

# Hyperparameter Optimization using Hyperopt

Yassine Alouini - Paul Coursaux  
03/11/2016

[@qucit](#)

[@YassineAlouini](#)

# About us

---



## Yassine

- Data Scientist @ Qucit
- Centrale Paris & Cambridge
- Quora's Top Writer 2016



## Paul

- Data Scientist @ Qucit
- Centrale Paris
- Market finance in London
- Horse riding

# Outline

---

1. Hyperparameters in Machine Learning
2. How to Choose Hyperparameters ?
3. Tree-structured Parzen Estimation Approach
4. Live-coding Example

# 1. Hyperparameters in Machine Learning

---

# What are hyperparameters ?

---

**Parameters:**

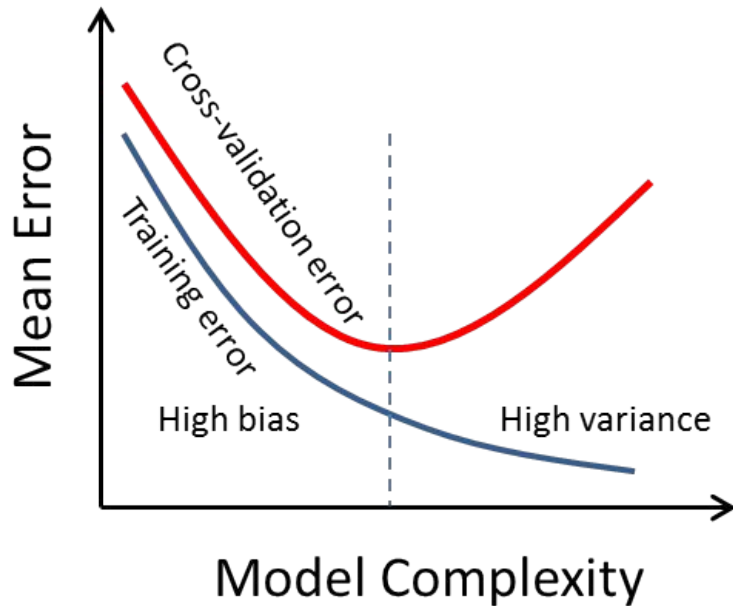
Rent =  $a_1 \times$  surface +  
 $a_2 \times$  distance to city center +  
...

**Hyperparameters:**

$$\text{RMSE}_{\text{LASSO}} = \text{RMSE} + \alpha \times (|a_1| + \dots)$$

# The impact of hyperparameters

---



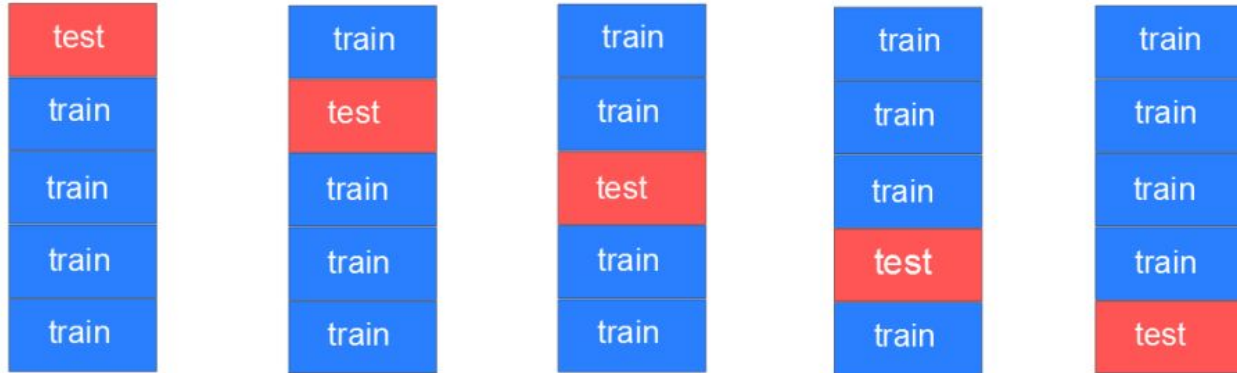
## 2. How to choose hyperparameters ?

---

# Cross validation

---

Enable to choose the hyperparameter(s) with the best generalization capabilities making an efficient use of the data



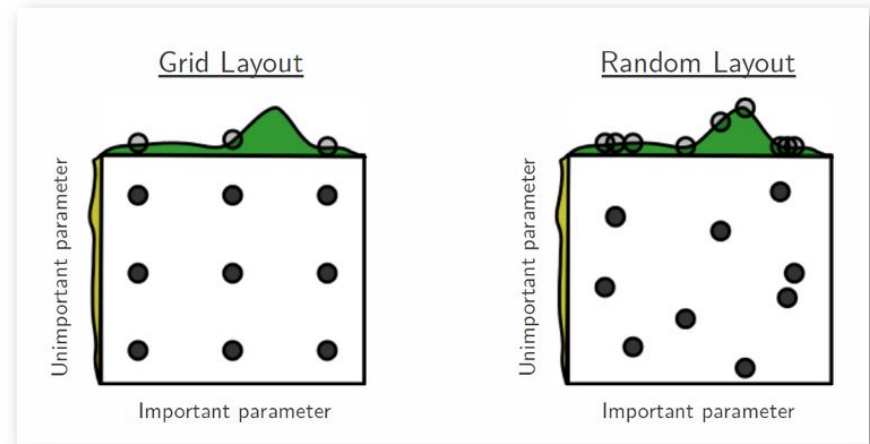


# How to choose the points to cross-validate?

## Grid search



## Random search

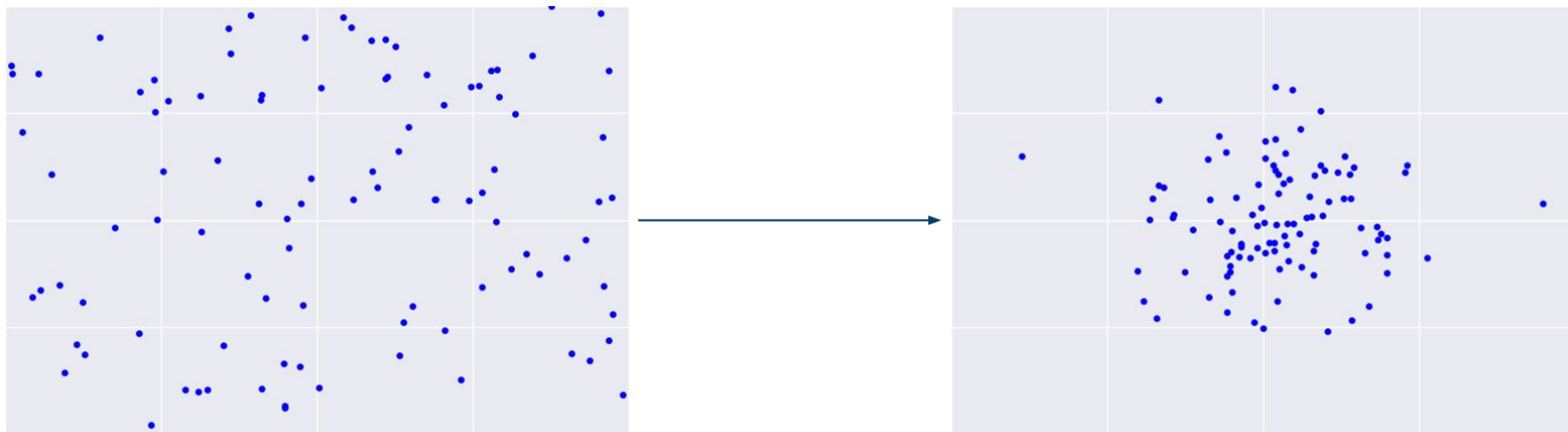


# 3. Tree-structured Parzen Estimation Approach

---

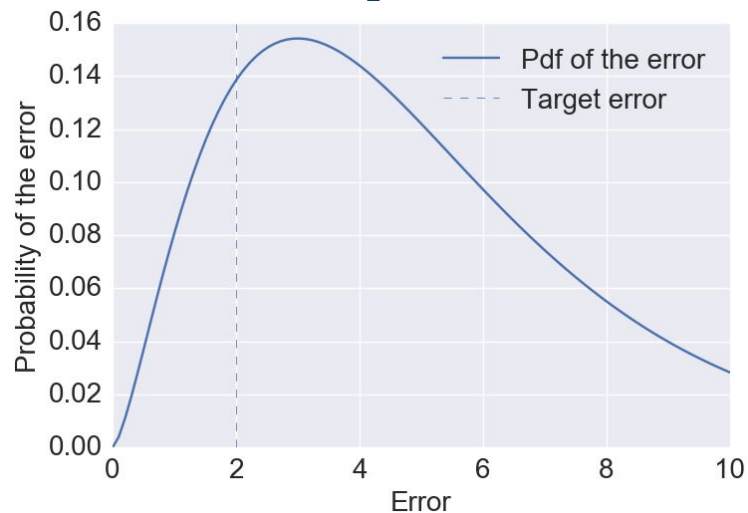
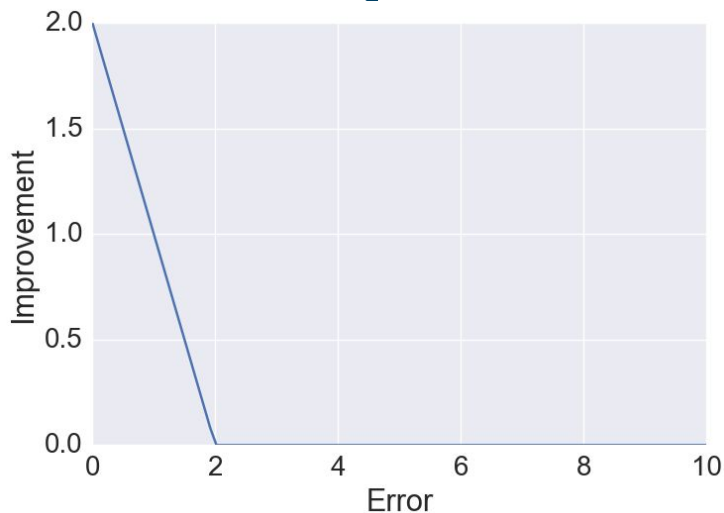
# Sequential Model-based Global Optimization

---



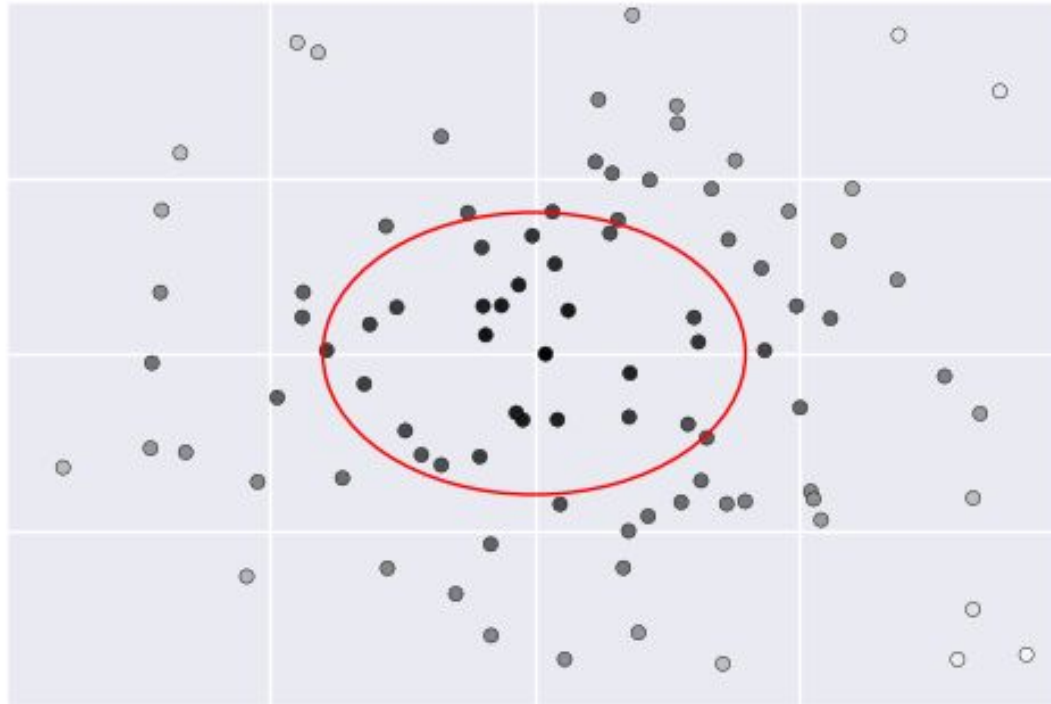
# The Expected Improvement

$$EI_{\varepsilon^*}(a) = \int \max(\varepsilon^* - \varepsilon, 0) p_M(\varepsilon|a) d\varepsilon$$



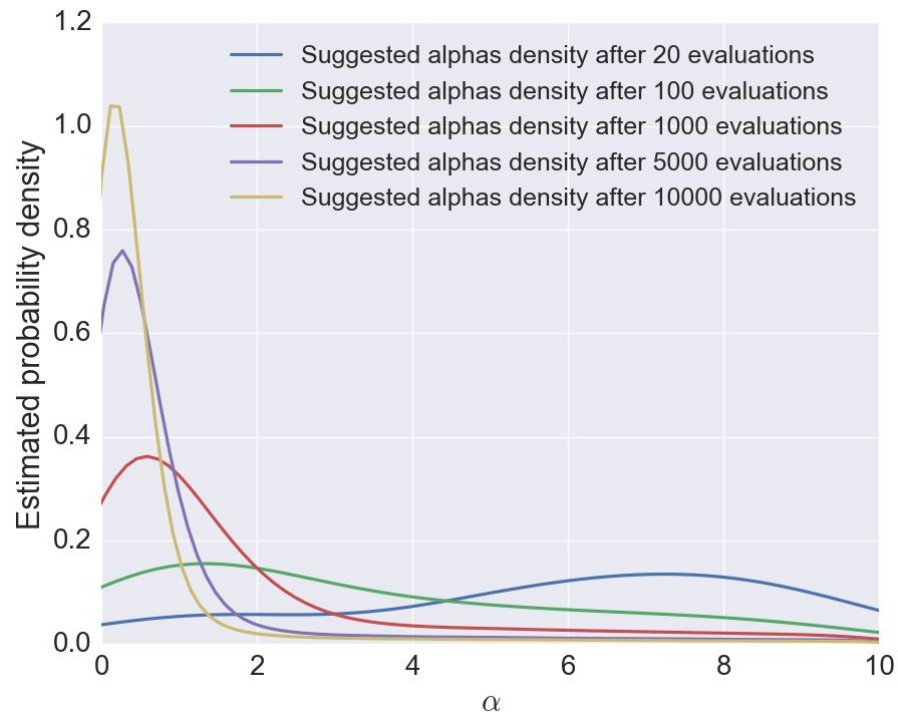
# How to Optimize the EI ? (1)

---



# How to Optimize the EI ? (2)

- Lasso model on the Boston Housing Dataset
- Distribution of the suggested  $\alpha$ s

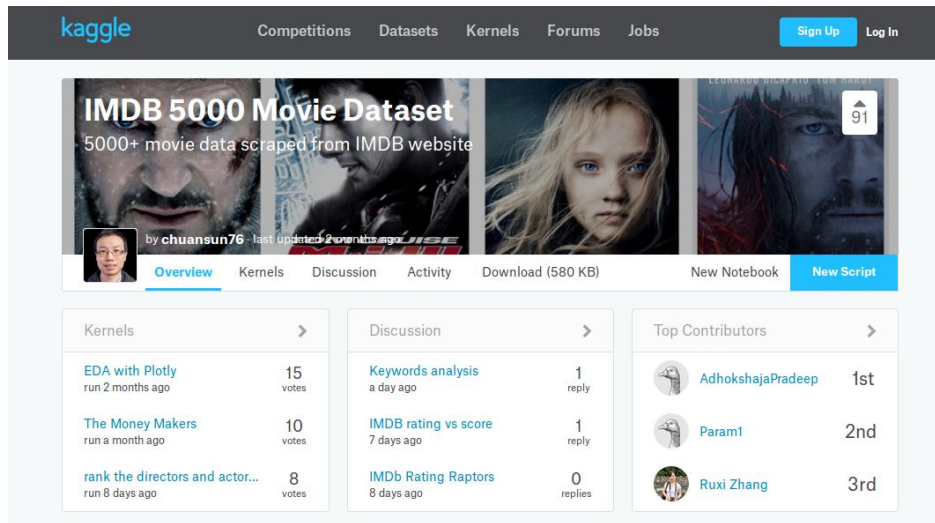


# 4. Live-coding Example

---

# Description of the dataset

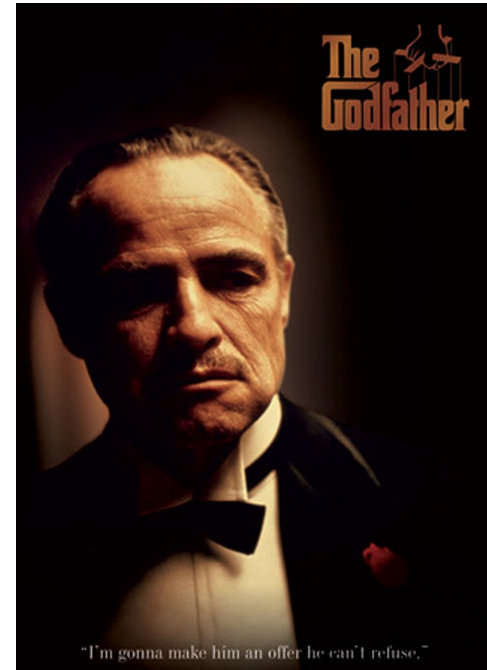
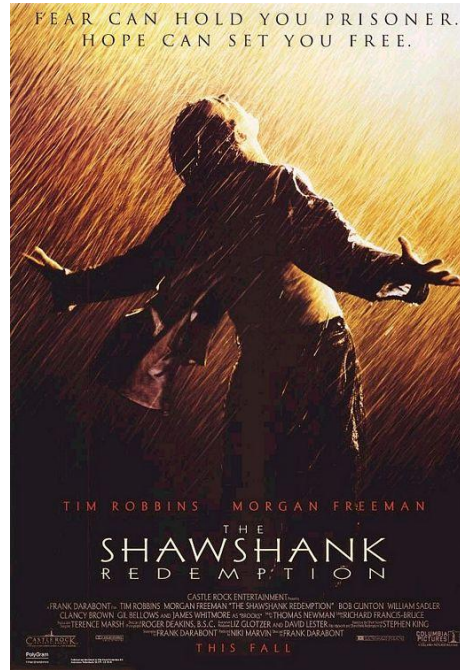
- IMDb dataset
- Dataset publicly available (from Kaggle)



The screenshot shows the Kaggle interface for the 'IMDb 5000 Movie Dataset'. The page header includes the Kaggle logo and navigation links for Competitions, Datasets, Kernels, Forums, and Jobs. A 'Sign Up' button and a 'Log In' link are also present. The main content area features a large banner for the dataset, titled 'IMDb 5000 Movie Dataset', with a subtitle '5000+ movie data scraped from IMDb website'. The dataset is attributed to 'chuansun76' and was last updated '2 months ago'. Below the banner, there are tabs for 'Overview', 'Kernels', 'Discussion', 'Activity', 'Download (580 KB)', 'New Notebook', and 'New Script'. The 'Overview' tab is selected. Below the tabs, there are three sections: 'Kernels', 'Discussion', and 'Top Contributors'. The 'Kernels' section lists three kernels: 'EDA with Plotly' (15 votes, run 2 months ago), 'The Money Makers' (10 votes, run a month ago), and 'rank the directors and actor...' (8 votes, run 8 days ago). The 'Discussion' section lists three discussions: 'Keywords analysis' (1 reply, a day ago), 'IMDb rating vs score' (1 reply, 7 days ago), and 'IMDb Rating Raptors' (0 replies, 8 days ago). The 'Top Contributors' section lists three contributors: 'AdhokshajaPradeep' (1st), 'Param1' (2nd), and 'Ruxi Zhang' (3rd).

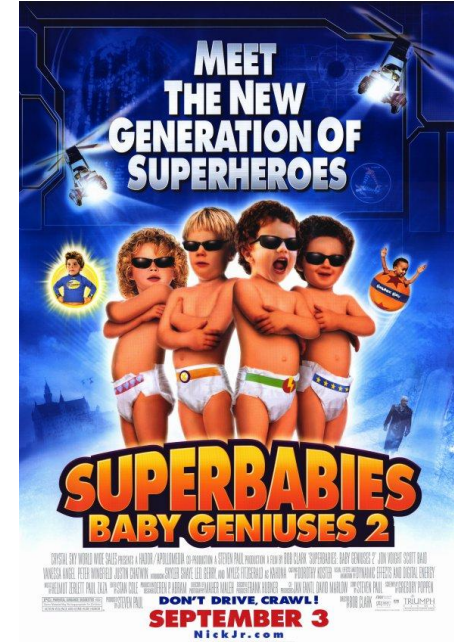
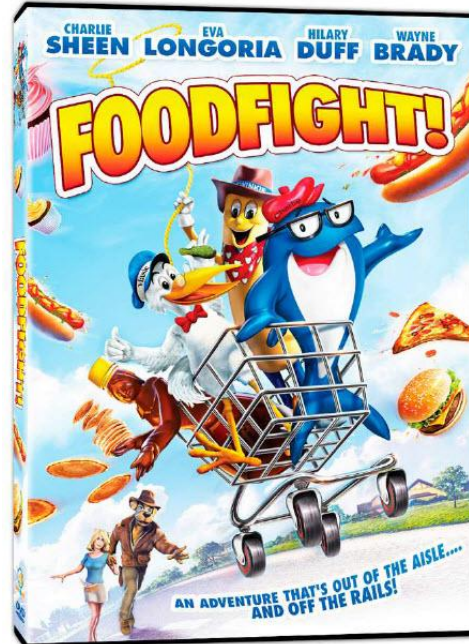
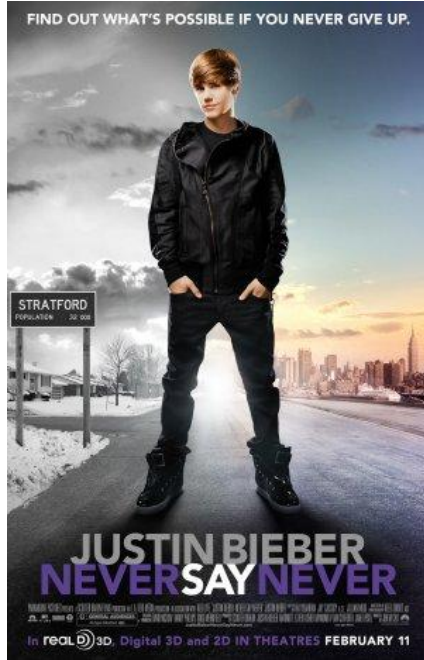


# Movies having the best score



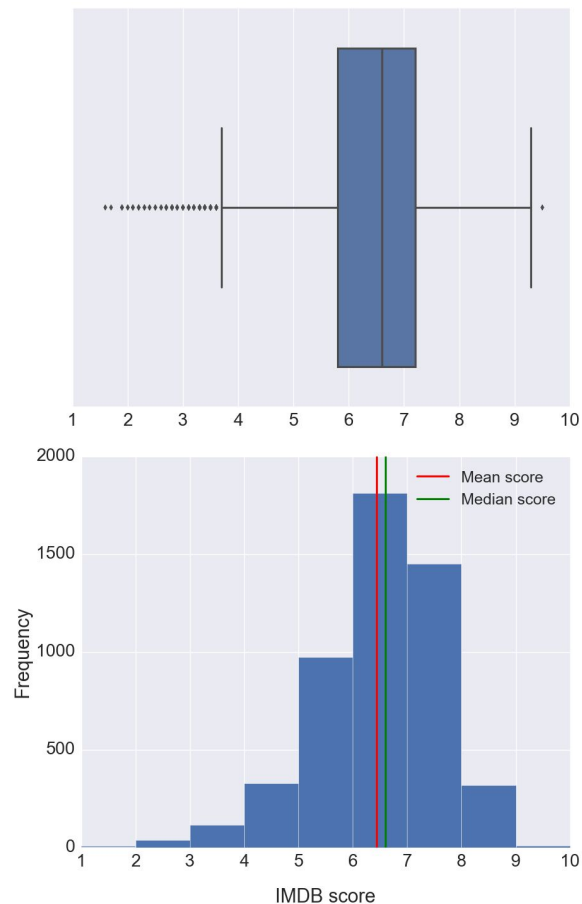
# Movies having the worst score

---



# Task

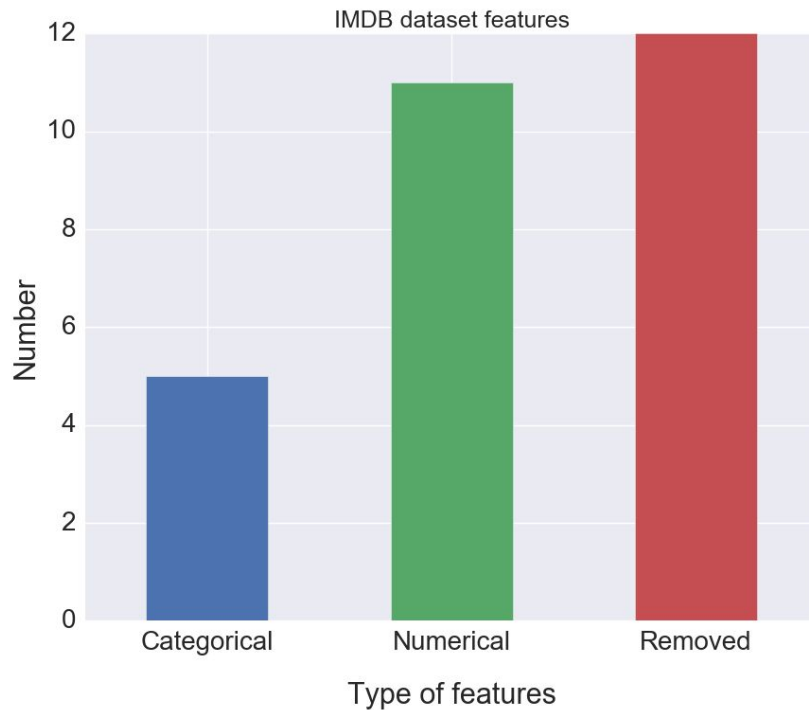
- Predict the IMDB movie score
- Gradient Boosting algorithm (XGBoost package)
- 3 hyperparameters optimization strategies
  - A *naive* grid search
  - An *expert* grid search (\*)
  - The TPE algorithm (hyperopt package)



# Features description

---

- **28 features:**
  - 14 movie-related
  - 4 review-related
  - 10 cast-related
- **16 kept:**
  - 11 numerical
  - 5 categorical
- **12 removed**



# Live demo

---

Our code is available here:

<https://github.com/yassineAlouini/hyperparameters-optimization-talk>



# Conclusion

---

- Outperforms the standard methods in most cases
- Search space matters
- Other Python libraries: Spearmint, BayesOpt, Scikit-Optimize
- Distributed optimization (using MongoDB)

---

Thanks for your attention.  
Question time

Qucit is hiring!

**qu**cit

# References

---

- <https://papers.nips.cc/paper/4443-algorithms-for-hyper-parameter-optimization.pdf>
- [https://conference.scipy.org/proceedings/scipy2013/pdfs/bergstra\\_hyperopt.pdf](https://conference.scipy.org/proceedings/scipy2013/pdfs/bergstra_hyperopt.pdf)
- <https://github.com/scikit-optimize>
- <http://jaberg.github.io/hyperopt/>
- <https://github.com/JasperSnoek/spearmint>
- <https://github.com/fmfn/BayesianOptimization>
- <http://xgboost.readthedocs.io/en/latest/>
- [http://www.cs.ubc.ca/~hutter/papers/13-BayesOpt\\_EmpiricalFoundation.pdf](http://www.cs.ubc.ca/~hutter/papers/13-BayesOpt_EmpiricalFoundation.pdf)