POLITECNICO
MILANO 1863

Project of Data Intelligence Application

Matteo Costantino (10494637)

Part 3 - Online Advertising

Index

Project Description

1. Choose a product to advertise by means of digital tools. Provide a brief description.
2. Imagine 5 advertising sub-campaigns. Imagine an average daily budget/clicks curve (providing, for every value of daily budget, the number of daily clicks) aggregating the curves of three different classes of users. Notice that, in order to define the curves, it is necessary the definition of probability distributions. Provide a description of the three classes of users. Note: the definition of the classes of the users must be done by introducing features and different values for the features (e.g., gender, interests, age).
3. Be given a cumulative daily budget constraint. Be also given a discretisation for the daily budget values. Apply the Combinatorial-GP-TS to the aggregate curve (we are implicitly assuming that the bidding is performed automatically by the advertising platform) and report how the regret varies in time.
4. Focus on a single sub-campaign. Report the the average regression error of the GP as the number of samples increases. The regression error is the maximum error among all the possible arms.
5. Suppose to apply, the first day of every week, an algorithm to identify contexts and, therefore, to disaggregate the curves if doing that is the best we can do. And, if such an algorithm suggests disaggregating the curve at time t, then, from t on, keep such curves disaggregate. In order to disaggregate the curve, it is necessary to reason on the features and the values of the features. Apply the Combinatorial-GP-TS algorithm and show, in a plot, how the regret and the reward vary in time (also comparing the regret and the reward of these algorithms with those obtained when the algorithms are applied to the aggregate curve).

Product Presentation

In this project we study the online advertising for a mobile application.

Specifically we focus on pay-per-click advertising, meaning advertising where the advertiser pays only when a user clicks on their ad.

An advertising campaign is characterized by a its sub-campaigns each with a (potentially) different pair ad/targeting and by a cumulative daily budget (also called spending plan). In this type of advertising campaign, the advertiser takes part into an auction specifying a bid and a value for the daily budget for each sub-campaign. The goal of advertisers is to select these variables to maximize the expected revenue they get from the advertising campaign. According to the literature, common methods used to face this problem are the second price auction (GSP) and the Vickrey-Clarke auction (VCG). The proposed method uses combinatorial bandit algorithms techniques to face the problem in an online fashion.

The application we choose to analyze has different target users, we decided to split them based on their age and gender.

- 60% of the users are male and 40% are female
- 70% of the users are young (Under 30) and 30% are old (30+)

Considering these classes we have that each sub-campaign is composed of 4 combined classes each with a different probability distribution..

Our campaign has 5 sub-campaigns.

Subcampaign and Aggregate Curves

Model:

• N number of sub campaigns composing the advertising campaign C

• C = {C1, ..., CN } advertising campaign

• T time horizon (expressed in days)

• t $\in$ T time (expressed in days)

• Xj bids (Not Considered in this project)

• Yj budget

• B spending plan

• vj value per click for sub-campaign j $\in$ N

• nj (yj) number of clicks for sub-campaign j $\in$ N given value of budget yj $\in$ Yj

Probability Distribution of each class:

Logistic Function: $n_j(x) = \frac{max}{1+e^{-g(x-x_0)}}$

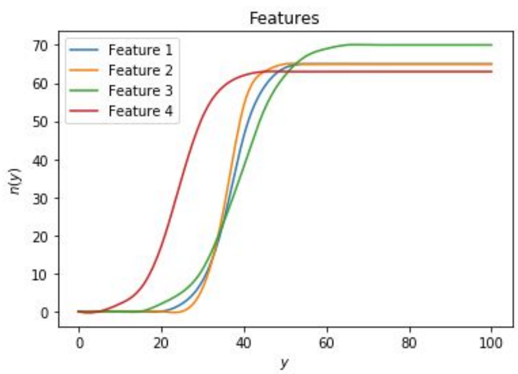
X = budget, max = Max value of clicks for this class, g = growth rate, x0 = midpoint value

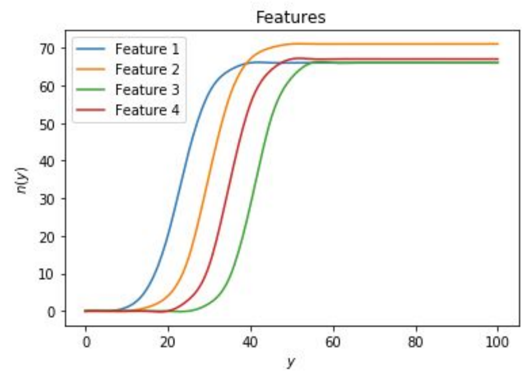
Specifications:

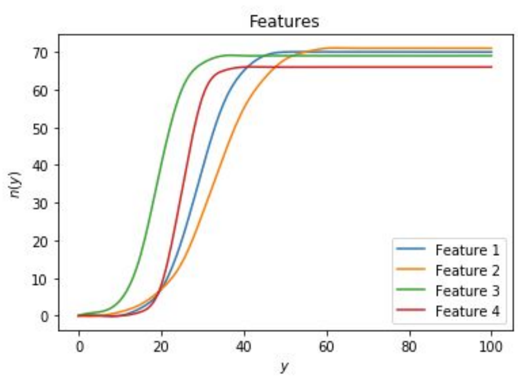
In this project run the experiment over a time horizon of T = 100 days, with N = 5 sub-campaigns and linearly spaced values of the budget yj $\in$ [0, 100] and |Yj | = 21, $\forall$j $\in$ N. We do not take into account the bids Xj. We also assume the values per click vj = 1.0, to be constant for each sub-campaign j $\in$ N. In addition, we set y j,t = [0,100]

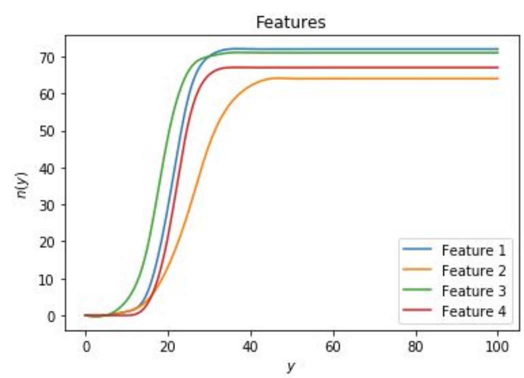
Graphs:

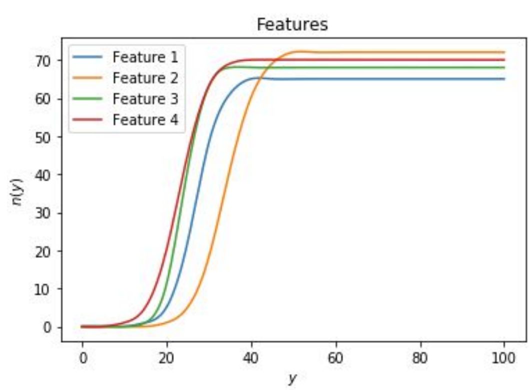
Features = (Male,Young), (Male,Old), (Female,Young), (Female,Old)

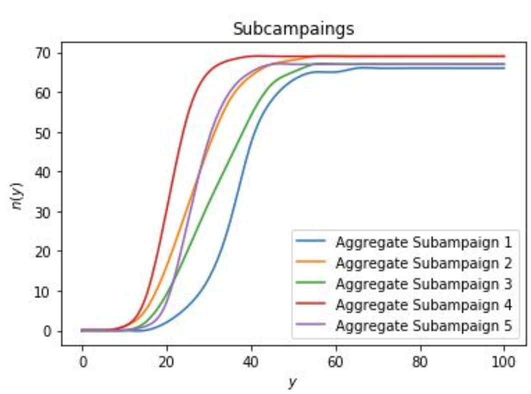












Combinatorial GPTS

Cumulative budget constraints of 100

Budget discretisation: 21 equispaced value from 0 to 100

First we define an algorithm to compute the clairvoyant baseline for the experiment:

Multichoice Knapsack Optimization Problem

Mathematical Formulation

- $\max_{x_{j,t},y_{j,t}} \sum_{j \in N} v_j n_j(x_{j,t}, y_{j,t})$
  s.t.
- $\sum_{j \in N} y_{j,t} \leq \bar{y}_t, \forall t \in T$
- $\underline{x}_{j,t} \leq x_{j,t} \leq \bar{x}_{j,t}, \forall j \in N, \forall t \in T$
- $\underline{y}_{j,y} \leq y_{j,t} \leq \bar{y}_{j,t}, \forall j \in N, \forall t \in T$

This algorithm takes in input all the subcampaigns, the budget discretization, and a spending plan and returns the optimal allocations of each budget for each subcampaign.

For our project we decided to use a constant value per click of 1, a spending plan of [0,100] for each subcampaign and do not take into account the bids, but only the budget.

This are the results and the clairvoyant aggregate solution:

	0.0	5.0	10.0	15.0	20.0	25.0	30.0	35.0	40.0	45.0	50.0	55.0	60.0	65.0	70.0	75.0	80.0	85.0	90.0	95.0	100.0
C0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
C1	0.0	0.0	0.0	0.0	2.0	6.0	13.0	27.0	47.0	58.0	63.0	65.0	65.0	66.0	66.0	66.0	66.0	66.0	66.0	66.0	66.0
C2	0.0	0.0	1.0	5.0	16.0	31.0	46.0	58.0	64.0	67.0	68.0	69.0	69.0	78.0	93.0	105.0	116.0	122.0	127.0	130.0	132.0
C3	0.0	0.0	1.0	5.0	16.0	31.0	46.0	58.0	64.0	67.0	68.0	69.0	78.0	90.0	101.0	112.0	120.0	126.0	129.0	132.0	134.0
C4	0.0	0.0	1.0	8.0	30.0	54.0	65.0	68.0	69.0	70.0	85.0	100.0	112.0	123.0	129.0	132.0	135.0	136.0	144.0	155.0	166.0
C5	0.0	0.0	1.0	8.0	30.0	54.0	65.0	68.0	69.0	70.0	85.0	102.0	114.0	125.0	130.0	133.0	135.0	148.0	160.0	172.0	183.0

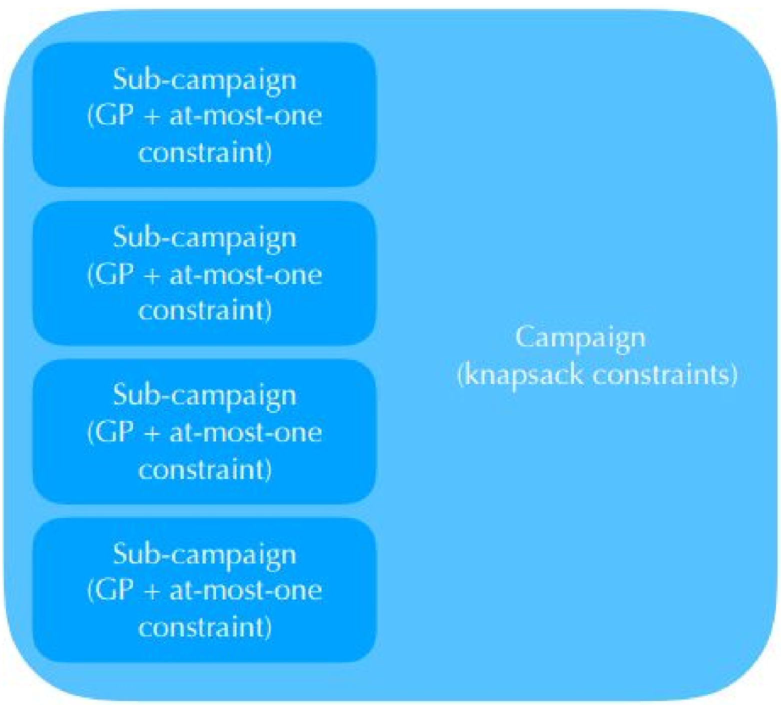
```
Max value: 183.0
Best super arm: [0.0, 7.0, 0.0, 6.0, 7.0]
```

Combinatorial Gaussian Process Thompson Sampling

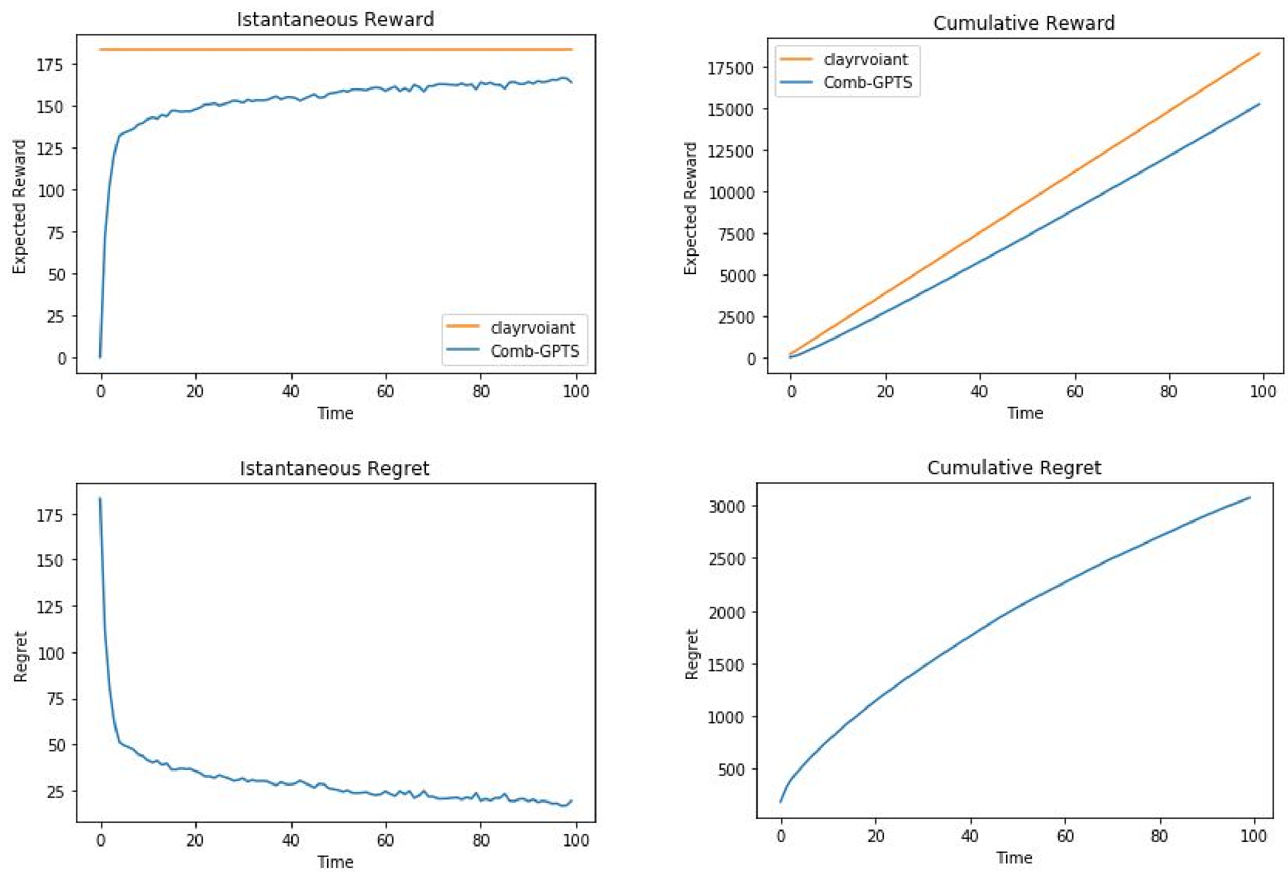
This algorithm uses Gaussian Processes applied with Thompson Sampling to learn the subcampaigns functions. Gaussian processes work well with bandit algorithms and allows us to have a measure of uncertainty of the regression.

The combinatorial variant use GPTS to learn each subcampaign and then applies the Knapsack Optimization to obtain the best allocation for the budget.

Parameters: sets $\{X_j\}_{j=1}^{N}$ of bid values, sets $\{Y_j\}_{j=1}^{N}$ of budget values, prior model $\{\mathcal{M}_j^{(0)}\}_{j=1}^{N}$, spending plan $\mathcal{B}$, time horizon T
for $t \in \{1, \dots, T\}$ **do**
  for $j \in \{1, \dots N\}$ **do**
    if $t = 1$ **then**
      $\mathcal{M}_j \leftarrow \mathcal{M}_j^{(0)}$
    else
      Get $(\tilde{n}_{j,t-1}, \tilde{c}_{j,t-1}, \tilde{r}_{j,t-1}, \tilde{v}_{j,t-1})$
      $\mathcal{M}_j \leftarrow$ Update $(\mathcal{M}_j, (\hat{x}_{j,t-1}, \hat{y}_{j,t-1}, \tilde{n}_{j,t-1}, \tilde{c}_{j,t-1}, \tilde{r}_{j,t-1}, \hat{v}_{j,t-1}))$
    $X_{j,t} \leftarrow X_j \cap [\underline{x}_{j,t}, \overline{x}_{j,t}]$
    $Y_{j,t} \leftarrow Y_j \cap [\underline{y}_{j,t}, \overline{y}_{j,t}]$
    $(n_j(\cdot,\cdot), v_j) \leftarrow$ Sampling $(\mathcal{M}_j, X_{j,t}, Y_{j,t})$
  $\{(\hat{x}_{j,t}, \hat{y}_{j,t})\}_{j \in N} \leftarrow$ Optimize $(\{n_j(\cdot,\cdot), v_j, X_{j,t}, Y_{j,t}\}_{j \in N}, \overline{y}_t)$
  Set $(\{(\hat{x}_{j,t}, \hat{y}_{j,t})\}_{j \in N})$



This are the results when we applied this algorithm to our project, with 50 experiment each long 100 time steps:



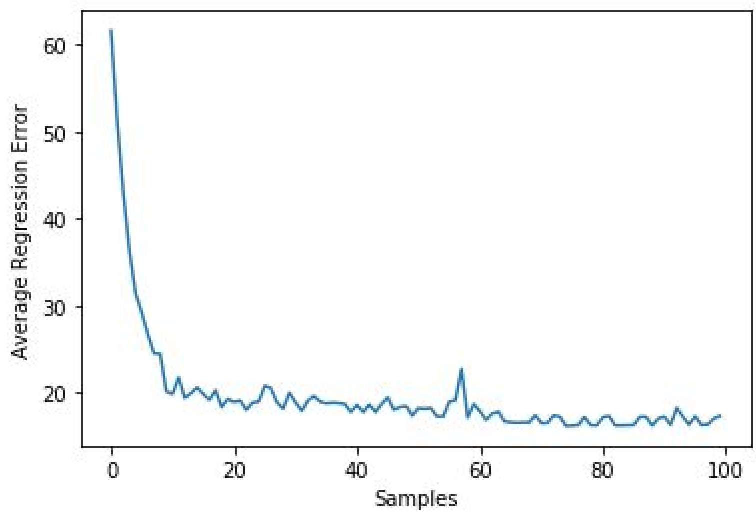
The results are as expected. The more time passes, the better the GPTS “learns” the campaign functions and so the regret is lower.

Regression Error

We focus on the first subcampaign and we compute the regression error for the GP of the last step.

The regression error is the maximum error among all the possible arms.

This are the results:



The more sample we have the lower the regression error tends to be. After a fast decline there is some noise due to the fact that the corresponding GP has probably “pull” a non-significant arm thus increasing the error for that sample.

Context Generation

We know apply a context generation algorithm:

Notation

- t time
- l features, we assume each feature can have a binary value (e.g. for feature young, we have that a customer is either young or not)
- $F \subseteq \{0,1\}^l$ space of attributes
- $c \subseteq F$ context
- $\mathscr{P} = \{c : c \subseteq F, c \cup \varnothing, c \cap F\}$ context structure

Value of a context structure $v = \sum_{c \in \mathscr{P}} p_c J_c^*(y)$

Where:

$J_c^*(y) = \sum_{j=1}^{N} v_j^* n(y_j^*)$

(J_c is the objective function of the MCK problem for context c)

Split condition $\underline{p}_{c_1} \underline{J}_{c_1}^*(y) + \underline{p}_{c_2} \underline{J}_{c_2}^*(y) \geq \underline{J}_{c_0}^*(y)$

Where:

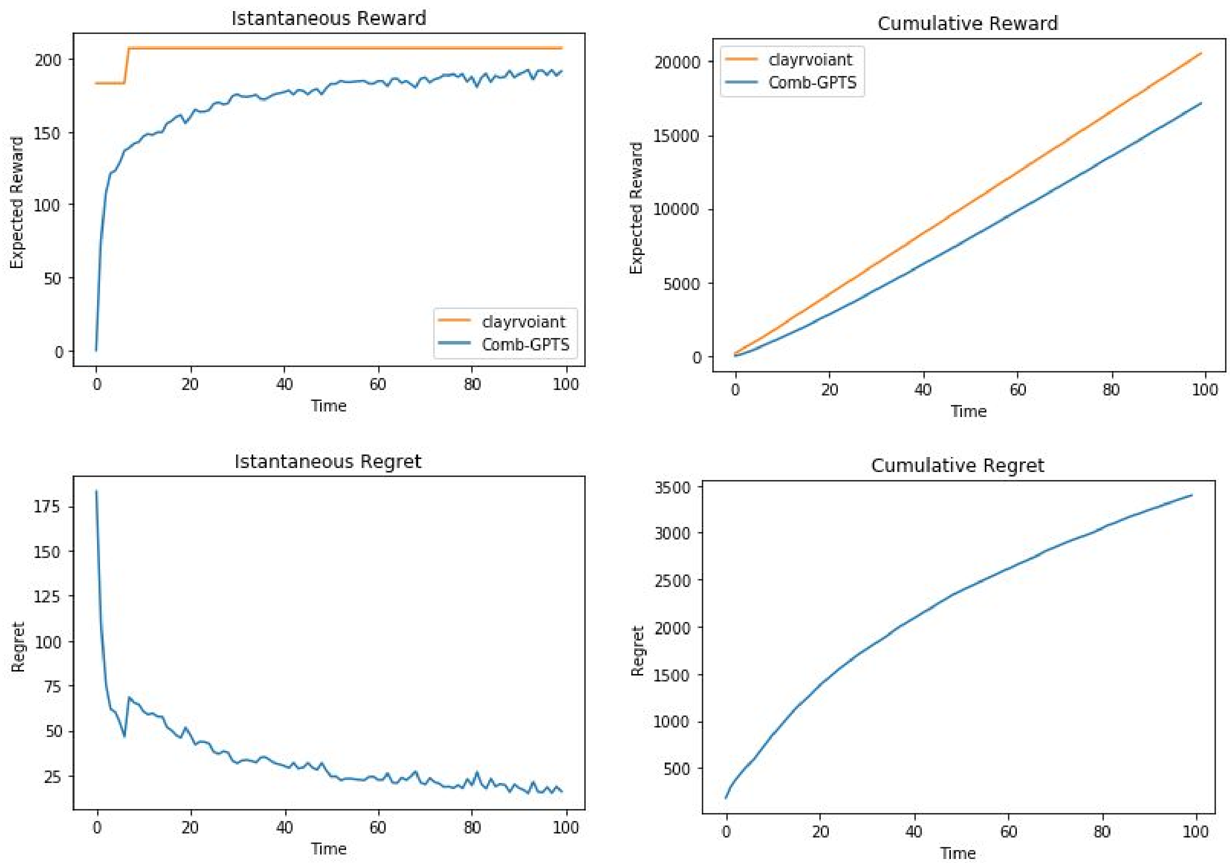
$\underline{p}_{c_i}$ is the lower bound of the probability of context c_i and,

$\underline{J}_c^*(y)$ is the lower bound of the (optimal) objective function of the MCK problem defined before in context c_i: $\underline{J}_c^*(y) = \sum_{j=1}^{N} \underline{v}_j^* \underline{n}(y_j^*)$

Every start of the week we sample 100 new users for our campaign, and we apply the algorithm.

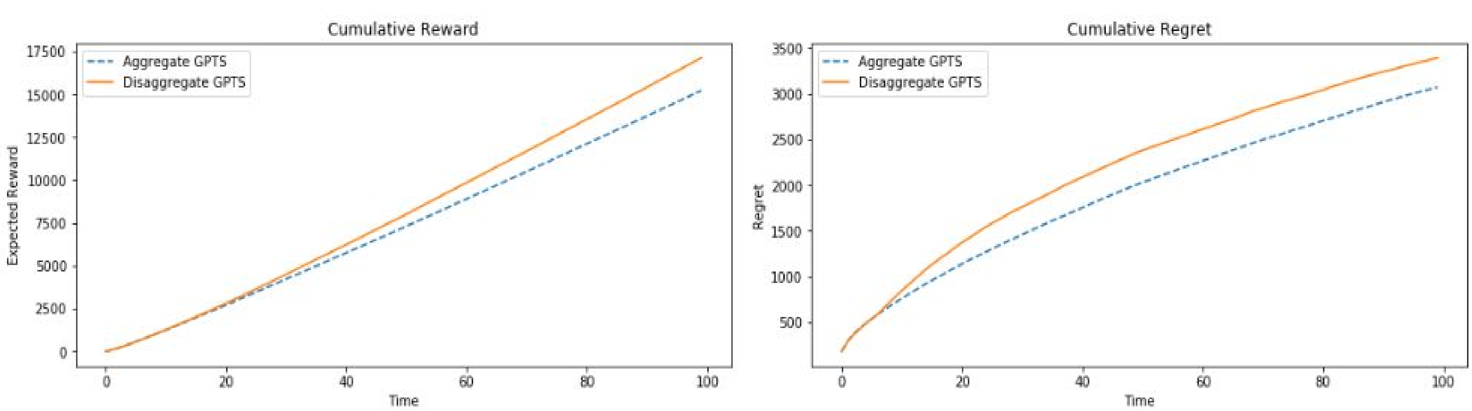
This algorithm tells us if, based on the users, it’s best to disaggregate the considered function and instead consider a single class instead of the aggregate sub campaign.

This are the results:



As we can see in the initial phase there is a context switch that increases the clairvoyant solution.

We notice that with this new context the GPTS has a better reward w.r.t. the aggregate one.



References

- Course slides and lectures
- Nuara, Alessandro & Trovò, Francesco & Gatti, Nicola & Restelli, Marcello. (2018). A Combinatorial-bandit Algorithm for the Online Joint Bid/Budget Optimization of Pay-per-click Advertising Campaigns.