From Regression to Classification

Evgenii Egorov

Plan

+ We know how to solve linear regression problem

Plan

+ We know how to solve linear regression problem
+ There is two restriction in beauty of our knowledge
  + How to fit *linear model* to the *continuous* response
+ Let’s deal today with constraint on *continuous* response

Plan

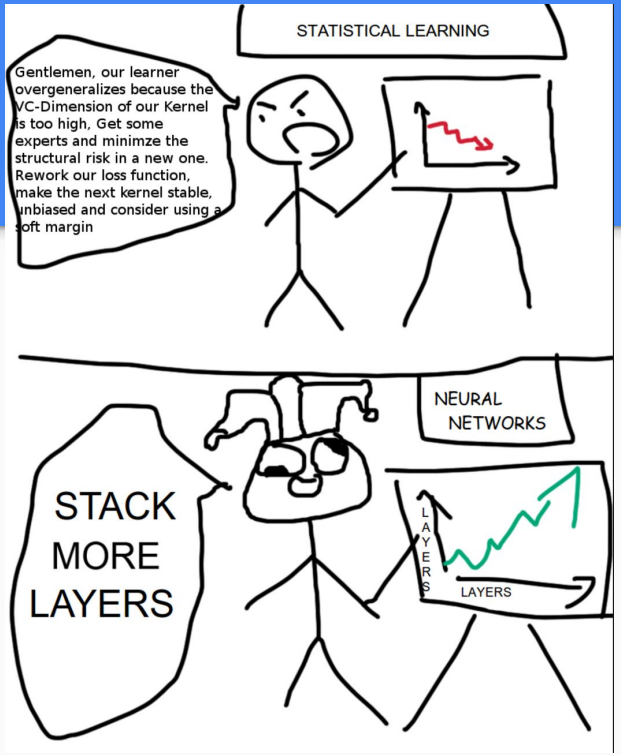
- Let’s deal today with constraint on *continuous* response
- Why is this important:
  - Classification problems, where Y can be 0 or 1, or 1,...,K.
  - Count problems, where output is a number of events:
    - Earthquakes, disorders, number of infected people
- If we can deal with this, generalization to non-linear models is simple step
- Besides, linear models are already useful

Plan

+ Let’s deal today with constraint on *continuous* response
+ Two roads to solution:
    1. Probabilistic point of view -> general framework

Plan

- Let's deal today with constraint on *continuous* response
- Two roads to solution:
  1. Probabilistic point of view -> general framework
  2. ML point of view -> Ad-Hoc adjusting optimization problem



Plan

- Let's deal today with constraint on *continuous* response
- Two roads to solution:
  1. Statistical point of view -> general framework
  2. ML point of view -> Ad-Hoc adjusting optimization problem
- We will discuss both
  - Statistical is general :) in cost of formality :(
  - ML has geometrical sense :) in cost of Ad-Hoc :(
    - Actually probabilistic has its own rich geometry :)
      - But for more much cost :(:(