

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



Algoritmos de Machine Learning

No hay un consenso general de cómo agrupar los algoritmos de Machine Learning en términos de su función o de cómo trabajan. Se mostró un método útil de agrupación, que no es perfecto y ni exhaustivo en los grupos y algoritmos. Hay algoritmos que pueden encajar en varias categorías como Learning Vector Quantization que es a la vez un método inspirado en una red neuronal y un método basado en instancia. También hay categorías que tienen el mismo nombre que describen el problema y la clase de algoritmo como Regression y Clustering. Se podría manejar estos casos listando los algoritmos dos veces o insertando en el grupo al que subjetivamente se ajusta mejor. Se utiliza este último enfoque de no duplicar algoritmos.



Algoritmos de Machine Learning

- Algoritmos de regresión

Los Algoritmos de Regresión (*Regression Algorithms*) modelan la relación que existe entre variables, se mejora iterativamente utilizando una medida de error en las predicciones hechas por el modelo. Los métodos de regresión son herramientas de las estadísticas que se han adoptado en el aprendizaje de la máquina. Se pueden utilizar el término regresión para referirse a la clase de problema y también a la clase de algoritmo. Estos algoritmos se implementan en problemas de aprendizaje supervisado.

- Algoritmos basados en instancia

El modelo de aprendizaje Basado en Instancia (*Instance-Based Learning*) es un problema de decisión con instancias o ejemplos de datos de entrenamiento que se consideran importantes o requeridos para el modelo. Estos métodos típicamente construyen una base de datos con ejemplos y los compara con los nuevos datos utilizando una medida de similaridad para así encontrar la mejor coincidencia y hacer la predicción. El enfoque se pone en la representación de las instancias almacenadas y las medidas de similaridad utilizadas entre instancias. Estos algoritmos se implementan en problemas de aprendizaje supervisado.



Algoritmos de Machine Learning

- Algoritmos de regularización

Los Algoritmos de Regularización (*Regularization*), comprenden una extensión hecha a otros métodos (típicamente a los métodos de regresión). Penaliza los modelos basándose en sus complejidades, favoreciendo modelos más simples que también son mejores de generalizar. Son populares, potentes y en general simples modificaciones de otros métodos. Estos algoritmos se implementan en problemas de aprendizaje supervisado y no supervisado.

- Algoritmos de árboles de decisión

Los métodos de Árboles de Decisión (*Decision Tree*) construyen un modelo de decisiones hechas en base a los valores de atributos en los datos. Las decisiones se bifurcan en la estructura del árbol hasta que se tome una decisión de predicción para un registro dado. Los árboles de decisión son entrenados en los datos para problemas de clasificación y regresión. Los árboles de decisión son a menudo rápidos y precisos y son un gran favorito en aprendizaje automático. Estos algoritmos se implementan en problemas de aprendizaje supervisado.



Algoritmos de Machine Learning

- Algoritmos bayesianos

Los métodos Bayesianos (*Bayesian*) son los que aplican explícitamente el teorema de Bayes para problemas tales como la clasificación y la regresión. Estos algoritmos se implementan en problemas de aprendizaje supervisado.

- Algoritmos de agrupación

La Agrupación (*Clustering*) así como también la regresión describen la clase de problema y la clase de método. Los métodos de agrupación suelen estar organizados según el enfoque del modelado, tales como los basados en centroides y los jerárquicos. Todos los métodos atañen a la utilización de las estructuras inherentes en los datos, para organizar dichos datos de la mejor manera posible en grupos de máxima uniformidad. Estos algoritmos se implementan en problemas de aprendizaje no supervisado.



Algoritmos de Machine Learning

- Algoritmos de aprendizaje de reglas de asociación

Los métodos de aprendizaje de Reglas de Asociación (*Rule System*) extraen reglas que mejor explican las relaciones observadas entre variables en los datos. Estas reglas pueden descubrir asociaciones importantes y comercialmente útiles, en grandes conjuntos de datos multidimensionales que pueden ser explotados por una organización. Estos algoritmos se implementan en problemas de aprendizaje supervisado.

- Algoritmos de redes neurales artificiales

Las Redes Neuronales artificiales (*Neural Networks*) son modelos inspirados en la estructura y/o función de las redes neuronales biológicas. Son una clase de búsqueda de patrones que se utilizan comúnmente para problemas de regresión y clasificación. Es realmente un enorme subcampo compuesto de cientos de algoritmos y variaciones para todo tipo de problemas. Se ha separado el aprendizaje profundo de las redes neuronales debido a su enorme crecimiento y popularidad. Estos algoritmos se implementan en problemas de aprendizaje supervisado, no supervisado y por refuerzo.



Algoritmos de Machine Learning

- Algoritmos de aprendizaje profundo

Los métodos de Aprendizaje Profundo (*Deep Learning*) son una moderna actualización de las redes neuronales artificiales que explotan el abundante y barato poder de computación. Se ocupan en construir redes neuronales mucho más grandes y complejas y, muchos métodos se refieren a problemas de aprendizaje semi-supervisados donde grandes conjuntos de datos contienen muy pocos datos etiquetados. Estos algoritmos se implementan en problemas de aprendizaje supervisado y no supervisado.

- Algoritmos de reducción de dimensionalidad

Al igual que los métodos de agrupación, la Reducción de la Dimensionalidad (*Dimensionality Reduction*) busca y explora la estructura inherente en los datos, pero en este caso de una manera no supervisada o en orden a resumir o describir los datos utilizando menos información. Esto puede ser útil para visualizar datos dimensionales o para simplificar datos que luego se pueden utilizar en un método de aprendizaje supervisado. Muchos de estos métodos pueden ser adaptados para su uso en clasificación y regresión. Estos algoritmos se implementan en problemas de aprendizaje no supervisado.



Algoritmos de Machine Learning

- Algoritmos ensemble

Métodos de Ensamble (*Ensemble*) son modelos compuestos por múltiples modelos más débiles, que son entrenados independientemente y cuyas predicciones son combinadas de alguna manera para hacer la predicción general. Mucho esfuerzo se pone en qué tipos de aprendices débiles combinar y las formas en que hay que combinarlos. Esta es una clase de técnica muy poderosa y como tal es muy popular. Estos algoritmos se implementan en problemas de aprendizaje supervisado y no supervisado.

- Algoritmos no lineales

Uno de los mas importante es Support Vector Machines. Estos algoritmos se implementan en problemas de aprendizaje supervisado y no supervisado.



Algoritmos de Machine Learning

Algoritmos de aprendizaje por refuerzo

Programación dinámica:

- Programación dinámica con Policy Iteration.
- Programación dinámica con Value Iteration.
- Programación dinámica con Generalized Policy Iteration (GPI).
- Asynchronous DP.
- Bootstrapping.



Algoritmos de Machine Learning

Método de Monte Carlo

- On-Policy Monte Carlo Control.
- Off-Policy Monte Carlo Control.

Aprendizaje por Diferencias Temporales

- Sarsa: On-Policy TD Control.
- Q-learning: Off-Policy TD Control.
- Actor-Critic Methods.
- R-Learning for Undiscounted Continual Tasks.

