Explaining SSD Failures using Anomaly Detection

Chandranil Chakraborttii, Heiner Litz
University of California, Santa Cruz

Introduction

- We analyze telemetry data from over 30,000 SSDs running live applications in Google's datacenters over a span of six years, for predicting and explaining SSD failures using machine learning techniques.
- We propose the use of 1-class isolation forest and autoencoder-based anomaly detection techniques for predicting previously unseen SSD failure types with high accuracy
- We show that ignoring the minority class for training can improve the performance by up to 9.5% and if adaptability to dynamic environments is required, by up to 13%
- We utilize 1-class autoencoders to enable model interpretability
- We deploy a set of powerful feature selection techniques that improve the model performance by up to 1.3× and reduce training times by up to 1.8×

Dataset

- Collected at a Google data center from 3 MLC drive models spanning 6 years
- 30,000 unique drives
- Approximately 40,000,000 reports, total with 4000 failures
- 21 features collected including drive characteristics, status flags and error counts
- Some failed drives were put back into service upon repair
  Treated as a separate failed case

Drive Model	Number of failures	%age of drives failed
MLC -A	734	6.95
MLC-B	1565	14.27
MLC-C	1580	12.51
Total	3789	11.3

Prior Work

- Narayanan, et al.in "SSD Failures in Datacenters: What? When? and Why?"**[1]** and Meza et al. in " A Large-Scale Study of Flash Memory Failures in the Field" **[2]** studied SSD failures and its correlation with UECC errors
- Alter et. Al in "SSD failures in the field: symptoms, causes, and prediction models" **[3]** used Random Forest, Logistic Regression, k-NN, SVM, Decision Tree to failure prediction

- Issues
  - Low accuracy for the failed drives
  - No adaptivity to dynamic environments
  - No interpretability (black box models)

Our Approach

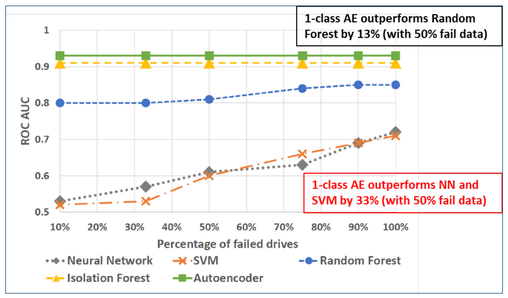
- A new 1-Class Autoencoder Model that
  - Improves accuracy by up to 13 %
  - Enables adaptivity to unseen failures
  - Enables interpretability of the results (white box)
- Key Idea
  - Training models on the majority class while ignoring the minority class
  - Prevents overfitting
  - Improves generalizability (unseen failure types)
- Additional Contributions
  - New feature selection technique
  - Extensive Hyperparameter tuning

Performance Comparison

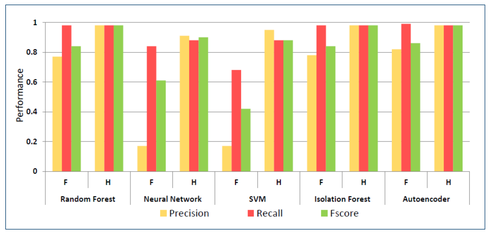
ROC is chosen as a metric as it a good measure of separation for imbalanced datasets

Baselines:
Random Forest, SVM and Neural Network [3]

Improvement
7.6% over best baseline technique



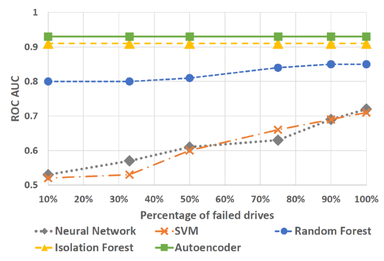
- 1-class Autoencoder, outperforms Random Forest by 9.5% while 1-class Isolation Forest outperforms baselines by 7%.
- Random Forest performs considerably worse for failed drives



Predicting unseen failures

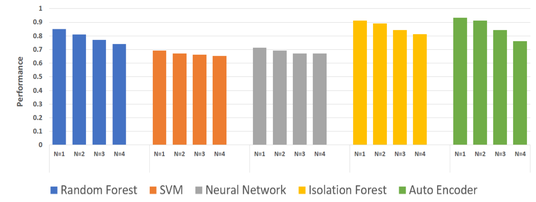
Baselines:
Random Forest, SVM and Neural Network [3]

Improvement:
For e.g., with 50% failed drive data, 1-class autoencoder technique outperforms random forest by 13% and NN and SVM by 33%.



Predicting ahead

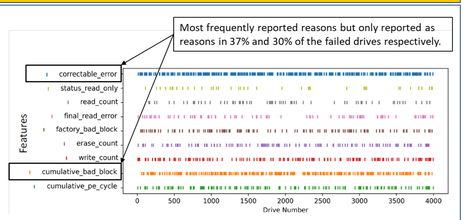
- **N** = number of days predicting ahead
- **Baselines:** Random Forest, SVM and Neural Network [3]



For N = 4, 1-class isolation forest performs best outperforming random forest baseline by 6% and SVM and NN by 11% and 13%

Interpreting SSD failures

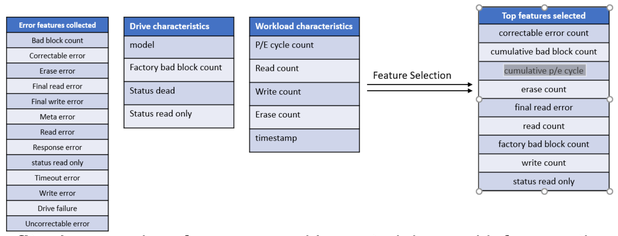
- Used reconstruction error per feature to compute reasons for failure
- If error more than avg_error/feature, it is listed as a reason for failure



- Combination of several features enables accurate failure prediction

Feature Selection

- Implemented eight different algorithms using three different feature selection methods (Filter, Embedded, and Wrapper techniques)



- **Benefits**: Improved performance and lower training and inference time

Conclusions

- We show that our approaches based on 1-class machine learning models outperform prior work by 9.5% ROC AUC score by significantly improving on the prediction accuracy for failed drives.
- For dynamic environments, our 1-class techniques improve over the baselines by 13%
- We show that 1-class autoencoders enable interpretability of model predictions

Contact:

Chandranil Chakraborttii
1156 High Street,
UC Santa Cruz
cchakrab@ucsc.edu
https://www.linkedin.com/in/chandranil-chakraborttii-a7552ba2

Heiner Litz
1156 High Street,
UC Santa Cruz
hlitz@ucsc.edu
https://www.linkedin.com/in/heiner-litz-3a332713