

No-Code ML Dataset Evaluator: Empowering Experts with Intelligent Model Selection Using Web-Based Statistical Analysis

Mr S. Sadeesh¹, Trisha Shiny J², Vedhalakshmi M³, Vigneswari D⁴

¹Assistant professor, Dept of AI&DS, Saranathan College of Engineering, Trichy, Tamilnadu, India

^{2,3,4} UG Scholar, Dept. of AI&DS, Saranathan College of Engineering, Trichy, Tamilnadu, India

Emails: sadeesh7141@saranathan.ac.in¹, Trishashiny2@gmail.com², vedhalakshimamani13@gmail.com³, Vigneswari11173@gmail.com⁴

Abstract

This paper presents a comprehensive web-based system that automates machine learning dataset quality assessment and algorithm recommendation for non-expert practitioners. The platform implements a novel six-component trainability scoring framework (0-100 scale) that evaluates data completeness, feature quality, sample size adequacy, duplicate detection, outlier prevalence, and class balance to quantify dataset readiness. The system performs dual-layer statistical analysis: descriptive metrics (mean, median, standard deviation, skewness, kurtosis) for distribution characterization, and inferential tests (Pearson correlation, normality assessment) for relationship identification and transformation guidance. A multi-task evaluation engine independently scores datasets for classification, regression, and clustering suitability, while a context-aware recommendation algorithm matches datasets to specific models based on detected quality signatures—suggesting XGBoost with class weighting for imbalanced data or Huber loss regressors for outlier-heavy datasets. Users receive three to five prioritized improvement recommendations with quantified impact predictions, comparing baseline versus projected performance metrics. Interactive visualizations render statistical insights through histograms, correlation heatmaps, and quality dashboards. Implemented entirely client-side using JavaScript, Chart.js, and PapaParse, the system ensures data privacy while providing instant feedback, successfully.

Keywords: Machine Learning, Dataset Quality Assessment, Algorithm Recommendation, Trainability Scoring, Statistical Analysis, Data Profiling, Automated Model Selection, Web-Based Tool, Data Quality Metrics, Interactive Visualization.

1. Introduction

High-quality datasets are the backbone of trustworthy machine learning systems, as the performance, generalization, and interpretability of models are highly dependent on the quality of the training datasets. Real-world datasets, however, tend to be noisy, imbalanced, inconsistent, or ill-structured, making it challenging to conclude whether a dataset is fit for a given task. Many users, particularly newcomers or domain specialists lacking in-depth knowledge of data science, find it challenging to assess datasets before model training, which may result in erroneous outputs, unnecessary computational expenditure, and misleading conclusions. Thus, there is an increasing need for an automated approach to analyse datasets, measure their readiness, and suggest suitable tasks, models, and optimizations before model training.

1.1. Challenges in existing data evaluation system

Inadequate Overall Quality Assessment - The conventional approach is limited to detecting missing data or duplicated entries, without assessing the quality of underlying properties like relevance, distribution, or predictive value. [1]

Difficulty in Manual Model Selection – Selecting an appropriate model is often a complex process that demands extensive expertise, experimentation, and computational resources.

Ambiguity in Target Variable Selection – It is often difficult for users to identify the appropriate target variable, especially when there are several potential variables in a dataset.

Insufficient Task Suitability Information – Conventional approaches lack the ability to suggest

whether a particular dataset is more appropriate for classification, regression, clustering, or other machine learning tasks. [2]

Lack of Feedback on Improvements – Current analysis tools lack the capability to offer feedback on dataset improvement or demonstrate the effectiveness of improvements made.

1.2.1.2. Why an Intelligent Dataset Evaluation System?

With the advent of automated analytics, statistical modelling, and intelligent recommendation systems, it is possible to build systems that can holistically assess datasets and assist users in the preprocessing and model selection process. An intelligent dataset evaluation system can improve decision-making in the following ways:

- Calculate a trainability score that quantifies the readiness of the dataset for machine learning. [3]
- Evaluate the dataset for various tasks and recommend the most suitable task.
- Recommend the most suitable target variable based on statistical correlation and predictive power.
- Compare various algorithms and recommend the most suitable model.
- Assist in understanding the dataset through visualizations and statistical analysis.
- Evaluate performance metrics before and after dataset improvement.
- Recommend improvements in terms of imbalance, noise reduction, or optimal feature distribution.

1.3. Objectives of proposed system

The proposed system is intended to:

- Automate the assessment of datasets and prevent manual trial-and-error analysis.
- Offer interpretable scores to quantify the quality and readiness of datasets. [4]
- Aid users in identifying the best machine learning tasks, targets, and models.
- Improve the reliability of datasets with actionable improvement suggestions.
- Allow for multiple uploads of datasets and provide separate reports for each.
- Present statistical, graphical, and

performance-driven insights in a single interface.

2. Problem Statement

In contemporary machine learning, data scientists commonly face challenges while trying to establish whether a given dataset is appropriate for developing a trustworthy predictive model. The raw dataset may have missing entries, noisy data points, unbalanced classes, redundant attributes, or an unclear target concept, which can cause a substantial drop in model performance. This poses the following challenges:

- **Excessive experimentation time:** The need to continually validate several models and data preprocessing techniques.
- **Model selection ambiguity:** The absence of clear guidance causes the selection of suboptimal models.
- **Poor prediction performance:** Low-quality data hurts model training performance.
- **Ineffective resource utilization:** Wasted computational resources on inappropriate data or tasks.
- **Lack of interpretability:** unclear statistical properties make it hard to understand data behavior.

A systematic and automated process for evaluating datasets is required to overcome the limitations of current manual analysis techniques, which tend to offer only elementary results like basic statistics or simple plots. Current manual analysis techniques do not evaluate the readiness of the dataset, the suitability of the task, or the effect of dataset improvement on model performance. Moreover, most current systems do not provide suggestions for target variables or machine learning models. To overcome these limitations, the intelligent dataset evaluation system will analyse dataset properties systematically, provide scores, provide suggestions for suitable tasks and models, and offer improvement plans for more efficient machine learning development. [5]

3. Literature Review

Over the years, several approaches have been proposed for evaluating the suitability of datasets for machine learning. Traditional statistical analysis is one of the most popular methods of evaluating

datasets, which involves analyzing the dataset using parameters such as mean, variance, and distribution. Although these approaches are very helpful in understanding the nature of a dataset, they are primarily based on surface-level characteristics and are often ineffective in determining whether a dataset is actually suitable for machine learning model training. Several data quality assessment frameworks have also been proposed by researchers, which involve evaluating parameters such as completeness, consistency, validity, and redundancy. These methods are very helpful in identifying problems such as missing values, noise, and formatting errors, but they are generally only useful for pointing out problems and not for providing recommendations on how to improve the dataset or which machine learning models to use. More advanced model-aware assessment methods have also been proposed, which involve correlating dataset properties with machine learning model performance, providing more in-depth information; however, these methods are generally only feasible after running machine learning models, which can be time-consuming and resource-intensive. To further ensure consistency in the evaluation of the dataset, automated scoring tools have been incorporated. While these tools make it easier to compare, most current solutions only take into account a few aspects of quality and do not have a well-rounded view that encompasses task appropriateness, relevance, and compatibility. Thus, the need for a comprehensive framework that incorporates statistical analysis, predictive assessment, and improvement suggestions still exists. To fill this void, the proposed system will provide a well-rounded platform that can assess multiple datasets, provide structured scores, and make suggestions for improvement. [6]

4. Proposed Solution

To overcome the shortcomings of the conventional dataset analysis techniques, we propose an Intelligent Dataset Evaluation System that combines automated statistical analysis, machine learning-based scoring, visualization, and recommendation tools into a single framework. The proposed system aims to assess the quality of the dataset, identify suitability for a task, identify the best target variables, suggest suitable models, and suggest ways for improvement. The

proposed system will thus enable faster, more accurate, and more reliable dataset preparation for machine learning tasks.

4.1. System Overview

The Proposed system consists of the following key components:

4.1.1. Multi-Dataset Upload and Processing

- Users can upload one or multiple datasets simultaneously.
- Each dataset is validated for format, structure, and schema consistency.
- Independent processing pipelines are assigned so results are generated separately for each dataset.

4.1.2. Automated Data Quality Assessment

- The system scans datasets for missing values, duplicates, noise, outliers, and inconsistencies.
- It computes quality metrics such as completeness score, redundancy score, and distribution balance.
- Initial dataset health indicators are generated before preprocessing.

4.1.3. Statistical and Structural Analysis

- Descriptive statistics such as mean, variance, and distribution patterns are calculated.
- Correlation and dependency measures identify relationships between features.
- Inferential statistical techniques detect significant patterns affecting learning performance

4.1.4. Feature and Target Evaluation

- Each feature is evaluated for importance, variance, and predictive strength.
- Multiple possible target variables are analysed and ranked.
- The system recommends the most suitable target for modelling tasks.

4.1.5. Task Suitability Detection

- Based on dataset structure and labels, the system determines whether the dataset best fits classification, regression, clustering, or other machine learning tasks.
- Each task is assigned a compatibility score.

4.1.6. Model Recommendation Engine

- Candidate machine learning models are

assessed based on dataset properties and simple simulations.

- Performance metrics like accuracy, robustness, and error behaviour are employed for model comparison.
- The model recommendation engine recommends the most suitable machine learning model for the dataset.

4.1.7. Trainability Score Generation

- All derived scores are combined into a single trainability score.
- The trainability score is an indicator of the overall readiness of the dataset for machine learning.

4.1.8. Dataset Improvement and Re-evaluation

- The system recommends techniques for dataset improvement, including normalization, encoding, balancing, feature selection, and noise reduction.
- The dataset is re-evaluated for comparison of performance metrics before and after improvement.

4.2. System Architecture & Data Flow

The Architecture is designed as a modular pipeline architecture with self-contained processing layers for scalability and flexibility.

4.2.1. User Interface Layer

- Allows uploading datasets, specifying parameters, and displaying reports.
- Displays results, scores, and recommendations.

4.2.2. Processing Engine

- Executes preprocessing, statistical analysis, feature evaluation, and scoring tasks.
- Allows for concurrent processing of multiple datasets.

4.2.3. Evaluation Core

- Holds algorithms for task detection, model scoring, and trainability scoring.
- Combines results from all analysis modules.

4.2.4. Data Storage and Report Module

- Saves datasets, intermediate results, and evaluation reports.
- Offers structured summaries and downloadable reports. Figure 1 shows System

Workflow

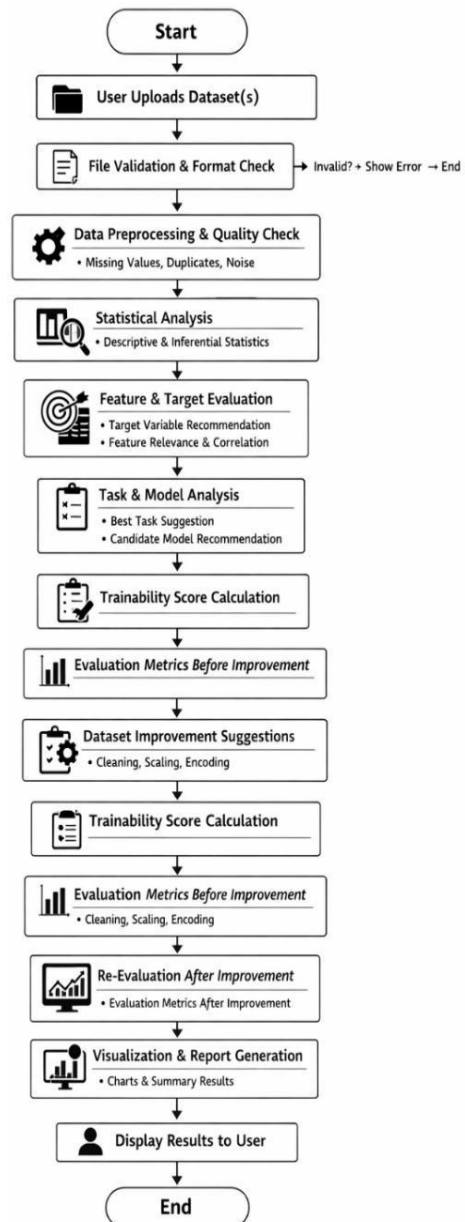


Figure 1 System Workflow

4.3. Key Technologies & Algorithms Used

4.3.1. Statistical Analysis Techniques

- Measures central tendency, variance, correlation, and distribution patterns.
- Detects imbalance, noise, and anomalies in data.

4.3.2. Feature Selection and Ranking Methods

- Correlation analysis, variance thresholds, and

entropy-based scoring determine feature importance.

- Redundant features are identified and flagged.

4.3.3. Task Identification Algorithms

- Rule-based and heuristic methods analyse label types and feature structures to determine suitable learning tasks.

4.3.4. Model Suitability Estimation

- Lightweight model testing and meta-learning principles estimate algorithm compatibility without full training.

4.3.5. Scoring Framework

- Weighted scoring formulas combine quality metrics into a unified trainability index for decision support.

5. Tables

Table 1 Evaluation Metrics Used

Metric Type	Purpose	Outcome
Completeness Score	Measures missing data ratio	Indicates dataset reliability
Correlation Score	Evaluates feature-target relationships	Identifies predictive strength
Distribution Score	Detects imbalance or skewness	Ensures fairness and stability
Noise Score	Finds duplicates/outliers	Improves data clarity
Feature Relevance	Measures importance of attributes	Helps feature selection
Model Performance	Estimates expected accuracy	Guides algorithm choice
Trainability Score	Combines all metrics	Overall dataset readiness

Table 2 Comparative Results

Parameter	Traditional Dataset Analysis	Proposed Dataset Evaluation System
Processing Speed	Manual and time-consuming	Fully automated and fast
Quality Detection	Limited (missing values only)	Multi-dimensional quality analysis
Feature Assessment	Manual Inspection	Automatic Feature Scoring
Model Selection	Trial-and-error	Intelligent recommendation
Dataset Handling	Single dataset at a time	Supports multiple datasets simultaneously

Error Detection	Basic validation	Advanced anomaly and noise detection
Predictive Suitability	Not evaluated	Automatically estimated
Visualization	Minimal charts	Detailed visual analytics
Preprocessing Feedback	Not available	Before–after comparison
Decision Support	Depends on user expertise	System- generated insights

6. Advantages Of Proposed System

6.1. Automated and Faster Dataset

Assessment The manual assessment of datasets is often a slow and error-prone process, involving repeated evaluation and testing. The proposed system will automate this process by rapidly assessing datasets and providing structured scores and results, facilitating faster decision-making and development. (Table 1) (Table 2)

6.2. Intelligent Model and Task Recommendation

The selection of appropriate models and tasks is often a trial-and-error process. The system will assess the characteristics of the dataset and make intelligent recommendations for suitable tasks and algorithms, increasing accuracy and avoiding unnecessary testing.

6.3. Comprehensive Quality Evaluation

Most current systems are limited to basic checks such as missing values. The proposed system will comprehensively assess the quality of datasets, including distribution, relevance, noise, and predictive power, providing a more complete understanding of dataset quality.

6.4. Before-and-After Performance Comparison

It is often difficult to assess the impact of preprocessing operations. The system will assess datasets before and after improvements, enabling users to easily view the results of data cleaning and optimization efforts

6.5. Scalable Multi-Dataset Analysis

Most systems currently available are limited to analyzing one dataset at a time. The proposed system will enable the analysis of multiple datasets

at once, providing separate reports for each dataset, saving time and facilitating easy comparison.

7. Challenges And Future Work

7.1. Computational Complexity

Analysing large amounts of data using various evaluation techniques may consume many resources. Future development will concentrate on optimizing techniques and parallel processing.

7.2. Dataset Diversity

Various domains demand different evaluation techniques. Future releases will contain domain-specific modules for optimizing the evaluation of specialized datasets.

7.3. Model Estimation Accuracy

As recommendations are generated without complete model training, prediction accuracy may occasionally be compromised. Future upgrades will implement adaptive learning algorithms to improve accuracy.

7.4. Scalability for Big Data

Efficiently processing large amounts of data is still a challenge. Future development will incorporate scalable and parallel processing platforms.

7.5. Interpretability of Scores

Overall scores may be hard to understand for end-users. Future upgrades will implement interpretable techniques to provide insights into score and recommendation calculations.

8. Results And Discussion

8.1. Results

The proposed Dataset Evaluation System was experimented on various datasets of different sizes and types to test the efficiency of the system in evaluating the quality of the dataset, suggesting models, and improving trainability. The experimental results showed that the system was able to effectively

identify the weaknesses of the dataset, provide suggestions for preprocessing, and improve model trainability, thus facilitating accurate and efficient machine learning. Figure 2 shows Trainability Score, Figure 3 shows Target Variable Prediction, Figure 4 shows Evaluation Metrics Before Improvement and After Improvement, Figure 5 shows Best Model Suggestion.

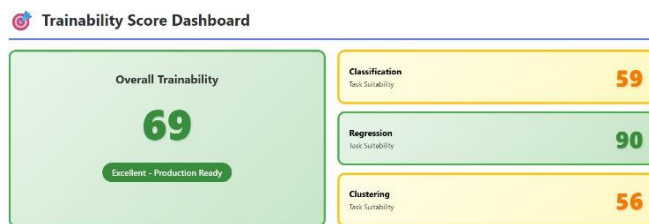


Figure 2 Trainability Score

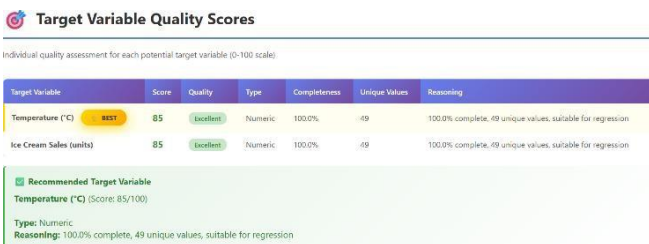


Figure 3 Target Variable Prediction

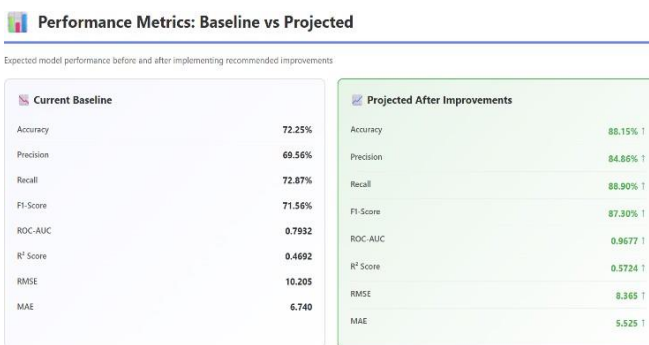


Figure 4 Evaluation Metrics Before Improvement and After Improvement



Figure 5 Best Model Suggestion

8.2. Discussion

The experimental evaluation reveals that the proposed Dataset Evaluation System is useful for evaluating datasets of different domains and structures, and it gives significant insights into the quality of the dataset and the appropriateness of the model. The results obtained from the study indicate that the automated scoring of attributes such as completeness, balance, noise, and feature relevance can play an important role in revealing the hidden limitations of the data, which may otherwise affect the performance of the model. The system enables the linking of the attributes of the dataset to the suggestions of the algorithm, thus removing the need for trial and error in the selection of the model. In addition, the improvements brought about by the preprocessing validation further confirm that the evaluation of the dataset can have a direct effect on trainability and predictability.

Conclusion

To sum up, the proposed Dataset Evaluation System In conclusion, the proposed Dataset Evaluation System assists users in understanding how good their dataset is before developing any machine learning model. Rather than training models directly and waiting for positive outcomes, this system first verifies critical aspects such as missing values, data distribution, usefulness of features, noise, and whether the dataset is appropriate for the task at hand. In this way, users can correct errors before they occur and prevent poor model performance. The system was tested on various types of datasets, and it was found that proper evaluation and improvement of the dataset can significantly improve the efficiency of training and accuracy of predictions. It also facilitates model selection by recommending appropriate algorithms according to the characteristics of the dataset. Most importantly, this project emphasizes the importance of a simple fact: the best AI models begin with the best data. The proposed system makes data analysis easier, faster, and more accurate for both novices and experts.

Acknowledgements

The authors would like to thank the organizing committee of the international conference for giving them the opportunity to present their research work. We would like to thank our mentors and faculty

members for their expert advice and guidance in the development of this research work. Their contributions have played an important role in improving the methodology, analysis, and presentation of the proposed system. We would like to thank our institution for giving us the facilities and support necessary to complete this research work successfully. We would like to thank our peers and reviewers for their suggestions that have helped us in improving the quality of this work. Finally, we would like to thank our families for their support and encouragement throughout this research journey.

2021.

References

- [1].S. Mazurek and M. Wielgosz, "Assessing Dataset Quality Through Decision Tree Characteristics in Autoencoder-Processed Spaces," arXiv-Machine Learning, 2023.
- [2].C. Wang, Q. Dong, X. Wang, and Z. Sui, "Statistical Dataset Evaluation: Reliability, Difficulty and Validity," Natural Language Processing, 2024.
- [3].S. Mohammed, L. Budach, M. Feuerpfeil, N. Ihde, A. Nathansen, N. Noack, H. Patzlaff, F. Naumann, and H. Harmouch, "The Effect of Data Quality on Machine Learning Performance on Tabular Data," Information Systems, 2025.
- [4].D. Bhaduria, A. Sierra Múnera, and R. Krestel, "Statistical Dataset Evaluation: A Case Study on Named Entity Recognition," in Proc. Ninth Workshop on Noisy and User-generated Text (W-NUT), 2024, pp
- [5].G. I. Winata, D. Anugraha, E. Liu, A. F. Aji, S.-Y. Hung, A. Parashar, P. A. Irawan, R. Zhang, Z.-X. Yong, J. C. B. Cruz, N. Muennighoff, S. Kim, H. Zhao, S. Kar, K. E. Suryoraharjo, M. F. Adilazuarda, E.-S. A. Lee, A. Purwarianti, D. T. Wijaya, and M. Choudhury, "Datasheets Aren't Enough: DataRubrics for Automated Quality Metrics and Accountability," arXiv-Audio and Speech Processing, 2025.
- [6].H. Patel, F. Ishikawa, L. Berti-Equille, N. Gupta, S. Mehta, S. Masuda, S. Mujumdar, S. Afzal, S. Bedathur, and Y. Nishi, "Data Quality Assessment for Machine Learning," in Proc. KDD Workshop on Data Quality,