Online Learning

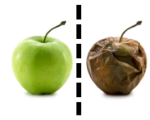
Alexander Mey, Tom Viering

TU Delft

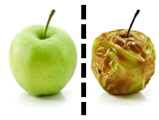
Statistical Learning

Example: buy apples on the market and classify them.

1. Observe labeled examples (training)



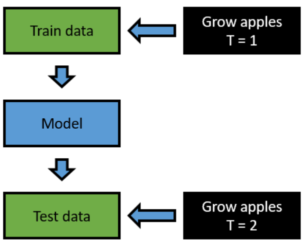
2. Build model
3. Evaluate on new data (testing)



Assumptions of Statistical Learning

Assumptions:
Data generation process

- is ‘fixed’, does not change
- is independent of ‘us’



When is this assumption NOT OK?

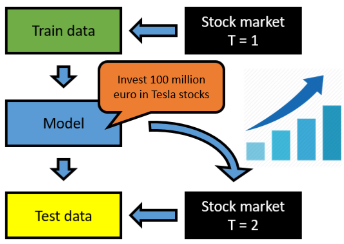
Investment

Consider investment:

- we build a model that predicts which stocks to buy / sell.
- our model is used by a very large hedgefund

Problem:

- Actions of our model influence the data.
- Data generation process changes over time.



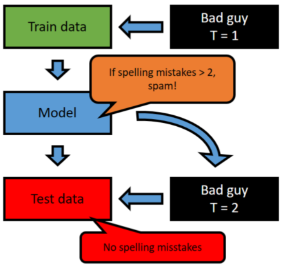
Spam Filter

Consider spam filtering:

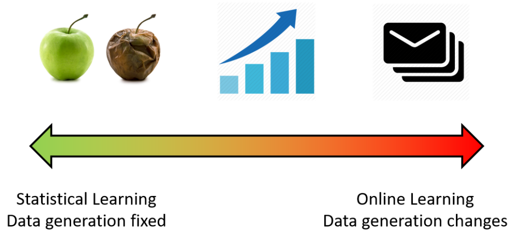
- we build a spam filter for GMail
- Model predicts: spam / no spam.

Problems:

- Actions of our model **heavily** influences data generation process.
- Data changes a lot over time.



Online Learning: Motivation



- We want no assumptions on the data generation process.
- We do not want training/testing separation.
- We want to be able to deal with dynamic/streaming data.

Online Learning: General Idea

We consider learning a game with 2 players.

- We make prediction, adversary generates data.

We will discuss two settings in Online Learning:

1. Online Learning with Expert Advice (Tom)
2. Online Convex Optimization (Alexander)

Online Learning with Expert Advice: Example

2 stock brokers (experts) each day t tell us what stocks we should buy and sell.

$t = 1$ Our turn: choose confidence in experts $p_t = (0.5, 0.5)$.

- Higher confidence in expert $i \rightarrow$ higher value of p_t^i.

$t = 1$ Adversary turn: choose $z_t = (100, 0)$.

- z_t^i is **the loss of expert** i.

$t = 1$ **Our loss** in this round: $l_d(p_t, z_t) = p_t^T z_t = 50$.

- $l_d(p_t, z_t)$ is the "dot loss", other loss also possible.
- This could mean we have lost 50 euros on day $t = 1$.

$t = 2$ Our turn: choose p_t

$t = 2$ Adversary turn: choose z_t

$t = 2$ Our loss $l_d(p_t, z_t)$

Online Learning with Expert Advice

We have d experts, each time step t give some advice.

Online Learning Procedure

- For $t = 1, \dots, n$
  - Our turn: we choose a distribution $p_t \in \Delta_d$ on d experts.
  - Adversary turn: Adversary chooses $z_t \in \mathcal{Z}$ (losses of d experts).
  - Our loss is $l(p_t, z_t)$.

- $p_t \in \Delta_d$ means $p_t = (p_t^1, .., p_t^d) \in [0,1]^d$ and $\sum_{i=1}^d p_t^i = 1$.
- $\mathcal{Z}$ is specified later.
- We will consider two losses: l_m (mix loss), l_d (dot loss).

- If we completely believe expert k on day t, then $p_t = e_k$.
- Example ($d = 3$), then $e_1 = (1, 0, 0)$, $e_2 = (0, 1, 0)$, etc...

Expert Regret

- We cannot guarantee small loss.
- Instead, we analyze the expert regret R_n^E

$$R_n^E = \sum_{t=1}^n l(p_t, z_t) - \min_i \sum_{t=1}^n l(e_i, z_t) \tag{1}$$

- Expert regret: loss compared to best (fixed) expert.

Goal: algorithm that chooses p_t that guarantees small R_n^E.

- 'Small regret' means R_n^E does not grow linearly with n

Mix Loss Setting

The mix loss (investment) is:

$$l_m(p_t, z_t) = -ln\left(\sum_{i=1}^d p_t^i \; e^{-z_t^i}\right) \tag{2}$$

- In this setting $z_t^i \in (-\infty, \infty]$, $\mathcal{Z} = (-\infty, \infty]^d$.
- If $p_t = e_i$, then $l_m(p_t, z_t) = z_t^i$.

For mix loss the expert regret simplifies:

$$R_n^E = \sum_{t=1}^n l_m(p_t, z_t) - \min_i \sum_{t=1}^n l(e_i, z_t)$$

$$R_n^E = \sum_{t=1}^n l_m(p_t, z_t) - \min_i \sum_{t=1}^n z_t^i$$

Example: what can go wrong? (mix loss)

$t = 1$ we choose $p_1 = (\frac{1}{2}, \frac{1}{2})$.
$t = 1$ adversary chooses $z_1 = (0, 1)$.
$t = 1$ we get loss $l_m(p_1, z_1) = 0.38$.
$t = 2$ we choose $p_2 = (1, 0)$ since e_1 performed pretty good.
$t = 2$ adversary turn, $z_2 = (\infty, 0)$.
$t = 2$ we get loss $l_m(p_2, z_2) = z_2^1 = \infty \to R_n^E = \infty$!

Mix Loss: Aggregating Algorithm (AA)

- How to guarantee small regret against any adversary?
  - Answer: 'spread our chances' / be conservative.
- Define cumulative loss of expert i up to time t:

$$L_t^i = \sum_{s=1}^{t} z_s^i \tag{3}$$

- AA strategy:

$$p_t^i = \frac{e^{-L_{t-1}^i}}{C_{t-1}} \tag{4}$$

where $C_{t-1} = \sum_{j=1}^{d} e^{-L_{t-1}^j}$ ensures $\sum_i p_t^i = 1$ (normalization).
- In first round, $p_t^i = 1/d$ for all experts i, since $L_0^i = 0$ for all i. Note $C_0 = d$.

Aggregating Algorithm: Theoretical Guarantee

Theorem (AA, mix loss)

For any adversary, for any $n > 0$, if l is the mix loss l_m, we have for AA that

$$R_n^E \leq ln(d) \tag{5}$$

Proof: see lecture.

Note that the expert regret does not grow at all if n increases!

Dot Loss: Exp Strategy

- Dot loss:

$$l_d(p_t, z_t) = p_t^T z_t \tag{6}$$

- For the dot loss we have to assume $z_t^i \in [0,1]$, $\mathcal{Z} = [0,1]^d$.
- Exp Strategy for dot loss. Very similar to AA
- We require as input a learning rate $\eta \in (0, \infty)$.
- Strategy:

$$p_t^i = \frac{e^{-\eta L_{t-1}^i}}{\sum_{j=1}^{d} e^{-\eta L_{t-1}^j}} \tag{7}$$

Exp Strategy: Theoretical Guarantee (Dot Loss)

Theorem (Exp Strategy, Dot Loss)

For any adversary, for any $n > 0$, if l is the dot loss l_d, we have for the Exp Strategy that

$$R_n^E \leq n\frac{\eta}{8} + \frac{ln(d)}{\eta} \tag{8}$$

Proof: Chapter (2.2) in [1].

If we set $\eta = \sqrt{\frac{8 log(d)}{n}}$ then

$$R_n^E \leq \sqrt{\frac{n}{2} log(d)} \tag{9}$$

and then R_n^E does not grow linearly in n!

[1] S. Bubeck (2011). "Introduction to online optimization". In: *Lecture Notes*, pp. 1–86