

Introducción al Aprendizaje Automático

Grado en Ingeniería Informática (Esp. Computación)

Práctica 2

Regresión

Objetivos

El objetivo de esta primera práctica es implementar un algoritmo básico de regresión lineal múltiple, comparando los resultados de nuestra implementación con los que producen otras librerías comerciales, tanto a nivel de resultado como de eficiencia.

Para llevar a cabo la práctica, además de usar nuestra propia implementación, utilizaremos dos librerías de Python muy populares en el ámbito del machine learning: *Turi Create*¹ y *scikit-learn*².

Desarrollo de la práctica

1. Codifica una función que aplique el algoritmo de regresión lineal múltiple utilizando los dos últimos “trucos” o criterios que se explicaron en clase, y que coinciden con la aplicación del método del gradiente descendente sobre la función de error absoluto medio y error cuadrático medio.
2. Desarrolla un programa principal que haga uso de estas dos funciones, y que las aplique sobre el siguiente conjunto de datos:

x1	x2	y
1	2	1,03
1	3	-1,44
2	3	4,53
2	4	2,24
3	2	13,27
3	5	5,62
4	1	21,53

Compara los resultados obtenidos con ambas versiones del algoritmo.

3. Utiliza ahora las funciones propias de *Turi Create* para resolver el mismo problema y compara los resultados. Compara también la eficiencia de este método con el que has implementado, midiendo varias veces la ejecución de este y comparando con tu implementación.

¹ <https://github.com/apple/turicreate>

² <https://scikit-learn.org/stable>

4. Repite la cuestión anterior utilizando *scikit-learn* y compara los resultados, tanto a nivel de resultados como de eficiencia.
5. Ahora vas a modificar el código desarrollado en los ejemplos anteriores para que ambas funciones de error incluyan, además del error cuadrático, un término de regularización L1 y L2. Es decir, vas a implementar los métodos de regresión lasso y ridge, y aplicarás la función sobre los datos del conjunto diabetes³.
6. Utiliza ahora las funciones propias de Turi Create y Scikit-learn para realizar las regresiones lasso, ridge y elastic net.

³ <https://www4.stat.ncsu.edu/~boos/var.select/diabetes.tab.txt>