

# DA Drive Analyzer Upgrades Predictive Mechanism with LightGBM for Enhanced Drive Failure Forecasting

*ULINK DA Drive Analyzer's upgrade will soon make predictions one day after receipt of the data.*

SANTA CLARA, CALIFORNIA, UNITED STATES, September 13, 2023 /EINPresswire.com/ -- [DA Drive Analyzer](#)'s drive failure prediction mechanism is about to get an upgrade. Predictions will soon be made the day following receipt of the data. The new AI will support drive failure predictions for SATA, NVMe, and SCSI drives. It will also consume a host of new predictors without sacrificing speed. The basic algorithm underlying this upgrade is called Light Gradient-Boosting Machine, or LightGBM.

How does LightGBM work?

LightGBM is a machine learning framework commonly used for ranking, and classification. Its advantages like high speed, accuracy, low memory usage, and ability to handle large-scale data make it ideal for use in [ULINK](#) DA Drive Analyzer.

LightGBM is an open-source distributed gradient-boosting framework based on decision tree algorithms. Before creating the trees, LightGBM creates histograms with the data, which groups values into bins to yield great advantages on both efficiency and memory consumption. The histograms help the model to identify optimal split points for each leaf. It then grows trees leaf-wise, choosing the leaf it believes will yield the largest decrease in loss. And when the trees are grown, the prediction of each tree is weighted and then summed up to get a final prediction.

The LightGBM algorithm also utilizes two interesting techniques called Gradient-Based One-Side Sampling (GOSS) which helps when training with class-imbalanced data (which is the case for drive failures) and Exclusive Feature Bundling (EFB) which helps the model to train with less memory.

Gradient-Based One-Side Sampling (GOSS)

GOSS is a method that calculates the loss gradient for each data record and then orders the records by the loss gradients. The records with the largest loss gradients are all retained. On the other hand, the records with smaller loss gradients are randomly downsampled. This helps the model to learn more from the records that it has previously performed poorly on. This leads to

better learning of the imbalanced data.

## Exclusive Feature Bundling (EFB)

EFB is a smart way to reduce the number of features (i.e., predictors) that the model has to consider. The way this works is if the model notices that two or more features are mutually exclusive, such as a column for whether a drive is an HDD or SSD, and another column for whether the drive disk rotation speed is high or low, it can summarize the information in both the columns into a single column. This feature reduction process makes it faster for the model to work with a large number of features.

## Key Benefits of LightGBM

The LightGBM algorithm boasts a range of significant advantages that make it a popular choice for various machine learning tasks. Here's a breakdown of its key benefits:

**Higher Speed and Accuracy:** LightGBM is designed with efficiency in mind. Its unique leaf-wise tree growth strategy significantly accelerates the training process. This swift approach to building decision trees allows LightGBM to process data faster than many other gradient boosting algorithms. Despite its speed, LightGBM maintains a high level of accuracy, ensuring reliable and precise predictions.

**Lower Memory Usage:** LightGBM's histogram-based learning approach enables it to convert continuous feature values into discrete bins. This methodology reduces memory consumption during training and inference compared to other algorithms. The smaller memory footprint makes it more accessible for deployment on resource-constrained devices or in environments with limited memory capacity.

**Parallel and Distributed GPU Learning:** LightGBM supports parallel computing and GPU acceleration, enabling the algorithm to leverage the power of multiple processors or graphics processing units. This parallelization accelerates both training and prediction, making LightGBM suitable for handling large datasets and complex models.

**Handling Large-Scale Data:** LightGBM is particularly adept at handling large-scale datasets. Its leaf-wise tree growth strategy and efficient histogram-based approach ensure that the algorithm scales well to substantial volumes of data without sacrificing performance or accuracy. This scalability is crucial for applications dealing with massive datasets.

We hope you can [try out our new model soon](#).

ULINK Technology  
DA Drive Analyzer  
[email us here](#)

Visit us on social media:

[Facebook](#)

[Twitter](#)

[LinkedIn](#)

---

This press release can be viewed online at: <https://www.einpresswire.com/article/655514483>

EIN Presswire's priority is source transparency. We do not allow opaque clients, and our editors try to be careful about weeding out false and misleading content. As a user, if you see something we have missed, please do bring it to our attention. Your help is welcome. EIN Presswire, Everyone's Internet News Presswire™, tries to define some of the boundaries that are reasonable in today's world. Please see our Editorial Guidelines for more information.

© 1995-2023 Newsmatics Inc. All Right Reserved.