**Diseño una red MLP que permita predecir series de**

**Tiempo**

**Perceptrón multicapa para calcular la función Seno**

**Multilayer perceptron to calculate the Sine function**

**Autor 1: Ángelo Gutiérrez Correa**

**Autor 1: Juan David Gallego Rangel**

*Universidad Tecnológica de Pereira, Risaralda, Colombia*

[angutierrez@utp.edu.co](mailto:angutierrez@utp.edu.co)

fbdavid37@utp.edu.co

***Resumen*— Gran cantidad de información económica es recopilada con fines de análisis. ¿Qué son las series de tiempo? Es un registro metódico a intervalos de tiempo fijos de las características de una variable, o su observación numérica.**

**Herramientas estadísticas: Se usan para describir y analizar series de tiempo, Describir: Resumir en forma concisa, mediante gráficas y parámetros, la información.**

**Analizar: inferir de la información, características de toda la población. Para inferir, muchas veces es útil “descomponer” la serie de tiempo por sus principales componentes:**

**· Tendencia y ciclo: representa el movimiento de largo plazo de la serie.**

**· Estacionalidad: representa efectos de fenómenos que ocurren o se reproducen periódicamente.**

**· Irregularidad: movimientos impredecibles o aleatorios. Tendencia y ciclo y estacionalidad, conforman la parte determinística de la serie, mientras que la irregularidad representa la parte no-determinística o estocástica de la serie.**

***Palabras clave—* MLP, algoritmos, Backpropagation, redes neuronales, series de tiempo, predicción, entrenamiento, validación, perceptrón, multicapa.**

***Abstract— A great deal of economic information is collected for analysis purposes. What are time series ?: Methodical record at fixed time intervals of the characteristics of a variable, or its numerical observation.***

***Statistical tools: They are used to describe and analyze time series. Describe: Summarize the information in a concise way, using graphs and parameters.***

***Analyze: infer from the information, characteristics of the entire population. To infer, it is often useful to “decompose” the time series by its main components:***

***· Trend and cycle: represents the long-term movement of the series.***

***· Seasonality: represents effects of phenomena that occur or are reproduced periodically.***

***· Irregularity: unpredictable or random movements. Trend and cycle and seasonality make up the deterministic part of the series, while the irregularity represents the non-deterministic or stochastic part of the series.***

***Keywords— MLP, algorithms, Backpropagation, neural networks, time series, prediction, training, validation, perceptron, multilayer***

**INTRODUCCIÓN**

Toda institución, ya sea la familia, la empresa o el gobierno, tiene que hacer planes para el futuro si ha de sobrevivir y progresar. Hoy en día diversas instituciones requieren conocer el comportamiento futuro de ciertos fenómenos con el fin de planificar, prever o prevenir. La planificación racional exige prever los sucesos del futuro que probablemente vayan a ocurrir. La previsión, a su vez, se suele basar en lo que ha ocurrido en el pasado. Se tiene pues un nuevo tipo de inferencia estadística que se hace acerca del futuro de alguna variable o compuesto de variables basándose en sucesos pasados. La técnica más importante para hacer inferencias sobre el futuro con base en lo ocurrido en el pasado es el análisis de Series de Tiempo. Son innumerables las aplicaciones que se pueden citar, en distintas áreas del conocimiento, tales como, en economía, física, geofísica, química, electricidad, en demografía, en marketing, en telecomunicaciones, en transporte, etc.

**El OBJETIVO DEL ANÁLISIS DE UNA SERIE DE TIEMPO**

Los principales conceptos y características que se tienen en cuenta para el análisis de una serie de tiempo son: 1. Obtener modelos estadísticos que describen la estructura pasada de las observaciones que generan la serie; y/o estudiar modelos que explican la variación de una serie en términos de series (explicativas) conocidas. 2. Suponer que la estructura pasada de la serie de interés o de las series explicativas se conserva y bajo este supuesto, pronosticar valores futuros de la serie bajo estudio. 3. Analizar la significancia de los efectos que políticas o intervenciones pasadas causaron en la estructura de la serie.4. Simular valores futuros de la serie, bajo condiciones o restricciones definidas por políticas o criterios nuevos, para así supervisar y controlar los cambios que se producen en la serie. Para una mayor comprensión de la estructura, el comportamiento y el análisis de una serie de tiempo se debe tener en cuenta como primera medida conocer y entender algunos términos que se describirán a continuación.

**COMPONENTES DE UNA SERIE DE TIEMPO**

Una Serie de Tiempo puede tener las siguientes componentes:

1. Tendencia: Esta componente representa la trayectoria suavizada que define la serie en el rango de variación del índice y se halla observando la forma funcional de la gráfica de la serie z vs t) (t a lo largo del tiempo. La tendencia puede ser: constante, línea, cuadrática, exponencial, etc.

2. Componente Estacional: Esta componente se presenta cuando la serie tiene patrones estaciónales que se repiten con una frecuencia constante, produciendo en su grafica un efecto periódico. Los patrones estaciónales se presentan por fenómenos climáticos, recurrencia en los pagos, costumbres y/o agrupamiento, afectan las observaciones que generan la serie.

3. Componente Aleatoria: Esta componente se representa los cambios que sufre la serie ocasionados por fenómenos externos no controlables.

4. Componente Cíclica: Esta componente se presenta en series que son afectadas por fenómenos físicos o económicos que ocurren con una periodicidad variable.

**ENTRADAS Y SALIDAS.**

Las variables de entrada y salida pueden ser binarias (digitales) o continuas (analógicas), dependiendo del modelo y aplicación. Por ejemplo, un perceptrón multicapa o MLP (**Multilayer Perceptrón)** admite ambos tipos de señales. Así, para tareas de clasificación poseería salidas digitales {O, +1}, mientras que, para un problema de ajuste funcional de una aplicación multivariable continua, se utilizarían salidas continuas pertenecientes a un cierto intervalo.

Dependiendo del tipo de salida, las neuronas suelen recibir nombres específicos. Así, las neuronas estándar cuya salida sólo puede tomar los valores 0 o 1 se suelen denominar genéricamente neuronas de tipo McCullochPitts, mientras que aquellas que únicamente pueden tener por salidas -1 o + 1

**REGLA DE PROPAGACION**

La regla de propagación permite obtener, a partir de las entradas y los pesos el valor del potencial postsináptico hI de la neurona. La función más habitual es de tipo lineal, y se basa en la suma ponderada las entradas con los pesos sinápticos.



que formalmente también puede interpretarse como producto escalar de los vectores de entrada y los pesos.

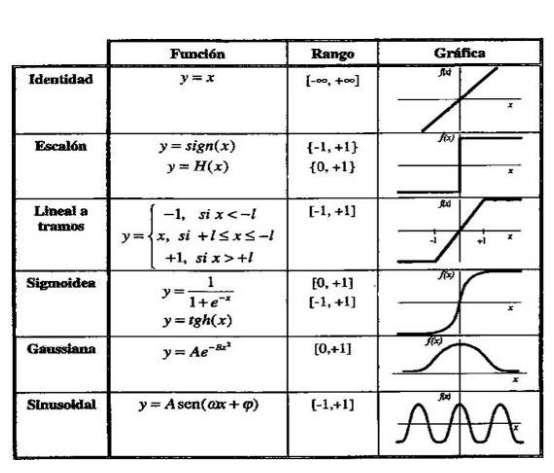
**FUNCIÓN DE ACTIVACION**

La función de desempeño o de transferencia proporciona el estado de actual aI(t) a partir del potencial postsináptico hI (t) y del propio estado de anterior aI(t-1). Sin embargo, en muchos modelos se considera que el estado actual de neurona no depende de su es estado anterior, sino únicamente del actual.

La función de desempeño se suele considerar determinista, y en la mayor parte de los modelos es monótona creciente y continua, como se observa habitualmente en las neuronas. Las funciones de activación más empleadas en muestra en la figura 1.1

En ocasiones los algoritmos de aprendizaje requieren que la función de desempeño cumpla la Condición de ser derivable. Las más empleadas en este sentido son las funciones de tipo sigmoideo.

Para explicar porque se utilizan estas funciones de activación se suele emplear la analogía a la aceleración de un automóvil. Cuando un auto inicia su movimiento necesita una potencia elevada para comenzar a acelerar. Pero al ir tomando velocidad, esta demanda un menor incremento de dicha potencia para mantener la aceleración. Al llegar a altas velocidades, nuevamente un amplio incremento en la potencia es necesario para obtener una pequeña ganancia de velocidad. En resumen, en ambos extremos del rango de aceleración de un automóvil se demanda una mayor potencia para la aceleración que en la mitad de dicho rango.



**APRENDIZAJE DE UNA RED NEURONAL**

El aprendizaje es el proceso por el cual una red neuronal modifica sus pesos en respuesta a una información de entrada. Los cambios que se producen durante el mismo se reducen a la destrucción, modificación y creación de conexiones entre las neuronas. Una red neuronal debe aprender a calcular la salida correcta para cada constelación (arreglo o vector) de entrada en el conjunto de ejemplos.

Este proceso de aprendizaje se denomina: Proceso de Entrenamiento o Acondicionamiento. El conjunto de datos (o conjunto de ejemplos) sobre el cual este proceso se basa es, por ende, llamado: Conjunto de datos de Entrenamiento.

Durante el proceso de aprendizaje, los pesos de las conexiones de la red sufren modificaciones, por lo tanto, se puede afirmar que este proceso ha terminado (la red ha aprendido) cuando los valores de los pesos permanecen estables, si los pesos cambian y sus valores son iguales a 0, se dice que la conexión de la red se ha destruido.

De esta manera, se dice que los pesos se han adaptado, ya que sus valores son distintos de 0 y su derivada es igual a 0. En otras palabras, el aprendizaje es el proceso por el cual una red neuronal modifica sus pesos en respuesta a una información de entrada.

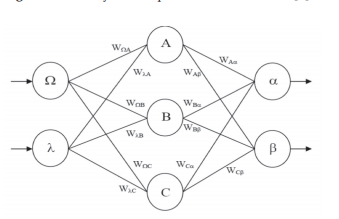
Los cambios que se producen durante el mismo se reducen a la destrucción, modificación y creación de conexiones entre las neuronas.

**EL ERCEPTRON MULTICAPA (MLP)**

se ha visto las limitaciones del perceptrón simple, ya que con él tan solo se puede discriminar patrones que pueden ser separados por un hiperplano (una recta en el caso de dos neuronas de entrada). Una manera de solventar estas limitaciones del perceptrón simple es por medio de la inclusión de capas ocultas, obteniendo de esta forma una red neuronal que se denomina perceptrón multicapa.

La Figura a continuación muestra las regiones de decisión que se obtienen para distintas arquitecturas de redes neuronales considerando dos neuronas en la capa inicial. Así por ejemplo para una arquitectura de perceptrón simple la región de decisión es una recta, mientras que el perceptrón multicapa con una única capa de neuronas ocultas puede discriminar regiones convexas.

Por otra parte, el perceptrón multicapa con dos capas de neuronas ocultas es capaz de discriminar regiones de forma arbitraria.



**REDES NEURONALES EN LA PREDICCION**

En la actualidad, las Redes Neuronales Artificiales, se reconocen como una de las herramientas matemáticas de uso computacional que mejores resultados presenta a la hora de modelar un problema, ya sea de aproximación de funciones: como el diagnóstico y control de maquinaria, en el control del piloto automático de un avión, o en el control de un robot. En problemas de clasificación: como el diagnóstico médico, reconocimiento de caracteres, detección de fraude o clasificación de riesgo crediticio. En el procesamiento de datos: como en el filtrado de ruido o en el encripta miento de información

**EL MODELO**

Un perceptrón multicapa (MLP, por sus siglas en inglés) es un tipo de red neuronal que imita la estructura masivamente paralela de las neuronas del cerebro. Desde un punto de vista matemático, es un modelo que puede aproximar cualquier función continúa definida en un dominio compacto con una precisión arbitraria previamente establecida (Hornik y Stinchcombe, 1989; Cybenko, 1989; Funahashi, 1989). En la práctica, los MLP se han caracterizado por ser muy tolerantes a información incompleta, inexacta o contaminada con ruido por lo que han sido usados en la modelación empírica de series temporales no lineales.

**LA ESPECIFICACION DEL MLP**

El proceso de especificación de un MLP es difícil debido a la gran cantidad de pasos metodológicos que requiere (la selección de las entradas, la determinación de las neuronas de la capa oculta, la estimación de los parámetros (Ortiz et ál., 2007), etc.), a los criterios subjetivos en cuanto a cómo abordar cada paso, y a que los resultados obtenidos en cada paso son críticos.

Uno de los aspectos que dificultan el proceso de especificación es la falta de identificabilidad estadística del modelo. Esto se refiere a que los parámetros óptimos no son únicos para una configuración específica de un modelo (número de entradas, cantidad de neuronas en la capa oculta, funciones de activación, etc.), y un conjunto de datos dado, debiéndose a que:

* Se pueden obtener múltiples configuraciones idénticas en comportamiento cuando se permutan las neuronas de la capa oculta, manteniendo vinculadas las conexiones que llegan a dichas neuronas.
* -Cuando las neuronas de la capa oculta tienen funciones de activación simétricas alrededor del origen, la contribución neta de la neurona a la salida de la red neuronal se mantiene igual si se cambian los signos de los pesos que entran y salen ella.
* -Si los pesos de las conexiones entrantes a una neurona oculta equivalen a cero, es imposible determinar el valor del peso de la conexión de esa neurona oculta a la neurona de salida.
* -Si el peso de la conexión de una neurona oculta hacia la neurona de salida es cero, es imposible identificar los valores de los pesos de las conexiones entrantes a la neurona oculta. Las consideraciones anteriores sobre la identificabilidad del modelo son el punto inicial para plantear modificaciones sobre la especificación del MLP de manera tal que se obtengan nuevas configuraciones que puedan modelar y pronosticar series no lineales de una forma más simple y permitan obtener mejores resultados en comparación con otros modelos.

**ANALISIS DE LAS SERIE**

Para el desarrollo del análisis de la serie se deben escalan los datos, esto se debe a las funciones de transferencia que se pretenden usar en el entrenamiento de las redes, las funciones utilizadas requieren entradas en el **rango [-1, 1].** Para este fin se utilizan las herramientas de Preprocesamiento de datos por medio de la función **SENO.**

**CODIGO**

Se organizan los datos en 2 columnas, 3 columnas y 4 columnas. Esta organización se realiza con el fin de manejar topologías con 1, 2 y 3 neuronas en la capa de entrada y variar de 2 a 3 en la capa oculta y hasta 2 capas ocultas, esto último tomando en cuenta el número de ejemplos de entrenamiento que se tiene disponible de la forma que se escaló y el numero optimo definido.

**LIBRERIAS UTILIZADAS Y CREACION DE LA CLASE PRINCIPAL**

Texto

Descripción generada automáticamente

**CREANDO LA RED NEURONAL**

Texto

Descripción generada automáticamente

**CREANDO LOS DATOS CON BASE A LOS PARAMETROS INICIALES**

**Texto

Descripción generada automáticamente**

**CREANDO EL ENTRENAMIENTO SUPERVISADO CON EL METODO DE CALCULO DEL GRADIENTE**

**Captura de pantalla de un celular

Descripción generada automáticamente**

**CALCULO DE LA TASA DE ERROR INICIAL**

**Texto

Descripción generada automáticamente**

**TASA ERROR INICIAL OBTENIDA**

[**Computacion-Blanda\Parcial 2\Datos generados.xlsx**](../Desktop/UTP/COMPUTACION%20BLANDA/Computacion-Blanda/Parcial%202/Datos%20generados.xlsx)

Texto

Descripción generada automáticamente

**Gráfico, Gráfico de dispersión

Descripción generada automáticamente**

**Gráfico, Gráfico de líneas

Descripción generada automáticamente**

**REPRESENTACIONES GRAFICAS**

**Texto

Descripción generada automáticamente**

**Interacción #1**

**Gráfico, Gráfico de dispersión

Descripción generada automáticamente**

**Interacción #2**

**Gráfico, Gráfico de dispersión

Descripción generada automáticamente**

**Interacción #3**

**Gráfico, Gráfico de dispersión

Descripción generada automáticamente**

**Interacción #4**

**Gráfico, Gráfico de dispersión

Descripción generada automáticamente**

**Interacción #5**

**Gráfico, Gráfico de líneas, Gráfico de dispersión

Descripción generada automáticamente**

**Interacción #6Gráfico, Gráfico de líneas, Gráfico de dispersión

Descripción generada automáticamente**

**Interacción #7**

**Gráfico, Gráfico de líneas, Gráfico de dispersión

Descripción generada automáticamente**

**DATOS ORDENADOS Y OPTIMIZADOS, RESULTADO FINAL**

**Interacción #8**

**Gráfico

Descripción generada automáticamente**

**CONCLUSONES**

* Gracias a las características de flexibilidad yadaptabilidad las Redes Neuronales Artificiales nos presentan una alternativa, no lineal, para analizar y predecir las series de tiempo.
* El problema de la predicción es un tema que ha abarcado muchos escenarios de la vida diaria, como se vio, las redes neuronales son una herramienta poderosa para el análisis de estos fenómenos, la aparente fiabilidad que tiene debido a los muchos ejemplos de predicción en los cuáles se aplica, da seguridad suficiente para creer en los resultados, en algunos casos se puede llegar a tener más fiabilidad en las RNA que en los métodos clásicos de predicción.
* La predicción través de redes neuronales al parecer es un método totalmente paralelo a los métodos estadísticos clásicos para la predicción, pero ha sido mediante la estadística que las redes neuronales se han fortalecido como métodos estándar, en este trabajo se expone el modelo del aproximador general de funciones, esto se demuestra a través de métodos y teoremas estadísticos.
* El número de datos de entrenamiento y de datos de validación condiciona en gran forma los resultados que se obtienen cuando se entrena una red neuronal, si se tiene un gran número de datos se debe controlar que no haya sobre entrenamiento, pero al tener pocos, no se sabe claramente lo que se pude llegar a obtener, la reglas expuestas de como escoger un óptimo número de datos de validación intenta solucionar este problema, pero aun así este problema es resuelto más basado en la experiencia de éxito de otros modelos en problema similares, por lo tanto las redes neuronales siempre van a tener una parte aleatoria que pude llegar a condicionar a la parte teórica.

**BIOGRAFIAS**

1. **Metodologías de diseño de redes Neuronales:** [**https://www.google.com/url?sa=t&rct=j&q=&esrc=s&source=web&cd=&cad=rja&uact=8&ved=2ahUKEwjUw\_PbvdLxAhWmaDABHb9sBbsQFjANegQIFxAD&url=https%3A%2F%2Faccedacris.ulpgc.es%2Fbi tstream%2F10553%2F17473%2F4%2F0724684\_00000\_0000.pdf&usg=AOvVaw3pQOwI2OyUVh\_E5kdpCvLs**](https://www.google.com/url?sa=t&rct=j&q=&esrc=s&source=web&cd=&cad=rja&uact=8&ved=2ahUKEwjUw_PbvdLxAhWmaDABHb9sBbsQFjANegQIFxAD&url=https%3A%2F%2Faccedacris.ulpgc.es%2Fbi%20tstream%2F10553%2F17473%2F4%2F0724684_00000_0000.pdf&usg=AOvVaw3pQOwI2OyUVh_E5kdpCvLs)
2. **Aplicación de redes Neuronales Artificiales en entornos virtuales inteligentes:** [**https://scielo.conicyt.cl/pdf/infotec/v25n5/art15.pdf**](https://scielo.conicyt.cl/pdf/infotec/v25n5/art15.pdf)
3. **Diseño de redes Neuronales artificiales mediante algoritmos evolutivos:** [**https://scielo.conicyt.cl/pdf/infotec/v25n5/art15.pdf**](https://scielo.conicyt.cl/pdf/infotec/v25n5/art15.pdf)
4. **Redes neuronales con la librería Pybrain:** [**https://www.youtube.com/watch?v=WDP0X\_JqWDY**](https://www.youtube.com/watch?v=WDP0X_JqWDY)
5. **Arquitectura de un Perceptrón multicapa:** [**https://www.diegocalvo.es/perceptron-multicapa/**](https://www.diegocalvo.es/perceptron-multicapa/)
6. **Que es un perceptrón:** [**https://es.wikipedia.org/wiki/Perceptr%C3%B3n\_multicapa**](https://es.wikipedia.org/wiki/Perceptr%C3%B3n_multicapa)
7. [**https://www.youtube.com/watch?v=Fz4Uo8\_uUfg**](https://www.youtube.com/watch?v=Fz4Uo8_uUfg)