Machine learning Algorithms

蔡苗[1]

2019-04-18

[1]Department of Epidemiology and Biostatistics, College for Public Health and Social Justice, Saint Louis University. Email: miao.cai@slu.edu

感谢我的家人的支持。

Acknowledgement

I want to thank my mentor.

目录

表格

插图

Preface

This book works as a notebook to summarize the algorithms used in Bayesian inference and machine learning.

第一章 Introduction

第二章　Optimization

2.1 Discrete optimization

The **objective function** allows us to measure how “good” any given solution to the problem is. We seek to maximize or minimize the objective function.

Derivative/gradient based methods keep going “uphill” until they are at the top of the h

2.1.1 Heuristic and metaheuristic methods

> “a **heuristic** is a technique designed for solving a problem more quickly when classic methods are too slow, or for finding an approximate solution when classic methods fail to find any exact solution.”
>
> Wikipedia

Heuristic methods do not guarantee to find the global optimal solution (best solution)! Instead, they seek to find **a best available solution, given the resource spent looking for it**. A **heuristic method** is **geared towards a specific problem**.

> a **metaheuristic** is a higher-level procedure or heuristic designed to find, generate, or select a heuristic (partial search algorithm) that may provide a sufficiently good solution to an optimization problem, especially with incomplete or imperfect information or limited computation

> capacity. Metaheuristics sample a set of solutions which is too large to be completely sampled. Metaheuristics may make few assumptions about the optimization problem being solved, and so they may be usable for a variety of problems
>
> – Wikipedia

A metaheuristic method is like a heuristic, but generalizable to a broad class of problems.

1. Genetic Algorithms (Holland – 1975)

- Natural selection / genetics based. Popular method.

2. Simulated Annealing (Kirpatrick – 1983)

- Metallurgy annealing, find lowest energy level!

3. Particle Swarm Optimization (Eberhart Kennedy - 1995)

- Based on insect behavior, swarming towards optimal location (food). Less common in discrete spaces. originally proposed for continuous spaces.

4. Tabu Search (Al-Sultan – 1999)

- Search for best neighborhood solution, then find new neighborhood. Prior neighborhoods are forbidden (tabu)

General meta-heuristics traits

- Evaluate many potential optimal solutions.
- Evaluate the fitness of each solution based on a cost (objective) function.
- Use some concept of stochastic (random) movement to generate new solutions from the parameter space.
- Use some set of rules to determine where to move next in the parameter space.
- Declare convergence once some set of criteria has been met. Perhaps no improvement for X iterations.

2.1.2 Genetic algorithm and simulated Annealing as examples

Genetic algorithm: need to explore large portions of the parameter space at random. Concept of "neighbor" is vague.

A nice shiny app[1]

An GA example: Since a new treatment for Hep C has become available, where is the optimal place to locate limited new Hep C resources, considering where our patients live?

The problem become intractable with large number of locations and resources: How many combinations of patients and clinics can I calculate the full feature space for to find a maximum?

- Exact Solution is NP-Hard
- Calculations = $n^{\sqrt{k}}$
- I conveniently stopped my analysis at 6 sites with ~5k patients, requiring 1,149,712,053 distance calculations (I have a big server)
- The "k-center" problem

Simulated Annealing:

- The concept of a 'neighbor' is strong.
- Can be sensitive to parameter choice, or algorithm gets stuck in global minima!
- Generally, you should try both to see what works best. Hard to guess up front.

2.1.3 Constrains

Hard constraints

- If this constraint is violated, we have invalid solution.

[1] https://toddwschneider.com/posts/traveling-salesman-with-simulated-annealing-r-and-shiny/

- Labor laws, number of nurses available, etc

Soft Constraints

- These are nice to meet if possible (included in cost function somehow), but if they are not met the solution is still valid.
- Nurse prefers to only work X night shifts per month.
- Leave requests.

第三章　Introduction