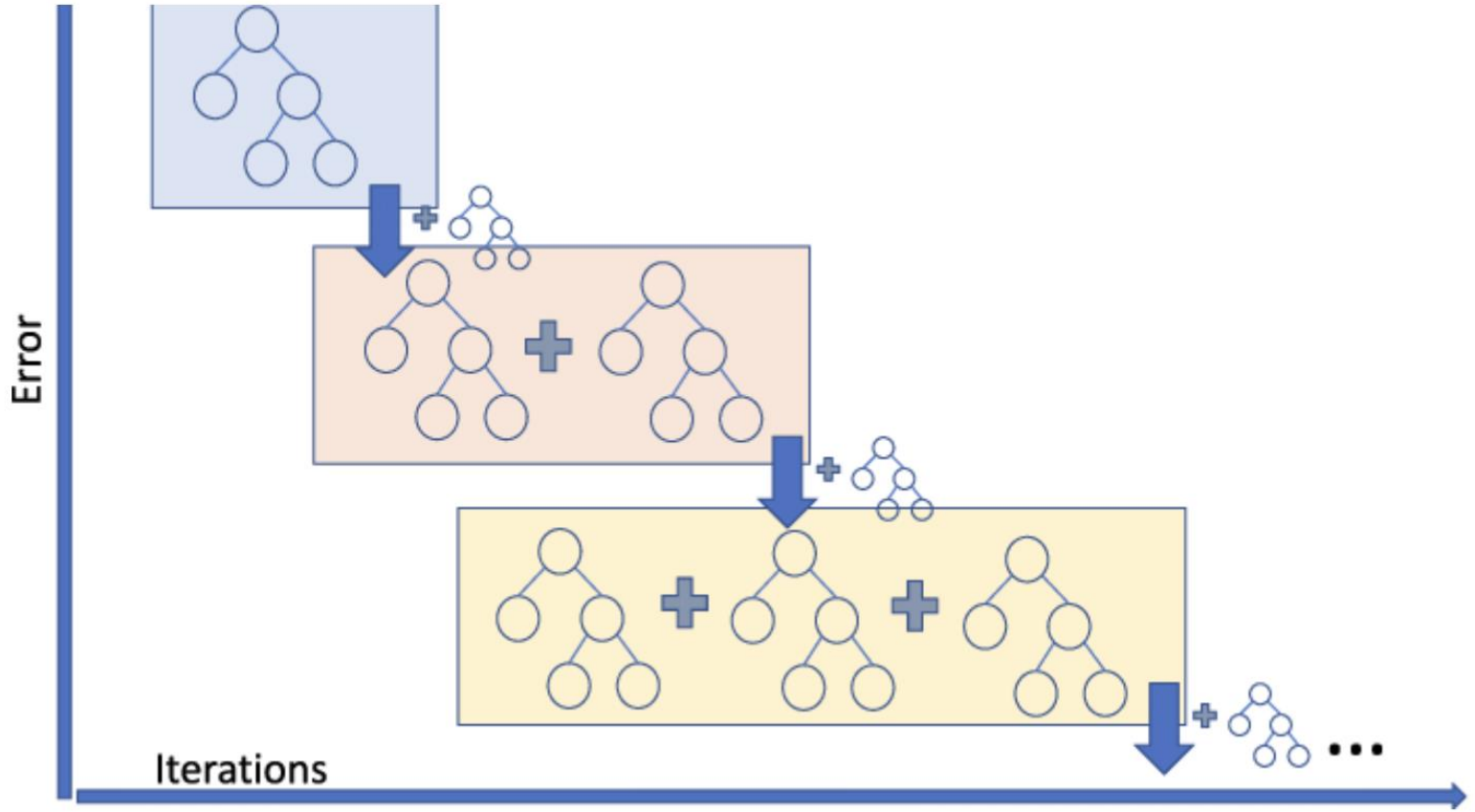
LG boost – Regression

It is an ensemble learning framework that uses gradient boosting method which constructs a strong learner by sequentially adding weak learners in a gradient descent manner.

LG boost Regression diagram



- **Initial Model**: Start with a single decision tree (top left) that predicts the target variable with some error.
- **Residual Calculation**: Identify the errors (residuals) from the initial prediction.
- **Add New Tree**: Build a new decision tree to predict the residuals, and add it to the previous model (middle section).
- **Iterate**: Repeat the process—calculate new residuals from the updated model and add more trees (bottom section).
- **Combine Models**: Combine all trees to improve the overall prediction, reducing error with each iteration.

The process continues over multiple iterations, gradually minimizing the error.

The formula for Gradient Boosting (e.g., Gradient Boosted Regression Trees) can be expressed as:

$$F_m(x) = F_{m-1}(x) + \eta \cdot h_m(x)$$

Where:

- $F_m(x)$: The model after m iterations (final prediction).
- $F_{m-1}(x)$: The model from the previous iteration.
- η: The learning rate (a small positive number that controls the contribution of each tree).
- $h_m(x)$: The new decision tree (weak learner) trained to predict the residuals (errors) of $F_{m-1}(x)$.