ANALYSIS OF ENSEMBLE METHODS OF MACHINE LEARNING ALGORITHMS

Pulatov Giyos
PhD student, "TIIAME" National Research University

https://doi.org/10.5281/zenodo.18888990

ARTICLE INFO

Received: 01st March 2026
Accepted: 04th March 2026
Online: 06th March 2026

KEYWORDS

Ensemble learning is a machine learning approach aimed at achieving more accurate and stable results compared to individual models by combining several base models

ABSTRACT

This section examines ensemble learning methods and their importance in improving prediction accuracy and stability over single machine learning models. It focuses on the main ensemble approaches—bagging, boosting, and stacking—and explains their working principles and structural differences. The analysis also highlights homogeneous and heterogeneous ensembles, sequential and parallel strategies, and the growing role of ensemble deep learning in intelligent prediction systems.

Ensemble learning is a machine learning approach aimed at achieving more accurate and stable results compared to individual models by combining several base models.

Bagging is an abbreviated term formed from the combination of *bootstrapping* and *aggregating*. In bootstrapping, ensemble models are trained on bootstrap copies of the training data. In aggregation, the final result is obtained by combining the predictions of the models through majority voting.

The bagging function is expressed by Formula 1:

$$f(x) = \frac{1}{B}\sum_{b=1}^{B} f_b(x) \qquad (1)$$

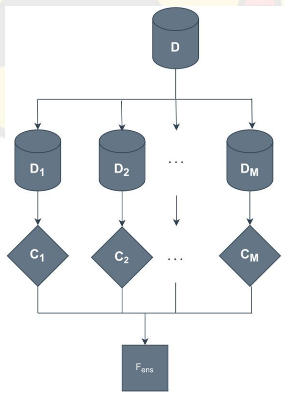

Figure 1. Block diagram of the bagging method.

In boosting, several weak learners are built sequentially and in a highly adaptive manner. When each model in the sequence is fitted, greater importance is given to the observations that previous models failed to handle well.

The boosting function is expressed by Formula 2:

$$f(x) = \sum_{t} \alpha_t h_t(x) \qquad (2)$$

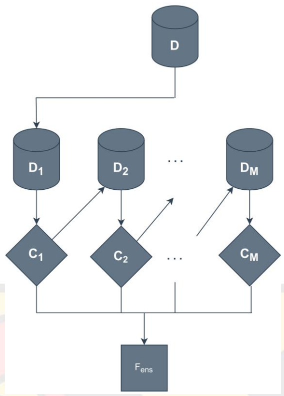

Figure 2. Block diagram of the boosting method.

The stacking method is an ensemble approach that combines information obtained from several predictive models to create a new model, namely a meta-model.

The architecture of a stacking model usually consists of two or more base models. The outputs of the base models are fed as inputs to the meta-model. In a classification problem, these may be probability values or class labels.

The stacking function is expressed by Formula 3:

$$f_s(x) = \sum_{i=1}^{n} a_i f_i(x) \qquad (3)$$

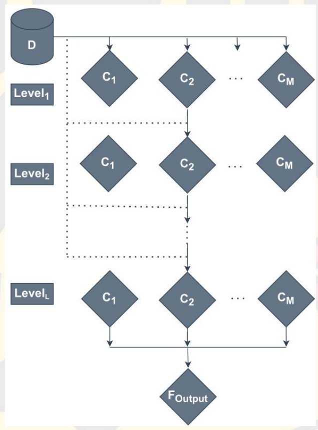

Figure 3. Block diagram of the stacking method.

Then, based on this ensemble method, the output prediction is given by Formula 4:

$$y_i = \varphi(x_i) = G(c_1, c_2, ..., c_k) \qquad (4)$$

The general abstract scheme of ensemble learning is shown in Figure 4. Any ensemble consists of a set of base classifiers trained on the input data, which produce predictions, and these predictions are then combined to obtain an overall prediction.

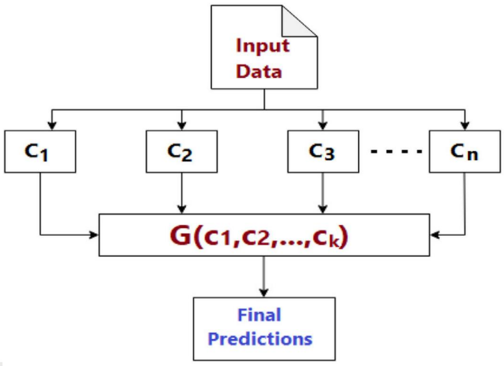

Figure 4. General structure of the ensemble model.

Ensemble strategies differ depending on how the base classifiers are selected and trained. There are two types of diversity in terms of the nature of base classifiers: **homogeneous** and **heterogeneous** ensembles (Figure 5).

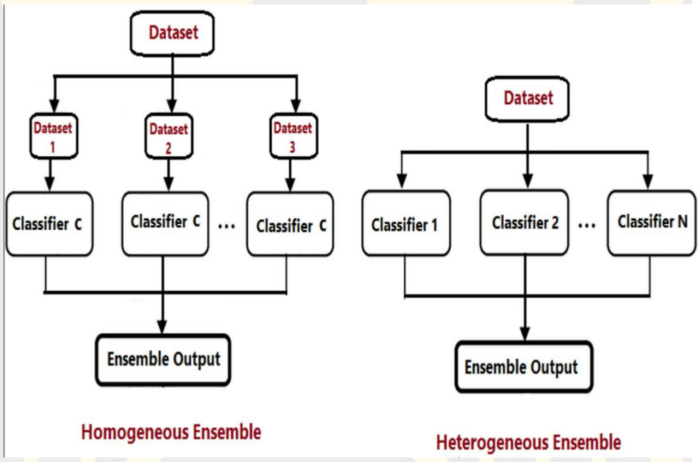

Figure 5. General structure of homogeneous and heterogeneous ensembles.

A homogeneous ensemble consists of base classifiers of the same type, but each classifier is built on different data. The main difficulty in the homogeneous approach is generating diversity while using a single learning algorithm.

A heterogeneous ensemble consists of different types of base classifiers, and each classifier is built on the same data. In heterogeneous classifiers, the feature selection method may differ even for the same training data.

In a sequential ensemble method, different learners are trained one after another, since there is dependence in terms of data. Therefore, the errors made by the first model are corrected sequentially by the second model (Figure 6).

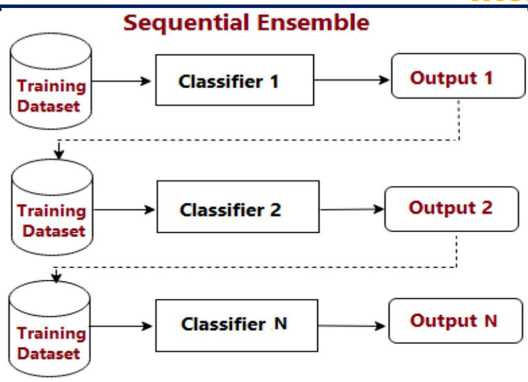

Figure 6. General structure of a sequential ensemble.

In a parallel ensemble method, however, the base learners are created simultaneously because there is no data dependency. Therefore, the data for each base learner are formed independently (Figure 7).

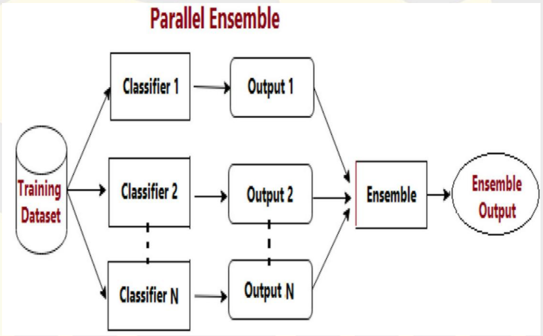

Figure 7. General structure of a parallel ensemble.

Although ensemble deep learning systems are powerful in improving prediction quality, most of the literature mainly focuses on using the majority voting method because of its simplicity.

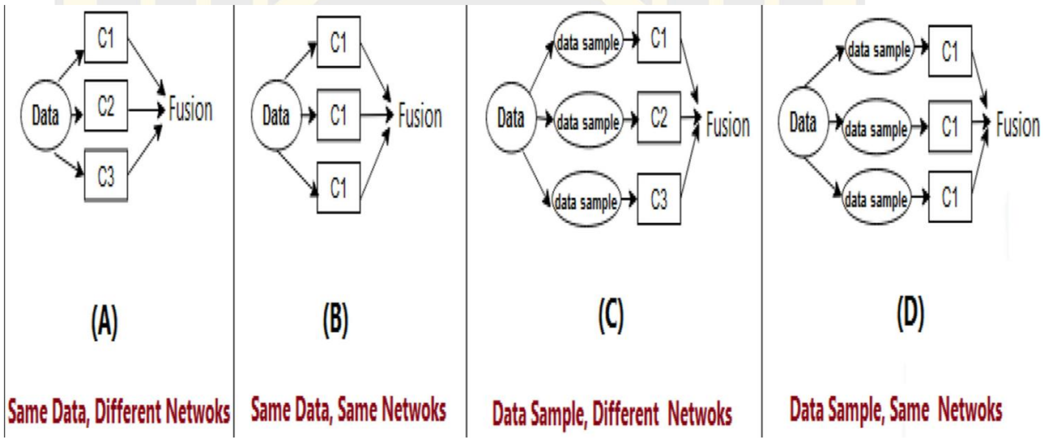

Figure 8. Different cases of ensemble deep learning.

A comparison of these strategies shows that strategies A and C are suitable for both deep learning models and traditional learning methods. Strategies B and D, however, are specific only to deep learning models and cannot be applied in traditional learning methods; this makes ensemble deep learning strategies even more diverse. In addition, strategies B and D make it possible to reduce the hyperparameters of the base deep models in the ensemble through different structures of the same base model.

References:

1.Ammar Mohammed, Rania Kora. “A comprehensive review on ensemble deep learning: Opportunities and challenges”. Journal of King Saud University - Computer and Information Sciences 35 (2023) 757–774. https://doi.org/10.1016/j.jksuci.2023.01.014

2.M.A. Ganaiea, Minghui Hub. “Ensemble deep learning: A review”. Preprint submitted to Elsevier August 9, 2022.

3.Pulatov Giyos, Pulatova Gulkhayo. “Ensemble-based decision making for forecasting”. International conference on engineering & technology. https://doi.org/10.5281/zenodo.18872437