Machine Learning and Pattern Recognition — IMT4612
Gjøvik University Collage, Spring 2014

Author: Ole Henrik Paulsen
Student number: 130572

Assigment 2

This document was compiled on March 4, 2014

Contents

Abstract

Assigment 2 in Machinelearning

1 Learning as a Search

1.1 Global optimal solutions

1.2 Genetic Algorithm(GA)

1.3 Gradient Descent method

1.4 Performance domain

2 Statistical Learning

2.1 Computer program

I did the programming in Python with the library Numpy and Time. You need to install Numpy, but time is a core library of Python. Numpy are used to handle sqrt, max, min and mathematical functions on arrays. I also used Numpy to read in data from the txt files. Numpy also have functions for Euclidean and chebyshev, but they are NOT used in my program.

K Neaarest Neighbor are also programmed by from the bottom instead of using a library.

K nearst Neighbor code

The program are scaled to handle large amount of input, both train and validation data, and 13 attributes takes under 3 secounds to handle.

2.2 Read input files

The train.txt and validation.txt are read into the program with Numpy's genfromtxt. *genfromtxt code*

I save the input data in a masked arrays. and split out the label into a own array.

Splitting code

I confirm the input is 120 train samples and 10 validation samples in the output.

2.3 Radar and Area plot

2.4 Distance Algorythms

The follow three algorithms are included into the program:

Euclidean:

Squar Euclidean:

Chebyshev:

As you can see they got different output from each other

2.5 Output

The output of the program is as follow:

Output from the program

References

[1] Wikipedia. Lorem ipsum — wikipedia, the free encyclopedia, 2013. [Online; accessed 20-October-2013].