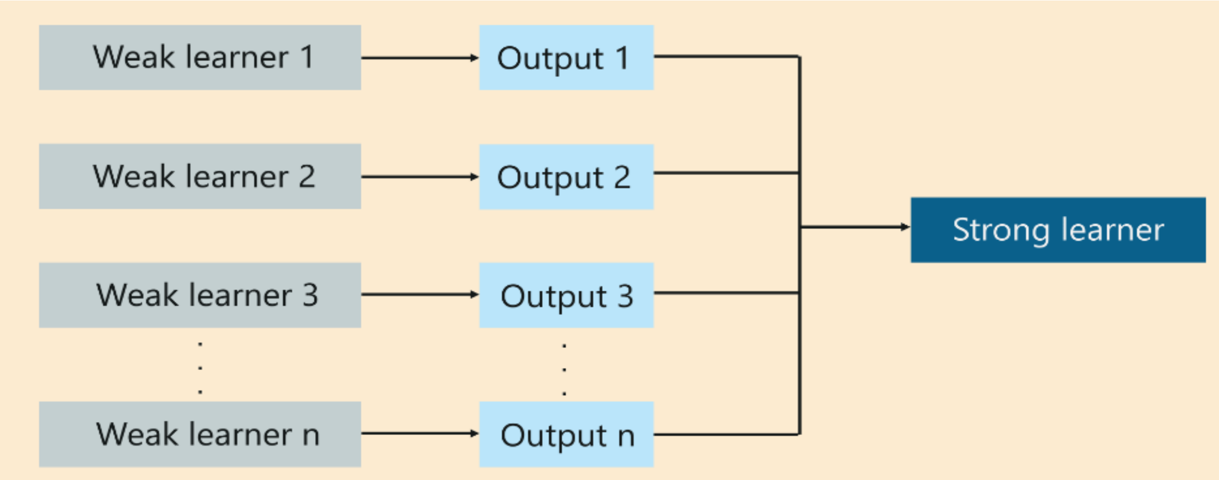
# Boosting

It is an ensemble modeling technique that attempts to build a strong classifier from the number of weak classifiers.



Weak learner 1

Weak learner 2

Weak learner 3

Weak learner n

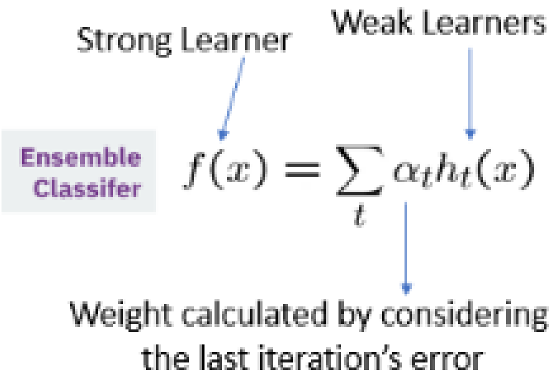
Output 1

Output 2

Output 3

Output n

Strong learner



Strong Learner

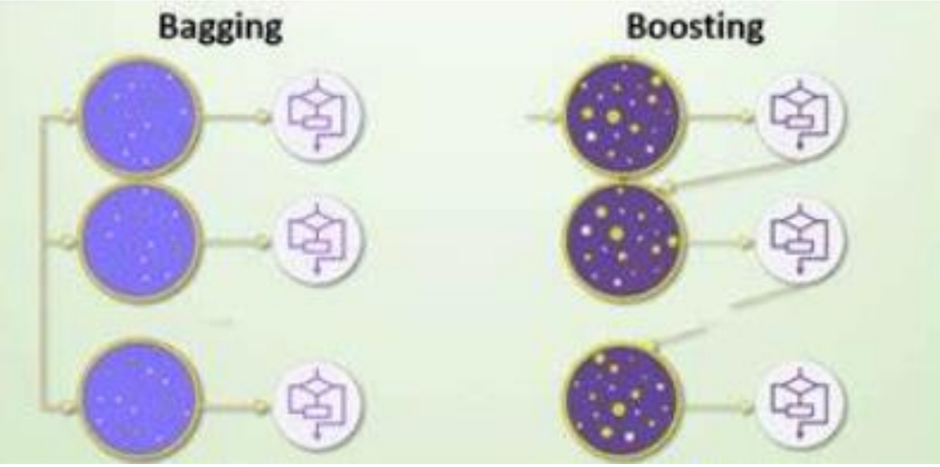
Weak Learners

Ensemble Classifer

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

Weight calculated by considering the last iteration's error

# Boosting Vs Bagging:



Bagging:
1.Resampling
2.Uniform distribution
3.Parallel style

Boosting:
1.Reweighting
2.Non-uniform distribution
3.Sequential style

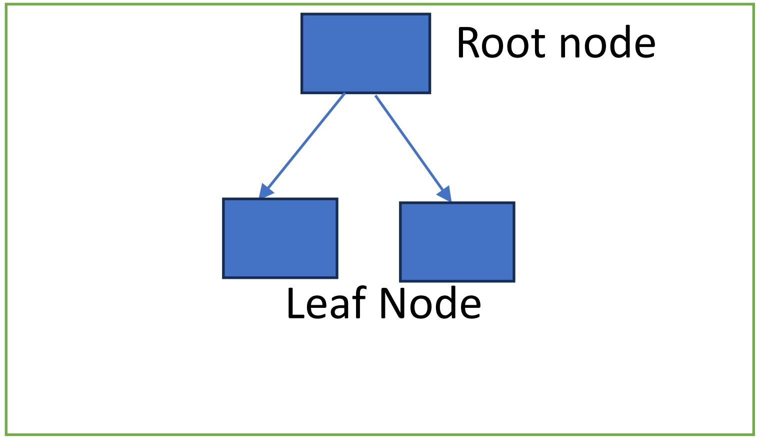
# Types of Boosting:

- Ada Boost
- Gradient Boost
- XG Boost
- Cat Boost
- Light GBM/ LG Boost

# Ada Boost:

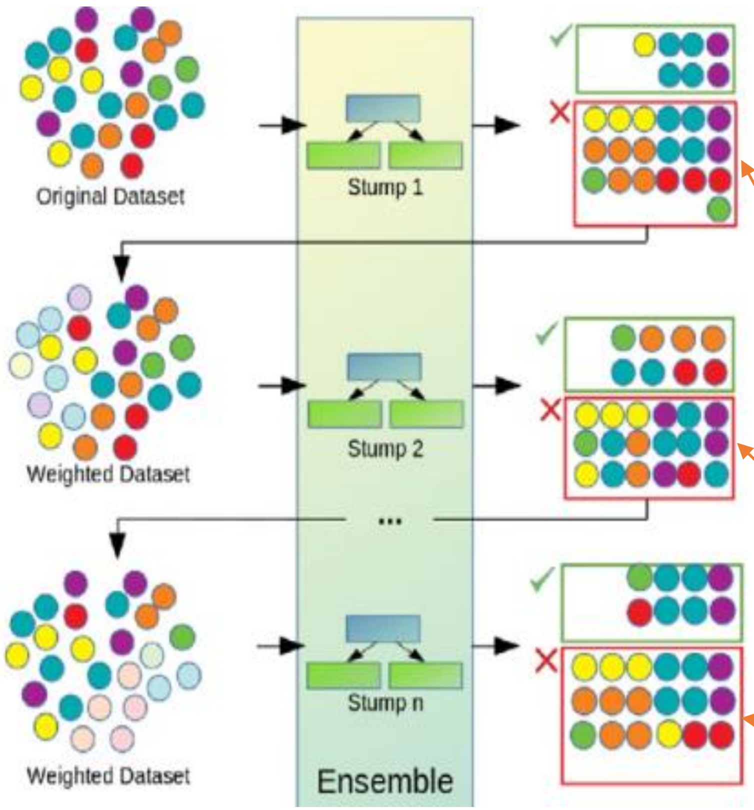
- Ada Boost or Adaptive Boosting
- It works same as normal Boosting algorithm.
- Transforms weak learners to strong learners
- Reassign weight to each instance, higher weight for incorrectly classified. This reduces bias and variance.
- Used for both classification and regression problems.

Stump



No fixed depth, AdaBoost takes only stumps

# 3 ideas behind Ada Boost:



Stumps can use only one variable to make a decision so it is a **Weak Learner**.

3 ideas of Ada Boost:

1. **Ada Boost** combines lot of weak learners to make classification. Weak learners are always **stumps**.
2. Some **stumps** get more say in classification than others.
3. Each **stump** is made by taking the previous **stump's** mistakes into account.

Misclassified data / Incorrectly Classified data

# Advantages of Ada Boost:

- Improved Accuracy
- Easy Implementation
- Handles complex data
- Few Tunable parameters

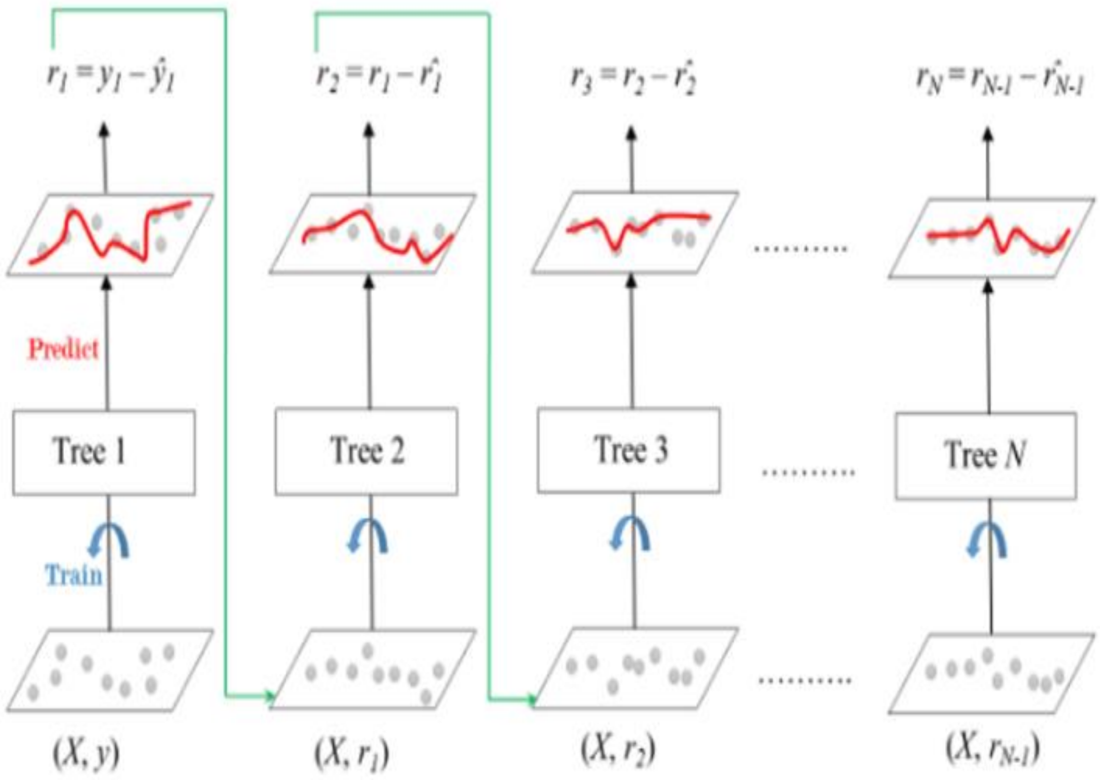
# Disadvantages of Ada Boost:

- Sensitivity to Noisy data
- Overfitting Risk
- Dependency on Weak classifiers
- Computationally expensive

# Gradient Boosting

- This algorithm combines several weak learners into strong learners.
- This model improves sequentially but not incremental weight concept.
- Optimize the loss function from previous learners.
- It is an additive model which regularize the loss function.

# Process of Gradient Boosting:



This model works sequentially.
From previous tree it takes the output and optimizes the value using loss function

# Advantages of Gradient Boosting:

1. **High predictive accuracy**: Often provides accuracy that cannot be beaten.
2. **Works well with heterogeneous features**: Handles both numerical and categorical features.
3. **Handles missing data impressively**: Imputation is unnecessary.
4. **No data pre-processing required**: Category and numeric values perform well as-is.

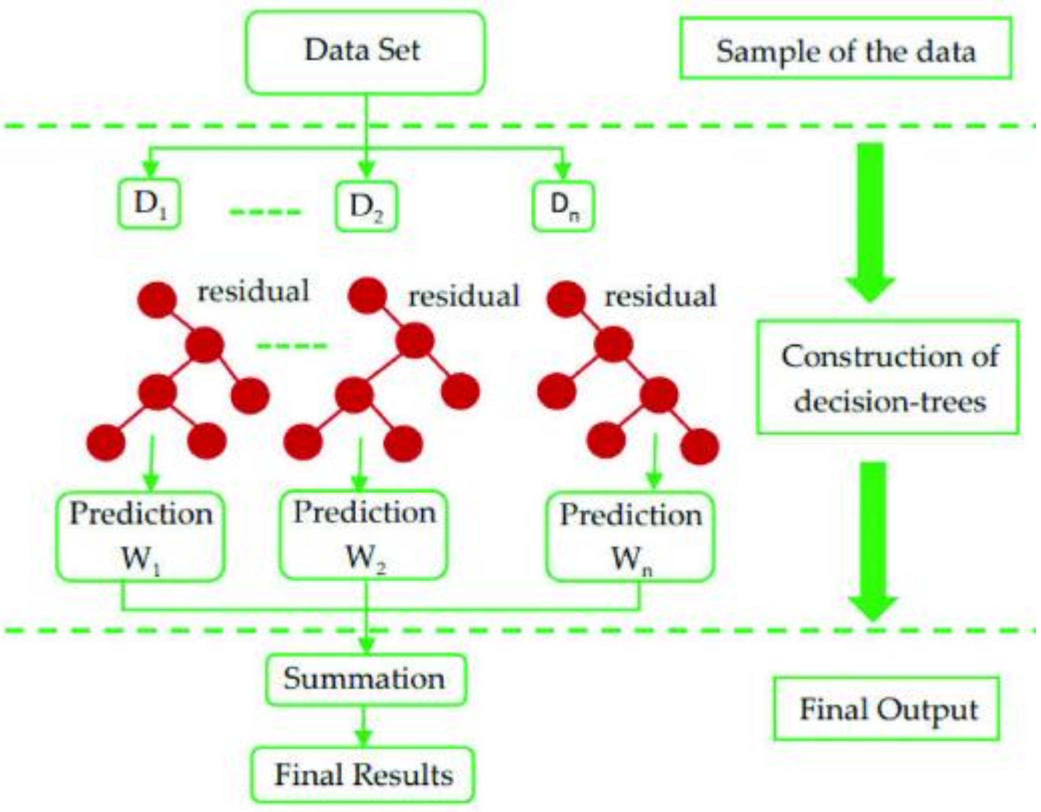
# Disadvantages of Gradient Boosting:

- It is **difficult to scale this algorithm** as every estimator is dependent on its predecessor. It is somewhat slower to train compared to algorithms like bagging.

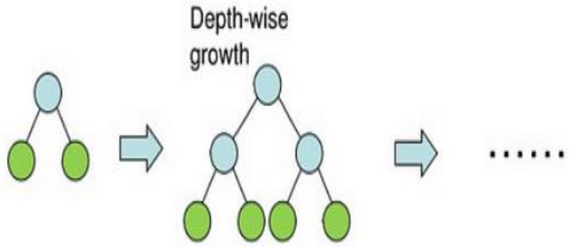
# XG Boost:

- It is an extreme Gradient Boosting.
- It is same as gradient boosting but has additional functionality.
- Distributed machine learning process.
- It is more computational speed and model efficiency.
- Gradient boosting is sequentially slow when compared to XG boosting.

# Process of XG Boost:



XGBoost architecture



This single dataset is distributed to multiple dataset and does the function. So the function is done in more efficiently and speed

# Advantages of XG Boost:

1. **Speed and scalability**: XGBoost is known for its speed and scalability, making it a good choice for large datasets.
2. **Performance**: XGBoost has been shown to outperform other machine learning algorithms in many benchmark tests.
3. **Flexibility**: XGBoost can be used with a variety of data types and can handle missing data.
4. **Interpretability**: XGBoost provides feature importance scores, which can help with model interpretation.
5. **Regularization**: XGBoost applies regularization to prevent overfitting and improve generalization.

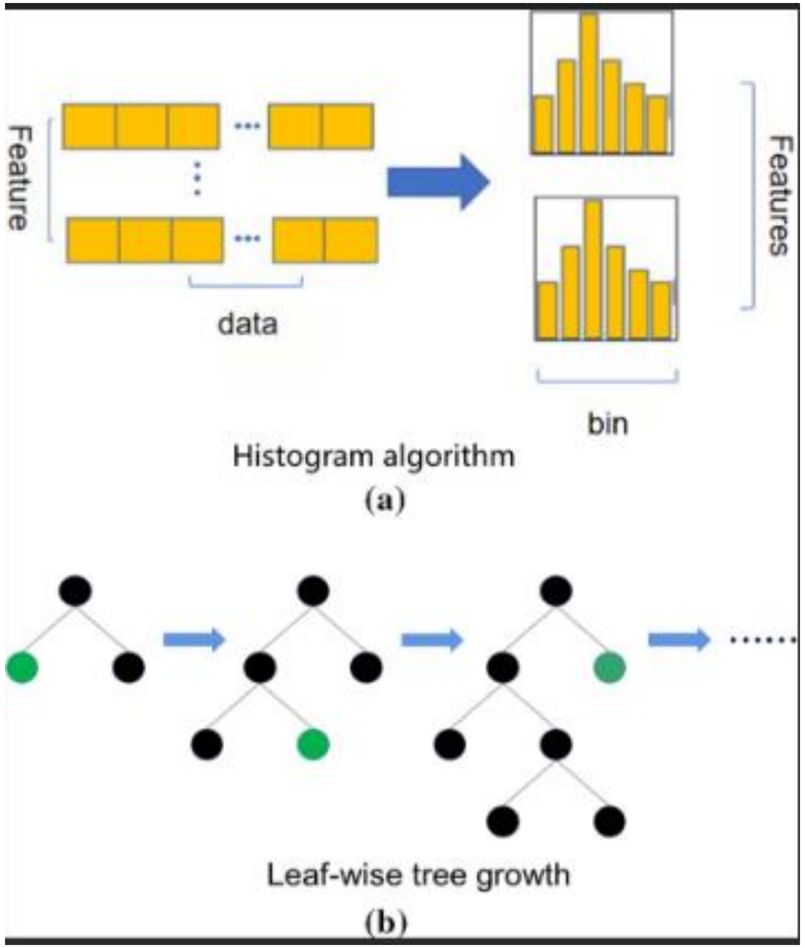
# Disadvantages of XG Boost:

1. **Complexity**: XGBoost is a complex algorithm that requires some understanding of its underlying mechanics.
2. **Computational resources**: XGBoost can require significant computational resources, especially when using large datasets or many iterations.
3. **Hyperparameter tuning**: XGBoost has several hyperparameters that need to be tuned for optimal performance.

# Light GBM/LG Boosting/Light Gradient Boosting Machine

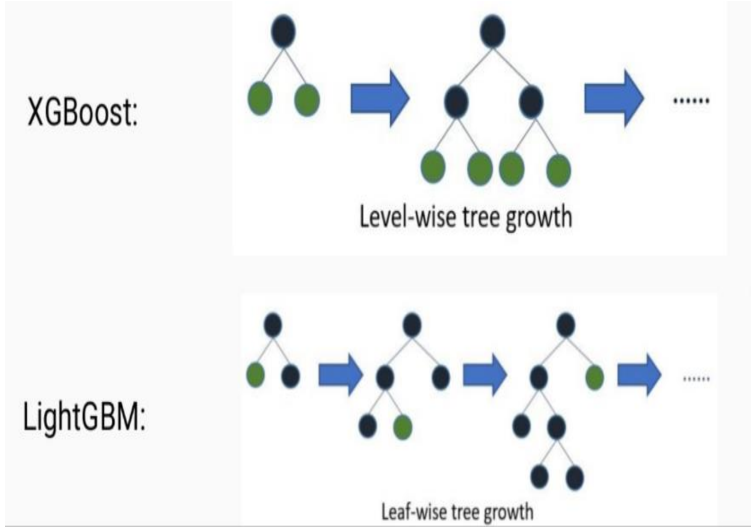
- It can handle large amount of data easily.
- Poor in handling small data.
- It uses histogram based method for selecting best fit
- For continuous values splits up into bins or buckets

# Process Of LGBM:



Histogram algorithm
(a)

Leaf-wise tree growth
(b)



Here XG Boost checks level wise and expands the tree but LGBM checks Leaf wise and expand that leaf

This model converts the data into histogram method to find the best fit

# Advantages of LGBM:

- **Efficiency and Speed**: Light GBM is known for its fast training speed and efficient performance. It can handle large volumes of data with ease.
- **High Accuracy**: Light GBM provides accurate predictions, making it suitable for various tasks.
- **Flexibility**: It works well for both classification and regression problems.
- **Robust to Missing Data**: Light GBM has built-in capabilities for handling missing values

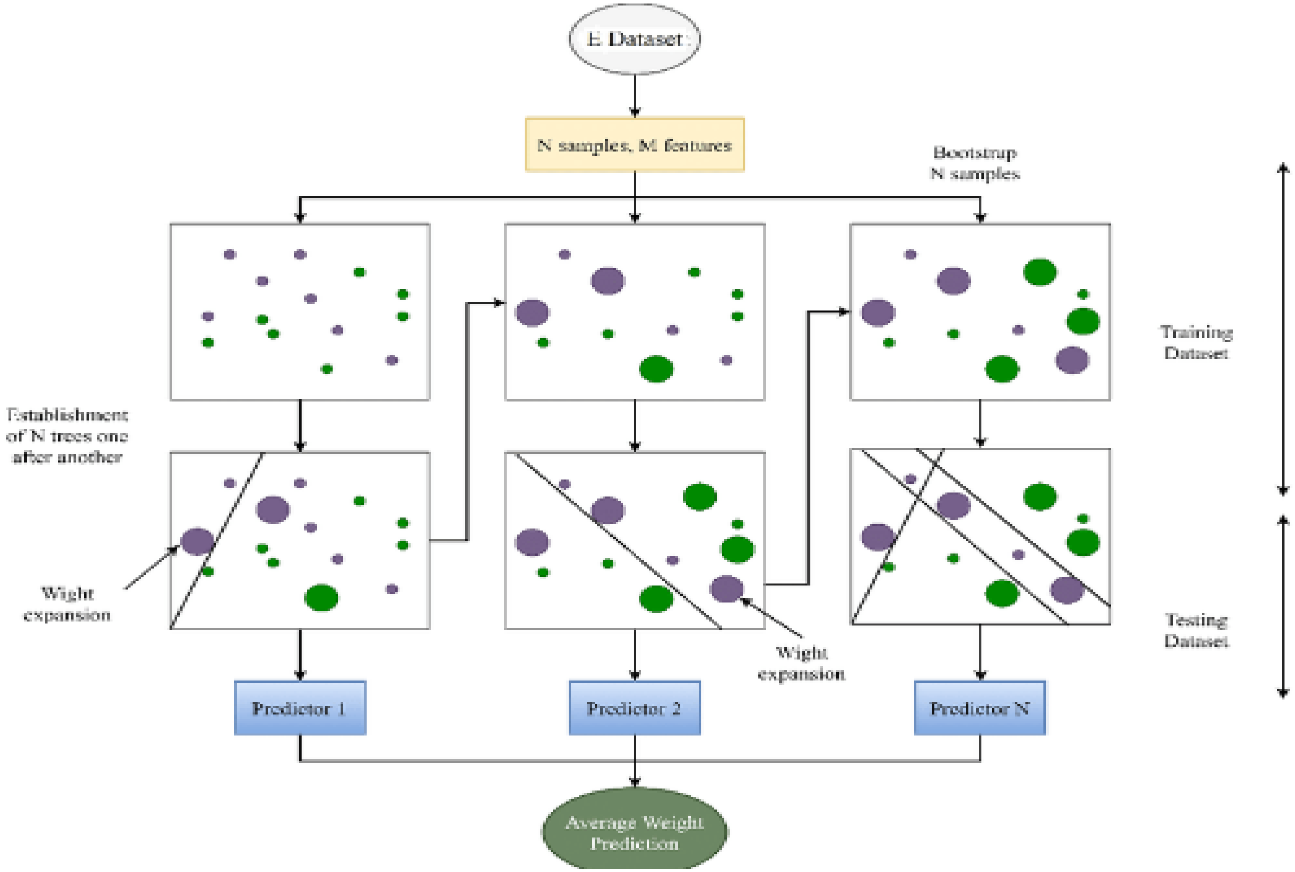
# Disadvantages of LGBM:

- **Computational Intensity**: As the model complexity grows, the computational cost increases. This can be a limitation when dealing with very large datasets.
- **Not Suitable for Small Data Points**: Light GBM may not perform well when the number of data points is small.

# Cat Boost:

- Cat Boost or Categorical Boosting
- It is designed for solving a wide range of machine learning tasks
- Cat boost stands out for its speed, accuracy, and ease of use in dealing with structured data.
- Most machine learning algorithm fails categorical data.
- Cat boost is good with categorical data.
- Various statistics and numerical operations.

# Process of Catboost:



To prevent overfitting, regularization strategies are also included. When generating predictions, Catboost combines the forecasts from every tree, producing incredibly dependable and precise models

# Advantages of Cat Boost:

- Performance
- Handling Categorical features automatically
- Robust
- Easy-to-use

# Disadvantages of Cat Boost:

- The model sizes are quite big which makes it difficult while deploying in a production environment.
- It is a memory intensive algorithm.