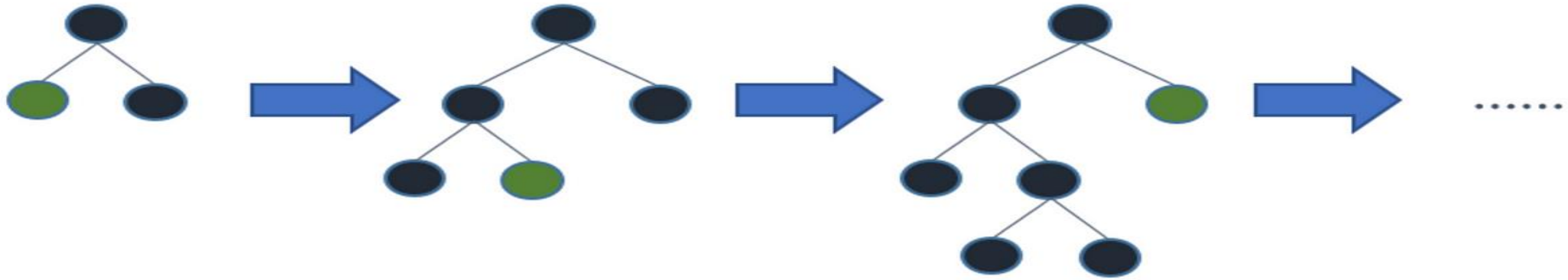
LG BOOSTING

LightGBM uses a novel technique of Gradient-based One-Side Sampling (GOSS) to filter out the data instances

LightGBM will by default consider model as a regression model. learning_rate: This determines the impact of each tree on the final outcome.

GBM works by starting with an initial estimate which is updated using the output of each tree.

The learning parameter controls the magnitude of this change in the estimates.



Leaf-wise tree growth

How it differs from other boosting algorithms?

- **Light GBM** splits the tree **leaf-wise** with the **best** fit whereas other boosting algorithms split the tree depth-wise or level-wise rather than leaf-wise. In other words, Light GBM grows trees vertically while other algorithms grow trees horizontally.
- **Advantages of Light GBM**
- **Faster training speed and higher efficiency:** Light GBM uses a histogram-based algorithm i.e it buckets continuous feature values into discrete bins which fasten the training procedure.
- **Lower memory usage:** Replaces continuous values to discrete bins which results in lower memory usage.
- **Better accuracy than any other boosting algorithm:** It produces much more complex trees by following leaf wise split approach rather than a level-wise approach which is the main factor in achieving higher accuracy.
- **Compatibility with Large Datasets:** It is capable of performing equally well with large datasets with a significant reduction in training time as compared to XGBoost.

Disadvandages of LG boosting

Disadvantages of Light GBM

- **Overfitting:** Light GBM split the tree leaf-wise which can lead to overfitting as it produces much complex trees.
- **Compatibility with Datasets:** Light GBM is sensitive to overfitting and thus can easily overfit small data.