

ULINK TECHNOLOGY

3120 De La Cruz Blvd
Suite 117
Santa Clara, CA 95054
408.446.8455

Technology Brief: Machine Learning with ULINK Drive Analyzer

May 20, 2019

What is ULINK Drive Analyzer?

ULINK Drive Analyzer is a software service developed by ULINK Technology that will monitor various storage drive health indicators and utilize a machine learning model (a type of artificial intelligence) to notify the user when their drive health is at risk of failure.

The model will utilize a deep neural network, which involves creating multiple layers of drive health predictors, with each predictor in each upper layer being derived from a pattern of predictors in the layer below it (the bottom layer consists of the original health indicators). Initially, this model will be trained from a 500+ gigabyte pool of drive health indicators coming from various drives. This will give our initial model the proper weights on each predictor to successfully predict drive failure. The weighting process is called prediction and the observed data. Over time, drives to ensure that it predicts our client's drives' performance with a high degree of accuracy. During the development process, we will also systematically tune non-data-driven hyperparameters (such as learning rates) in order to optimize the model's prediction accuracy.

Advantages of a Machine Learning Approach

Previously, the health of storage drives was determined by reading drive health indicators called Self-Monitoring Analysis and Reporting Technology (SMART) attributes off of drives. SMART attributes serve as sometimes-critical indicators of drive health. However, judging whether or not a drive will fail based on individual SMART attributes doesn't offer much predictive accuracy -even the most predictive SMART attribute shows up only 44.8% of the time when a drive fails¹. Furthermore, SMART variables alone do not offer very specific timing information about drive failure. For example, there



is work looking at the correlation between SMART variables and *annual failure rates*².

A machine learning approach offers the potential to be vastly more accurate as well as more specific about the time of failure. For example, recent attempts to use machine learning to predict optical network equipment failure have achieved accuracies of 95% or greater^{3 4}. Similarly, machine learning attempts to predict pump equipment failures have achieved greater-than 95% accuracy⁵. These failures were predicted on the order of *days*.

¹ Backblaze has published articles detailing the failure rates of several models of storage drives. For example, see <https://www.backblaze.com/blog/what-smart-stats-indicate-hard-drive-failures/>

² and <https://www.backblaze.com/blog/2018-hard-drive-failure-rates/>.

³ Wang, Z., Zhang, M., Wang, D., Song, C., Liu, M., Li, J., ... & Liu, Z. (2017). Failure prediction using machine learning and time series in optical network. *Optics express*, 25(16), 18553-18565.

⁴ Shahkarami, S., Musumeci, F., Cugini, F., & Tornatore, M. (2018, March). Machine-learning-based soft-failure detection and identification in optical networks. In *2018 Optical Fiber Communications Conference and Exposition (OFC)* (pp. 1-3). IEEE.

⁵ https://valiancesolutions.com/wp-content/uploads/2018/03/IEEE_Equipment-Reliability.pdf

Storage Drive Failure

As of 2018, individual drives from some companies still have an annual failure rate of greater than 2%. If you are a business or consumer whose data is stored on one of these drives, then you run the risk of losing valuable work. If you rely on policies based on individual SMART predictors to tell you when to replace storage drives, you may end up either replacing the drives too early or too late. Replacing too early is costly from a financial perspective, and replacing too late is costly from a productivity perspective. Timely and predictive replacement is ideal. Machine learning is a hot new tool that can enable this.

The ULINK Advantage

At ULINK, we are well-positioned to develop the machine learning model needed to predict drive failure. We are a company that has performed storage testing for companies such as Intel, Seagate, and Western Digital. By leveraging our experience in storage testing, for HDD's and SSD's, we know how to collect a wide variety of drive data, not simply SMART attributes. This gives us a wide pool of drive health predictors to work with. In creating a machine learning model, having the right data is extremely important. Today, at ULINK, we have the ability to make machine learning a practical solution tailored to all kinds of drives, including yours.

With data becoming more important than ever to businesses and consumers, can people afford to be uncertain about the fate of their storage drives? ULINK Drive Analyzer, driven by machine learning technology, is our answer.