Mikailalsys Journal of
Advanced Engineering International

e-ISSN : 3046-9694
p-ISSN : 3046-8914

Index : Harvard, Boston, Sydney University, Dimensions, Lens, Scilit, Semantic, Google, etc

https://doi.org/10.58578/MJAEI.v3i2.9185

Development of ANFIS-Based Hard Drive Failure Prediction Model for Cloud Platforms Using Intelligent Techniques

I. I. Ahmad, J. D. Jiya, MA. Baba
Abubakar Tafawa Balewa University, Bauchi, Bauchi State, Nigeria
auabubakarphysics@gmail.com

Article Info:

Submitted:	Revised:	Accepted:	Published:
Feb 1, 2026	Apr 29, 2026	May 11, 2026	May 16, 2026

Abstract

Hard drive failures remain a critical reliability concern in large-scale cloud data centres because they can lead to data loss, service downtime, and increased operational costs. Traditional threshold-based monitoring techniques often fail to capture nonlinear relationships among hard drive health indicators and may produce high false-positive rates. This study presents a conceptual framework for developing an Adaptive Neuro-Fuzzy Inference System (ANFIS)-based hard drive failure prediction model using selected Self-Monitoring, Analysis, and Reporting Technology (SMART) attributes. It further examines the potential impact of key SMART indicators on predictive performance. By integrating fuzzy logic reasoning with neural network learning, the proposed framework is designed to improve predictive accuracy while maintaining interpretability. The study concludes that an ANFIS-based prediction framework can support proactive maintenance strategies for cloud service providers by enabling earlier identification of potential hard drive failures. This framework contributes to the development of intelligent predictive maintenance systems in cloud computing environments and offers practical implications for improving system reliability, reducing downtime, and enhancing operational efficiency.

Keywords: ANFIS; Cloud Computing; Hard Drive Failure Prediction; Predictive Maintenance; SMART Attributes

Introduction

Cloud computing has become the backbone of modern digital infrastructure, offering scalable and on-demand computing resources (Islam et al., 2023). However, cloud storage systems rely heavily on large-scale hard disk drives (HDDs), which are prone to mechanical degradation and operational failure over time (Tomer et al., 2021). Even though disk failures are statistically rare events, their impact is costly and disruptive.

Self-Monitoring, Analysis, and Reporting Technology (SMART) provides internal drive health indicators such as reallocated sector counts, temperature, and read error rates. While these metrics support basic monitoring, traditional threshold-based prediction models are insufficient for capturing complex failure dynamics (Li & Huang, 2024).

Recent advances in machine learning have improved predictive maintenance in cloud environments (Ganesh et al., 2022). However, many models—such as Random Forest, Support Vector Machine (SVM), and Deep Neural Networks—lack interpretability, which is essential for operational decision-making (Ozah et al., 2023). The Adaptive Neuro-Fuzzy Inference System (ANFIS) offers a hybrid intelligent framework that combines learning capability with explainable fuzzy rules (Chopra et al., 2021).

This study conceptually addresses two objectives:

1. To develop an ANFIS-based hard drive failure prediction model for cloud platforms.

2. To investigate the impact of selected SMART attributes on the predictive performance of the ANFIS model.

Literature Review

Hard drive failure prediction has evolved from statistical modelling to advanced machine learning techniques. Early regression and tree-based approaches demonstrated strong classification performance but struggled with nonlinear complexity (Li et al., 2014). More recent ensemble models such as Random Forest and Gradient Boosting have shown improved prediction accuracy (Ozah et al., 2023).

Deep learning models, including Long Short-Term Memory (LSTM) networks, have been applied to predict remaining useful life (RUL) of disks using long-term SMART datasets (Mohapatra et al., 2023). While effective, these models often require large computational resources and lack interpretability.

ANFIS, introduced by Jang, integrates artificial neural networks with fuzzy logic systems to model nonlinear and uncertain systems (Mazandarani & Li, 2020). It has demonstrated effectiveness in fault diagnosis, time-series prediction, and industrial system monitoring (Gheibi et al., 2023).

Hakim et al. (2022) showed that ANFIS can outperform traditional statistical models in equipment failure prediction by effectively handling mixed numerical inputs. Similarly, Amekraz and Hadi (2022) enhanced workload prediction using a hybrid neuro-fuzzy system.

Despite these advances, limited research has specifically explored ANFIS for cloud-scale hard drive failure prediction with optimized SMART feature selection. This gap motivates the present conceptual framework.

Objective One: Development of an ANFIS-Based HDD Failure Prediction Model

This involved using Adaptive Neuro-Fuzzy Inference Systems (ANFIS) to anticipate hard drive (HDD) failures by analysing SMART (Self-Monitoring, Analysis, and Reporting Technology) attributes. It's particularly useful in reliability engineering because HDD failures often involve nonlinear relationships and uncertain thresholds in SMART data.

Conceptual Architecture

The proposed framework consists of five major layers:

1. Data Collection

SMART telemetry extracted from cloud service provider datasets.

2. Data Preprocessing involves the following.

a) Cleaning

b) Imputation

c) Normalization

d) Outlier removal

3. Feature Selection for the Different input was used.

Recursive feature elimination and sensitivity analysis (Ganesh et al., 2022).

4. ANFIS Modelling based on the illustrated procedures

i) Fuzzification

ii) Rule evaluation

iii) Normalization

iv) Defuzzification

5. Prediction Output

Binary classification (Healthy vs. Failure Risk).

ANFIS Structure

ANFIS operates using a five-layer Sugeno-type architecture:

i. Layer 1: Fuzzification
ii. Layer 2: Rule generation
iii. Layer 3: Normalization
iv. Layer 4: Consequent parameter estimation
v. Layer 5: Output aggregation

This structure allows modelling of nonlinear relationships between SMART indicators and failure events (Chopra et al., 2021).

Advantages of ANFIS in HDD Prediction

ANFIS offers several advantages:

i. Handles nonlinear dependencies (Gheibi et al., 2023)
ii. Provides interpretable fuzzy rules
iii. Suitable for rare event prediction
iv. Combines learning and reasoning

These characteristics makes the Model particularly suitable for cloud storage environments where both accuracy and explainability are required.

Objective Two: Impact of SMART Attributes on Predictive Performance

The impact is significant, especially in areas like monitoring & evaluation, data mining, and educational analytics.

Selected SMART Attributes

The model incorporates eight critical SMART indicators:

i. SMART 5 – Reallocated Sector Count

ii. SMART 7 – Seek Error Rate

iii. SMART 13 – Soft Read Error Rate

iv. SMART 198 – Offline Uncorrectable Sectors

v. SMART 203 – Run Out Cancel

vi. SMART 228 – Power-Off Retract Cycle

vii. SMART 231 – Temperature

viii. SMART 250 – Read Error Retry Rate

These indicators reflect mechanical degradation, data integrity issues, and thermal instability (Li & Huang, 2024).

Feature Importance and Sensitivity

Mechanical integrity indicators such as SMART 198 and SMART 203 are strong predictors of imminent disk failure (Ganesh et al., 2022). Thermal indicators such as SMART 231 influence long-term degradation patterns (Pecht & Elburn, 2021).

Feature selection improves:

a) Model accuracy

b) Computational efficiency

c) Generalization capability

Recursive feature elimination has been shown to improve model robustness in predictive maintenance tasks (Ganesh et al., 2022).

Practical Implications

This conceptual ANFIS Hard Drive Failure Prediction Model supports:

a. Early disk failure detection

b. Reduced false positives

c. Optimized maintenance scheduling

d. Reduced operational costs

For cloud service providers, this translates into improved Service Level Agreement (SLA) compliance and enhanced system reliability (Islam et al., 2023).

Conclusion

This study presented a conceptual ANFIS-based hard drive failure prediction framework for cloud platforms. By integrating intelligent feature selection and fuzzy inference reasoning, the model enhances both predictive performance and interpretability. Future empirical validation should evaluate the performance using real-world cloud datasets and benchmarking against deep learning and ensemble models.

References

Amekraz, S., & Hadi, Y. (2022). Hybrid neuro-fuzzy approach for workload prediction in cloud computing environments. *Journal of Cloud Computing, 11*(1), 45–60.

Chopra, V., Sharma, R., & Gupta, A. (2021). Adaptive neuro-fuzzy inference systems: A review of applications and advancements. *Artificial Intelligence Review, 54*(4), 3211–3245.

Ganesh, R., Kumar, S., & Babu, P. (2022). Handling imbalanced SMART datasets for hard disk failure prediction using ADASYN and feature extraction. *IEEE Access, 10*, 88921–88935.

Gheibi, A., Motahari, S., & Ahmadianfar, I. (2023). Neuro-fuzzy modelling for predictive maintenance in industrial systems. *Engineering Applications of Artificial Intelligence, 120*, Article 105865.

Hakim, M., Rahman, T., & Sada, A. (2022). ANFIS-based equipment failure prediction in data center infrastructure. *Computers & Electrical Engineering, 99*, Article 107745.

Islam, M. S., Hassan, M., & Raza, A. (2023). Reliability challenges in large-scale cloud storage infrastructures. *Future Generation Computer Systems, 144*, 54–66.

Li, J., Wang, X., & Pecht, M. (2014). Classification and regression trees for hard disk drive failure prediction. *Reliability Engineering & System Safety, 130*, 1–11.

Li, Y., & Huang, K. (2024). Failure characteristics analysis of large-scale hard disk drives in cloud data centers. *IEEE Transactions on Cloud Computing, 12*(1), 210–223.

Mazandarani, M., & Li, X. (2020). A comprehensive review of ANFIS architectures and learning strategies. *Neurocomputing, 401*, 196–213.

Mohapatra, P., Roy, S., & Sahoo, B. (2023). LSTM-based prediction of remaining useful life of hard disk drives using SMART data. *IEEE Transactions on Reliability, 72*(2), 480–492.

Ozah, F., Ahmed, M., & Bello, T. (2023). Comparative evaluation of regression models for hard drive failure forecasting using SMART attributes. *Journal of Big Data, 10*(1), Article 134.

Pecht, M., & Elburn, T. (2021). Temperature effects on hard disk drive reliability in large-scale storage systems. *Microelectronics Reliability, 117*, Article 113673.

Tomer, S., Singh, R., & Sharma, P. (2021). Hard disk failure analysis using SMART data in data centers. *IEEE Access, 9*, 134578–134589.