

Pronóstico de la demanda en empresas retail

Técnica basada en Business Intelligence y Machine Learning

Raúl Benítez - Alberto Garcete

Tutores: PhD. Diego P. Pinto Roa - Ing. Aditardo Vázquez

Universidad Nacional de Asunción - Facultad Politécnica

Agosto 2018



Support Vector Machine - SVM

El vector w determina la orientación de un plano discriminante. El escalar b determina el desplazamiento del plano a partir del origen. Asumimos que los dos conjuntos son linealmente separables, esto es, existe un plano que correctamente clasifica todos los puntos en dos conjuntos. Así, SVM busca encontrar el hiperplano óptimo que separa dos clases, mientras maximiza el margen entre los puntos de las clases más cercanos.



Tabla de decisiones

Una tabla de decisión es una forma de modelar un conjunto de reglas complejas y sus correspondientes acciones. En el caso de la clasificación, las tablas modelan un conjunto de reglas con sus correspondientes clases.

Dada una instancia no etiquetada, un clasificador basado en tabla de decisiones busca por coincidencias exactas en la tabla de decisión utilizando solo las características en el esquema. Si no encuentran instancias coincidentes, la clase con mayor ocurrencia (la clase mayoría) en la tabla de decisión es retornada; caso contrario, la clase con mayor ocurrencia de todas las instancias coincidentes es retornada.



Clasificadores bayesianos

Este enfoque está basado en cuantificar las compensaciones (tradeoffs) entre varias decisiones de clasificación utilizando las probabilidades y los costos que acompañan a tales decisiones. Este tipo de enfoque hace uso del Teorema de Bayes.



Clasificadores basados en instancias

- El aprendizaje en estos algoritmos consiste simplemente en almacenar los datos de entrenamiento presentados. Cuando una nueva instancia de consulta es encontrada, se procede recuperar un conjunto de instancias relacionadas similares de la memoria y se las utiliza para clasificar la nueva instancia [Mit97]. De acuerdo a [WF05], los algoritmos basados en instancias consisten en el almacenamiento de ejemplos de entrenamiento y el uso de una función de distancia con el propósito de determinar cuál miembro del conjunto de entrenamiento es el más “cercano” a una instancia desconocida. Una vez que el ejemplo (o instancia) de entrenamiento más cercano ha sido localizado, se predice la clase para la instancia desconocida.

