CAS ETH Machine Learning in Finance and Insurance. Mini-exercises - Lecture 7.

Andrea Ferrario

Version of May 23, 2024

Ensembling, bagging, random forests and boosting

1. Explain briefly what "wisdom of the crowd" means and how it is currently applied to domains such as finance and insurance.

2. **Averaging i.i.d. random variables reduces variance**: Remember that the variance of a random variable X is the quantity $\text{Var}[X] = \mathbb{E}[(X - \mathbb{E}[X])^2]$. Let $X_1, \dots, X_n$ be independent and identically distributed random variables, with $\text{Var}[X_i] = \sigma^2 < \infty$. Show that

$$\text{Var}[\sum_{i=1}^{n} X_i] = \frac{\sigma^2}{n}.$$

3. **The correlation of i.d. random variables limits the reduction of variance by averaging**: The correlation of two random variables, let us say X and Y, is the quantity $\text{Cor}(X, Y) = \mathbb{E}[(X - \mathbb{E}[X])(Y - \mathbb{E}[Y])]$. Let $X_1, \dots, X_n$ be identically distributed random variables, with $\text{Var}[X_i] = \sigma^2 < \infty$ and $\text{Cor}(X_i, X_j) = \rho \geq 0$, for all $i \neq j$. Show that

$$\text{Var}[\sum_{i=1}^{n} X_i] = \rho\sigma^2 + \frac{1-\rho}{n}\sigma^2.$$

4. Access the article Breiman, L. (1996). Bagging predictors. Machine learning, 24, 123-140. Describe the machine learning procedure used by the author to arrive at the results shown in table 2.

5. Explain the main differences between bagging and random forests.

6. Explain briefly what boosting is and the main difference between random forests and boosting.

7. Explain the role of weak learners in the AdaBoost algorithm. Why are decision tree stumps often used as weak learners, and how does AdaBoost transform them into a strong learner?

8. Describe the weight update formula in AdaBoost in terms of correctly classified and misclassified data points.

9. Explain the logic behind the gradient boosting machine algorithm and the steps necessary to fit a tree to residuals and update the boosting machine output. You can access this article for all details: Jerome H. Friedman. "Greedy function approximation: A gradient boosting machine.." Ann. Statist. 29 (5) 1189-1232, October 2001. Consider, in particular, "Algorithm 1" at pag. 1193.