Online Convex Opitmization: Project.

Olivier Wintenberger: `olivier.wintenberger@upmc.fr`

Recommendations: Groups of at most 3 students working on a chosen article among the articles available at
`https://dropsu.sorbonne-universite.fr/s/9kR297ZRYkKpZci`

The choice of the article is validated by emailing me at
`olivier.wintenberger@sorbonne-universite.fr`
Note that each group should choose a different article and the first to choose has the priority. The project must be done in the R or Python programming language. The results presented in a pdf file together with the codes ready to run should be sent by email for the 30th of December 2022.

The aim of the project is to implement online gradients methods for training linear SVM on MNIST dataset.

The performances of the algorithms, all trained on the same train set, will be quantified using their accuracy scores on the test set.

1 Preliminary

1. Download MNIST dataset in csv at `https://pjreddie.com/projects/mnist-in-csv/`.

2. Consider the binary classification 0 vs other digits. Create labels b_i and normalize the variable a_i so that the grey scale takes value in $[0, 1]$. Add an intercept.

2 Gradient Descent

1. Implement the unconstrained GD as seen during OCO lectures.

2. Implement also the ℓ^1-ball projection and implement the projected GD.

3. Discuss the choice of the hyperparameters λ and z, the parameter of the ℓ^2-regularization and the radius of the ℓ^1-ball, respectively.

3 Stochastic Gradient Descent

1. Implement the unconstrained and the projected SGD as seen during OCO lectures.

2. Compare the accuracy and the running time of the stochastic versus non-stochastic (projected) GD.

4 Regularized Follow The Leader

1. Implement the Stochastic Mirror Descent and compare with the projected SGD.

2. Implement the Stochastic Exponentiated Gradient +/- and compare with the projected SGD.

3. Implement the Stochastic AdaGrad and compare with the projected SGD.

5 Online Newton Step

1. Implement the Stochastic Online Newton Step.

2. Compare its performances in term of accuracy and running time with the other methods.

6 Exploration methods

1. Implement the Stochastic Randomized Exponentiated Gradient +/- and the Stochastic Bandit Exponentiated Gradient +/-.

2. Discuss their convergence. Is it in accordance with the regret bounds established during OCO lectures?

7 Other method from the chosen article

1. Implement the method from the chosen article.

2. Summarize the motivations for the introduction of the algorithm and the associated theoretical guarantees provided in the article.

3. Compare the performances with the ones of the algorithms from OCO lectures.

4. Were the performances in accordance with those described in the article?