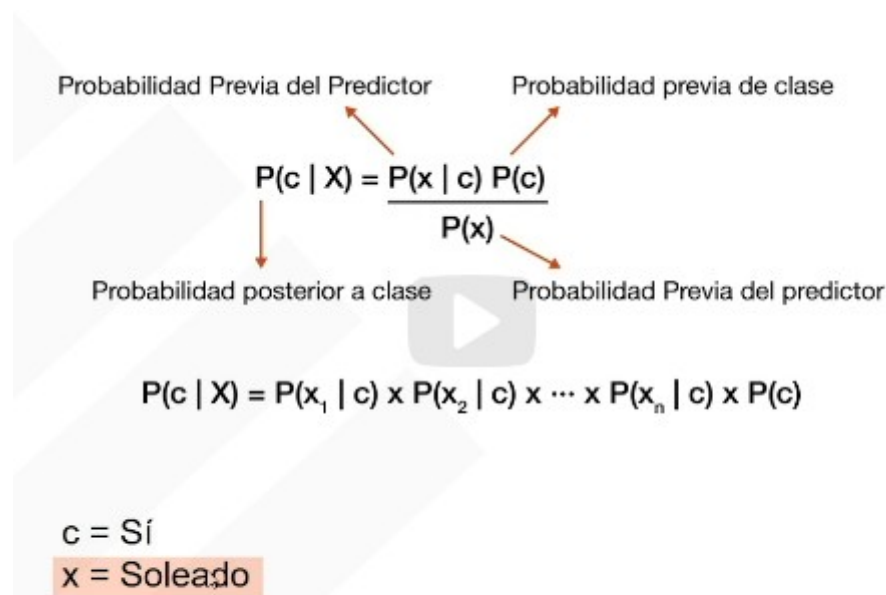
Figura 21: En el cálculo de probabilidad se usa el 14 porque es el número total de eventos

Probabilidad posterior de la clase Si

4. Lo que sigue es **calcular la probabilidad posterior de cada clase con la Ecuación de Bayes**.

En el caso ejemplo: “Los jugadores **jugarán** si está **soleado**” ¿Es correcta ésta afirmación?

Para averiguar eso se usa la fórmula de Bayes:



$$P(\text{Sí} | \text{Soleado}) = P(\text{Soleado} | \text{Sí}) * P(\text{Sí}) / P(\text{Soleado})$$

Figura 22: Fórmula de Bayes para el ejemplo de SI se juega cuando está Soleado

Lo que se calcula es la probabilidad de que si se juegue dado que está soleado

Aquí tenemos:

$$P(\text{Soleado} | \text{Sí}) = 3/9 = 0.33$$

$$P(\text{Soleado}) = 5/14 = 0.36$$

$$P(\text{Sí}) = 9/14 = 0.64$$

Entonces:

$$P(\text{Sí} | \text{Soleado}) = 0.33 * 0.64 / 0.36 = 0.60$$

| Tabla de Probabilidad | | | | |
|-----------------------|-------|-------|-------|-----|
| Clima | No | Si | | |
| Nublado | | 4 | =4/14 | 0.2 |
| Lluvioso | 3 | 2 | =5/14 | 0.3 |
| Soleado | 2 | 3 | =5/14 | 0.3 |
| Total | 5 | 9 | | |
| | =5/14 | =9/14 | | |
| | 0.36 | 0.64 | | |

Figura 23: Probabilidades de Si se juega estando soleado

Probabilidad posterior de la clase NO

4. Ahora Lo mismo se hace para la clase NO, y así poder compararlos

c= No

x= Soleado

$$P(\text{No} \mid \text{Soleado}) = P(\text{Soleado} \mid \text{No}) * P(\text{No}) / P(\text{Soleado})$$

Aquí tenemos:

$$P(\text{Soleado} \mid \text{No}) = 2/5 = 0.4$$

$$P(\text{Soleado}) = 5/14 = 0.36$$

$$P(\text{No}) = 5/14 = 0.36$$

Entonces:

$$P(\text{No} \mid \text{Soleado}) = 0.4 * 0.36 / 0.36 = 0.40$$

| Tabla de Probabilidad | | | | |
|-----------------------|-------|-------|-------|-----|
| Clima | No | Si | | |
| Nublado | | 4 | =4/14 | 0.2 |
| Lluvioso | 3 | 2 | =5/14 | 0.3 |
| Soleado | 2 | 3 | =5/14 | 0.3 |
| Total | 5 | 9 | | |
| | =5/14 | =9/14 | | |
| | 0.36 | 0.64 | | |

Figura 24: Cálculo de probabilidad de que no jueguen si está soleado

Resultado de la predicción

- Ahora lo que sigue es **Mostrar el resultado**. La clase con la predicción más alta es el resultado de la predicción

En el ejemplo:

Problema:

Los jugadores **jugarán** si está **soleado**.
¿Esta afirmación es correcta?

P (Sí | Soleado) = 0.60

P (No | Soleado) = 0.40

Respuesta: La afirmación es **correcta**.
Los jugadores jugarán si está soleado.

Figura 25: Ejemplo de mostrar el resultado