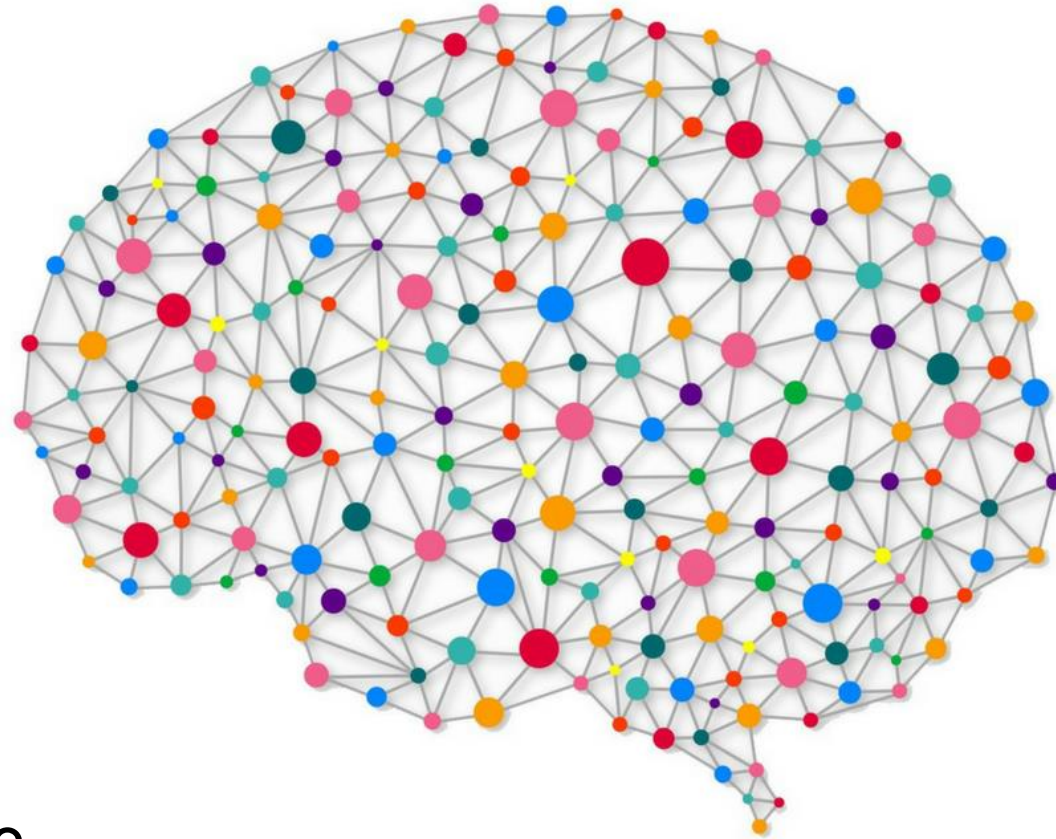


# PRONOSTICO CON REDES NEURONALES



Profesor: Daniel Osorio



# DATOS SECUENCIALES

- Datos **no secuenciales**:
  - Independencia
    - No hay relación entre registros (filas)
    - No hay relación entre variables (columnas)
  - Si se cambia el orden de registros o variables, un modelo aprendido no se afecta
- Datos **secuenciales**:
  - Texto: letras, palabras
  - Ventas semanales, mensuales
  - Secuencias de cromosomas
  - Relación de **orden** (eje dimensional), cambia el significado:
    - 10, 9, 8, 7, 6 vs 6, 7, 8, 9, 10
    - “El vino que me cae bien” vs. “Me cae bien el que vino”

Edad	Ingresos	# Hijos	Tiene carro?
24	3'200.000	0	NO
28	4'500.000	2	SI
45	6'500.000	1	SI
32	4'600.000	0	NO

	Item	Feb	Mar	Abr	May	
	001	15.5	16.1	16.8	16.7	
	002	8.2	8.6	8.7	8.9	
	007	10.2	10.1	9.6	8.7	
	021	4.6	4.5	4.6	4.5	
						egresos
						ingresos



# DATOS SECUENCIALES

- **Tareas** de modelos con data secuencial:
  - Forecasting univariado y multivariado
    - Pronóstico de valor de acciones
    - Pronóstico de tiempo
    - Pronóstico de inventario
  - Reconocimiento de habla
  - Traducción de texto
  - Clasificación de texto/audio
  - Generación de texto/audio
  - Modelamiento de proteínas
  - ...



# SECUENCIAS EN REDES NEURONALES DENSAS

Los datos de una serie de tiempo univariada o multivariada son entonces la base de una creación de un **nuevo dataset** teniendo una neurona en la capa de entrada por cada **variables predictoras**:

- Valores de los últimos periodos (**lag features**) que denoten la tendencia de los últimos datos (i.e., ventana móvil con datos de las últimas 8 semanas)
- Valores estacionales de diferentes periodicidades (i.e. datos de los últimos 4 trimestres, datos de los últimos 3 años)
- Valores agregados sobre periodos (i.e. moving average)
- Incluir información de la variable a predecir, pero también de otras covariables predictoras.

Con redes neuronales artificiales tradicionales, que solo involucran **capas densas**, no hablamos propiamente de pronóstico (forecasting) sino de **regresión**. Así hayan correlaciones en las variables predictoras, el modelo las considera como independientes, sin ninguna noción de orden o secuencialidad.

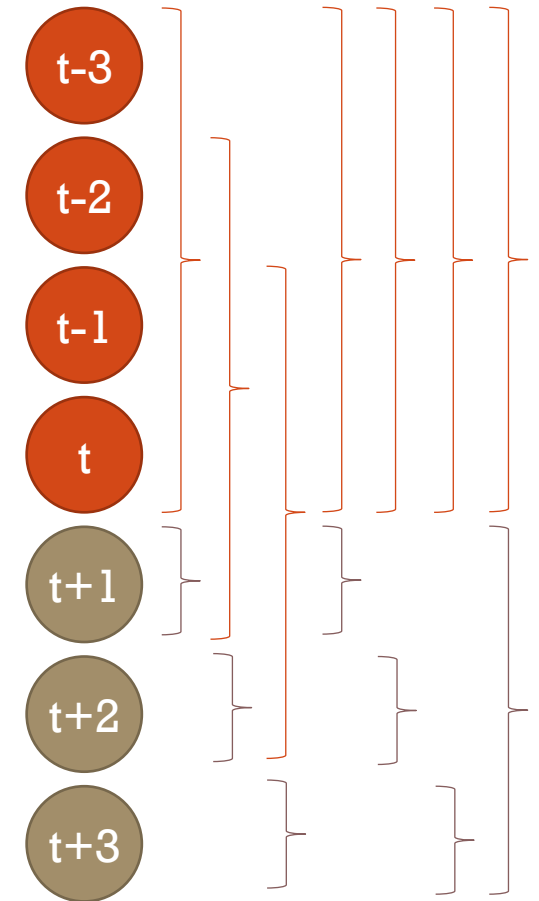
Igualmente en la creación del dataset es necesario definir él o los **valores futuros a predecir**, teniendo una neurona para cada uno en la capa de salida.



# SECUENCIAS EN REDES NEURONALES DENSAS

Las redes tradicionales con capas **densas** permiten realizar tareas de pronóstico sobre datos **univariados**. Es necesario definir la arquitectura de los modelos, estableciendo:

- El número de periodos sobre los cuales se basará el pronóstico: determina el número de neuronas en la **capa de entrada**
- El número de periodos a pronosticar en el futuro, que determina el número de neuronas en la **capa de salida**. En el caso de pronóstico de múltiples periodos, hay que definir la estrategia a seguir:
  - Crear un solo modelo: capa de salida con una neurona de salida para un único periodo a pronosticar (**iterativo**), que se desplaza auto-regresivamente.
  - Crear un modelo independiente para cada período en el futuro (**directo**): tantos modelos como periodos a pronosticar, con una neurona en la capa de salida.
  - Crear un solo modelo: capa de salida con tantas neuronas de salida como periodos a pronosticar (**salida de vector**).



# PREPARACIÓN DE DATOS PARA PREDICCIÓN

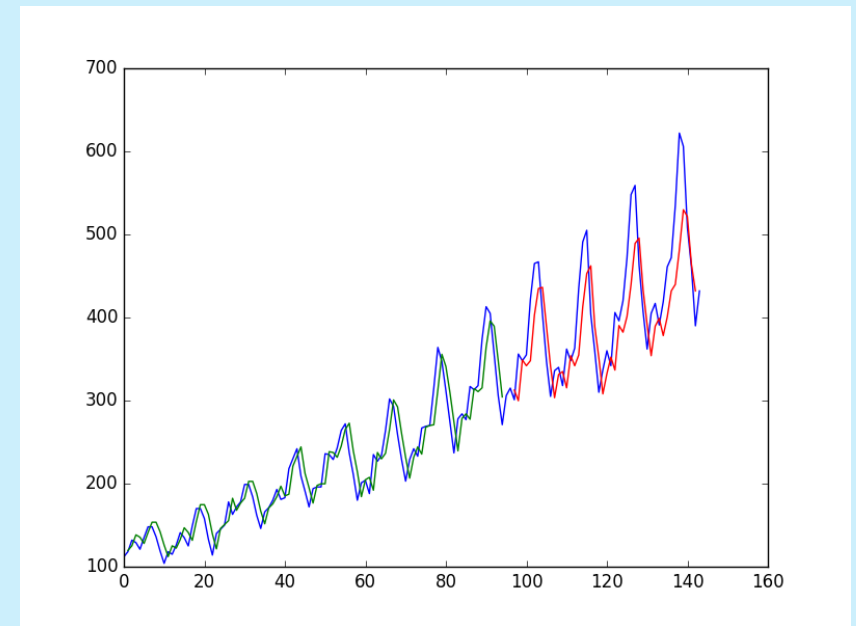
- Exploración y visualización de datos secuenciales
  - Identificación de estacionalidades y ciclicidades
  - Imputación de valores faltantes según la frecuencia de datos (considerar tendencias y estacionalidades)
  - Normalización y/o transformaciones (opcional, a determinar si mejoran resultados)
- Creación de variables de corte transversal:
  - Lag features: valores anteriores, a diferentes frecuencias (considerar estacionalidades y ciclicidades)
  - Agregaciones móviles: conteos, sumas, promedios, desviaciones, etc.
  - Variables temporales: día de la semana, día del mes, mes, festivos, etc.
  - Variables exógenas: regiones, sucursales, vendedor, etc.
- Particionamiento temporal train/test (barajar los de train).



# FORECASTING CON REDES DENSAS TRADICIONALES

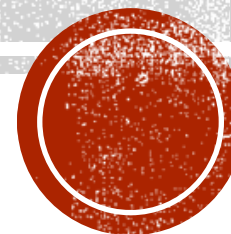
Utilizar una red neuronal tradicional para crear un modelo de forecasting:

- Creación de dataset por ventanas de tiempo
- Modelo de red neuronal artificial tradicional para realizar pronósticos a partir de regresión





**GRACIAS**





# REFERENCIAS

- *Deep Learning for Time Series Forecasting: Predict the Future with MLPs, CNNs and LSTMs in Python*, Jason Brownlee, 2019
- *Machine Learning with PyTorch and Scikit-Learn*, Sebastian Raschka, Yuxi Liu & Vahid Mirjalili, Packt, 2022
- Statistical and Machine Learning forecasting methods: Concerns and ways forward, Makridakis et al., 2018
- *Deep Learning for Time-Series Analysis*, John Cristian Borges Gamboa, 2017
- *Dive into Deep Learning*, Aston Zhang, Zachary C. Lipton, Mu Li & Alexander J. Smola, <https://d2l.ai/>, 2022

