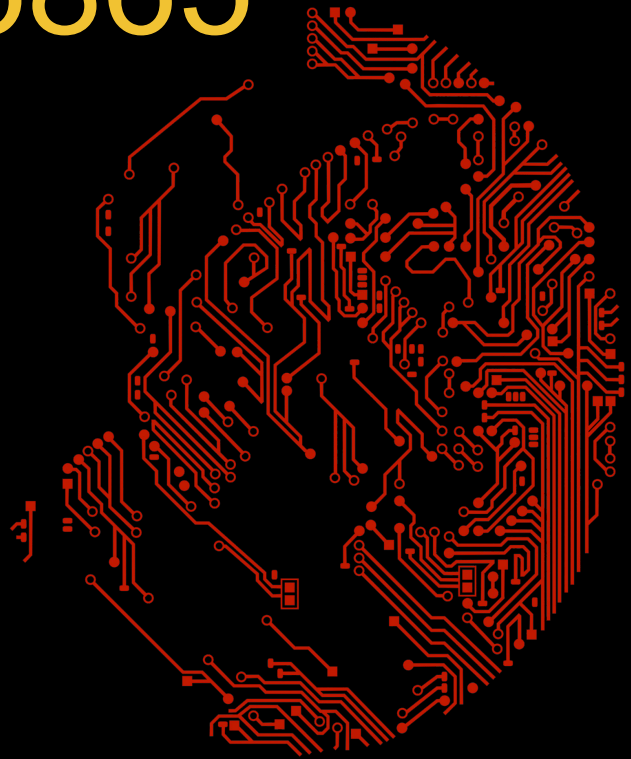
MAC0459/MAC5865

Ciência e Engenharia de Dados



link do Slido https://app.sli.do/event/lesnu1wt

Aula 21

Roberto Hirata Jr
hirata@ime.usp.br

Ensembles of Classifiers

- Who would you trust most, an expert or an ensemble of experts?

Ensembles of Classifiers

- Who would you trust most, an expert or an ensemble of experts?
- Who would you trust most, an expert or an ensemble of not so good experts?

Ensembles of Classifiers

- **Ideia muito antiga**
- **Conhecida como teorema do juri de Condorcet ("Essay sur l'applicacion de l'analyse à la probabilités dés decisions" de 1785).**
- **O marquês de Condorcet se pergunta: quantas pessoas são necessárias num juri para este tomar uma decisão correta?**
- **Se um juri tem que decidir entre duas decisões, uma delas correta, se cada pessoa escolher independentemente a decisão correta com probabilidade p, se p > ½, o adicionando mais membros ao juri, aumentam-se as chances que o juri se decida pela decisão correta.**

Ensembles of Classifiers

- **Advantage:** better accuracy.
- **Disadvantage:** hard to interpret the result (understand how the ensemble of classifiers reached the decision).

Why do they work (Dietterich 2002)?

- ***The Statistical Problem:*** hypothesis space is too large for the amount of available data.
- ***The Computational Problem*****:** learning algorithm can not guarantee finding the best hypothesis.
- ***The Representational Problem:*** hypothesis space does not contain any good approximation of the target classes.

Modified from: Evgueni Smirnov – Ensemble of Classifiers

Independently Constructing Ensembles

- Force a learning algorithm to construct multiple hypotheses
- Run the algorithm several times and provide it with somewhat different data in each run.

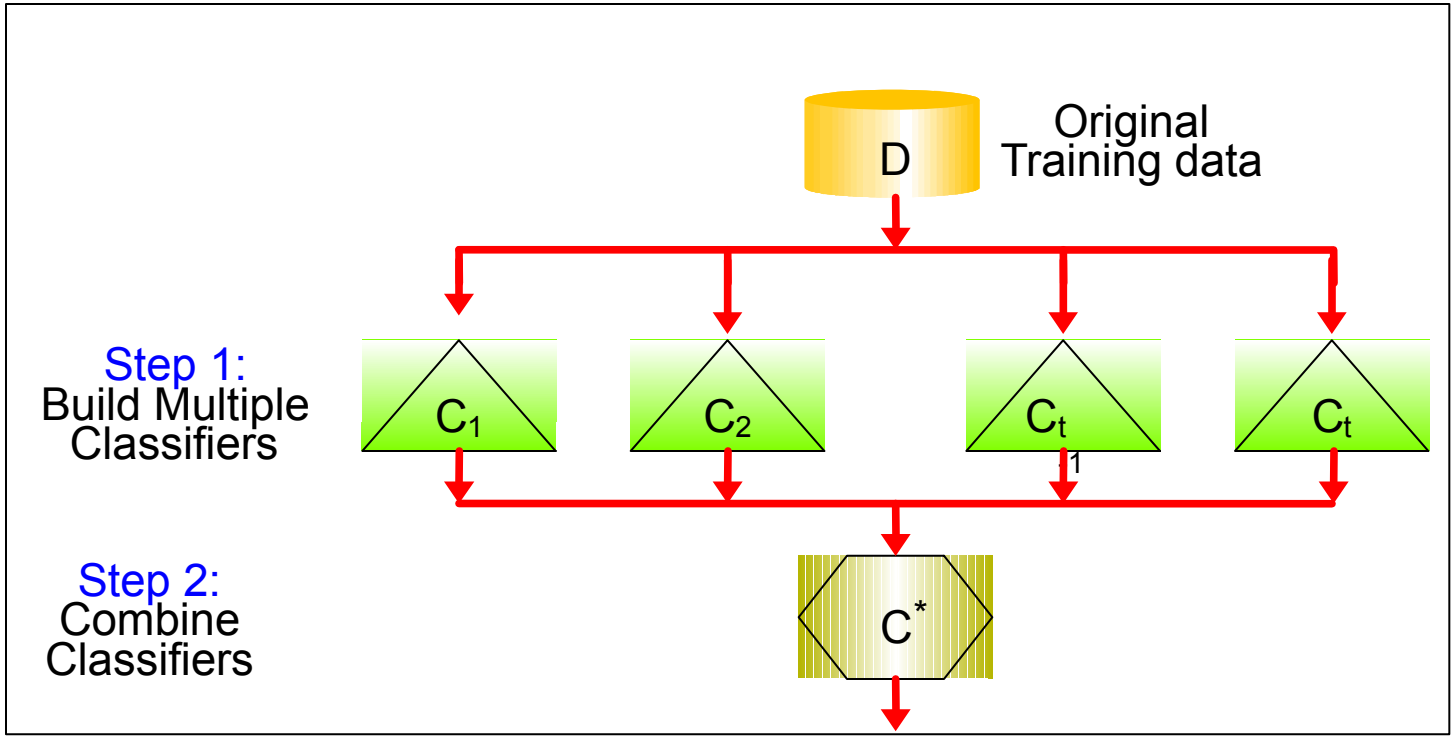
Modified from: Evgueni Smirnov – Ensemble of Classifiers

Independently Constructing Ensembles

- ***Majority Voting***
- ***Bagging***
- ***Randomness Injection***
- ***Feature-Selection Ensembles***
- ***Error-Correcting Output Coding.***

Modified from: Evgueni Smirnov – Ensemble of Classifiers

Majority Vote



Modified from: Evgueni Smirnov – Ensemble of Classifiers

Why Majority Vote works?

- Suppose there are 25 base classifiers
  - Each classifier has error rate 0.35
  - Assume errors made by classifiers are uncorrelated
  - Probability that the ensemble classifier makes a wrong prediction:

$$P(X \geq 13) = \sum_{i=13}^{25} \binom{25}{i} \varepsilon^{i} (1-\varepsilon)^{25-i} = 0.06$$

Modified from: Evgueni Smirnov – Ensemble of Classifiers

Bagging

- O nome "boosting" (incremento) aparece no artigo de Michael Kearns: "Thoughts on Hypothesis Boosting".
- Kearns introduz o conceito de classificadores fracos, ou hipóteses fracas ("weak learners", ou "weak hypothesis")

Bagging

- Classificadores “fracos” têm performance um pouco melhor que um classificador aleatório
- Ele apresenta algumas ideias de como juntar classificadores fracos para criar um classificador melhor.
- Combinação de classificadores ("ensemble learning").

Bagging

- Employs simplest way of combining predictions that belong to the same type.
- Combining can be realized with voting or averaging
- Each model receives equal weight

Modified from: Evgueni Smirnov – Ensemble of Classifiers

Bagging

- "Idealized" version of bagging:
  - Sample several training sets of size *n* (instead of just having one training set of size *n*)
  - Build a classifier for each training set
  - Combine the classifier's predictions
- This improves performance in almost all cases if learning scheme is *unstable* (i.e. decision trees)

Why does bagging work?

- Bagging reduces variance by voting or averaging, thus reducing the overall expected error
  - In the case of classification there are pathological situations where the overall error might increase
  - Usually, the more classifiers the better

Modified from: Evgueni Smirnov – Ensemble of Classifiers

Random Forests

- What if the ensemble is based on different feature sets?

Random Forests

- What if the ensemble is based on different feature sets?
- Who would you trust most, an expert or an ensemble of not so good experts?
- Shimizu’s slides

Gradient Boost Machine

- GBM notebook