

## C10

### BUSINESS INTELLIGENCE EN LA GESTIÓN DE STOCKS PARA ALMACENAMIENTO ÓPTIMO DE MEDICAMENTOS

Luque, A. ([aluque2@us.es](mailto:aluque2@us.es)); Aguayo, F. ([faguayo@us.es](mailto:faguayo@us.es)); Lama, J. R. ([jrlama@us.es](mailto:jrlama@us.es)); Peralta, E. ([mperalta1@us.es](mailto:mperalta1@us.es)); Córdoba, A. ([acordoba1@us.es](mailto:acordoba1@us.es)); de las Heras, A. ([adelasheras@us.es](mailto:adelasheras@us.es)).

TEP022: Diseño Industrial e Ingeniería del Proyecto y la Innovación

#### RESUMEN

En este trabajo se muestra la viabilidad de utilizar técnicas de análisis de series temporales para el problema de estimación de la demanda. Los trabajos realizados incluyen aspectos iniciales de preparación de datos, elección de las herramientas de procesado y análisis mediante técnicas de series temporales. Quedan pendientes de desarrollo futuro la aplicación de técnicas de minería de datos y su comparación con los resultados de esta primera fase.

**Palabras clave:** *business intelligence, gestión de stocks, optimización*

#### ABSTRACT

In this paper the feasibility of using techniques of time series analysis to the problem of estimating demand is shown. Works carried out include aspects of initial data preparation, choice of processing tools and analysis using time-series techniques. Application of data mining techniques and comparison with the results of this first phase are possible future works.

**Keywords:** *business intelligence, stocks management, optimization*

#### INTRODUCCIÓN Y OBJETIVOS

El objeto del presente documento es la presentación de los avances realizados durante el primer año de desarrollo en el proyecto demostrador “Business Intelligence para la Gestión de Stocks (BIGSTOCKS)” para su valoración por parte de la Comisión Evaluadora de los Proyectos Cátedra Telefónica “Inteligencia en la Red” de la Universidad de Sevilla.

En la literatura pueden encontrarse referencias sobre trabajos similares que han servido de inspiración y punto de partida de nuestro análisis (Olsen, C., 2012; Torreblanca, Tejera, García, del Prado Llergo, Cantarero & Camacho, 2012; Govindarajan, Perelló-Juncá, Parès-Marimòn, Serrais-Benavente, Ferrandez-Martí, Sala-Robinat, ... & Gonzalez-Estrada, 2013)

Los trabajos realizados durante este primer año de desarrollo abarcan: aspectos iniciales de preparación de datos; elección de las herramientas de procesado; y análisis mediante técnicas de series temporales. Para la segunda etapa del proyecto

quedan pendientes de desarrollo futuro: la aplicación de técnicas de minería de datos; y su comparación con los resultados de esta primera fase.

## **METODOLOGÍA**

Los datos analizados provienen del Servicio de Farmacia del Hospital Universitario Reina Sofía de Córdoba y cubren el período comprendido entre el lunes 2 de enero de 2012 y el domingo 11 de agosto de 2013, es decir, un total de 88 semanas.

La información se recibe como una base de datos (FarmaTools v.2.4) en formato Access y contiene 260 tablas con más de un millar de campos. Para el presente estudio se transforma y simplifica dicha información generando un archivo en formato de hoja de cálculo Excel conteniendo el consumo semanal para 10 fármacos objetos de estudio.

La estructura del archivo Excel generado es la siguiente:

- La primera fila contiene las cabeceras de cada columna.
- La columna 1 contiene un indicador de la semana (un valor de 1 a 88).
- Las columnas 2 a 11 contienen los consumos semanales de cada uno de los 10 fármacos.

Para poder operar con cada una de las técnicas de predicción, se dividirán los datos de cada serie en tres zonas (Guyon, 1997):

- Datos de entrenamiento. Ocupan el 70% del total de los datos disponibles (las 60 primeras semanas). Un modelo de una serie temporal suele depender de uno o varios parámetros que son estimados usando estos datos.
- Datos de validación. Ocupan el 15% del total de los datos disponibles (las 13 semanas siguientes). En ocasiones, en vez de usar un único modelo de cada serie, se obtendrán varios modelos de una misma familia. La elección del elemento concreto de la familia se realizará en base a estos datos de validación.
- Datos de prueba. Ocupan el 15% del total de los datos disponibles (las últimas 13 semanas). Se utilizan para comprobar la efectividad de la técnica de predicción elegida.

## **RESULTADOS Y DISCUSIÓN**

### ***Método del hospital***

La primera de las técnicas que vamos a considerar es la que viene siendo habitualmente utilizada por el propio hospital. Ésta consiste en hacer la predicción de una semana como la media de los consumos del año anterior (de las 52 últimas semanas).

### *Predicción por persistencia*

Quizás la forma más simple de realizar la predicción de valores de una serie temporal sea tomar el valor anterior, una técnica ampliamente usada en muchos campos y, especialmente, en el de la predicción meteorológica (Wilks, 2011). Este método, denominado de persistencia, se suele tomar de base para comprobar la mejora de otros métodos de predicción más elaborados.

### *Predicción por media móvil óptima*

Tanto el método de la persistencia como el método del hospital, se pueden considerar casos particulares de un método más general: el de la media móvil (Thomopoulos, 2016). Según este método la predicción se calcula como la media de los  $M$  valores anteriores.

### *Predicción por alisado exponencial*

Los métodos de media móvil, como los 3 analizados anteriormente, consideran para la predicción los  $M$  valores precedentes, “olvidando” el  $M + 1$  y anteriores. Una forma menos brusca de “olvido” consiste en darle más peso en la predicción a los valores más recientes y menos a los más antiguos. Una forma de conseguir este efecto es mediante el denominado método del alisado exponencial (Montgomery, Jennings & Kulahci, 2015; Gardner, 2006) que se puede expresar de la siguiente forma

$$\hat{x}_n = \alpha \cdot x_{n-1} + (1 - \alpha)\hat{x}_{n-1}$$

siendo  $0 \leq \alpha \leq 1$ . La predicción para un instante se toma en función del valor en el instante anterior y de la predicción para dicho instante anterior. Esta última predicción será por analogía

$$\hat{x}_{n-1} = \alpha \cdot x_{n-2} + (1 - \alpha)\hat{x}_{n-2}$$

por lo que, sustituyendo, tenemos

$$\hat{x}_n = \alpha \cdot x_{n-1} + (1 - \alpha)[\alpha \cdot x_{n-2} + (1 - \alpha)\hat{x}_{n-2}]$$

$$\hat{x}_n = \alpha \cdot x_{n-1} + \alpha(1 - \alpha)x_{n-2} + (1 - \alpha)^2\hat{x}_{n-2}$$

De nuevo podemos escribir

$$\hat{x}_{n-2} = \alpha \cdot x_{n-3} + (1 - \alpha)\hat{x}_{n-3}$$

y sustituyendo obtener

$$\hat{x}_n = \alpha \cdot x_{n-1} + \alpha(1 - \alpha)x_{n-2} + (1 - \alpha)^2[\alpha \cdot x_{n-3} + (1 - \alpha)\hat{x}_{n-3}]$$

$$\hat{x}_n = \alpha \cdot x_{n-1} + \alpha(1 - \alpha)x_{n-2} + \alpha(1 - \alpha)^2x_{n-3} + (1 - \alpha)^3\hat{x}_{n-3}$$

$$\hat{x}_n = \alpha \cdot [x_{n-1} + (1 - \alpha)x_{n-2} + (1 - \alpha)^2x_{n-3}] + (1 - \alpha)^3\hat{x}_{n-3}$$

Con carácter general (considerando  $\hat{x}_1 = x_1$ )

$$\hat{x}_n = \alpha \cdot [x_{n-1} + (1 - \alpha)x_{n-2} + (1 - \alpha)^2x_{n-3} + \dots + (1 - \alpha)^{n-2}x_2] + (1 - \alpha)^{n-1}x_1$$

o, lo que es lo mismo,

$$\hat{x}_n = (1 - \alpha)^{n-1}x_1 + \alpha \sum_{k=1}^{n-1} (1 - \alpha)^{k-1}x_{n-k}$$

### ***Predicción por modelo ARIMA***

Una secuencia de valores tomados en intervalos temporales, como el consumo de fármacos que nos ocupa, se denomina una serie temporal. Un método para la predicción de valores de una serie temporal es suponer que dicha serie se comporta de acuerdo con un determinado modelo estocástico. En este apartado exploramos un tipo de modelo ampliamente utilizado (Box, Jenkins, Reinsel & Ljung, 2015): el modelo AutoRegresivo Integrado de Media Móvil (*ARIMA: AutoRegressive Integrated Moving Average*).

### ***Comparación por fármaco***

En los apartados anteriores se han descrito diversas técnicas de predicción del consumo y se han obtenido las métricas que producen cada una de ellas. Sin embargo la visión que se ha ofrecido hasta ahora no deja de ser parcial y

fragmentaria. Conviene analizar esa misma información desde otros puntos de vista que permitan la comparación de las técnicas propuestas.

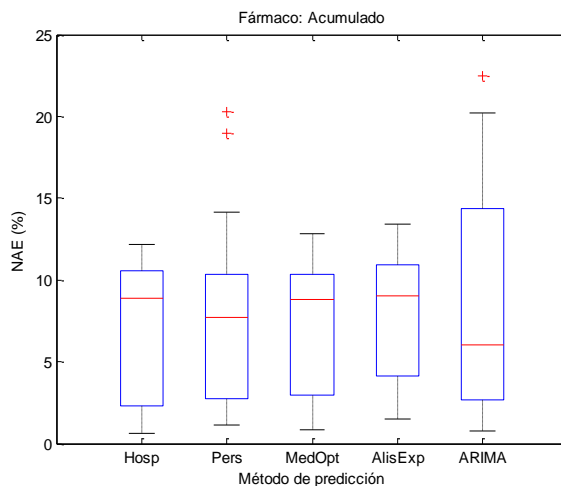
En primer lugar es pertinente realizar una comparación de técnicas, para determinar cuál de ellas se comporta mejor para cada fármaco. En este sentido, un indicador de la calidad de la predicción es el error que se comete. Y como lo que importa es la magnitud del error y no su signo, nos fijaremos en el error absoluto (*AE: Absolute Error*) de cada predicción definido como

$$AE_k \equiv |\varepsilon_k| = |\hat{x}_k - x_k|$$

A efectos de que posteriormente resulte más fácil la comparación con otros fármacos, se definirá el error absoluto normalizado (*NAE: Normalized Absolute Error*) de cada predicción como

$$NAE_k \equiv \frac{AE_k}{\bar{x}} = \frac{|\hat{x}_k - x_k|}{\frac{1}{N} \sum_{i=1}^N x_i}$$

Con estas definiciones es claro que tendremos tantos valores de  $NAE_k$  como predicciones se realicen. La siguiente figura representa en un gráfico de cajas la distribución de estos valores para el consumo de fármacos acumulado y para cada método de predicción. La caja azul comprende los valores entre el percentil 25 y el 75; línea roja representa la mediana; las líneas negras señalan los valores máximo y mínimo (no considerados anómalos); y las cruces rojas trazan los valores considerados anómalos (aquellos que superan en un 50% el rango intercuartil).



**Figura 1:** NAE para cada método de predicción

En esta gráfica observamos que, para el consumo acumulado, el método ARIMA es el que presenta una mediana del error más baja, aunque es la que tiene una mayor dispersión de valores del error. Esta dispersión puede provocar rupturas del stock más acusadas y la consiguiente necesidad de incrementar globalmente el stock para evitar estas rupturas. Es pues conveniente disponer de métricas más directas y simples para la comparación de técnicas.

### Comparación global

Como queda claramente de manifiesto en los gráficos anteriores, cada fármaco ofrece unos resultados distintos en cuanto a la mejor técnica de predicción. Esto no debe sorprender, ya que el consumo de cada uno de ellos toma la forma de una serie temporal con muy diversos comportamientos. Es lógico que cada técnica se adapte mejor (ofrezca mejores predicciones) a ciertas evoluciones de la serie.

No obstante, a menos que se conozca a priori la forma de evolución de cada fármaco (cosa que no ocurre en este caso), deberíamos seleccionar una técnica de predicción que tuviera un buen comportamiento para el conjunto de los fármacos. Es decir, una técnica que presente un buen compromiso para cualquiera que sea la forma de evolución de los consumos.

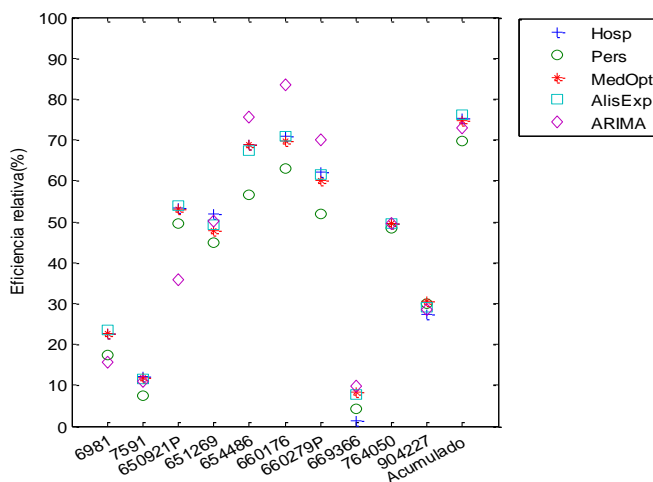


Figura 2: Eficiencia relativa

En el gráfico anterior podemos distinguir los mejores métodos para cada fármaco (con el criterio de la eficiencia) pero es difícil hacer una comparación global de métodos de predicción.

## CONCLUSIONES

En el trabajo desarrollado a lo largo de este primer año se han acometido las siguientes tareas:

- Se ha realizado una depuración y preparación de los datos suministrados para disponer la información del consumo de fármacos en forma de series temporales.
- Se han segmentado los datos en tres zonas de trabajo: entrenamiento, validación y prueba.
- Se han definido y formalizado diferentes métricas que permiten obtener indicadores cuantitativos del éxito de la predicción.
- Se ha propuesto una política simple de aprovisionamiento que permite medir las técnicas de predicción en función del volumen de stock al que conducen.
- Se han sistematizado y analizado las principales técnicas clásicas de predicción en series temporales:
  - El método utilizado por el hospital.
  - La persistencia.
  - La media móvil óptima.
  - El alisado exponencial.
  - Los modelos ARIMA.
- Se han diseñado indicadores y gráficas que permiten la comparación de técnicas:
  - Individualmente para cada fármaco.
  - De manera global para todos los fármacos.

Quedan abiertas diversas líneas de continuación de este trabajo que se pretenden abordar en el próximo futuro:

- Incorporación de nuevas técnicas de predicción del ámbito de la minería de datos.
- Análisis de sensibilidad de los parámetros de los modelos, midiendo su influencia en la calidad de la predicción.
- Análisis de la influencia los datos de consumo en un instante, sobre el consumo en otros instantes (autocorrelación).
- Análisis de la influencia de los consumos de un fármaco sobre los consumos de otro fármaco (covarianza).

- Elección y determinación de un modelo único de predicción para el conjunto de los fármacos que sustituya al actual método utilizado en el hospital.

## BIBLIOGRAFÍA

Vázquez Minguela, J. (2009). Cátedras Universidad-Empresa: una apuesta por la formación en la Universidad del siglo XXI. *Capital Humano*, (to), 54-58.

Box, G. E., Jenkins, G. M., Reinsel, G. C., & Ljung, G. M. (2015). *Time series analysis: forecasting and control*. John Wiley & Sons.

Dima, I. C., & Man, M. (2015). *Modelling and Simulation in Management: Econometric Models Used in the Management of Organizations*. Springer.

Gardner, E. S. (2006). Exponential smoothing: The state of the art—Part II. *International journal of forecasting*, 22(4), 637-666.

Govindarajan, R., Perelló-Juncá, A., Parès-Marimòn, R. M., Serrais-Benavente, J., Ferrandez-Martí, D., Sala-Robinat, R., ... & Gonzalez-Estrada, J. (2013). La gestión por procesos en la Farmacia Hospitalaria para la mejora de la seguridad del paciente. *Revista de Calidad Asistencial*, 28(3), 145-154.

Guyon, I. (1997). A scaling law for the validation-set training-set size ratio. *AT&T Bell Laboratories*, 1-11.

Hyndman, R. J., & Koehler, A. B. (2006). Another look at measures of forecast accuracy. *International journal of forecasting*, 22(4), 679-688.

Montgomery, D. C., Jennings, C. L., & Kulahci, M. (2015). Exponential Smoothing Methods. In *Introduction to time series analysis and forecasting* (pp. 233-326). John Wiley & Sons.

Olsen, C., (2012). Hospital Pharmacy Management. *Managing Access to Medicines and Health Technologies*, MDS-3, chap. 45.

Thomopoulos, N. T. (2015). *Demand forecasting for inventory control*. Springer International Publishing.

Thomopoulos, N. T. (2016). Forecasting. In *Elements of Manufacturing, Distribution and Logistics* (pp. 1-27). Springer International Publishing.

Torreblanca, J. M. M., Tejera, B. I., García, M. I. F., del Prado Llergo, J. R., Cantarero, T. Á., & Camacho, E. F. (2012). Análisis y minimización del riesgo de rotura de stock aplicado a la gestión en farmacia hospitalaria. *Farmacia Hospitalaria*, 36(3), 130-134.

Wilks, D. S. (2011). *Statistical methods in the atmospheric sciences* (Vol. 100). Academic press.