

TP INICIAL

“IA, Machine Learning & Deep Learning”

Laboratorio de construcción de software

❖ Alumnas:

- **Benítez Micaela**
- **Benítez Yamila**

❖ Profesores:

- **Juan Carlos Monteros**
- **Francisco Orozco De La Hoz**
- **Leandro Dikenstein**

Introducción

En este trabajo vamos a centrarnos en el campo de la Inteligencia Artificial (IA), el Machine Learning y el Deep Learning. Estas tecnologías están cambiando la manera en que las computadoras pueden aprender y mejorar su rendimiento en tareas complejas. Exploraremos desde los conceptos básicos hasta cómo se aplican en la vida real, implementando un modelo de ML que nos permita analizar un conjunto de datos relacionado al área de la salud.

Objetivo

Según la Organización Mundial de la Salud (OMS), el accidente cerebrovascular es la segunda causa de muerte en todo el mundo, responsable de aproximadamente el 11% del total de muertes.

Por lo tanto, nuestro objetivo a lograr es predecir si una persona corre el riesgo de sufrir un ataque cardiovascular (ACV) en función de las variables de entrada como el sexo, la edad, diversas enfermedades y el tabaquismo.

Lo llevaremos a cabo utilizando un [Dataset](#) (extraído de la plataforma web Kaggle) con factores de riesgo que pueden causar un accidente cerebrovascular.

El modelo de machine learning que utilizaremos es el supervisado ya que nosotras le proporcionaremos la información a la máquina con datos etiquetados.

IA (Inteligencia artificial)

La inteligencia artificial es un campo de la ciencia relacionado con la creación de computadoras y máquinas que pueden razonar, aprender y actuar de una manera que normalmente requeriría inteligencia humana o que involucre datos cuya escala exceda lo que los humanos pueden analizar.

La IA es un campo amplio que abarca muchas disciplinas diferentes, incluidas la informática, el análisis de datos y las estadísticas, la ingeniería de hardware y software, la lingüística, la neurociencia y hasta la filosofía y la psicología.

Algunas de las actividades de máquinas con inteligencia artificial están diseñados para incluir:

- **Reconocimiento de voz:** Convertir automáticamente una frase hablada en un texto escrito.
- **Reconocimiento de imágenes:** Identificar y categorizar diversos aspectos de una imagen.
- **Traducción:** Traducir palabras escritas o habladas de un idioma a otro.
- **Modelado predictivo:** Extraer datos para prever resultados específicos con altos niveles de detalle.
- **Analítica de datos:** Encontrar patrones y relaciones en los datos para la inteligencia empresarial.
- **Seguridad cibernética:** Analiza de forma autónoma redes en busca de ciberataques y amenazas.

Beneficios de la IA:

a) Automatización:

La IA puede automatizar flujos de trabajo y procesos, o trabajar de forma independiente y autónoma de un equipo humano. Por ejemplo, la IA puede ayudar a automatizar aspectos de la seguridad cibernética mediante la supervisión y el análisis continuos del tráfico de red. De manera similar, una fábrica inteligente puede tener decenas de tipos diferentes de IA en uso, como robots que usan visión artificial para

navegar por las fábricas o inspeccionar productos en busca de defectos, crear gemelos digitales o usar analítica en tiempo real para medir la eficiencia y la producción.

b) Reduce errores humanos:

La IA puede eliminar errores manuales en el procesamiento de datos, las estadísticas, el ensamblaje en la fabricación y otras tareas a través de automatización y algoritmos que siguen los mismos procesos cada vez.

c) Elimina las tareas repetitivas:

La IA se puede usar para realizar tareas repetitivas, lo que libera al capital humano a fin de que trabaje en los problemas de mayor impacto. La IA se puede usar para automatizar procesos, como verificar documentos, transcribir llamadas telefónicas o responder preguntas sencillas de los clientes, como “¿A qué hora cierran?”. Con frecuencia, los robots se usan para realizar tareas “aburridas, sucias o peligrosas” en lugar de que las haga un ser humano.

d) Rápido y preciso:

La IA puede procesar más información de forma más rápida que un ser humano, mediante la búsqueda de patrones y el descubrimiento de relaciones en datos que el humano podría no detectar.

e) Disponibilidad infinita:

La IA no tiene limitaciones en términos de horarios, necesidad de descansar ni ningún otro factor que pueda interrumpir la labor de un ser humano. Cuando se ejecutan en la nube, la IA y el aprendizaje automático pueden estar “siempre activos”, y trabajar continuamente en las tareas asignadas.

f) Investigación y desarrollo acelerados:

La capacidad de analizar grandes cantidades de datos con rapidez puede acelerar los avances en investigación y desarrollo. Por ejemplo, la IA se usó en el modelado predictivo de nuevos tratamientos farmacéuticos potenciales o para cuantificar el genoma humano.

Machine Learning

El Machine Learning (aprendizaje automático) es un campo científico y, más particularmente, una subcategoría de inteligencia artificial. Consiste en dejar que los algoritmos descubran patrones recurrentes, en conjuntos de datos. Esos datos pueden ser números, palabras, imágenes, estadísticas, etc. Todo lo que se pueda almacenar digitalmente puede servir como dato para el Machine Learning. Al detectar patrones en esos datos, los algoritmos aprenden y mejoran su rendimiento en la ejecución de una tarea específica. Los algoritmos de Machine Learning aprenden de forma autónoma a realizar una tarea o hacer predicciones a partir de datos y mejorar su rendimiento con el tiempo. Una vez entrenado, el algoritmo podrá encontrar los patrones en nuevos datos.

A grandes rasgos, suelen usarse tres tipos de modelos de aprendizaje en el aprendizaje automático:

- a) Aprendizaje supervisado:** Es un modelo de aprendizaje automático que asigna una entrada específica a un resultado mediante datos de entrenamiento etiquetados (datos estructurados). Los datos se etiquetan para indicar a la máquina qué patrones tiene que buscar. El sistema se entrena sobre un conjunto de datos etiquetados, con la información que se supone que tiene que determinar. Es posible que los datos ya estén clasificados de la forma en que se supone que debe hacerlo el sistema.
- Este método requiere menos datos de entrenamiento que los demás y facilita el proceso de entrenamiento, ya que los resultados del modelo se pueden comparar con los datos ya etiquetados. Sin embargo, etiquetar los datos puede resultar caro. Un modelo también puede estar sesgado debido a los datos de entrenamiento, lo que afectará a su rendimiento más adelante cuando procese nuevos datos.

Existen dos tipos de datos que pueden ser introducidos en este modelo:

- **Clasificación:** clasifica un objeto dentro de diversas clases. Por ejemplo, para determinar si un paciente está enfermo o si un correo electrónico es spam.

Hay varias técnicas de machine learning que podemos usar en problemas de clasificación. Podemos destacar:

- regresión logística (logistic regression)
- máquinas de vectores de soporte (support vector machines)
- árboles de decisión (decision trees)
- bosques aleatorios (random forests)
- redes neuronales y aprendizaje profundo (deep learning)

- **Regresión:** predice un valor numérico. Sería el caso de los precios de una casa al escoger diferentes opciones o la demanda de ocupación de un hotel.

Hay varias técnicas de machine learning que podemos usar en problemas de Regresión. Podemos destacar:

- regresión lineal y regresión no lineal
- máquinas de vectores de soporte (support vector machines)
- árboles de decisión (decision trees)
- bosques aleatorios (random forests)
- redes neuronales y aprendizaje profundo (deep learning)

b) Aprendizaje no supervisado: Es un modelo de aprendizaje automático que aprende patrones en función de datos no etiquetados (datos no estructurados). A diferencia del aprendizaje supervisado, el resultado final no se conoce con anticipación. En cambio, el algoritmo aprende de los datos y los clasifica en grupos en función de diversos atributos. En lugar de automatizar decisiones y predicciones, este enfoque ayuda a identificar patrones y relaciones que los humanos pueden pasar por alto en los datos. Esta técnica no es muy popular porque es menos fácil de aplicar. Sin embargo, es cada vez más popular en el campo de la ciberseguridad.

Existen dos tipos de algoritmos para Machine Learning no supervisado:

- **Clustering:** clasifica en grupos los datos de salida. Es el caso de las segmentaciones de clientes según qué hayan comprado.
- **Asociación:** descubre reglas dentro del conjunto de datos. Por ejemplo, aquellos clientes que compran un coche también contratan un seguro, por lo que el algoritmo detecta esta regla.

c) Aprendizaje por refuerzo: Es un modelo de aprendizaje automático que se puede describir en términos generales como “aprender haciendo”. Un "agente" aprende a realizar una tarea definida mediante prueba y error (un ciclo de reacción) hasta que su rendimiento está dentro de un rango deseado. El agente recibe un refuerzo positivo cuando realiza la tarea de forma correcta y un refuerzo negativo cuando tiene bajo rendimiento. Un ejemplo de aprendizaje por refuerzo sería enseñarle a una mano robótica a recoger una pelota.

Deep Learning

El Deep Learning (aprendizaje profundo) es una rama del machine learning y hoy en día es la más utilizada. Los algoritmos de Deep Learning se aplican a redes neuronales artificiales, que imitan a la de los humanos, y que están estructuradas en forma de capas:

- Input layer (entrada)
- Hidden layer (oculta)
- Output layer (salida)

Los datos entran por la primera capa (input layer), en la que hay varias neuronas artificiales que se activan o no dependiendo de los datos. Cuando se activa, esos datos pasan por las capas intermedias (hidden layer) y se transforman en una información. Finalmente, la capa de salida (output layer) determina exactamente el tipo de información (una forma, una letra, un número, etc).

Esta técnica se llama red de neuronas profundas. Esa profundidad corresponde al gran número de capas de nodos de cálculo que constituyen estas redes y que trabajan conjuntamente para tratar los datos y hacer predicciones.

Este tipo de aprendizaje impulsa muchos servicios y aplicaciones de inteligencia artificial (IA) que mejoran la automatización, realizando tareas analíticas y físicas sin intervención humana. La tecnología de deep learning reside detrás de muchos productos y servicios de uso cotidiano (como los asistentes digitales, los controles de TV habilitados por voz y la detección de fraudes con tarjeta de crédito), así como de tecnologías emergentes (como los automóviles autónomos).

La forma de entrenar a un algoritmo de deep learning es suministrarle cantidades masivas de datos. Cuantos más analice, más preciso se vuelve. Por ello, todas estas tecnologías también están relacionadas con el big data y tienen un papel relevante en sus aplicaciones, que buscan extraer significado, tal y como haría un humano, de enormes cantidades de datos.

El Deep Learning es un campo que está en pleno crecimiento y cada vez se descubren nuevas aplicaciones, incrementando la expectativa de hasta dónde podrá llegar la inteligencia artificial.

1. Reconocimiento de imágenes y escenas.
2. Procesamiento del lenguaje natural.
3. Visión artificial.
4. Sistemas de recomendación.

5. Monitorización del sentimiento de las personas.
6. Monitorización de la fidelización de los clientes.
7. Prevención del fraude.
8. Detección de incidentes de ciberseguridad.

Herramientas:

- Repositorio GitHub: <https://github.com/micabenitez/TP-Inicial-PP1>

Bibliografía

IBM (<https://www.ibm.com/es-es/topics/deep-learning>)

Iberdrola (<https://www.iberdrola.com/innovacion/deep-learning>)

Datascientest

(<https://datascientest.com/es/machine-learning-definicion-funcionamiento-usos>)

Ceupe (<https://www.ceupe.com/blog/machine-learning.html>)

Google Cloud

(<https://cloud.google.com/learn/what-is-artificial-intelligence?hl=es-419#section-4>)

Arimetrics (<https://www.arimetrics.com/glosario-digital/ia>)

hubspot (<https://blog.hubspot.es/marketing/inteligencia-artificial-esta-aqui>)

Linkedin(<https://es.linkedin.com/pulse/qu%C3%A9-es-deep-learning-y-c%C3%B3mo-nos-puede-beneficiar->)

DataSet (<https://www.kaggle.com/datasets/fedesoriano/stroke-prediction-dataset>)

IA (<https://www.iartificial.net/clasificacion-o-regresion/>)

