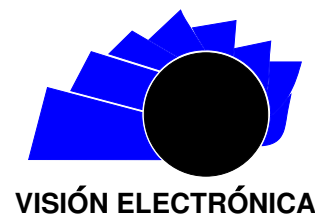




Visión Electrónica

Más que un estado sólido

<http://revistas.udistrital.edu.co/ojs/index.php/visele/index>



VISIÓN INVESTIGADORA

Comparativo de técnicas de inteligencia artificial aplicadas a pronósticos

Benchmarking among artificial intelligence techniques applied to forecast

Cristhian Johnatan Izquierdo Ortiz^a, Carlos Enrique Montenegro Marin^b

INFORMACIÓN DEL ARTÍCULO

Historia del artículo:

Enviado: Septiembre 2014

Recibido: Octubre 2014

Aceptado: Febrero de 2015

Palabras clave:

Pronóstico de la demanda

Algoritmos genéticos

Redes neuronales artificiales

Métodos de pronóstico



Keywords:

Demand forecasting

Genetic algorithms

Artificial neural networks

Forecasting methods

RESUMEN

La creación de un espacio de pruebas para varias técnicas de pronóstico de demanda corresponde a espacio de desarrollo de software en donde, además de probar varios algoritmos sobre una misma base de datos, estos pueden compararse unos con otros. El presente artículo, describe la construcción de una herramienta que permite generar pronósticos para que sean aprovechables en la toma de decisiones en las compras de compañías distribuidoras como la compra de productos eléctricos (cable, iluminación y accesorios), comunicaciones, tecnológicos, ropa, calzado, entre otros. Igualmente, se comparan algunas técnicas simples de pronóstico como la Media Móvil (MM) y Último Periodo con otras técnicas como las Redes Neuronales Artificiales (RNA) y los Algoritmos Genéticos (AG), la comparación se realiza teniendo en cuenta los criterios de error de los pronósticos generados y el tiempo de procesamiento de los métodos. Se explican: el análisis, diseño, desarrollo e implementación de los métodos antes mencionados y su integración con la herramienta.

ABSTRACT

The article is about creating a space for multiple tests of demand forecasting techniques, this space is a software development where besides to testing the algorithms on the same database, these code routines can be compared with each other, this tool allows generate forecasts to be usable in decision making on purchases of Distribution Companies. Besides comparing forecasting some simple techniques like Moving Average (MM) and Last Period with other techniques such as Artificial Neural Networks (ARN) and genetic algorithms (GA), the comparison is made taking into account the error criteria of generated forecasts and the processing time of the methods. Throughout the article explains the design, development and implementation of the above methods and their integration with the tool.

^aIngeniero de Sistemas. Universidad Distrital "Francisco José de Caldas". Planeador de la Demanda Melexa S.A. e-mail: cizquierdo@melexa.com

^bPhD.En Ing. De sistemas. Universidad de Oviedo. Docente Facultad de Ingeniería, Universidad Distrital "Francisco José de Caldas", Colombia. e-mail: cemontenegrom@udistrital.edu.co

1. Introducción

El crecimiento corporativo de compañías distribuidoras depende en gran medida en optimizar el proceso de balanceo entre las cantidades vendidas con las cantidades compradas, minimizando el inventario que se mantiene en las bodegas. De este proceso de balanceo se encarga la planeación y el análisis del inventario [1]. Cuando se tiene un desarrollo efectivo de estos procesos aumenta la capacidad de servicio de la empresa representado en un aumento de disponibilidad en el inventario y una disminución en la cobertura (tiempo de permanencia del inventario) [2].

El paso inicial para una planeación efectiva de las compras es pronosticar con la mayor exactitud posible las ventas esperadas en los periodos posteriores al análisis con esto se obtiene un valor de referencia que permite a la empresa prepararse para la demanda a través de sus compras. Para mayor información sobre la importancia empresarial de los pronósticos se sugiere el artículo [3].

El presente estudio, realizará pronósticos específicamente sobre los datos de una empresa colombiana -Melexa S.A.- dedicada a la distribución de materiales eléctricos y de comunicaciones, tiene su demanda ubicada en empresas institucionales, industriales, constructoras y ferreterías (sub distribuidores) a nivel nacional [4]. Como se espera de una compañía encargada de la comercialización y distribución de mercancía su proceso principal de negocio consiste en la compra, mantenimiento y venta de productos. Estas transacciones son comúnmente planeadas a través de un sistema ERP [5].

El artículo se divide en dos grandes partes: la primera describe el proceso de procesamiento de la información de la empresa, el diseño y construcción de una herramienta web que, utilizando estos datos como insumos, es capaz de recibir y utilizar segmentos de código que generan pronósticos (Métodos de Pronóstico) sobre los datos base de las transacciones de la empresa, esta herramienta permite la gestión, evaluación y aprovechamiento de estos pronósticos, además de comparar los Métodos utilizados. Estos métodos pueden ser creados por cualquier programador siguiendo unas reglas básicas. La segunda parte describe la configuración de varios métodos de pronóstico en los que sobresale un método que propone una implementación de Redes Neuronales Artificiales (RNA) y un método que utiliza Algoritmos Genéticos (AG). De cada uno de estos se muestra su configuración y los resultados de su implementación a través de la herramienta elaborada en la primera parte en comparación con otros métodos más simples de pronóstico.

2. Materiales y métodos

Como las transacciones son planeadas a través de un sistema ERP; la metodología más adecuada adopta 5 etapas, partiendo de la extracción y preparación de los datos base de los análisis hasta presentación de resultados y conclusiones de los comparativos realizados. En la primera etapa los datos son pre-procesados, incluye la transformación de las transacciones de venta en series temporales, la categorización y normalización de estas series. La segunda etapa se encarga de la construcción de un ambiente adecuado para la implementación y comparación de métodos de pronóstico. La tercera etapa comprende la adaptación de técnicas convencionales y de Inteligencia Artificial sobre el problema a resolver haciendo énfasis en la configuración de los algoritmos de entrenamiento. La cuarta etapa presenta los resultados de la experimentación considerando criterios de desempeño en tiempo y en exactitud. La última etapa sintetiza los resultados de los comparativos realizados y recoge elementos importantes en cada fase del proyecto.

2.1. Pre-procesamiento de la información, clasificación y normalización

En esta sección se describe el método de obtención, la clasificación y las transformaciones de los datos.

2.1.1. Construcción de las series temporales

Con las fechas de las ventas se crea un vector para cada producto, sumando las ventas conseguidas en cada semana a lo largo de toda su historia. Cada elemento del vector se conforma por las ventas conseguidas en una semana, el subíndice determina el año y el número de la semana. Para mayor claridad se da un ejemplo en tabla 1:

Tabla 1. Ejemplo de una serie temporal.

año	2010	2010	2010	2010	2010	2010	2010
semana	1	2	3	4	5	6	7
venta	854	386	512	215	201	786	592

Fuente: elaboración propia.

El vector anterior muestra que en la semana 3 del año 2010 se vendieron 512 unidades. Estos vectores son ingresados en la tabla serie tiempo (ver tabla 2) utilizando como claves el producto, el año y la semana.

Tabla 2. Series de tiempo.

Producto	Año	Semana	Demanda
X001	2010	5	201
X001	2010	6	786
X001	2010	7	592
X001	2009	1	40
X002	2009	2	5
X002	2009	3	20

Fuente: elaboración propia.

2.1.2. *Clasificación de los productos*

Una práctica común y actualmente incorporada en las soluciones para manejo de inventario es la clasificación de los productos con las letras ABC como muestra [6]. No existe una regla estricta para definir cuando un producto pertenece a una letra específica pero se ha estandarizado el uso de la letra “A” para marcar los productos más importantes, B los siguientes en importancia y C los demás. Siguiendo el principio de Pareto [7] se busca que los A sean muy pocos en comparación con los productos pertenecientes a las demás letras y C la mayor cantidad de productos.

Dado que se tiene la información organizada en series temporales se propone la siguiente clasificación:

- Artículos de alta rotación categoría A: Estos productos han presentado ventas en el 80% de las semanas en el último año. Con esta clasificación se encontraron 86 productos.
- Artículos de mediana rotación categoría B: Son productos que, sin ser de categoría A, presentan ventas en más del 80% de los meses en los últimos 2 años. A este grupo pertenecen 178 productos.
- Artículos de baja rotación categoría C: Son todos los demás productos que no pertenecen a las categorías A y B. Por su comportamiento atípico estos productos se excluyen del análisis.

2.1.3. *Normalización de las series*

La normalización es una técnica que se utiliza para comparar datos que parecen muy dispares utilizando la desviación estándar como regla [8]. Se comparan los datos individuales de las series con su media y se relativizan utilizando la siguiente fórmula.

$$z = \left(\frac{y - y'}{s} \right) \tag{1}$$

Donde y es el valor a normalizar, y' es la media aritmética de los valores de la serie y s es la desviación

estándar (poblacional) que rige la serie. El valor obtenido z se convierte en el valor normalizado. Algunas de las ventajas de este proceso es que permite:

- Usar diferente escalas.
- Usar diferentes unidades.
- Usar diferentes poblaciones.

La serie transformada tienen la misma distribución que la original.

Después de realizar el proceso de transformación de los valores de la serie, se incluye esta información en la tabla serieTiempo. Adicionalmente los valores de media y desviación se conservan como parámetros de cada producto, con esta información es posible realizar la desnormalización de las series con la expresión (2):

$$y = (z \times s) y' \tag{2}$$

Los 302 productos seleccionados son ingresados a la tabla producto (ver tabla 3), los datos almacenados son el código del producto, la categoría a la que pertenece y los estadígrafos media y desviación.

Tabla 3. Ejemplo tabla producto.

Producto	Cat	Media	Desviación
X001	A	1034	20
X002	A	12873	303
X003	B	56	10
X004	B	204	31
X005	B	5	1
X006	C	0	0

Fuente: elaboración propia.

2.2. *Diseño y desarrollo de la herramienta: Gestor de Pronósticos*

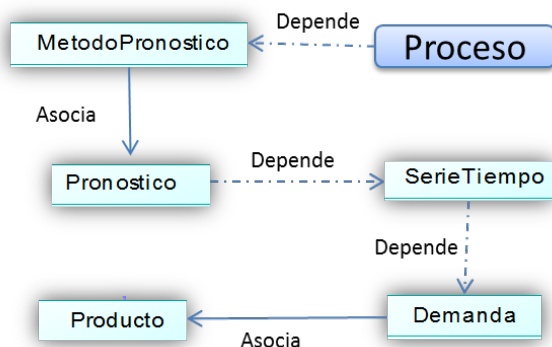
Uno de los requisitos para la utilización de la herramienta es que en el lugar donde se implemente se cuente con un sistema de información que logre almacenar en una base de datos la información de todas las transacciones de venta de la empresa. La tabla donde se almacena esta información será llamada “demanda” (ver tabla 4). Los datos que se encuentran en esta tabla son: el código del documento, los productos vendidos en la transacción, las cantidades y la fecha; es importante identificar en la transacción si se trata de una devolución o no, como se explicó en la sección 2, buscando extraer de esta tabla la información correspondiente a las ventas reales de los productos.

Tabla 4. Tabla demanda.

Doc.	Producto	Cantidad	Fecha	Dev
10043	X001	3	2010-02-15	
10044	X001	-3	2010-02-20	1
10045	X023	120	2011-03-13	
10046	X045	50	2011-04-30	
10047	X001	-120	2012-06-03	1
10048	X001	120	2013-03-08	1

Fuente: elaboración propia.

La categorización de los productos vendidos es un factor importante dentro del análisis que se espera realizar, las empresas clasifican sus productos en tres categorías: A, B y C. Las reglas para definir esta clasificación son propias de cada organización, en este proyecto se propone una (ver figura 1) acorde con la compañía base de análisis pero el proyecto es adaptable a cualquier forma de clasificación.

Figura 1. Diagrama de clases. Elementos de la herramienta.

Fuente: elaboración propia.

Los registros de la demanda se organizan en una serie temporal, la serie temporal es una sucesión de números en donde cada elemento en la sucesión representa el total de ventas logradas por el producto en un periodo de tiempo.

Para el efecto de comparar diferentes series temporales la herramienta se desarrolló utilizando las siguientes tecnologías:

- Lenguaje de programación en el servidor PHP.
- Motor de Base de Datos MySQL.
- Lenguaje de programación en cliente Javascript
- Tecnología de desarrollo web AJAX
- ThickBox tecnología para crear efectos de ventanas flotantes.

- Pcharts, tecnología para la graficación de datos.
- PCLZIP, tecnología trabajo con archivos comprimidos .zip
- Integración con Excel.

2.3. Métodos de pronóstico

Todos los Métodos de Pronostico que se requieran incluir en la herramienta deben cumplir con las siguientes características:

Estar escrito en el lenguaje de programación PHP: Dado que la herramienta “Gestor de Pronósticos” utiliza este lenguaje los llamados a los servicios y el paso de los parámetros deben ser entendidos por el Método.

El desarrollo del Método debe estar Orientado a Objetos: El Método debe estar compuesto por clases con métodos y atributos, además de utilizar objetos de estas clases.

El Método debe contener un fichero llamado Principal.php (ver figura 2): Este fichero debe contener una clase llamada Principal que será la encargada de enrutar todas las solicitudes enviadas por la herramienta.

Figura 2. Métodos y atributos de la clase Principal. Todos los métodos programados deben tener una clase con estos elementos.

Principal.php
-Zona -Número de método -Producto -Serie de tiempo -Año límite -Semana límite
+Principal(); + uso Actualización(); +entrenar Método(); +correr pronóstico(); +escribir Base Datos();

2.4. Paquete estándar de Métodos

Los siguientes métodos de pronósticos hacen parte de los algoritmos básicos presentes en la mayoría de aplicaciones de planeación de inventarios, se explicarán los algoritmos y la implementación en el proyecto.

2.4.1. Media móvil de 4 y 12 periodos

En el método se recupera de la serie los últimos 4 periodos (MM4), dependiendo si se trata de unidades de tiempo mensual o semanal, esto se identifica en el nombre del pronóstico. El pronóstico se calcula sumando la demanda de los 4 periodos y dividiendo entre 4, este valor pronosticado retorna a la herramienta para que ella lo almacene adecuadamente.

También se desarrollará un Método de media móvil de 12 periodos (MM12) de la misma forma que la MM4 ampliando el número de periodos a 12, el algoritmo MMN donde N es el número de periodos se muestra en Segmento de Código 1. En total se dispone de 4 métodos de media móvil (MM4 mensual, MM4 semanal, MM12 mensual y MM12 semanal).

Segmento de Código 1. Código media móvil *n* periodos (MMN).

```
public function correrPronostico($producto,$serieTiempo,$annoLimite,$semanaLimite){
    $sumaVentas=0;
    // se recorren las últimos $n semanas
    for($i=0;$i<$n;$i++){
        $auxVenta=$serieTiempo[$annoLimite][$semanaLimite];
        $sumaVentas=$sumaVentas+$auxVenta;
        $auxSemanaLimite=$auxSemanaLimite-1;
        //en caso de que los periodos cambien de año
        if($auxSemanaLimite==0){
            $auxSemanaLimite=$ultimaSemanaAnno;
            $auxAnnoLimite=$auxAnnoLimite-1;
        }
    }
    //calculo promedio
    $promedioVentas=$sumaVentas/4;
    return $promedioVentas;
}
```

Fuente: elaboración propia.

Estos y otros métodos se explican en detalle en [9].

2.4.2. Alisado exponencial

La clase principal en su método correrPronostico() recorre toda la serie temporal desde la primera semana hasta el límite incluido en los parámetros, cada uno de los periodos se multiplica por un 80% la demanda real del mes anterior y por un 20% el pronóstico del mes anterior, este cálculo se convierte en el pronóstico para el mes siguiente. Continúa el proceso hasta llegar al último periodo y termina con el cálculo del pronóstico solicitado el cual retorna a la herramienta de pronóstico, esta iteración se presenta en el segmento de código 2.

Segmento de Código 2. Código Alisado Exponencial.

```
public function correrPronostico($producto,$serieTiempo,$annoLimite,$semanaLimite,$parametro)
{
    // parámetro está entre 0 y 1 ejemplo 0.8
    $numeroSemanas=0;
    foreach ($serieTiempo as $annoSem=>$valor){
        if($numeroSemanas==0){//el valor de la primera semana
            $pronostico=$valor;
        }
        else{
            $pronostico=$valor*$parametro+(1-$parametro)*$pronostico;
        }
        $numeroSemanas++;
    }
    return $pronostico;
}
```

Fuente: elaboración propia.

Este método no requiere de ningún entrenamiento.

2.4.3. Uso de las ventas del último periodo

Este método recupera la venta del último periodo, es decir, el último elemento del vector de la serie temporal y lo replica como pronóstico para el siguiente. No utiliza ningún entrenamiento y la función usoActualizacion() es cero.

2.5. Métodos que utilizan Inteligencia Artificial

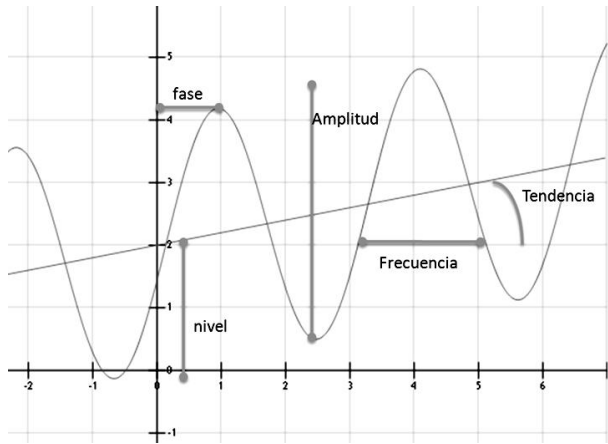
Ahora se presentan 2 métodos creados e implementados, son métodos que utilizan técnicas de Inteligencia Artificial (IA) y que buscan generar pronósticos más exactos adaptando los patrones de comportamiento de las ventas incluidos en las series temporales. Las técnicas de IA utilizadas son Algoritmos Genéticos y Redes Neuronales, de cada una se explicará el funcionamiento del algoritmo.

2.5.1. Aplicación del Algoritmo Genético

En este apartado se explica el funcionamiento del algoritmo a través de los objetos y funciones planteadas en el diseño. Una base teórica para abordar el tema de los AG se encuentra en [10].

El primer paso en la implementación de esta técnica es la representación del problema del pronóstico a través de un individuo. Un individuo se compone de la suma de 10 señales senoidales, cada señal se puede parametrizar con 6 variables: la fase, nivel, amplitud, frecuencia, tendencia y umbral, estas variables se muestran sobre la señal en la figura 4.

Figura 4. Señal que compone a los individuos del AG, cada individuo tiene 10 señales.



Fuente: elaboración propia.

En resumen, cada individuo tiene 60 genes correspondientes a los valores de las variables de cada una de las 10 señales.

Fase de Entrenamiento del Algoritmo Genético:

Al igual que los demás métodos de pronóstico el proceso de entrenamiento comienza con la actualización del Método dentro de la herramienta. Posterior a esta acción el usuario debe seleccionar una fecha límite para el entrenamiento en el calendario, es usual que el usuario seleccione una fecha cercana al periodo que se desea pronosticar. Después del evento disparador del entrenamiento la clase Principal del Método es instanciada y para cada serie de tiempo se ejecuta el entrenamiento, la figura 5 presenta el flujo grama con las diferentes fases del entrenamiento:

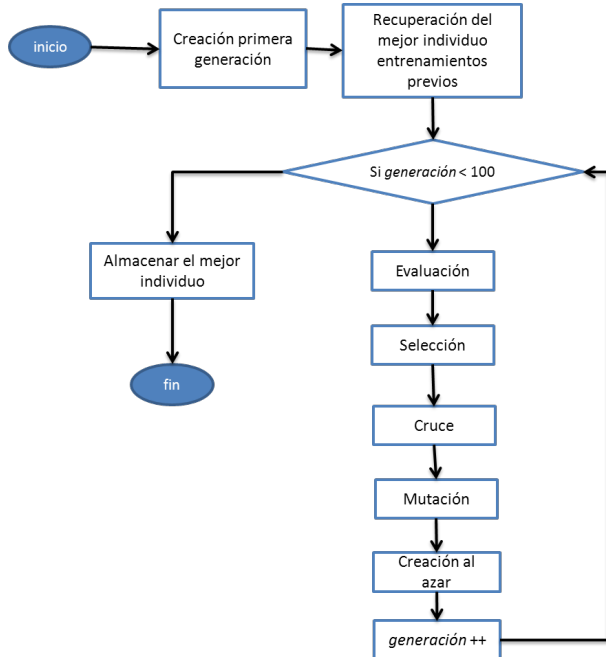
- 1) Creación de la población (primera generación): El entrenamiento inicia con la creación de una Población en la cual se generan 120 individuos nuevos al azar, este conjunto conforma la primera generación.
- 2) Búsqueda del mejor individuo de entrenamientos anteriores: La primera actividad del entrenamiento es consultar si existen entrenamientos anteriores. En caso de existir se reconstruye el mejor individuo de los entrenamientos pasados con la información de los genes almacenada en la base de datos. Este individuo se incluye dentro de la primera generación.
- 3) Evaluación de los individuos: Se aplica la evaluación a través del cálculo del error cuadrático medio del individuo a través de 104 semanas (2 años), para esto

es necesario que el algoritmo utilice la serie temporal relacionada en los parámetros de la función de entrenamiento. Una vez evaluado cada individuo la calificación se almacena en un atributo llamado fitness.

- 4) Ranking y selección: Se ordenan los individuos desde en mejor evaluado en adelante. La selección se realiza con un parámetro de elitismo determinado en 20 individuos, es decir se preservan los 20 mejores individuos los demás son eliminados. Cada individuo tiene un atributo de edad, si el individuo sobrevive una selección la edad aumenta en uno.
- 5) Cruce: Para la operación de cruce cada uno de los mejores 5 individuos de los sobrevivientes seleccionan parejas al azar de los demás sobrevivientes. Para fortalecer la preservación de los mejores genes o características de los individuos, el individuo número 1 tiene derecho a tener 5 parejas, el individuo 2 tiene 4 parejas y así sucesivamente. Finalmente se conforman 15 parejas diferentes. Una vez conformadas las parejas realizan los dos tipos de cruce: cruce de señales y cruce de coeficientes. De esta forma de cada pareja emergen 4 nuevos individuos, completando en total 60 individuos nuevos.
- 6) Mutación: Se escogen 5 individuos al azar de los sobrevivientes de la generación anterior, en cada uno de ellos se realiza la operación de mutación alterando uno de sus genes al azar. Creando así 5 individuos nuevos.
- 7) Creación al azar: Para completar la generación y agregar diversidad al algoritmo se crean nuevos individuos desde cero, el número de individuos nuevos se define por un parámetro que para esta implementación es 35.
- 8) Iteraciones: Una vez conformada una nueva generación de 120 individuos el proceso se repite desde el punto 3 con la evaluación de los individuos. El número de iteraciones los define el programador, en la implementación se realizaron 100 iteraciones.
- 9) Almacenamiento del mejor individuo: Al final de las iteraciones que se definan, el mejor individuo es almacenado en la base de datos, siendo este el producto final del entrenamiento.

Para facilitar la comprensión del algoritmo se presenta el siguiente diagrama de flujo:

Figura 5. Flujo de entrenamiento del Método AG.



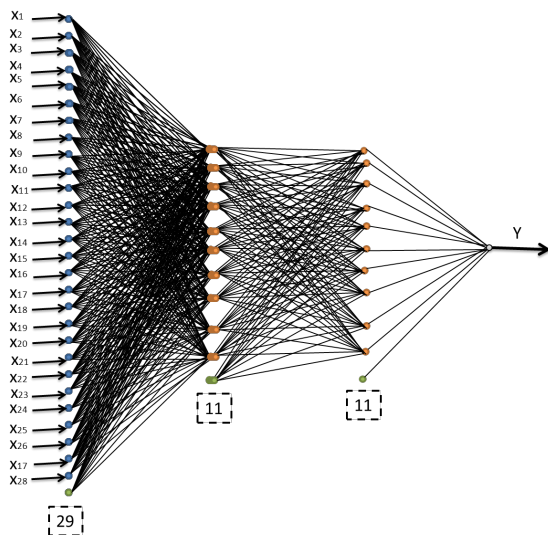
Fuente: elaboración propia.

2.6. Aplicación de la Red Neuronal Artificial

Las bases teóricas se pueden encontrar en el texto referenciado en [11].

La red neuronal construida tiene la apariencia de la figura 6.

Figura 6. Topología de la red neuronal.



Fuente: elaboración propia.

Esta red está compuesta por 4 capas; la capa inicial tiene 29 neuronas, las dos capas ocultas tienen 11 neuronas y una neurona en la capa final. En cada capa diferente a la capa final se observa la neurona bias [12].

Cada neurona realiza 2 cálculos para definir su salida: una regla de propagación y una función de activación.

Regla de propagación:

Definida por la siguiente fórmula (3):

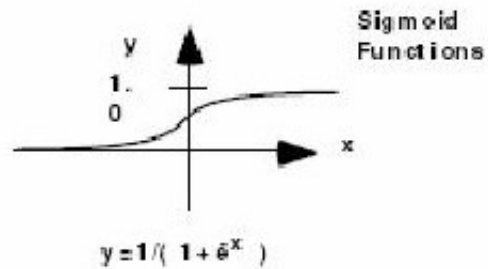
$$\text{Regla de propagación} = \sum_{i=1}^n X_i W_{ij} \quad (3)$$

Donde X_i es la salida de cada una de las neuronas de la capa anterior y W_{ij} es el valor de la sinapsis que une a las neuronas.

Función de activación:

La función de activación transforma el resultado de la Regla de propagación, la función utilizada en este proyecto será la función sigmoide (ver figura 7):

Figura 7. Función de activación sigmoide, entrega como salida valores entre 0 y 1.



Fuente: [11].

Para emplear la fórmula del gráfico la x representa el resultado de la regla de propagación y la salida de la neurona.

2.6.1. Capa inicial de la red

En la capa inicial se crean 29 neuronas, encargadas de tomar la información necesaria de la serie temporal (historia de ventas). En la figura 8 se observa el agrupamiento de las neuronas de esta capa.

Figura 8. El gráfico muestra la misión de cada neurona de la capa inicial.



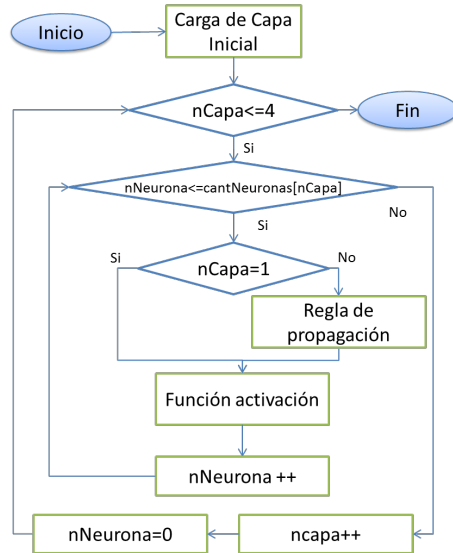
Fuente: elaboración propia.

El horizonte se utiliza para determinar cuántos periodos hacia el futuro se desea pronosticar.

Propagación hacia adelante

La propagación hacia adelante es la acción de cargar las neuronas de la primera capa con la información de la serie temporal y obtener de la red una salida que representa el pronóstico para la semana solicitada. La figura 9 muestra el flujo de las fases en el proceso de programación hacia adelante.

Figura 9. Flujo programación hacia adelante RNA.



Fuente: elaboración propia.

El proceso se describe así:

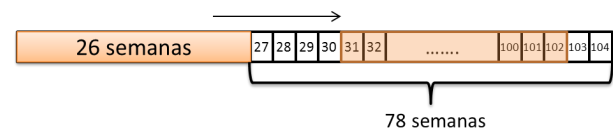
1. Se carga la primera Capa con la información de 26 periodos de la serie temporal, el número de la semana a pronosticar, el horizonte deseado y la entrada del bias que siempre es uno (1).
2. Después de cargar la primera capa, cada una de las neuronas que la componen realiza su función de activación (función signoidal), esta capa no utiliza la Regla de propagación debido a que solo tiene una entrada.

3. Las neuronas de la segunda capa hacen su parte; primero aplican la Regla de Propagación multiplicando cada salida de la capa anterior con la sinapsis correspondiente y realizando la sumatoria de estos productos.
4. En seguida cada neurona de la segunda capa aplica su función de activación, entregando un valor de salida entre 0 y 1.
5. Las neuronas de la tercera capa realizan los mismos procedimientos que las neuronas de la segunda capa explicado en los pasos 3 y 4.
6. Finalmente la neurona de la última capa realiza su regla de propagación y su función de activación para entregar una salida de la red.

2.6.2. Entrenamiento

El entrenamiento de la red es supervisado, sometiéndola a un conjunto de entrenamiento del que se conoce de antemano la respuesta correcta (venta real). A continuación se explica cuál es el conjunto de entrenamiento de la red y como esta modifica sus sinapsis interna en la búsqueda de ofrecer un mejor pronóstico: Conjunto de patrones para el entrenamiento: Para el entrenamiento se toma un segmento de 104 periodos tomado de la serie temporal original (ver figura 10). El segmento comienza desde la semana límite seleccionada por el usuario contando 104 semanas (2 años) hacia atrás. Dado que la entrada de la red toma un conjunto de 26 semanas (6 meses) el conjunto de entrenamiento está conformado por todos los grupos de 26 semanas consecutivas que se pueden formar a lo largo de 104 semanas, en total son 78 patrones de entrenamiento. Este número de patrones es fácil de deducir si se observa que $26 + 78 = 104$.

Figura 10. Conjunto de patrones para el entrenamiento.



Fuente: elaboración propia.

Cada patrón está compuesto por:

- Datos de entrada: La información de 26 semanas, el número de la semana siguiente (semana a pronosticar) y el horizonte
- Salida deseable: La salida esperada de la red es el valor de la serie temporal (venta real) de la semana siguiente al conjunto de las 26 semanas.
- Back-propagation señal de error.

Es probable que con la configuración inicial de la red, con pesos aleatorios, las respuestas otorgadas estén alejadas de los pronósticos deseables, entonces la red debe ajustarse. Este ajuste consiste en modificar el valor de las sinapsis entre cada par de neuronas de la red guiándose por las diferencias entre la salida y el valor esperado (ventas reales), a esta diferencia se le llama error de la red. Como resultado del entrenamiento se obtiene la configuración final de todas las sinapsis [13].

$$Error = y' - y \quad (4)$$

Donde y' es el valor esperado de la red dado en el patrón de entrenamiento y y es la salida de la red.

El aprendizaje se da a través del principio de optimización de descenso por gradiente buscando modificar los pesos según la participación de cada neurona en la salida de la red [14].

Se calcula la señal de error de cada capa a partir del error en la ecuación (4).

$$\tilde{E} = y \cdot (1 - y) \cdot (y - y') \quad (5)$$

Donde \tilde{E} es la señal de error de la neurona de salida [15]. La señal de error de las demás capas depende de la señal de error de las capas posteriores [15]

$$\tilde{E}_a = y \cdot (1 - y) \cdot \sum_{j=1}^m \tilde{E}_p W_{aj} \quad (6)$$

\tilde{E}_i es la señal de error en la capa i , W_{ij} es el vector de pesos sinápticos entre cada neurona de la capa i y la neurona j de la capa posterior p .

Finalmente cada sinapsis cambia según la siguiente función [15]:

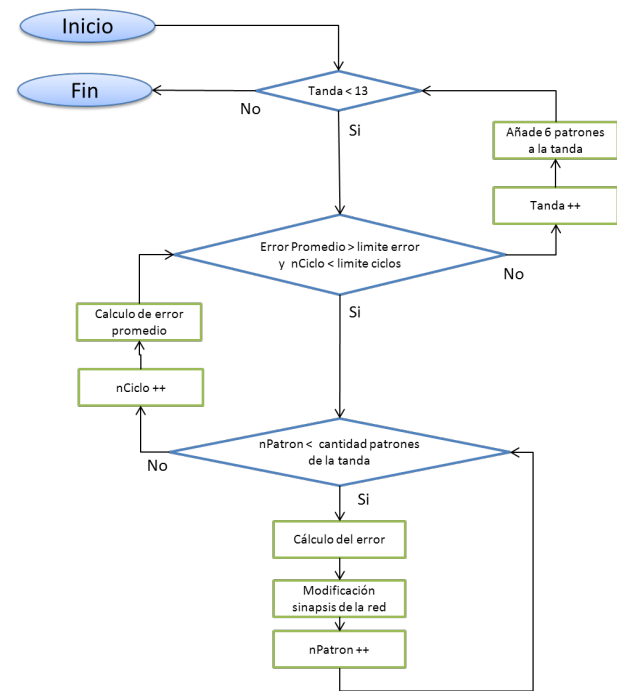
$$W_{ij}(t + 1) = W_{ij}(t) + \varepsilon \cdot x_i \cdot \tilde{E}_j \quad (7)$$

Donde $W_{ij}(t)$ es el peso entre la neurona i y j en el momento t , ε es la tasa de error del entrenamiento, x_i es la salida de la neurona i .

Programa de entrenamiento:

La red neuronal es sometida a los 78 patrones de entrenamiento de forma progresiva por tandas. La figura 11 presenta las fases del flujo de entrenamiento. En la primera tanda la red “aprende” los primeros 6 patrones, en la segunda tanda se añaden otros 6 patrones conformando un grupo de 12, lo mismo con las demás tandas hasta completar 13.

Figura 11. Flujo de entrenamiento de Método RNA.



Fuente: elaboración propia.

En cada tanda la red ajusta sus pesos sinápticos con la diferencia de su salida con el valor esperado de cada patrón, es decir, el error. Estos valores de error se conservan hasta completar todos los patrones de la tanda, al final, estos errores se promedian, si este promedio es superior a un límite definido como parámetro del algoritmo (límite error) la tanda debe repetirse. Cada vez que la red aprende de todos los patrones de una tanda cumple un ciclo.

$$Error\ Promedio = \frac{|y_t - y'_t|}{N} \quad (8)$$

N es el número de patrones, las variables y_t y y'_t son el valor esperado y la salida de la red con el patrón t .

3. Evaluación y resultados de los métodos aplicados

Al normalizar las series de tiempo se simplifica el proceso de evaluación de los pronósticos generados, esto se logra al poder realizar operaciones básicas entre los resultados del pronóstico y la demanda real utilizando la misma escala para todas las series.

Las mediciones de error que sirven para calificar los pronósticos son las siguientes:

- Error del pronóstico por producto:

$$Error\ pronóstico\ producto = |y - y'| \quad (9)$$

Donde y es la demanda real y y' es el valor pronosticado. Se utiliza el valor normalizado para las dos variables.

Este valor de error se almacena para cada pronóstico-producto en la tabla pronóstico-producto

- Error del pronóstico:

Esta medida se calcula para todos los productos, solo los productos A y solo los productos B. Para cada uno de estos conjuntos se utiliza la siguiente ecuación:

$$Error\ pronóstico = \frac{\sum Error\ pronóstico\ producto}{\# productos\ del\ conjunto} \quad (10)$$

Es el promedio de los errores de cada conjunto en cada pronóstico.

- Error del método:

Muestra el promedio de los errores de todos los productos de los pronósticos calculados a través de este método de pronóstico

$$Error\ método = \frac{\sum Error\ pronóstico\ conjunto}{\# productos\ del\ conjunto} \quad (11)$$

Es importante resaltar que con esta forma de medición el método o pronóstico mejor calificado es el que obtiene un menor valor en las mediciones de error descritas anteriormente.

4. Resultados en tiempo de procesamiento

El servidor donde se implementa la herramienta es una máquina con las siguientes características.

- Procesador: 1.73 GHz.
- Memoria: 1,25 GB

Esto es importante ya que los tiempos de procesamiento podrían variar según las características del equipo donde se “monte” el servidor.

Los tiempos de generación de un pronóstico y de entrenamiento, para cada serie temporal se muestran en la tabla 5.

Los anteriores tiempos fueron tomados al correr solo un proceso a la vez. Se recuerda que la herramienta puede ejecutar varios procesos simultáneamente lo que aumenta el tiempo de procesamiento de cada proceso.

Según los resultados anteriores los Métodos convencionales MM4, MM12 y “Último Periodo” son más rápidos en pronosticar que los métodos creados AG y

RNA. Aunque las diferencias no son significativas. Mientras que en el entrenamiento se ve una mayor diferencia ya que los primeros métodos no requieren entrenamiento mientras que los métodos creados tardan más de un día.

Tabla 5. Resultados de la comparación en tiempo.

Método de pronóstico	Cada serie Temporal		Todas las 264 series	
	Tiempo generación pronóstico	Tiempo entrenamiento	Tiempo generación pronóstico	Tiempo entrenamiento
MM4	2 seg	–	10 min	–
MM12	2 seg	–	10 min	–
Anterior	2 seg	–	10 min	–
AG	4 seg	6,4 min	18 min	28 h
RNA	5 seg	10,8 min	36 min	47 h

Fuente: elaboración propia.

4.1. Resultados en exactitud

El proceso para obtener la evaluación de los métodos de pronósticos fue el siguiente.

- Entrenamiento de los métodos de pronóstico:

Los métodos de entrenamientos RNA y AG se entrenan teniendo como parámetro la semana límite 2010-35 que es la primera semana de septiembre del año 2010. Se escoge esta fecha debido a que es un periodo más o menos regular de ventas, no tiene los picos de venta de diciembre ni las bajas de venta de medio año (esto aplica para la compañía fuente de los datos “Melexa S.A.”).

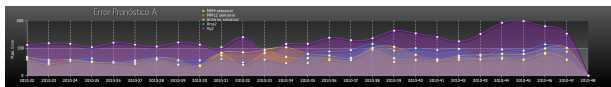
Los pronósticos evaluados antes de esta fecha serán pronósticos on-line y los pronósticos con fechas posteriores serán pronósticos off-line.

- Generación de pronósticos:

Utilizando los 5 métodos de pronóstico a evaluar se generan 26 pronósticos por cada método, desde la semana 22 hasta la semana 47, en total son 130 pronósticos generados a través de la herramienta “Gestor de Pronósticos”.

Luego a través de la misma herramienta es posible seleccionar los métodos y el rango de fechas desde la semana 22 hasta la 47 de 2010, en la pantalla aparece la información con las evaluaciones de los pronósticos relacionadas en la figura 12.

Figura 12. Flujo de entrenamiento de Método RNA.



Fuente: elaboración propia.

En la parte izquierda de la línea vertical de cada gráfica se muestran los resultados on-line y hacia la derecha se encuentran los resultados off-line. Se recuerda que esta clasificación aplica únicamente para los métodos creados RNA y AG debido a que son los únicos que requirieron entrenamiento. Los resultados compilados se muestran en la tabla 6.

Tabla 6. Ranking de los métodos.

Ranking	Método	General	A	B
1	MM12	0,53	0,57	0,50
2	MM4	0,55	0,59	0,53
3	Anterior	0,64	0,74	0,60
4	RNA	0,70	0,75	0,67
5	AG	1,35	1,33	1,36

Fuente: elaboración propia.

Como la comparación se da entre series normalizadas la unidad de medida utilizada son estas unidades normalizadas, estos valores se encuentran entre 0 y 2 siendo 0 la mayor exactitud.

Al revisar los datos y las gráficas se llegan a las siguientes afirmaciones:

- Los métodos de pronóstico creados no superan en exactitud a los otros métodos.
- Los dos métodos creados tiene un mejor comportamiento sobre los datos on-line que sobre los off-line y existe una tendencia a ser más inexacto entre más a futuro se realiza la predicción a partir de la fecha de entrenamiento.
- Los pronósticos sobre productos de categoría B tiende a ser más exactos que los A, excepto en el método AG.
- El mejor método calificado fue MM12 seguido de MM4 y el método llamado “Anterior”, lo que indica que en estos métodos simples entre más historia se tome el método funciona mejor.

5. Conclusiones

- La clasificación sugerida de productos ABC. es un aporte del proyecto, comúnmente las clasificaciones ABC están sujetas también al precio de los productos ó del impacto que tiene sobre las ventas, al utilizar solo la frecuencia de ventas se busca probar la exactitud de los pronósticos sobre grupos de productos con frecuencias de venta similares. Con esta clasificación se da mayor valor a los productos con demandas más frecuentes.
 - Entre todos los periodos evaluados, el periodo siguiente al límite de entrenamiento obtienen un mejor criterio de análisis, es el primer periodo dentro de los resultados off-line, corresponde a la semana 36/2010. La causa de este comportamiento es que al generar un pronóstico con mayor exactitud los entrenamientos de los métodos se realizan con la fecha límite más cercana al periodo a pronosticar, aprovechando la mayor historia reciente que se pueda. Como se demuestra en los resultados entre más alejado está el pronóstico de la fecha límite de entrenamiento más inexacto es el pronóstico, esto es visible para los métodos con entrenamiento.
 - A pesar de que los resultados del método AG no fueron mejores que los otros métodos programados, durante las pruebas anteriores de este algoritmo se probaron diferentes configuraciones como:
 - Diferentes cotas para las variables de las señales.
 - Diferentes números de señales, aumentando de 3 a 10.
 - Agregando una variable umbral a cada señal.
- Modificando los anteriores aspectos se logró mejorar los resultados de exactitud del método desde resultados alejados (más de 5 veces el error de los otros métodos) a máximo 3 veces y en los pronósticos cercanos a la fecha límite de entrenamiento a menos de 2 veces.
- El método RNA obtuvo mejores resultados que el método de AG, pero no superó los demás métodos. Su evaluación estuvo solo entre 2 y 3 décimas por encima del mejor considerando las tres evaluaciones (General, A y B). Este resultado fue el mejor después de diversas pruebas realizadas, donde se cambiaron las siguientes configuraciones:
 - Diferentes tasas de aprendizaje.
 - Diferentes límites de error y de ciclos.
 - Diferentes arquitecturas (números de neuronas por capas).

- Diferentes patrones y formas de entrenamiento (entrenamiento por tandas).
- Diferentes transformaciones de la respuesta.

Referencias

- [1] L. Mora, "Gestión Logística Integral, Gestión Logística - FESC", Bogotá: editorial ECOE, 2008, Capítulo 2 pp.118-150
- [2] El Control y Planificación de los Inventarios Unidad IV, Universidad Jose Carlos Mariategui, pp. 29-54, Julio 2005, [En línea] Disponible en: http://www.ujcm.edu.pe/bv/links/cur_comercial/LogisticaEmpresarial-11.pdf
- [3] T. Foster, "Forecasting, Demand Planning in a Difficult Economy, Global Logistics & Supply Chain Strategies", portal: Supply Chain Brain, pp 1-4, December 2008. [En línea]. Disponible en: <http://www.supplychainbrain.com/content/technology-solutions/forecasting-demand-planning/single-article-page/article/forecasting-demand-planning-in-a-difficult-economy/>
- [4] "Mercados Melexa", portal Melexa-Sonepar SAS, Julio 2013 [En línea]. Disponible en: <http://www.melexa.com/mercados>.
- [5] M. Becerra, "Enterprise Resource Planning (ERP) - presentación BAAN", Mind de Colombia, Bogotá, pp. 2-15, Colombia, Diciembre 2008,
- [6] T. Fucci, "El grafico abc como tecnica de gestión de inventarios", Universidad Nacional de Lujan, Buenos Aires, Argentina, pp. 1-6, Junio 1999.
- [7] Free-Logistics, the free supply chain portal, "ABC analysis Pareto principle", Free-Logistics, [En línea]. Disponible en: http://www.free-logistics.com/index.php?option=com_content&view=article&id=378:abc-analysis-pareto-principle&catid=67&Itemid=51.
- [8] C. Rojas, "La desviación estándar y los modelos normales", Universidad Autónoma de Chile, pp.5-24, Octubre 2011.
- [9] M. Oliveros, "Tema: Pronósticos", Universidad de los Andes, Bogotá, Colombia, pp. 10-21, Septiembre 2007.
- [10] A. De Peña, P. Truyol, and G. S. Aqs, "Algoritmos genéticos", Universidad Carlos III, Madrid, España, pp.1-8, Enero 2007.
- [11] I. Martínez, "Introducción a las Redes Neuronales", Universidad Complutense de Madrid, trabajo doctoral, pp. 1-19, Junio 2005.
- [12] A. M. Pedro Larrañaga, Iñaki Inza, "Tema 8. Redes Neuronales," in Redes Neuronales, U. del P. Vasco, p. 12-17. Ed. 2011.
- [13] M. A. V. Reyes, C. Y. Márquez, and L. P. S. Fernández, "Algoritmo Backpropagation para Redes Neuronales: conceptos y aplicaciones," INSTITUTO POLITÉCNICO NACIONAL, 2006.
- [14] J. A. P. Ortiz, "Modelos predictivos basados en Redes Neuronales recurrentes de tiempo discreto", Tesis Doctoral, Universidad de Alicante, España, pp. 35-63, Julio 2002.
- [15] G. Colmenares, "CAPÍTULO 2 TLU(s) MULTICAPAS," RedULA, Universidad de los Andes Venezuela, pp. 40-78. Noviembre 2004.
- [16] J. Casellas, "Apéndice A: Metodología para la evaluación del modelo de pronóstico meteorológico", Tesis Doctorales en Red. p. 3. Enero 2005.