Optimization

1. Describe machine learning as an optimization problem.
2. For classification: the optimization is to find a classification function which minimum the classification error according to some data set.
3. For linear regression, the optimization is to find the hypothesis function which can minimum the least-square error between hypothesis f and real f.
4. For SVM, the optimization is to find the classification which has maximum margin between support vector.
5. Random Walk
6. Write pseudocode for random walk search algorithm.

From initial position:

Randomly choose a neighbor whose state is higher than current state

Set current position to neighbor.

Keep updating until reach the maximum.

1. What is the drawback of this search algorithm?

Compare with steepest gradient method, it is slower than steepest gradient.

1. Steepest gradient descent
2. Write pseudocode for the steepest gradient descent search algorithm

From initial position:

choose a neighbor whose state is compare to current one has deepest gradient.

Set current position to neighbor.

Keep updating until reach the maximum.

1. What is the drawback of this search algorithm?

It does not guarantee to find global maximum state, stuck in local maximum state.

1. Simulated annealing search
2. Explain how the simulated annealing search algorithm works?

To solve this local maximum (or minimum) state problem, in the process of choose gradient descent:

Instead of choosing the best one, it picks a random one, if the move improves the state, then this move will always be accepted. Otherwise, this move will be accepted with a probability. This probability decreases exponentially with the extent of how badness the move it is and the probability also decrease as the execution time increases.

The algorithm stops eventually when the scheduled time reached.

1. Is it guaranteed to find the global minimum (or maximum)?

It doesn’t guarantee to find the global maximum state, but with enough long time, it can reach the global minimum with a probability which is approaching 1.

1. Genetic algorithms
2. Explain how the genetic algorithm work?
3. Randomly generate some ‘individual’ which is your first generation.
4. Using fitness function to evaluate those ‘people’
5. Then chose pairs from those people. This choosing process can be base on the associated probability which is calculated from proportion on each individual’s fitness.
6. In each of those chosen pairs, randomly do crossover. (exchange a part of each other)
7. After crossover, randomly select one ‘bit’ in the individual and change it randomly (mutation).
8. Keep going with this process when they have offspring which satisfy our goal.
9. Explain how a child is created from two parent using crossover and mutation?

Each individual is represented as chromosome.

During the crossover, the selected parents exchange part of it with each other based on crossover point.

Then with some low probability, introduce a mutation in child (one bit in chromosome changed).

1. What are the benefit and drawback of elitism?

It can make the converge happened faster than random version. However, it will make the offspring have less diversity.