

NOTAS DE CLASE

INTELIGENCIA ARTIFICIAL - 12 DE AGOSTO - SEMANA 2

Luis Alfredo González Sánchez
Escuela de Ingeniería en Computadores
Instituto Tecnológico de Costa Rica
Cartago, Costa Rica
2021024482 gonzal3z.luis@estudiantec.cr

Abstract—This document provides a concise summary of the key concepts and examples covered in week two of the artificial intelligence course. It begins with an overview of general AI concepts, followed by an explanation of the types and main approaches of machine learning. The document also explores various problem-solving paradigms and concludes with a brief review of the essential components of the machine learning pipeline. This summary aims to enhance understanding and reinforce the material presented in class.

Index Terms—AI, Machine learning, Pipeline.

I. INTRODUCTION

El presente documento busca brindar un breve resumen a los conceptos y ejemplos vistos en la clase de la semana 2 del curso de inteligencia artificial, comenzando por un breve recorrido a determinados conceptos generales de la IA, luego se procederá a explicar los tipos de aprendizaje en machine learning, sus principales enfoques, algunos paradigmas de resolución de problemas y por último un breve repaso a los conceptos del pipeline de machine learning.

II. CONCEPTOS GENERALES

Se presentará un breve repaso a conceptos vistos en clase.

A. Definición de modelos

- **Modelo determinista** : Es aquel modelo que, bajo una misma entrada se produce los mismos resultados. No incorpora aleatoriedad, por lo que los resultados obtenidos se vuelven predecibles, por ejemplo, bajo características médicas se puede determinar una enfermedad asociada.
- **Modelo estocástico** : Es aquel modelo en que, bajo una entrada determinada las salidas son diferentes debido a un conjunto de posibles resultados asociados a dicha entrada, existe control respecto al grado de aleatoriedad, por ejemplo, consultar el clima a la misma hora todos los días.

B. Definición de parámetros

- **Parámetros** : Son los datos que se le configuran al modelo, permiten controlar el comportamiento de sus algoritmos
- **Hiperparámetros** : Son todos los valores que el modelo aprende a partir de los datos de entrenamiento, le permiten al modelo identificar errores y ajustar sus algoritmos.

C. La jerarquía de conceptos en IA

Observe la figura 1.

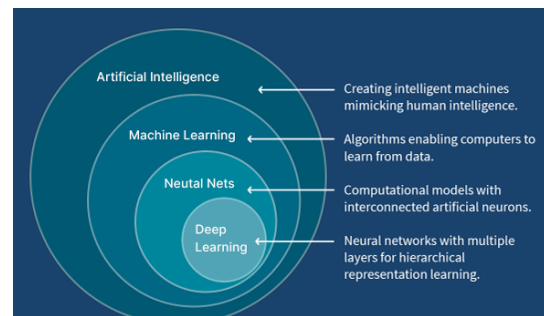


Fig. 1. Jerarquía de conceptos IA

D. ¿Qué es machine learning ?

Consiste en generar máquinas o programas a partir de algoritmos que sean capaces de resolver un problema a partir de inferencias. En este contexto se busca que, a partir de datos de entrada (y en ocasiones con sus resultados) donde el sistema sea capaz de resolver problemas en base a los datos que ya posee sin que se le explique como resolverlo. Considerando lo anterior, se describe a continuación los tipos de aprendizaje en machine learning.

III. TIPOS DE APRENDIZAJE

A. Aprendizaje supervisado

Tipo de aprendizaje que involucra un conjunto de datos con características y una respectiva etiqueta, esa etiqueta asocia dichas características con un valor. La etiqueta supervisa y corrige las aproximaciones dadas por la función planteada para obtener mejores resultados, una forma de visualizarlo es de la siguiente manera :

$$x = \{x_i, y_i\}$$

donde x representa mi característica, x_i representa la etiqueta y y_i , representa el resultado "supervisado" por la etiqueta.

B. Aprendizaje NO supervisado

Este tipo de aprendizaje es aquel que se entrena con datos sin etiquetas, es decir, sin información previa que le permita determinar cuál es el "resultado correcto". A diferencia del aprendizaje supervisado, donde las etiquetas sirven como una "supervisión" o guía para corregir y entrenar el modelo, en el aprendizaje no supervisado no hay una referencia explícita para evaluar o corregir. Un ejemplo mencionado en la clase, son los algoritmos de clúster, donde analiza en base a sus características si pertenece a un grupo u otro, sin hacer uso de etiquetas.

C. Aprendizaje semi-supervisado

En palabras sencillas es una combinación de los métodos de aprendizaje anteriores, en este tipo de algoritmo, se utilizan tanto datos etiquetados como datos no etiquetados para entrenar el modelo, es decir no siempre va a existir una etiqueta asociada a las características, es principalmente útil cuando se posee etiquetas en solo en una parte de los datos para entrenar, y en datos donde las etiquetas son difíciles de obtener, aprovecha la información de datos sin etiquetas para mejorar el rendimiento del modelo.

D. Aprendizaje auto-supervisado

Para este tipo de aprendizaje, el modelo no depende de datos con etiquetas "pre- insertadas" para el entrenamiento, en lugar de eso, genera automáticamente etiquetas o señales de supervisión a partir de los propios datos. Un ejemplo visto en clase es el uso en modelos de lenguaje, donde el sistema predice a partir de palabras previas cual es la palabra que sigue a dicha palabra, puede visualizarse la palabra previa como dato, pero a la vez como una misma etiqueta para poder definir la palabra siguiente.

E. Aprendizaje por refuerzo

El aprendizaje por refuerzo consiste en un aprendizaje con "premiación", existe un agente y dicho agente aprende a tomar decisiones a través de la interacción con un entorno, buscando maximizar una recompensa, se pueden penalizar por acciones que lleven a un resultado insatisfactorio. Ejemplo, observe la figura 2. En el videojuego representado, se puede entrenar al



Fig. 2. Videojuego de ejemplo

personaje (agente) para que cruce al otro lado basándose en si el salto se realiza correctamente (acción) se recompensa, y si la acción es insatisfactoria (caer al vacío y perder una vida) se castiga por dicha acción.

F. Otros tipos de aprendizaje

Se detallan a continuación algunos tipos de aprendizaje adicionales

- **Few-shot** : Método en el cual el modelo aprende a partir de breves ejemplos. Una posible aplicación es con respecto a los LLMs, si el LLM no sabe como realizar la tarea, se le brinda ejemplos y a partir de esos ejemplos es capaz de resolver la tarea.
- **One shot** : Método similar al anterior, solo que reduce la cantidad de ejemplos a 1, y a partir de ese único ejemplo logra resolver.
- **zero shot** : En este método el sistema resuelve en base a lo que "sabe", no existen ejemplos del que pueda tomar para aprender.

IV. ENFOQUES DEL MACHINE LEARNING

Existen dos principales enfoques, ciencia e ingeniería:

A. Machine learning: Ciencia

- **Se genera conocimiento** : Es la parte académica que se dedica a investigar y desarrollar nuevas ideas respecto a aprendizaje automático, incluye la elaboración de modelos desde cero y optimizar funciones, por lo general, está en manos de investigadores con un fuerte trasfondo matemático que les permite conocer a gran detalle los procesos que suceden durante el desarrollo de dichos modelos o algoritmos.
- **Métricas** : A partir de los datos recolectados, los data scientist generan métricas que permiten evaluar el desempeño del modelo o algoritmo.
- **Data scientist** : Son las personas encargadas de realizar la recolección y la preparación de los datos, de tal manera que sean aplicables para el entrenamiento de los modelos u algoritmos.
- **Research scientist** : A diferencia de los data scientist, los research scientist se encargan de investigar y proponer nuevos modelos, algoritmos o modelos.

B. Machine learning: Ingeniería

- **Puesta en producción del modelo** : Es necesario la continua monitorización del modelo antes de la puesta a producción y pasado un tiempo, se busca realizar correcciones, monitorizar si los resultados o salidas mantienen una distribución normal probabilística y como el progreso de entrenamiento luego de la puesta en producción mueve dicha distribución, es decir, dar seguimiento a las salidas.
- **Transformar el modelo** : Hace referencia a todas las estrategias y cambios requeridos para adaptar el modelo según los requerimientos especificados, por ejemplo, si se dispone de un modelo en python, cuáles son las estrategias para adaptarlo en otras plataformas como c++ o móvil, una de dichas estrategias es la reducción del peso del modelo.
- **Onnx** : Es una herramienta que facilita la portabilidad de modelos entre plataformas y dispositivos, optimizando y reduciendo el tamaño del modelo.

- **MLOps** : Los equipos de MLOps tiene la finalidad de monitorear y gestionar el modelo de manera detallada ,tanto con los datos como la calidad de los mismos, por ejemplo, en el seguimiento que se da en puesta en producción estan encargados de monitorear el comportamiento del modelo, en casos donde suceda algun cambio en los datos, deben de entrenarlo nuevamente.

V. PARADIGMAS DE RESOLUCIÓN DE PROBLEMAS

Se detallan a continuación ciertos paradigmas o estrategias para resolver problemas en el ámbito de la inteligencia artificial:

A. Problemas de búsqueda

: Son todos aquellos algoritmos que permiten encontrar la mejor solución, representada por el camino más barato o la ruta más corta, se detallan algoritmos como Dijkstra , DFS, BFS.

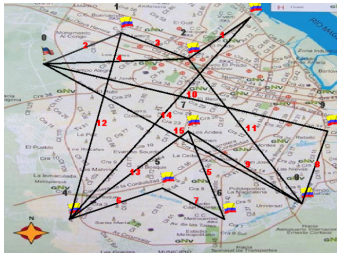


Fig. 3. Ejemplo de algoritmos de búsqueda

B. Problemas de optimización

Son todas aquellas soluciones que pretenden encontrar la solución más optima dentro de un grupo de soluciones.Existen 2 tipos:

- **Solución local** : Es aquella seleccionada dentro de un grupo o área específica, un ejemplo es el visto en clase respecto a bares, el mejor bar cercano al tec de cartago es la nave, pero esto es por que en el área cercana al tec es la mejor o la única opcion.
- **Solución local** : Se refiere a la mejor solución encontrada entre todas las posibles soluciones, tomando el ejemplo del bar, si se evalúa a nivel provincia , probablemente se encuentre un mejor bar que la nave.

C. Problemas de predicción y clasificación

En los problemas de predicción , existe una serie de datos respecto a un evento, comportamiento, entre otros y se desea encontrar patrones para realizar una predicción ,generalmente se representa con una función, que podrá generar una estimación de la relación de los datos con sus salidas, un ejemplo es el visto en clase , donde en base al motor del vehículo , se obtienen datos respecto al consumo de gasolina y se genera una función la cual será capaz de predecir cuál es el consumo de gasolina de otro vehículo. Estos datos son cuantificables o numéricos.

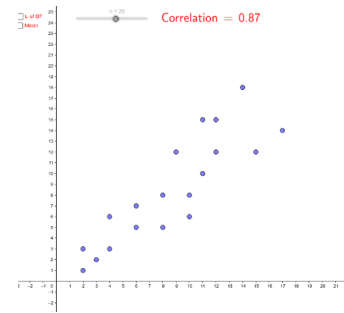


Fig. 4. Ejemplo de gráfica para visualizar problemas de predicción

D. Problemas de agrupamiento

Para los problemas de agrupamiento no existen etiquetas con las cuales entrenar, se busca encontrar pertenencia a un grupo dadas similitudes a características, las soluciones a estos problemas permite encontrar relaciones a preguntas que quizá no se tenían previamente.

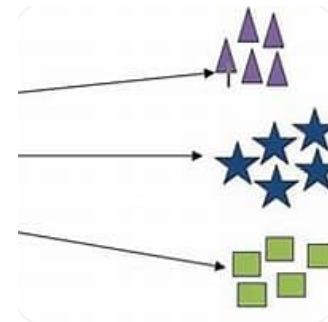


Fig. 5. Ejemplo de problemas de agrupamiento

VI. PIPELINE DEL MACHINE LEARNING

- **Data acquisition** :Una de las partes más importantes del piplene de machine learning, es todo el proceso de recolección de datos **Relevantes** del problema a resolver, se deben de resolver preguntas como ¿Cuál es la calidad de los datos que estoy tomando? ¿De dónde los saco? ¿Es mi fuente libre de sesgo?.
- **Data preparation** : La etapa de la preparación de los datos, considerando que el computador trabaja con números, se debe de resolver las pregunta ¿Cómo voy a representar los datos a mi algoritmo o modelo? , por ejemplo, si los datos son imágenes, le asigno valores puesto que los algoritmos son matemáticos. En esta etapa también se realizan otras labores como la revisión de datos duplicados y la búsqueda de datos faltantes, por ejemplo en un set de datos, con algún dato faltante, elimino dicho set o calculo la mediana y la introduzco como el dato faltante.
- **Freature engineering** : Los freatures son los atributos que proporcionan información útil de los datos, en esta etapa se pueden generar nuevos freatures a partir de los

existentes, seleccionar solo los relevantes para la solución o eliminar features que generen algún sesgo.

- **Model selection:** Es la etapa de selección del modelo más adecuado para cubrir la demanda de recursos y requisitos que genera el problema. Entra un concepto importante para la selección del modelo, la explicabilidad, si tengo que dar explicación del por qué de los resultados generados constantemente, quizá la mejor solución no es utilizar IA, si no , podría escoger un algoritmo que brinde una solución más elegante, un ejemplo es para un sistema de un banco que detecta si un usuario es elegible para un préstamo, entrenar una IA que determine esto puede no ser la mejor manera , si no , un algoritmo que con datos del usuario pueda determinar si es posible o no darle el préstamo.
- **Model training:** Es la fase de entrenamiento del modelo, los datos son divididos en datos de entrenamiento y datos de validación, en esta parte los algoritmos realizan optimizaciones en base a los datos de entrenamiento.
- **Model deployment:** En esta etapa el modelo se manda a producción en donde se debe de supervisar para garantizar un correcto funcionamiento.