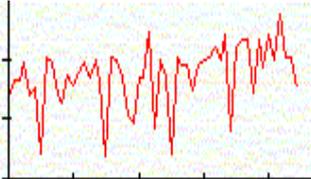


NOMBRE DEL PROYECTO	Aplicación de minería de datos para predicción de ventas
SECTOR	Gestión
OBJETIVO	Desarrollo de un modelo para predecir las ventas de un producto en un determinado mes, basándose en datos sobre las ventas en meses previos.
DESCRIPCIÓN	
<p>Uno de los campos de aplicación tradicionales de la minería de datos es la predicción de la evolución en el futuro de una variable (o conjunto de variables) a partir de datos históricos sobre su comportamiento en el pasado. Las técnicas de minería de datos constituyen una alternativa útil y eficaz a las aproximaciones matemáticas tradicionales, especialmente en el caso de variaciones muy irregulares, complicadas de modelar con los métodos clásicos.</p>  <p>La empresa Bayer mantiene un registro histórico de diferentes datos, entre ellos las cifras de ventas. Basándose únicamente en los datos de ventas de uno de sus productos, sin indicadores adicionales, pretende desarrollar un modelo del comportamiento de dicho producto en el mercado que le permita predecir las ventas del mismo con cierta anticipación. En concreto, se dispone de las cifras de los últimos 56 meses.</p> <p>Con esta información, los resultados obtenidos aplicando métodos de predicción tradicionales no son suficientemente precisos. Así, tomando como estimación para un mes el importe correspondiente al mes anterior, se obtiene un error del 25,6%, y del 14% si se usa el de ese mes en el año anterior. La calidad de la predicción mejora utilizando medias móviles, pero el error es aún del 11,8%. Debido a los altos porcentajes de error obtenidos, ninguno de estos métodos satisface las necesidades de la compañía.</p> <p>Para mejorar la precisión del modelo y conseguir la exactitud necesaria en las predicciones, se han aplicado técnicas de minería de datos. En primer lugar, se han analizado las características básicas de la serie. A simple vista, se observa que las ventas presentan una tendencia creciente en el tiempo que puede modelarse con medias móviles. También se observan oscilaciones estacionales, aunque estas regularidades no aparecen en todos los meses. Por ejemplo, si bien los valores de las ventas son siempre bajos en agosto, los de mayo presentan grandes variaciones. Esto puede significar que la serie incluye varios factores de influencia con distintos periodos. Estas observaciones se ven confirmadas por el análisis del espectro de frecuencia, que muestra varios máximos.</p> <p>Las conclusiones de los estudios preliminares sugieren la conveniencia de incluir en el modelo información no sólo de los valores de ventas en los meses previos sino también sobre la tendencia de la serie y sobre la temporada en cuestión, datos todos ellos contenidos en la propia serie.</p> <p>El sistema desarrollado, basado en una red neuronal, predice las ventas en un determinado mes partiendo únicamente de características extraídas de la serie de ventas en función del tiempo, sin indicadores adicionales. En concreto, los datos de entrada con los que se han obtenido los mejores resultados son:</p> <ul style="list-style-type: none"> • Valores de las ventas en los tres meses anteriores. • Ventas del mes a predecir en el año anterior. • Valor medio de las ventas durante los últimos 12 meses. • Identificador del mes. <p>Las predicciones del modelo obtenido aplicando minería de datos son notablemente más precisas, con un error medio del 6,4%, que con las técnicas clásicas, debido sin duda a que la serie considerada presenta un comportamiento altamente volátil, difícil de modelar con los métodos clásicos.</p> <p><i>Este sistema ha sido desarrollado por MIT GmbH y Bayer AG utilizando DataEngine. DataEngine es distribuido en España por DAEDALUS – Data, Decisions and Language, S.A. Nuestro departamento de Ingeniería está a su disposición para desarrollar proyectos como éste.</i></p>	