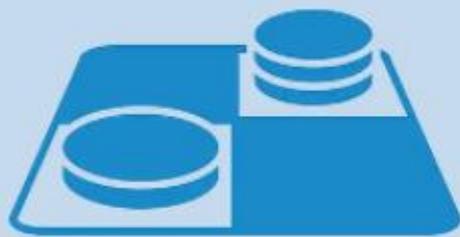


Comparación de metodologías de Deep Learning para pronósticos en series temporales

Rodrigo Del Rosso
rdelrosso@economicas.uba.ar

ARTIFICIAL INTELLIGENCE

Artificial Intelligence captures the imagination of the world.



MACHINE LEARNING

Machine learning starts to gain traction.



DEEP LEARNING

Deep learning catapults the industry.



1950

1960

1970

1980

1990

2000

2010

2020

2030

2040

XXII Jornadas Nacionales
y Latinoamericanas Actuariales

 **consejo** GESTIÓN
Y FUTURO

.UBA200

.UBAeconómicas

A Tour of AI Technologies in Time Series Prediction

Deep Learning: Subset of ML that enables machines to learn, train, and perform tasks by themselves by exposing on multilayered neural networks.

Machine Learning: A subset of AI that includes statistical algorithms that enable machines to improve learning over time.

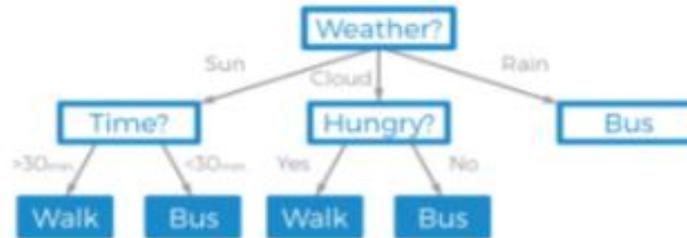
Artificial Intelligence: Techniques that enable computers to mimic human intelligence.

<https://www.soa.org/globalassets/assets/files/resources/research-report/2019/tour-ai-technologies.pdf>

Machine Learning



Input



Decision tree

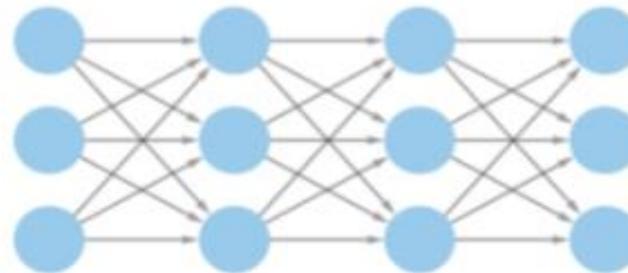


Output

Deep Learning



Input



Feature extraction + Classification



Output

¿Qué metodologías se utilizan para analizar las series de tiempo?

- **Metodologías tradicionales**

- Modelos $ARIMA(p,d,q)$
- Modelos $SARIMA(p,d,q)(P,D,Q)[m]$
- Modelos $GARCH(p,q)$ y sus variantes
- Suavizado Exponencial de Holt-Winters
- Transformada Rápida de Fourier
- Etc

- **Metodologías modernas**

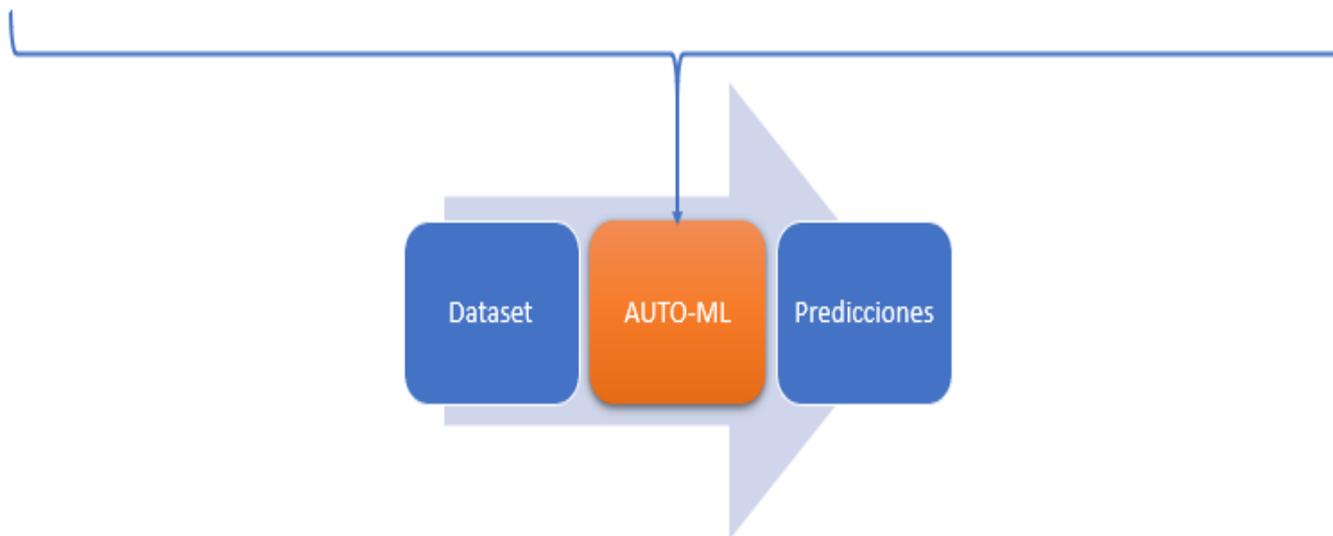
- Redes Neuronales Autorregresivas (NNAR)
- Long-Short Term Memory (LSTM - RNN)
- Algoritmo Prophet de Facebook (2017)
- Algoritmo Neural Prophet
- Aprendizaje Automático Automatizado (AutoML)
- Etc

El AutoML consiste en automatizar el proceso de aplicación de técnicas de data mining para resolver problemas reales y tiende a mecanizar la mayor cantidad de pasos, y reducir la intervención del usuario, sin comprometer el rendimiento del modelo (Pandey, 2019).

En una aplicación de *data mining* común, los usuarios cuentan con un conjunto de datos. A dichos datos se le deben aplicar técnicas de preprocesamiento, realizar ingeniería de atributos, extracción y selección de variables para hacer el conjunto de datos legible para el modelo a aplicar. Terminada dicha etapa, sigue la selección del modelo y la optimización de hiperparámetros para maximizar el rendimiento predictivo.

AutoML simplifica drásticamente estos pasos e incluye de manera muy amplia,

- Automatizar ciertas partes de la preparación de datos como, por ejemplo: imputación de variables, estandarización de variables, selección de atributos, etc.
- Ser capaz de generar varios modelos de forma automática y optimizarlos
- Obtener el mejor modelo de todos los modelos generados.



Ventajas y Desafíos de AutoML

Los principales beneficios de AutoML son:

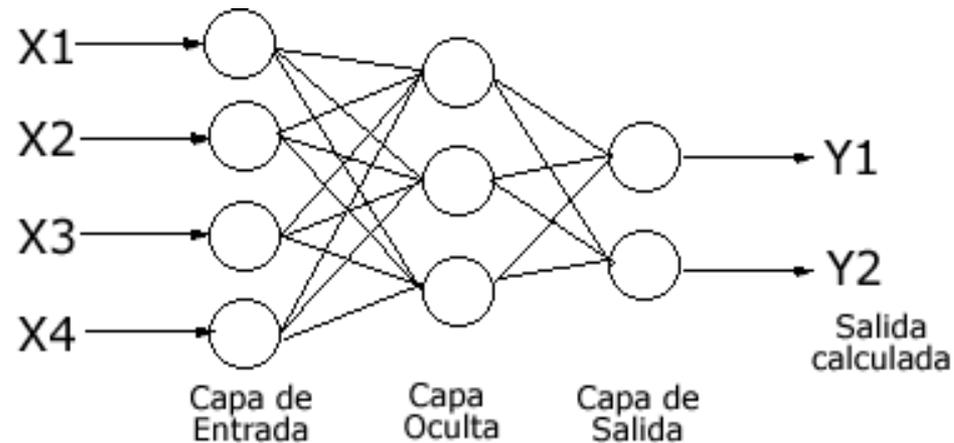
- **Eficiencia:** acelera y simplifica el proceso de data mining y reduce el tiempo de creación de modelos.
- **Ahorro de costos:** la consecuencia del paso anterior implica menor tiempo en tareas manuales y por lo tanto un ahorro de costos.
- **Accesibilidad:** tener un proceso más simple permite que la aplicación de herramientas de data mining sea posible a una gama más amplia de perfiles profesionales.
- **Rendimiento:** los algoritmos de AutoML también tienden a ser más eficientes que los modelos programados de forma manual.

Uno de los principales desafíos de AutoML es la tentación de verlo como un reemplazo del conocimiento humano. AutoML no puede entender los problemas de negocio, no puede entender los datos ni adquirirlos. AutoML está diseñado para realizar tareas de forma eficiente y con precisión, lo que permite que los usuarios se concentren en tareas más complejas o novedosas. Sin embargo, el usuario aún debe participar para evaluar y supervisar el modelo, pero ya sin la necesidad de intervenir en cada paso del proceso de data mining. (Lutkevich, 2020).

Redes Neuronales

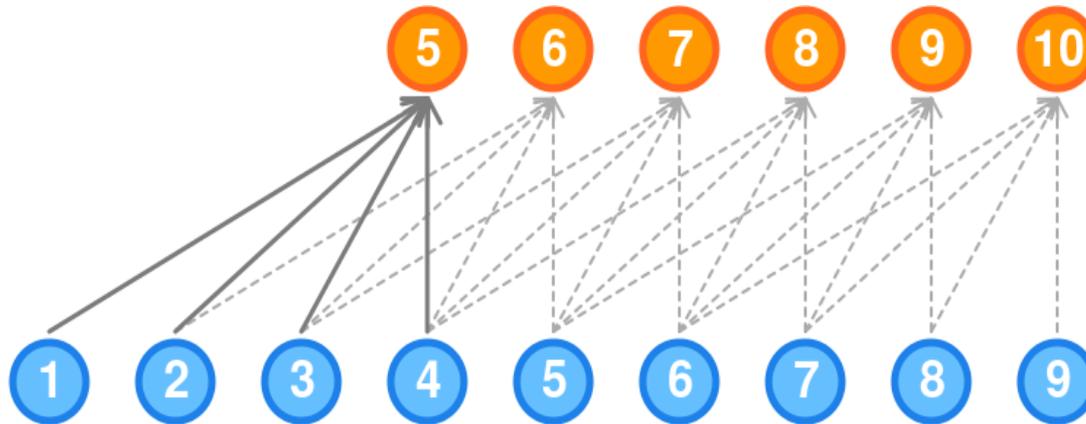
- Son un modelo computacional inspirado en el comportamiento observado en su homólogo biológico.
- Consiste en un conjunto de unidades, llamadas neuronas artificiales, conectadas entre sí para transmitirse señales.
- La información de entrada atraviesa la red neuronal (dónde es sometida a diversas operaciones) produciendo unos valores de salida. Entre varias aplicaciones de estas, cabe destacarse su utilización como métodos de pronóstico de series temporales.
- Cada neurona está conectada con otras a través de unos enlaces y el valor de salida de la neurona anterior es multiplicado por un valor de peso. Estos pesos en los enlaces pueden incrementar o inhibir el estado de activación de las neuronas adyacentes. Del mismo modo, a la salida de la neurona, puede existir una función limitadora o umbral, que modifica el valor resultado o impone un límite que se debe sobrepasar antes de propagarse a otra neurona. Esta función se conoce como función de activación (McCulloch & Pitts, 1943).

- Una red neuronal puede tener varias capas ocultas o no tener ninguna. Las conexiones sinápticas (las flechas que llegan y salen de las neuronas) indican el flujo de la señal a través de la red, y tienen asociadas un peso sináptico correspondiente. Si la salida de una neurona va dirigida hacia dos o más neuronas de la siguiente capa, cada una de estas últimas recibe la salida neta de la neurona anterior. La cantidad de capas de una red neuronal es la suma de las capas ocultas más la capa de salida. Un ejemplo sencillo de perceptrón multicapa (red neuronal con capas ocultas) se exhibe en la siguiente figura,



Redes Neuronales Autorregresivas (NNAR)

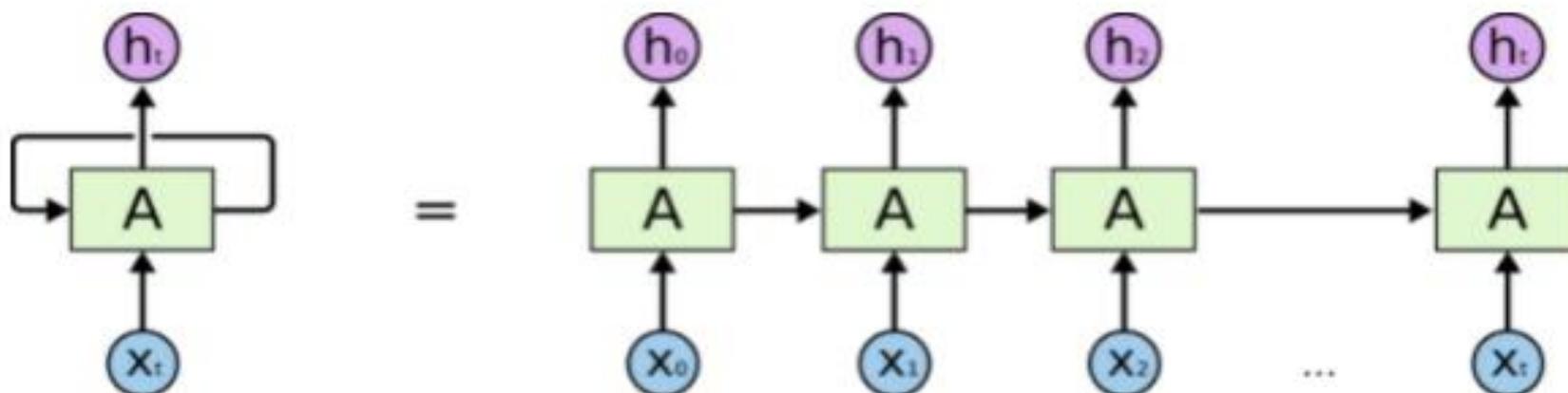
Un primer acercamiento al uso de redes neuronales para predecir valores futuros es emplear una Red Neuronal Autorregresiva.



La clave para el uso de este tipo de redes es considerar los datos ordenados de la serie temporal como entradas de la red, del mismo modo que se utilizan los valores rezagados (o desfasados) en un modelo autorregresivo tradicional. Cuando se trata de pronósticos mediante este algoritmo, la red se aplica de forma iterativa. Para pronosticar un paso adelante, simplemente utiliza las entradas históricas disponibles. Para pronosticar dos pasos adelante, emplea el pronóstico de un paso como entrada, junto con los datos históricos. Este proceso continúa hasta que se haya computado todos los pronósticos requeridos.

Redes de Memoria Largo-Corto Plazo (LSTM)

Las redes neuronales recurrentes (RNN) tienen la capacidad de reconocer y predecir secuencias de datos a lo largo del tiempo y se fundamentan en bucles que permiten que la salida de la red o de una parte de ella en un momento dado sirva como entrada de la misma en el siguiente momento.



En la medida en que se incorporan capas, su capacidad de modelado crecerá de manera que será capaz de reconocer mayores secuencias cada vez con menor error. Tienen la capacidad de aprender dependencias a largo plazo.

Prophet

- Es la popular herramienta de pronóstico desarrollada por Facebook y liberada en 2017 para pronosticar series de tiempo basado en un modelo aditivo en el que las tendencias no lineales se ajustan a la estacionalidad.
- Una de las particularidades es que se adapta mejor a series de tiempo que poseen fuertes efectos estacionales y varias temporadas de datos históricos.
- Soporta datos faltantes y cambios en la tendencia, y normalmente maneja bien los valores atípicos.
- El procedimiento es un modelo de regresión aditivo de tres funciones de tiempo (crecimiento, estacionalidad y días festivos / eventos) más un término de error (Taylor et al., 2017),

$$y(t) = g(t) + s(t) + h(t) + \varepsilon_t$$

Donde $g(t)$ el crecimiento, $s(t)$ la estacionalidad, $h(t)$ los días festivos / eventos y ε_t el término de error.

- Es una red neuronal que fuertemente inspirada en el algoritmo mencionado en el apartado anterior.
- El objetivo de la librería es combinar el poder de las redes neuronales y la interpretabilidad de los modelos autorregresivos tradicionales, que tienen la ventaja de que son interpretables e intuitivos, lo que significa que se puede exhibir cómo se ponderan los valores pasados y cómo esa entrada influye en las predicciones.
- La desventaja es que estos modelos son paramétricos (demasiado rígidos), no pueden tener en cuenta ninguna característica fuera de la autocorrelación y hacen un mal trabajo al escalar conjuntos de datos con un mayor número de regresores autocorrelacionados (Hyndman & Athanasopoulos, 2018).
- Las redes neuronales, por otro lado, no son paramétricas, lo que es importante para las series de tiempo dado que rara vez siguen una forma estándar única. Por lo general, pueden mapear cualquier función no lineal para aproximarse a cualquier expresión de índole continua. Para la librería en cuestión, la red neuronal se la conoce como AR-Net.

Conclusiones

- *Prophet* es conocido por sus hiperparámetros intuitivos y componentes interpretables, y *Neural Prophet* continúa con ambas características.
- AR-Net maneja la autorregresividad para acomodar una mayor cantidad de entradas, lo que permite una representación más precisa de la señal de la serie temporal.
- *Neural Prophet* utiliza las tendencias lineales por partes del *Prophet* original y la serie de Fourier para dar cuenta de múltiples estacionalidades, pero la incorporación de regresores adicionales se actualiza en mediante el uso de redes neuronales adicionales.

Referencias Bibliográficas

- Hyndman, R.J., & Athanasopoulos, G. (2018) Forecasting: principles and practice, 2nd edition, OTexts: Melbourne, Australia. OTexts.com/fpp2
- Lutkevich, B. (Octubre de 2020). Tech Target. Auto machine learning (AutoML). Obtenido de <https://searchenterpriseai.techtarget.com/definition/automated-machine-learning-AutoML>
- McCulloch, W.S., Pitts, W. A logical calculus of the ideas immanent in nervous activity. Bulletin of Mathematical Biophysics 5, 115–133 (1943). <https://doi.org/10.1007/BF02478259>
- Pandey, P. (14 de Octubre de 2019). Towards Data Science. Obtenido de <https://towardsdatascience.com/a-deep-dive-into-h2os-automl-4b1fe51d3f3e>
- Taylor SJ, Letham B. 2017. Forecasting at scale. PeerJ Preprints 5:e3190v2 <https://doi.org/10.7287/peerj.preprints.3190v2>
- Triebe, Oskar & Laptev, Nikolay & Rajagopal, Ram. (2019). AR-Net: A simple Auto-Regressive Neural Network for time-series.