



SKILLING  
CENTER

**TECMILENIO**



# Inteligencia artificial y machine learning

Modelos de predicción con  
base en redes neuronales



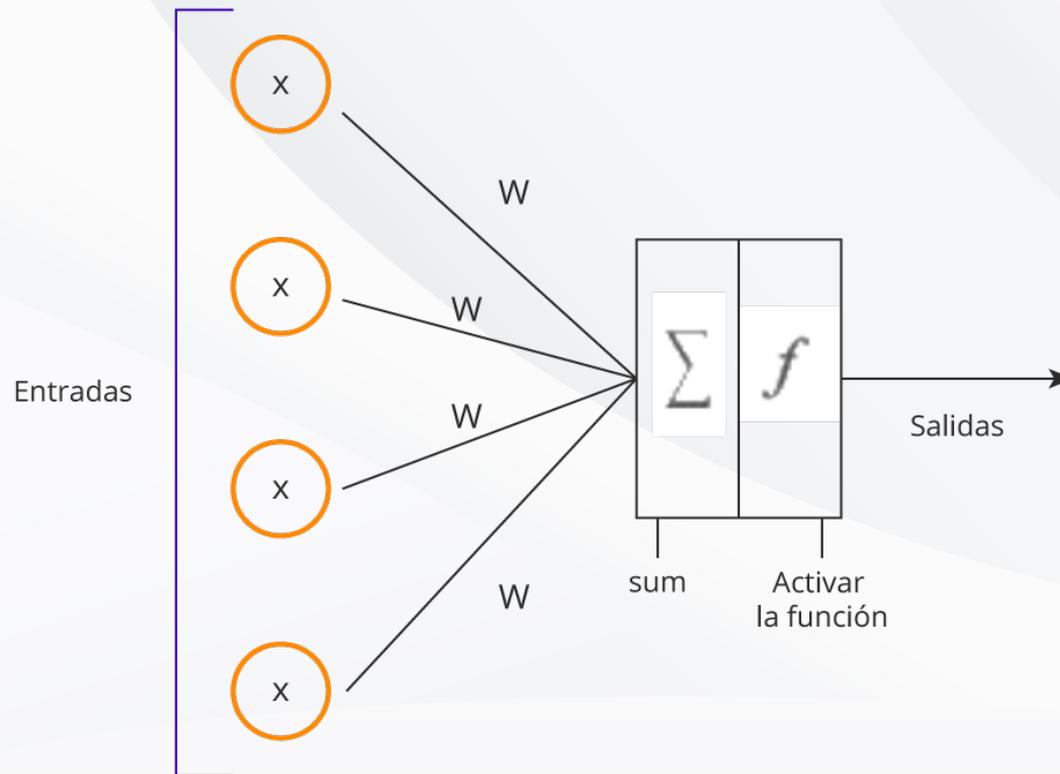


De acuerdo con Williems (2019), las redes neuronales son un tipo de modelos de aprendizaje automático inspirados en la estructura biológica del cerebro. La idea se basa en tomar un modelo simplificado de una neurona y representar computacionalmente las operaciones que realiza en conjunto con otras neuronas. Dentro de las finanzas se utilizan a través de la inteligencia artificial y del *machine learning*. Como ejecutivo de finanzas es importante conocer sobre estas tecnologías y adoptarlas con el fin de tomar decisiones más rápidas y certeras en procesos, como el otorgamiento de créditos a través de una óptima clasificación de los clientes o para agilizar los procesos de pago dentro de las organizaciones.

Dentro de esta experiencia educativa conocerás sobre los modelos basados en redes neuronales, la aplicación de ellas en el *deep-learning* (aprendizaje profundo) y la forma en la que se construye una proyección de estas variables a través de un ejemplo práctico.

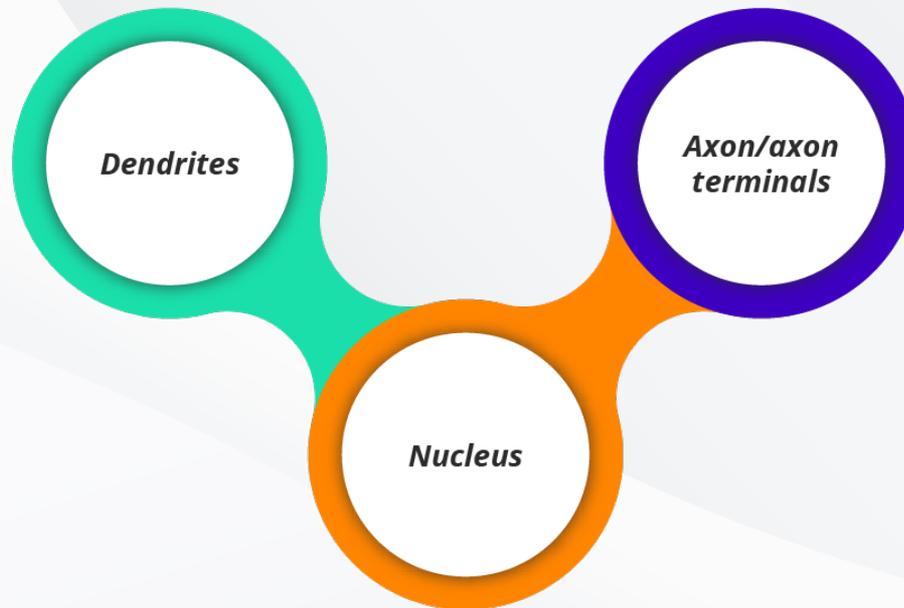


La idea de las redes neuronales se basa en tomar un modelo simplificado de una neurona y representar computacionalmente las operaciones que realiza en conjunto con otras neuronas. La siguiente figura compara la estructura biológica de una neurona individual con su representación computacional.





Estos son los componentes a considerar para crear un modelo computacional de una neurona:





El concepto de una neurona individual se conoce como un perceptrón. La idea de un perceptrón no es muy diferente a una regresión lineal. En esencia, se tiene la combinación lineal de las variables de entrada y una función de activación que se aplica sobre el resultado.

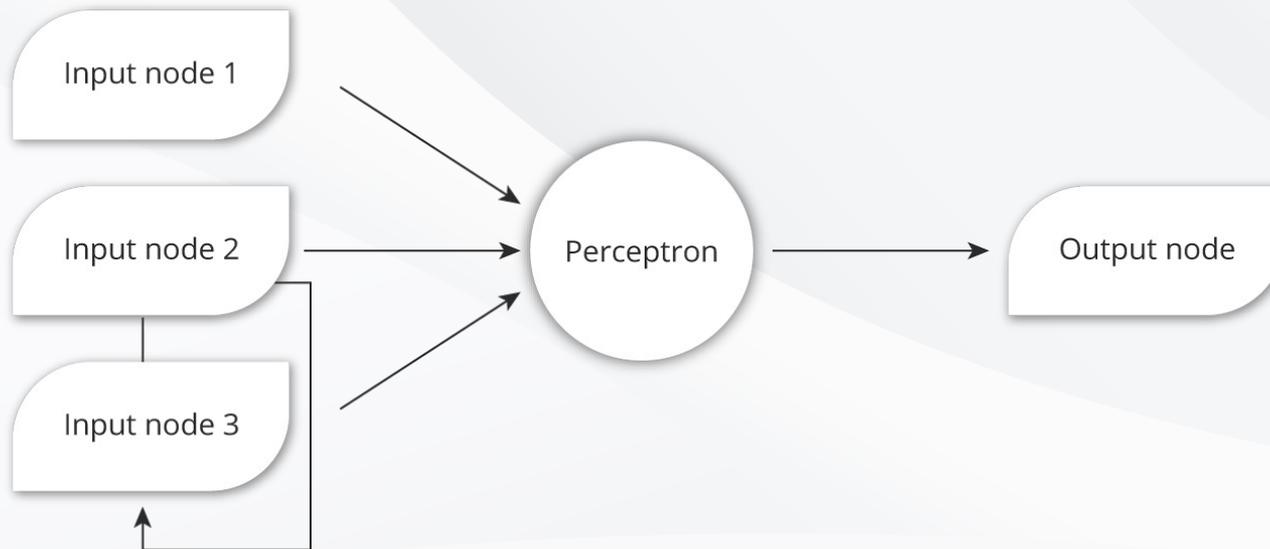
Si se usa la función identidad  $f(x) = x$  como la función de activación, se tiene la definición original de una regresión lineal.

De acuerdo con Baheti (2022), la función de activación permite al perceptrón emitir una señal. Es decir, le permite “activarse” al enviar información distinta de 0. Esto representa el equivalente biológico de estimular una neurona dado las condiciones adecuadas.



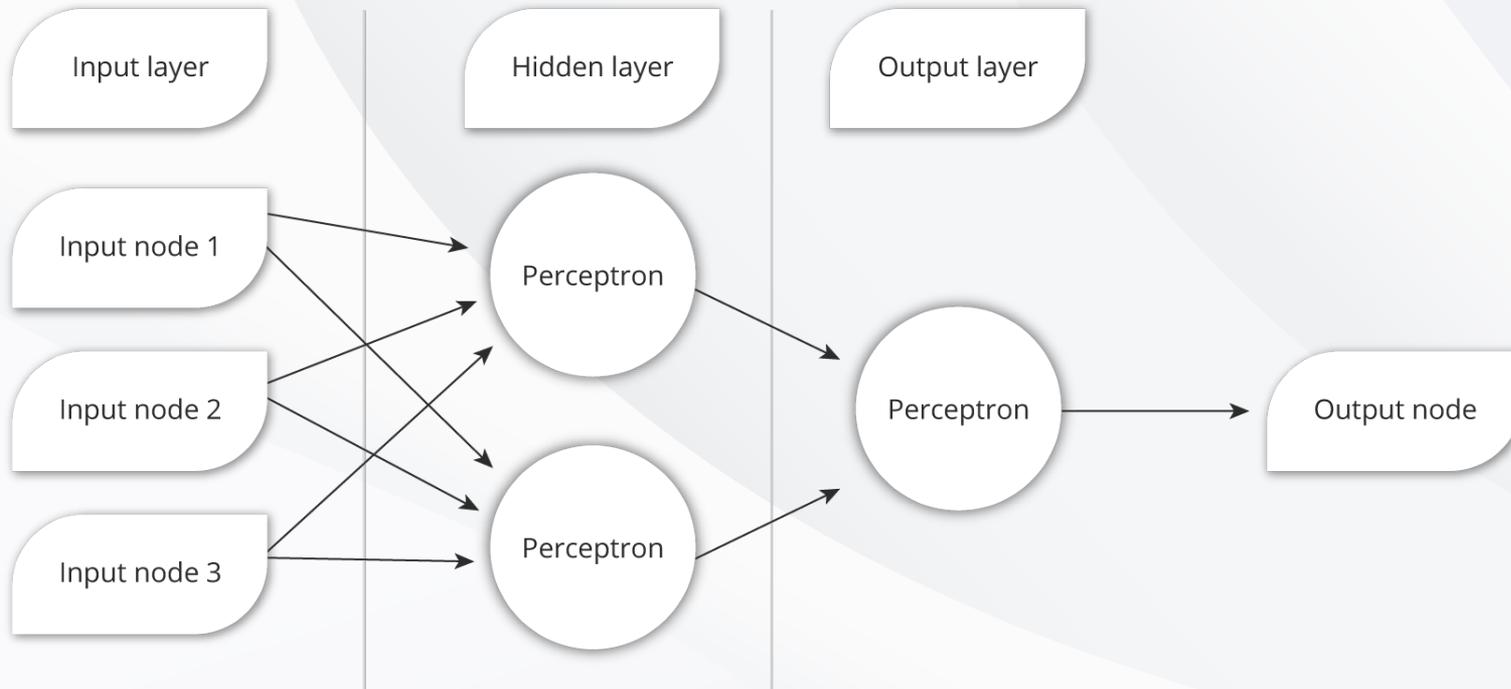
El perceptrón simple se compone de nodos de entrada, un núcleo y un nodo de salida. El procesamiento de la información que proveen los nodos de entrada se lleva a cabo en el núcleo, en donde se genera una combinación lineal de los datos y se aplica la función de activación. Para crear una red neuronal, se tienen que conectar los nodos de entrada de una neurona con el nodo de salida de otras neuronas. Un conjunto de perceptrones, organizados por capas, se conoce como un modelo de perceptrón multicapa.

## Single perceptron





## Multiplayer perceptron



Los modelos de perceptrón multicapa tienen una ventaja significativa sobre modelos lineales, dado que permiten utilizar varios perceptrones para procesar los datos, lo cual genera un efecto no-lineal que generalmente resulta en mejor capacidad predictiva.



Utilizar un software de alto o bajo nivel depende del tipo de problema que se esté resolviendo. En *Python*, se tienen varias implementaciones de código libre, entre las más populares están:





El modelo de perceptrón multicapa es considerado un modelo de *deep learning*. De acuerdo con Amini (2022), el *deep learning* se caracteriza por la utilización de modelos de redes neuronales para resolver problemas de predicción en donde la red neuronal se encarga de procesar y extraer toda la información relevante. En la industria financiera este tipo de modelos ha resultado de bastante utilidad para reemplazar modelos econométricos. Namini et al. (2018) publicaron un estudio en donde redujeron la función de error en promedio un 84 % al cambiar de un modelo econométrico (ARIMA) hacia un modelo de *deep learning* (LSTM).





Para la resolución de este problema se utiliza un conjunto de datos, dividido en tres particiones y disponible en este repositorio en la nube

(<https://gist.github.com/RHDZMOTA/406bb08e1a33469eceb66e5d6bf78e27/>)

La descarga de la información se realiza programáticamente con *Python* utilizando el siguiente código:

```
import os
import pandas as pd

url_base = (
    "https://gist.github.com/RHDZMOTA/"
    "406bb08e1a33469eceb66e5d6bf78e27/raw/a5f5d03c731df8a08e0574b150fe877f0abcc564/"
)
dataset_names = [
    "annual - income - category - low.json",
    "annual - income - category -medium.json",
    "annual - income - category - hight.json",
]
dataset = pd.concat ( [pd.read_json(f"{url_base}/{name}"). T for name in dataset_names] ) \
.sample(frac=1 , random_state=888). reset_index ( drop=true)

print(dataset.shape)
dataset.head()
(222416, 6)
```



	employment_length	home_ownership	zip_code	address_state	earliest_credit_line	annual_income
0	10	RENT	919xx	CA	1998	98000.0
1	10	MORTGAGE	623xx	IL	2001	49200.0
2	3	MORTGAGE	300xx	GA	1999	52000.0
3	6	OWN	431xx	OH	1995	44500.0
4	7	MORTGAGE	919xx	CA	1998	65851.4

Los siguientes enlaces son externos a la Universidad Tecmilenio, al acceder a ellos considera que debes apegarte a sus términos y condiciones.

Consulta la siguiente bibliografía:

- Mota, R. (s.f.). *Annual-income-category-high.json*. Recuperado de <https://gist.github.com/RHDZMOTA/406bb08e1a33469eceb66e5d6bf78e27/>
- Scikit learn. (s.f.). *1. Supervised learning*. Recuperado de [https://scikit-learn.org/stable/supervised\\_learning.html](https://scikit-learn.org/stable/supervised_learning.html)



Basándote en la información anterior, realiza el siguiente ejercicio:

1. Escoge una sola variable y haz una regresión lineal. Utiliza el conjunto de entrenamiento (*train*) para generar el modelo y responde:
  - a. ¿Cuál es el rendimiento predictivo del modelo de regresión con una sola variable? Calcula e interpreta el error cuadrado medio en ambos conjuntos: entrenamiento y prueba.
  - b. ¿Cuál fue la metodología para seleccionar la variable utilizada? Justifica con investigación propia.



Después de revisar la información presentada, es posible comprender la forma en la que las herramientas de *deep learning* apoyan en la proyección de modelos y se pueden usar para definir, por ejemplo, el comportamiento del tipo de cambio en el futuro. Es claro que estos modelos toman como base información publicada por organismos financieros, como el Banco de México y de esta manera se genera la proyección. Es decir, tomando en cuenta el comportamiento del pasado es posible predecir un comportamiento futuro, siempre y cuando el resto de las variables que pudieran afectar la economía permanezcan constantes.

A través de programas como *Python* y conociendo las variables de entrada, se pueden observar los cambios gráficamente y detallar aquellos puntos críticos que pueden afectar alguna inversión, ya sea en forma líquida, o bien, en especie.



- Alexander Amini. (2022, 11 de marzo). *MIT Introduction to Deep Learning (2022) | 6.S191* [Archivo de video]. Recuperado de <https://www.youtube.com/watch?v=7sB052Pz0sQ>
- Baheti, P. (2022). *Activation functions in neural networks [12 types & use cases]*. Recuperado de <https://www.v7labs.com/blog/neural-networks-activation-functions>
- Namini, S., Tavakoli, N., y Namin, A. (2018). *A comparison of ARIMA and LSTM in forecasting time series*. Recuperado de <https://par.nsf.gov/servlets/purl/10186768>
- Williams, K. (2019). *Keras tutorial: deep learning in python*. Recuperado de <https://www.datacamp.com/tutorial/deep-learning-python#introducing-artificial-neural-networks>



SKILLING  
CENTER

**TECMILENIO**



# Inteligencia artificial y machine learning

Modelos de predicción con  
base de árboles de decisión





Dentro de la industria financiera algunas de las actividades que cobran mayor relevancia al tomar una decisión son, sin duda alguna, el riesgo, la razón y la incertidumbre. En la mayoría de los casos las operaciones financieras implican algún tipo de riesgo, relacionados a su propia operación o al mercado al que pertenecen.

Tomar decisiones bajo incertidumbre es complicado ya que se hace necesario generar una combinación de variables para analizar el comportamiento de las mismas.

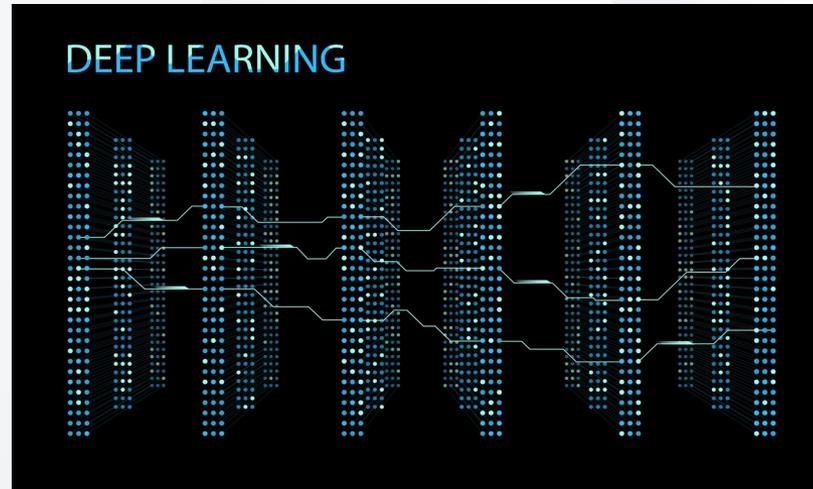
Conocerás sobre los árboles de decisión como modelos. Esto implica una relación de variables simultáneas a través del uso de herramientas de *machine learning*. Descubre lo interesante que esto puede resultar para la toma de decisiones.





Los árboles de decisión son uno de los algoritmos supervisados más utilizados en el área de *machine learning*. Son llamados algoritmos supervisados, ya que se indica el algoritmo a medida que se aprende. Algunos autores, como Grinsztajn et al. (2022) argumentan que los modelos con base en árboles de decisión tienen la facultad de demostrar consistentemente una mejor capacidad predictiva sobre un conjunto de datos tabular, en comparación con otros modelos no-lineales sofisticados tales como el *deep learning*.

En la industria financiera, el tipo de conjunto de datos predominante son los denominados tabulares, lo cual significa que tienen una representación estilo tabla, es decir arreglo de renglones y columnas.





Los árboles de decisión son el modelo más simple que dan pie a un conjunto de modelos más sofisticados. Estos árboles permiten resolver dos tipos de problemas en la categoría de aprendizaje supervisado:

## **Clasificación (*classification tree models*)**

- Resuelve problemas de clasificación en donde la variable objetivo es una variable binaria o categórica. Es muy utilizado para procesar solicitudes de crédito, diferencias, correo no deseado, etcétera.

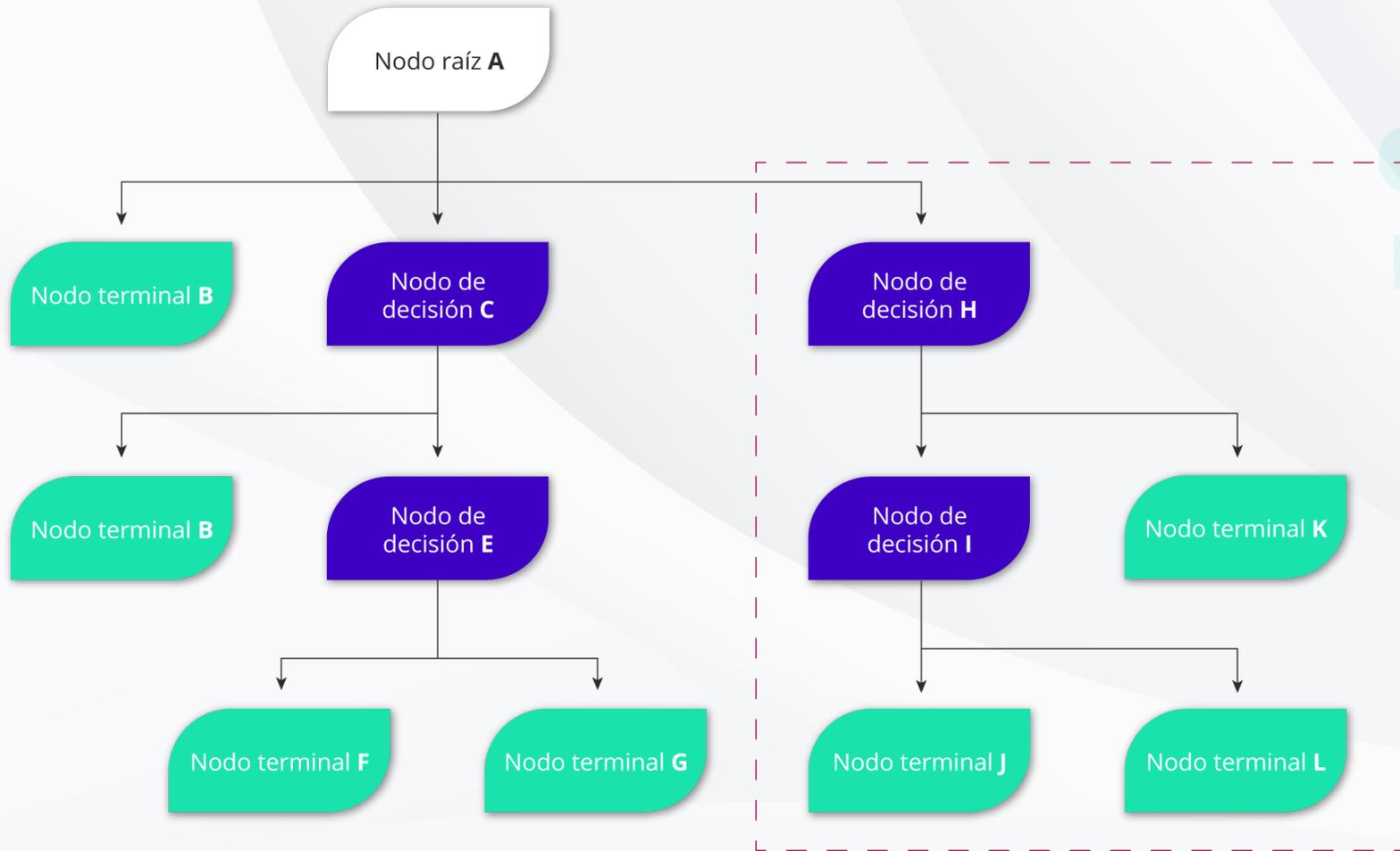
## **Regresión (*regression tree models*)**

- Permite resolver problemas de regresión en donde la variable objetivo es una variable numérica continua. Por ejemplo, determinar el nivel de riesgo crediticio de una persona en relación a su salario, edad, nivel de endeudamiento, entre otros.



Un árbol de decisión está compuesto por un conjunto de nodos entrelazados que se relacionan entre sí, formando una estructura de grafo acíclico dirigido en donde un nodo hijo no puede tener más de un padre. De acuerdo con Grinsztajn et al. (2022), la terminología relevante que tenemos que considerar para entender a los árboles de decisión es la siguiente:

- *Nodo raíz (root node)*: representa el primer nodo del árbol que contiene el conjunto de datos completo.
- *Dividir (splitting)*: proceso de dividir un nodo en subnodos.
- *Nodo padre, nodo(s) hijo(s) y nodo(s) hermano(s) o parent, child, and sibling node*: el nodo padre representa al nodo inmediato anterior, los nodos hijos representan los nodos resultantes de una división y los nodos hermanos representan aquellos nodos que comparten el mismo nodo padre.
- *Nodo de decisión (decision node)*: es el nodo que se divide en nodos adicionales, es decir, un nodo que tiene padre e hijos.
- *Nodo terminal u hoja (terminal node or leaf)*: representa un nodo sin hijos, considerado un nodo final.
- *Poda (pruning)*: es cuando se eliminan nodos del árbol de decisión.
- *Rama (branch)*: representa una sección del árbol (hasta el nodo terminal). También conocido como *subárbol*.





El siguiente paso es identificar cómo se puede hacer uso de esta estructura para resolver un problema de *machine learning* en la categoría de aprendizaje supervisado. Utilizando como referencia el trabajo de Cs.cmu.edu (s.f.), así como el de Boehmke y Greenwell (2020), es posible proponer el siguiente algoritmo recursivo:

1. Asignar todo el conjunto de datos al nodo raíz.
2. Para cada atributo en este conjunto de datos:
  - a) Dividir el conjunto de datos (observaciones) de acuerdo con los valores asociados al atributo.
  - b) Evaluar una función objetivo, la cual incorpora la información de la variable objetivo y los valores del atributo para generar un valor numérico que indica qué tan “bien” se puede utilizar este atributo para segmentar información del conjunto de datos y mejorar la capacidad predictiva del modelo.
    - Consideraciones del atributo:
      - En caso de contener valores categóricos, utilizar cada categoría como criterio de partición.
      - En caso de contener valores numéricos, dividir los datos en dos regiones (dividir en dos partes). Se selecciona el punto de corte que haga más sentido considerando la variable objetivo.
    - Consideraciones de la variable objetivo:
      - En el caso de un problema de regresión, usar el error cuadrado medio respecto al promedio de las observaciones.
      - Para un problema de clasificación, generalmente se utiliza una función de entropía o el índice de Gini.



3. Seleccionar el atributo que genere un mejor resultado en la evaluación de la función objetivo. Si la función objetivo resulta en un valor aceptable, entonces marcar el nodo actual como nodo terminal.
4. Utilizar los valores del atributo para dividir el conjunto de datos (observaciones o renglones) y generar nodos hijos para cada partición.
5. Para cada nodo hijo:
  - a) Si el nodo hijo contiene una partición con solamente un valor único de la variable objetivo, marcarlo como nodo terminal.
  - b) De lo contrario, aplicar el paso dos hacia el nodo hijo de forma recursiva.





Los árboles de decisión son elementos que ayudan a generar información a través de la relación de variables en un mismo plano. Imagina que como asesor financiero quieres generar un modelo de predicción sobre la relación que tienen las variaciones en el tipo de cambio y el pago exitoso de los créditos otorgados.

Si tuvieras que generar un modelo con base en esta información considerando un registro previo de datos, señala:

1. ¿Cuáles serían las variables a considerar?
2. ¿Cuál sería la variable objetivo de detonar?
3. ¿Cómo sería la construcción del modelo y de qué tipo sería (lineal, no lineal, multivariable)?
4. ¿De qué forma apoyarían los árboles de decisión para generar un resultado exitoso?
5. Investiga sobre un caso de éxito utilizando un modelo predictivo de árbol de decisión y compara tu análisis con el resultado del caso de éxito, ¿qué cambiarías en tu modelo para que fuera considerado un éxito?



Ahora conoces más sobre la forma en la que es posible utilizar los modelos de árboles de decisión para relacionar diferentes variables. También podrás elevar el nivel de confianza de tu decisión al otorgar algún tipo crédito, no solo en lo empresarial, sino también en lo personal.

Un árbol de decisión es el elemento más simple en su construcción y permite no solo resolver problemas de clasificación (binarios), sino que también es posible generarlos para dar solución a situaciones con variables numéricas continuas. También pueden utilizarse varios árboles de forma individual y posteriormente agruparse para dar mayor certeza al modelo.

Utilizando herramientas de *machine learning*, ahora puedes tomar mejores decisiones aminorando el riesgo.



- Boehmke, B., y Greenwell, B. (2020). *Hands-on machine learning with R*. Estados Unidos: CRC.
- Cs.cmu.edu. (s.f.). *Decision trees*. Recuperado de <https://www.cs.cmu.edu/~bhiksha/courses/10-601/decisiontrees/>
- Grinsztajn, L., Oyallon, E., y Varoquax, G. (2022). *Why do tree-based models still outperform deep learning on tabular data?* Recuperado de <https://arxiv.org/abs/2207.08815>

*Tecmilenio no guarda relación alguna con las marcas mencionadas como ejemplo. Las marcas son propiedad de sus titulares conforme a la legislación aplicable, estas se utilizan con fines académicos y didácticos, por lo que no existen fines de lucro, relación publicitaria o de patrocinio.*

---

*Todos los derechos reservados @ Universidad Tecmilenio*

*La obra presentada es propiedad de ENSEÑANZA E INVESTIGACIÓN SUPERIOR A.C. (UNIVERSIDAD TECMILENIO), protegida por la Ley Federal de Derecho de Autor; la alteración o deformación de una obra, así como su reproducción, exhibición o ejecución pública sin el consentimiento de su autor y titular de los derechos correspondientes es constitutivo de un delito tipificado en la Ley Federal de Derechos de Autor, así como en las Leyes Internacionales de Derecho de Autor. El uso de imágenes, fragmentos de videos, fragmentos de eventos culturales, programas y demás material que sea objeto de protección de los derechos de autor, es exclusivamente para fines educativos e informativos, y cualquier uso distinto como el lucro, reproducción, edición o modificación, será perseguido y sancionado por UNIVERSIDAD TECMILENIO. Queda prohibido copiar, reproducir, distribuir, publicar, transmitir, difundir, o en cualquier modo explotar cualquier parte de esta obra sin la autorización previa por escrito de UNIVERSIDAD TECMILENIO. Sin embargo, usted podrá bajar material a su computadora personal para uso exclusivamente personal o educacional y no comercial limitado a una copia por página. No se podrá remover o alterar de la copia ninguna leyenda de Derechos de Autor o la que manifieste la autoría del material.*