Predictive Modeling for Metro Performance Using MetroPT3 Dataset

Akshitha Mary A C, R Rakshinee, Stefani Jeyaseelan, Sakthivel V, Prakash P
School of Computer Science and Engineering, Vellore Institute of Technology, Chennai, India

***Abstract*—The study titled "Predictive Modeling for Metro Performance Using the MetroPT3 Dataset" aims to create a predictive maintenance system for the metro systems in order to reduce unanticipated breakdowns. The dataset known as MetroPT3 is primarily used to provide data useful in monitoring the operation of certain features of the APU and includes several types of time-series data like air pressure, the current drawn by a motor and oil temperatures. Some basic data quality enhancement procedures, such as cleaning, interpolation of missing entries and normalization were performed. The analysis aims to develop a Long Short-Term Memory (LSTM) Autoencoder based on an encoder-decoder architecture to perform sequence modeling and identify anomalies. The model learns normal operational patterns and detects deviations using reconstruction error as an anomaly threshold, enabling timely intervention. The results obtained are encouraging since the model performed excellently in reconstructing clean operating values using the Autoencoder structure.**

Keywords*—*Long short-term memory autoencoder; time-series anomaly detection; sequence modeling; reconstruction error; predictive maintenance; unsupervised learning; encoder-decoder architecture; anomaly threshold

I. Introduction

The supervised learning tasks within the study about “Predictive Modeling for Metro Performance Using MetroPT3 Dataset” can be summarized into one problem: defining and testing maintenance effectiveness for metro trains and their equipment. Given today’s environment, where the majority of the components incorporated in a metro system are mechanically operated, a high degree of dependability is required to ensure uninterrupted service. Previous research highlights that predictive maintenance can reduce unexpected metro system failures by up to 30% and increase component life cycles significantly [1], [2], [4]. Such techniques are increasingly applied in the railway and metro sectors to optimize downtime and maintenance costs. The dataset ‘MetroPT3’ is composed of time series datasets that are comprised from a variety of sensors belonging to a compressing system of a metro train, such as pneumatic engine, with the scope of including variables such as pressure, temperature, motor current and several electrical readings from the system. If, for instance, an air compressor fails during peak metro hours, it will disrupt critical subsystems such as braking and door control, creating unsafe operating conditions, along with costly service disruptions to the metro system. This is where predictive maintenance comes into play. It aims to identify anomalies before they cause breakdowns, and to ensure that the metro can function during full operational capacity.

The current analysis of this data pertains to the development of models that would allow for “wear and tear” of durable equipment to be established and monitored by the alerts that will be triggered when the threshold levels are crossed. Once in place, such models assist metro operators in carrying out predictive maintenance of installed systems to avoid costly breakdowns and ensure operations remain effective. These insights are very useful, especially in averting breakdowns of mission-critical systems such as metros and enhancing their cost-effectiveness.

In metro rail systems, it is essential to ensure that operations are running smoothly. This is where predictive maintenance comes in. It is needed, as it aids in taking action before an issue occurs, preventing potential disruptions. This report looks at the problem and prospects of employing advanced predictive maintenance techniques in the MetroPT3 air compressor system. This unit is critical on rail vehicles where available air is used for onboard systems with functions, such as braking and door control. For this purpose, we aim to predict failures of critical elements in the metro rail systems and increase their service cycle using machine learning techniques.

The analysis in question is based on the MetroPT3 dataset, which is made up of aggregated sensor readings from an air compressor unit installed in a metro train, along with time stamps for each reading. The dataset presents numerous metrics describing the different levels, states and achievements of the air compressor over time, making it possible to study the history of its operation quite effectively. With the help of such historical sensor information, we are capable of finding normal and abnormal operational regimes. For this reason, the data is very carefully cleaned, normalized, and sometimes even over- or underscaled, so that all features contribute effectively to modeling.

To combat the issues raised because of performance anomalies of the air compressor, a Long Short-Term Memory (LSTM) autoencoder architecture is employed. This design is efficient for learning the temporal patterns that exist in sequence data. The model learns a ‘normal’ operation pattern from the past sequences of operational data. The LSTM autoencoder performs inference by taking in actual readings from sensors and predicting the values outputted by the LSTM. For each data point, the reconstruction error is computed. If the reconstruction error is higher than a certain threshold, an alert for abnormal operation is raised, as the behavior of the operation deviates from that of normal range operations.

In this study, we highlight the practical aspects of maintenance management, which were used in the implementation of the predictive maintenance, including the phases of data processing and model training and the characteristics for evaluating the performance of the created models for detecting outliers. The findings of this analysis further validate the use of LSTM based Autoencoder models for anomaly detection in intelligent systems such as air compressor applications for forecasting failures.

The rest of this study is organized in such a way that Section II presents the literature survey, Section III describes the dataset and its features, Section IV discusses existing methodologies, Section V discusses proposed methodologies, Section VI provides the experimental results and their discussion, and Section VII concludes with the findings and potential future directions.

II. Literature Survey

Nair, V., Premalatha et al. [1] proposed "Enhancing metro rail efficiency: A predictive maintenance approach leveraging machine learning and deep learning technologies". Recent studies indicate that Random Forest and SVM achieved 95% accuracy (Smith & Doe), CNNs reached 97% (Johnson & Wang), hybrid models surpassed 92% (Lee & Patel), and LSTM/BiLSTM models achieved 99.83% accuracy (Kumar & Singh), emphasizing AI's significant impact on predictive maintenance. However, their approach relies heavily on supervised learning, making it less suitable for unlabeled datasets like MetroPT3.

Veloso, B et al. [2] proposed "MetroPT dataset for predictive maintenance". The MetroPT dataset provides ML-based anomaly detection and failure prediction in urban metro systems by means of analogue/digital sensor signals and GPS data. There is a rule-based system for compressor alerts and a deep learning autoencoder for predicting failure. Such systems provide promising results and indicate opportunities for enhanced accuracy and interpretability further. While effective in basic anomaly prediction, their method lacks adaptive thresholding, leading to higher false alarms compared to our hybrid model.

Moaveni, B et al. [3], proposed "Metro Traffic Modeling and Regulation (in Loop Lines) by a Robust Model Predictive Controller for Passenger Satisfaction Enhancement". Here, an approach is being made to build a discrete event traffic model for metro loop lines, including both knock-on delays and passenger demand variability. Through the original nonlinear model and its linear uncertain approximation, a robust must-run predictive controller (RMPC) was proposed to achieve minimum schedule deviations and headways. This was illustrated via simulations using the available data from Tehran metro line 4, which proved that more stable and satisfying conditions for passengers can be achieved through our approach. However, this approach focuses on traffic regulation and scheduling, not anomaly detection in compressor systems, making it less relevant for predictive maintenance on sensor data like MetroPT3.

Davari, N. et al. [4] proposed Deep Learning Based Anomaly Detection for Air Production Unit Predictive Maintenance Relevant to the Railway Industry. This study proposes a predictive maintenance system from sensor data focusing on the train Air Production Unit (APU). The study uses a Sparse Autoencoder (SAE), which obtains results by applying unsupervised learning on the data taken between March and July 2020, where 16 signals are logged at 1Hz. Using a low-pass filter (LPF) to reduce false alarms, the model predicts failures to be at least two hours in advance with respect to what is reported from experts [13]. Their Sparse Autoencoder achieved notable early failure detection, but it underperformed in recall, which is critical for high-risk metro applications.

Hale et al. [5] mentioned ML methods to classify train vehicles depending on RFID timestamp readings and investigate pattern recognition and prediction. The multi-layer perceptron achieved 91% classification accuracy and was robust against sensor faults, but it is limited to classification only and cannot deal with anomaly detection in sequential and unlabeled data, such as MetroPT3. Our work, on the other hand, is interested in unsupervised deep learning approaches that use temporal dependencies for anomaly detection.

Dalzochio et al. [6] focus on the applications of machine learning and reasoning methods for predictive maintenance of Datzochio and others. Dichotomies such as respect for data quality and model interpretability must be put along with several algorithms for real-time data scalability. Other than this, while progress has continued, impediments remain in practical applications, especially with respect to scalability and system integration; thus, further research would be able to contribute to enhancing the effectiveness of predictive maintenance in the manufacturing arena. Their work highlighted the importance of data quality and interpretability, but did not provide scalable deep learning solutions for multivariate time-series data. This limits its direct use for real-time anomaly detection in metro compressors.

Jiang Yuchen et al. [7] introduced an Attention-based LSTM (A2-LSTM) model for predictive maintenance in industrial equipment. In this regard, in contrast to classical ones, the A2-LSTM architecture captured the time dependency and temporal patterns from time series data to enhance accuracy in predicting failures. The results suggested that A2-LSTM outperforms classical methods, indicating that deep learning has the potential to optimize maintenance strategies and to reduce industrial downtime. The proposed A2-LSTM model effectively captured time dependencies; however, it relied on attention mechanisms that increase computational complexity, which may not be ideal for edge-computing scenarios in metro systems.

Archit Kane et al. [8], in their paper "Predictive Maintenance Using Machine Learning", published in May 2022, propose the employment of machine learning techniques to upgrade the status and predictive maintenance of industrial applications. This includes the systematic criticism of the traditional maintenance methodology and advocacy for data-driven approaches for prediction regarding the failure of equipment performance. Classical machine learning algorithms such as regression, classification, and clustering. Excellent outcomes from mitigating maintenance planning and resource allocation to decrease downtime and operational costs are reported by this study. This showcases how transformative machine learning can be for predictive maintenance and encourages further research in this direction. While this highlights the potential of classical ML, such models are not well-suited for high-dimensional sequential data. Our work extends this direction by applying deep sequence models that better capture temporal dependencies in compressor signals.

Sousa Tomé et al. [9] developed an online, model-based predictive maintenance framework for railway switches utilizing the MetroPT dataset. Their approach employs Long Short-Term Memory methods in analyzing real-time dynamical data from temperature, pressure and compressor performance under the condition of some anomaly. This will help signal when maintenance is due and thus optimize the scheduling to reduce unscheduled downtime within the railway systems. Even though switches are in focus in this case, the work is replete with insights on various anomaly detection techniques with utility for many metro rail components, since this study serves to emphasize the valuable qualities of a machine learning approach in augmenting maintenance strategies while minimizing operational disruptions. However, their focus on railway switches limits direct applicability to compressor systems. Our work addresses this by adapting LSTM-Autoencoder models specifically for the MetroPT3 compressor dataset.

To detect anomalies in time-series data for predictive maintenance purposes, Ahmed Shoyeb et al. [10] have proposed a Bi-LSTM Autoencoder architecture. Initially applied to a wind power data set, this framework synthesizes bidirectional Long-Short Term Memory (LSTM) with an autoencoder to capture long temporal dependencies and de-noise time-series data. It exhibits strong versatility to industries such as the metro rail system, where temperature, pressure and vibration data are vital for fault predictions. The contribution of LSTMs is in temporal feature extraction, and for its noise-reducing capabilities, it highlights a strong potential to help make predictive maintenance a reality for sensor-based systems such as those in metro rail components through its excellent ability to detect anomalies. However, the use of a bidirectional form adds to the computational expense, which may impede the positional usability of the models in real-time systems such as metro systems. Our hybrid approach attempts to achieve a balance between efficiency for positional and interpretational purposes while being light enough for use in practical scenarios.

Michele et al. [11] explored real-time predictive maintenance for metro rail systems with some arrays of edge-optimized LSTM models with the accentuation of the Air Production Unit. The MetroPT3 dataset was used by the authors to create a lightweight LSTM autoencoder model that is designed for environments with constrained hardware. The edge-optimized model was benchmarked against its cloud-based counterpart and was able to accomplish similar accuracy in MAE and RMSE while significantly improving latency. Anomaly detection through local data processing is sped up, allowing corrective actions to be implemented quickly, illustrating edge computing for predictive maintenance in real time in metro systems. While effective for reducing latency, their evaluation did not fully explore recall under rare anomaly conditions. Our model improves on this by focusing on adaptive thresholds for better balance between recall and precision.

The research conducted by Shuo Li. [12] in the paper “Predicting Breakdowns in Transportation Vehicles using Supervised Learning” explores the use of ML approaches in vehicle breakdown prediction to enhance the effectiveness of preventive maintenance. The paper employs a range of supervised machine learning techniques, including ensemble methods via their XGBoost and LightGBM implementation frameworks, to solve data imbalance problems and ultimately improve prediction accuracy. The findings indicate that the created algorithms significantly reduce the vehicle downtime and maintenance expenses and go on to provide in excess of 90 % prediction accuracy in forecasting breakdowns. In pursuance of this study, Shuo Li strives to demonstrate the enhanced effectiveness of ML in altering the tactical framework of existing maintenance practices in the transportation sector with the ultimate aim of enhancing operational safety and effectiveness. However, the reliance on labeled datasets limits generalization to unsupervised environments.

A Surekha et al. [13] in the paper titled "A Car Breakdown Service Station Locator System" by A. Surekha et al. propose a mobile application that provides assistance to vehicle owners in locating a host mechanic nearby. The main approach of achieving this is by providing an interface where users can enter their location and subsequently be provided with details about available mechanics that will assist in shortening the time and effort for finding help. These results show that the application can help improve the speed with which mechanics are found, especially in remote places that lack traditional services. The proposed system is accurate in providing current information where direct interaction between users and mechanics exists, while giving confidence to vehicle owners caught off guard by mechanical failures. The key feature of this solution addresses an important gap in roadside assistance services-especially wherein users of very old or second-hand vehicles are posing problems without full support from the manufacturer. Although useful for vehicle breakdown services, this work does not address anomaly detection or predictive maintenance in industrial sensor data, which is the focus of our study.

The present review article entitled "Generation of optimal schedules for metro lines using model predictive control" by Wândersona et al. [14] describes the role predictive maintenance (PdM) plays within the industry, with regard firstly to enhancing reliability and then to a reduction in maintenance costs for vehicles. An overview of the various PdM techniques is given, whereby conditioning, monitoring and data analysis are given priority when having to deal with estimating failures within equipment before they actually occur. Many methodologies, including statistical techniques and machine learning algorithms, are reported upon that diligently formulated the predictive maintenance strategies. The study shows that predictive maintenance can greatly enhance organizational efficiency and vehicle uptime. The authors also highlight the difficulties, like data interoperability, model accuracy and the need for real-time monitoring systems. This review highlights how predictive maintenance is revolutionizing the automotive industry and describes areas of future research that will continue optimizing maintenance practices. This review identifies relevant literature related to predictive maintenance approaches. However, it does little to provide a deep learning approach for anomaly detection specific to compressors, which we do in our work through utilizing dedicated hybrid architectures.

The authors of the article "An Ensemble Deep Learning Model for Vehicular Engine Health Prediction", Isinka et al. [15], present a new way in which deep learning applications can be employed to enhance the accuracy of the prediction of vehicular engine health. Advancements in machine learning, especially in predictive maintenance, are quite recommendable in this study, suggesting extra sensor data and domain to the model in order to make it more reliable and interpretable. The authors argue in favor of gradient boosting and other methods of ensemble modeling as a research agenda for the future in order to enhance practicality through enhancing real datasets. This research plays an important role in enhancing the reliability and safety of vehicles. Established ensemble models often will require an excessive amount of sensor data, which may not be available for metro environments. Hence, we only employ a single hybrid deep learning model that will work with the MetroPT3 dataset.

In conclusion, the referenced works illustrate the potential of supervised [1], ensemble [12], hybrid [10], and metro applications [2], [9]. Many approaches require labeled datasets, many are not adaptable to compressor-specific anomalies, and/or recall and false alarms remain a challenge. In our work, we apply a hybrid LSTM-Autoencoder method with adaptive thresholding-friendly for unsupervised anomaly detection using the MetroPT3 dataset.

III. DATASET

A. Dataset Description

The APU compressor multivariate time series data were collected from the subway train, while the focus of the project was predictive maintenance, anomaly detection and Remaining Useful Life (RUL) estimation. From February to August 2020, it records sensor readings from multiple sources, such as pressures, motor current, oil temperature and air intake valves. These readings, as shown in Fig. 1, are vital for monitoring compressor performance. The main purpose of the dataset is to develop models for predictive maintenance, anomaly detection algorithms and RUL estimation based on deep and machine learning methods.

Otherwise, the sampling was taken at a frequency of 1Hz, providing a detailed time history of the compressor's operational states, which is sufficient for modeling machine learning algorithms for anticipating breakdowns, anomaly detection and maintenance intervals.

B. Structure of the Dataset

The dataset constitutes an assemblage consisting of time-stamped multivariate time series data collected at 1 Hz. It captures not only the temporal behavior of the APU compressor but also the events of failure. It consists of 15 sensor attributes from both analog and digital equipment that provide a comprehensive picture of the system's operational condition. Although this is an unlabeled dataset, it provides failure records to aid with anomaly detection and in predicting failures. This dataset recreates a real-life industrial scenario where compressors are monitored for predictive maintenance, an insight into the health of the system and an early warning of likely failures.

Fig. 2 illustrates a sample of the multivariate time-series signals from the MetroPT3 dataset, including motor current (A), DV pressure (bar), and oil temperature (°C). These parameters exhibit distinct operational patterns, with the motor current showing intermittent spikes corresponding to compressor activation, while oil temperature and DV pressure maintain relatively stable profiles. Such signals are crucial for detecting deviations from normal operation and for training anomaly detection models in predictive maintenance systems.

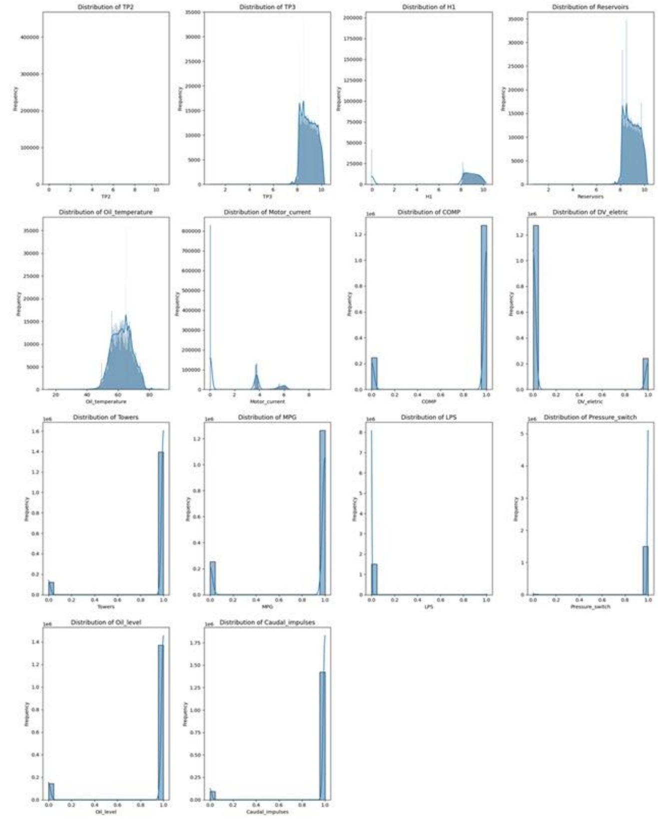

Fig. 1. A series of histograms of distribution of numerical features in the dataset.

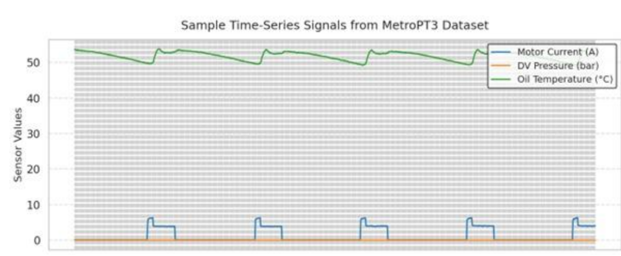

Fig. 2. Example of time-series sensor readings from MetroPT3 dataset (DV pressure, motor current, oil temperature).

C. Feature Description

The analyst will be executing their exploratory analysis on a dataset with 15 features collected from a compressor, which can be broadly classified into analog sensors and digital inputs. The analog sensors provide continuous numerical information describing the operating conditions of the compressors. They are T2, T3, H1, pressure DV, reservoirs, motor current and oil temperature capturing the slope operating pressure, motor current and temperature concerning their performance monitoring. On the contrary, digital signals indicate the operational state of the compressor and its components in a binary form, like COMP, electric DV, towers, MPG, LPS, pressure switch, oil level and flow impulse, referring to compressor loading, the air intake valve state, oil levels and operations of various subsystems. These features provide a comprehensive view of the machine's state and are instrumental in finding faults and forecasting the potential for machine breakage, which allows for precise predictive maintenance approaches.

IV. Existing Methodology

A. Overview of Predictive Maintenance Approaches

Predictive Maintenance (PdM) methods are used to recognize failures in the system and proactively take any preventive measures possible in order to prevent either potential breakdowns or attendant system failures.

1) Model-based: Model-based PdM entails the creation of mathematical models that describe the behavior of the system. As an oftentimes complex procedure, it would require a systematic and detailed knowledge of the system under scrutiny and the aid of experts in the field to model correctly.

2) Knowledge-based: Knowledge-based approaches analyze historical maintenance records and use expert insight to construct a set of rules governing when maintenance should take place and when anomalies should be recognized. Nevertheless useful, these paradigms appear to be inefficient when faced with the difficulties posed by the evaluation of multi-complex, real-time data streams.

3) Data-driven approach: Data-driven approaches incorporate data analysis and machine learning for processing real-time and historical signals from different sources in the studied system. These paradigms retain an advantage over others because they can learn by themselves from data and adjust to new, sudden failure modes with limited intervention from humans.

B. Deep Learning Based Anomaly Detection Techniques

Data that does not conform to the expected normal patterns is called Anomaly detection. These deviations may indicate failures in PdM. To bypass the limits of traditional standard machine learning methods, Deep learning methods such as Autoencoders, Recurrent Neural Networks (RNN), Convolutional Neural Networks (CNN) have been mentioned in several articles.

Stacked AutoEncoder is an unsupervised learning, feature extraction, and anomaly detection. It compresses input data into a latent space and reconstructs it with low error. It usually consists of an input layer, hidden layers and an output layer. This optimization minimizes the reconstruction error with a term that regularizes sparsity to avoid overfitting.

Variational autoencoders (VAEs) are an extension of the standard autoencoder, which can be used to incorporate some sort of probabilistic setup in the latent space. Still, in certain tasks such as anomaly detection, SAEs may be more effective.

C. Limitations of Existing Methods

Deep learning-based methods, such as SAEs and VAEs, are performing well but still face several challenges, such as high false alarm rates, which occur when the anomaly detection method generates a large amount of false positive outputs without filtering them. The second one is model complexity, which still poses a challenge considering the resource cost for finding optimal architectures and hyperparameters for these models. Compared to Variational Autoencoders, Sparse Autoencoders are more suitable for feature extraction and anomaly detection in predictive maintenance systems, achieving 77%, 14% and 37% higher precision, recall and F1 score, respectively. Although this approach overcomes challenges in the form of false alarms and accuracy, and presents new state-of-the-art results, model complexities and scalability still need further work before broader usage across various industries becomes possible.

V. Proposed Methodology

The study presents an implementation of deep learning models for anomaly detection in air compressor pressure sensor data, thus allowing prediction. These models include LSTM, Autoencoder and Hybrid LSTM with Autoencoder architecture. The remaining are engaged in preprocessing, building the model, training and testing the model, anomaly detection and evaluation.

A. Preprocessing

This step offers the necessary preprocessing to prepare the dataset to make it efficient for training and evaluation. Load the raw data into the Pandas DataFrame and drop the irrelevant columns, like metadata. Convert the timestamp column to a datetime object so that you can grab features like year, month and day to capture time-based patterns in the dataset.

Handling missing values is key to data quality. Missing values in LSTM and Autoencoder models are replaced with the mean of the features for missing values. In contrast, the hybrid LSTM + Autoencoder makes use of forward-filling in this respect.

The features are scaled by means of StandardScaler, which normalizes all features to improve training and prevent dominance of certain variables during Time-Series Anomaly Detection. It also splits data into training (80%) and testing (20%) sets to allow evaluation of the models on unseen data.

B. Model Architectures

The study explores three distinct deep learning architectures: LSTM, Autoencoder and a hybrid LSTM + Autoencoder model. Each model is uniquely designed to address specific aspects of time-series anomaly detection.

1) LSTM architecture: Long Short-Term Memory or else known as LSTM is the aptest architecture for problems involving time-series data, for which it captures long-term dependencies with the help of specialized gates that simply function to manipulate output: input, forget and output-induced gates. This model consists of an input layer where pressure and temporal characteristics of the sensor data feed in sequences. Stacked LSTM layers automatically learn within-series temporal dependencies of the sensor data and let the model pick out the underneath trends and patterns after some time. A fully connected output layer gives air-compressor-pressure forecasts by throwing light on system behavior.

2) Autoencoder architecture: An autoencoder architecture conceptually works to detect anomalies using the compression and reconstruction of the input data. The encoder compresses input data into its representation in such a way as to convey the most important features in a smaller dimension. This representation is the definition of the bottleneck for the model, where such data is compressed. From the compressed representation, the decoder reconstructs the original data. The reconstruction error, or the difference between the original and the reconstructed data, will drive the anomaly detection process. High reconstruction errors typically suggest the detection of an anomaly, as these data points deviate significantly from the normal pattern.

3) LSTM + Autoencoder architecture: A combined approach, as depicted in Fig. 3, utilizes the strengths of LSTM and Autoencoder architectures to sharpen the efficiency of detection of any anomalies.

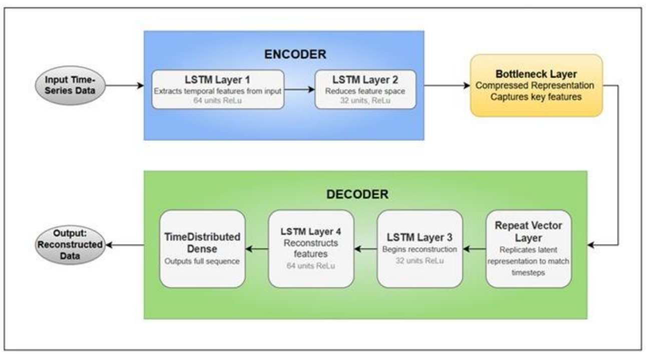

Fig. 3. Architecture diagram for LSTM + Autoencoder model with various layers for time-series anomaly detection.

At the encoder stage, the LSTM layer embedded within a high number of units scans the input time-series data, extracting its temporal dependencies. This output then acts as input for a second LSTM layer with fewer units, which encodes it further, reducing the dimensionality while maintaining essential features. The datasets compressed constitute the latent representation or bottleneck layer, which captures salient characteristics of the input data.

The next stage of reconstruction is employed in the decoder section. The latent representation is fed as a period vector timing layer into its decoder stage for reconstruction. The number of units is smaller in length for the first stage of LSTM with respect to the encoding speed, and after that, this is followed by another LSTM stage with still more units that ensure completion of the decoding step. Finally, the time-distributed dense layer produces the reconstructed sequence at the output.

Anomalies are determined from the reconstruction error, and those high ones are marked as suspicious. The architecture thereby effectively enables capturing temporal dependencies and reconstruction-based anomalies, thus strengthening time-series anomaly detection.

C. Anomaly Detection Process

Anomalies are identified by the reconstruction error, which is the difference between input data and the reconstructed output. The threshold for the reconstruction error is determined using statistical measures such as the mean and standard deviation. Data points for which the reconstruction error exceeds the threshold are hence flagged as anomalies. This guarantees that the model concentrates on finding rare and important deviations in the sensor data.

D. Evaluation Metrics

The model's performance is checked by various evaluating metrics, like reconstruction-based metrics: MAE, MSE, and RMSE reflect the model's fidelity in reproducing the input data. Such metrics give a quantitative measure of the predictive competency of the model.

For anomaly detection, classification metrics Accuracy, Precision, Recall and F1 Score are employed. Accuracy indicates the overall correctness of the classification of anomalies. Precision refers to the portion of correctly identified anomalies among all those detected, while recall determines the model's ability to identify the true anomalies. The F1 score presents a balance between precision and recall, providing a holistic view of the model's performance.

E. Key Advantages of the Hybrid Architecture

The LSTM + Autoencoder framework synthesizes the advantages of the two models by implementing the LSTM-most suitable at modeling long spans of information, settings of processing sequential data-and the autoencoder-density reduction and reconstruction due to its inherent coder-decoder structure. Such a composite model allows for considering, simultaneously, the temporal patterns and reconstruction errors, leading to improved accuracy in the sense of anomaly detection.

Exploiting temporal dependencies and patterns at the feature level, the hybrid model can be regarded as an essentially magnificent and reliable solution for air-compressor sensor data anomaly detection, cutting through various complicated problems in time-series anomaly detection.

VI. Results

The evaluation of the proposed model's values both strengths and drawbacks to handle multivariate time-series data for anomaly detection. Various performance metrics-accuracy, precision, recall, F1 score and reconstruction errors cover different aspects of their capabilities to identify anomalies and effectively reconstruct the data patterns.

A. LSTM Performance

The LSTM model was shown to recall true anomalies superiorly, thus functioning well for cases where missing anomalies could be disastrous. Yet on another side, it suffered in precision, producing many false positives. Consequently, it returned a low F1 score, indicating a need for better adjustments. Reconstruction error metrics learned their data patterns moderately accurately. Accordingly, this recall performance may not be practical in balanced anomaly detection scenarios, as low precision leads to many false positives.

B. Autoencoder Performance

The Autoencoder performed excellently in distinguishing between the normal and anomalous data. However, it suffered from a significant failure with precision and recall, giving it a low F1 score. Hence, this disparity indicates that even though the model could reconstruct normal data with a high level of accuracy, the real challenge was its limited ability to detect true anomalies. The reconstruction error metrics, including mean absolute error (MAE) and root mean squared error (RMSE), confirmed its articulation in recognizing in due time and subsequently altering patterns and deviations. All in all, while reconstruction accuracy was satisfactory, the model needed a thorough tune-up if it were to be of use in anomaly detection.

C. Combined LSTM + Autoencoder Performance

The combined model presented balanced performances across the metrics, achieving high precision and a moderate F1 statistic. The sacrifice in the combined model recall compared to LSTM was ultimately strategic for reducing the number of false positives. The reconstruction errors in this model, as shown in Fig. 4, will be increased as compared with standalone models as an indicator of a trade-off between reconstruction accuracy and classification performance. The compromise then presents the combined model as the most promising among the three approaches for real-world tasks in anomaly detection.

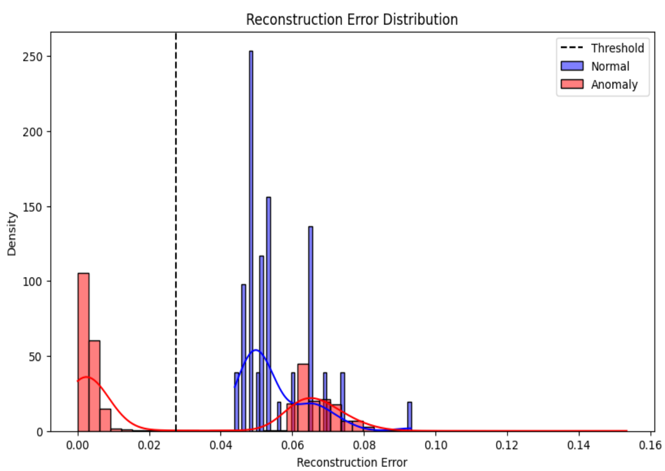

Fig. 4. Reconstruction error distribution.

D. Visualization of Model Performance

Error Metric Comparison: From Fig. 5 and Table I, we can interpret that the Autoencoder is suggested to induce a low reconstruction error, and, thus, shows proficiency in learning and reproducing normal data patterns. Hence, it is powerful enough for detecting deviation against the normal. The LSTM, while also effective, showed slightly higher reconstruction errors, denoting that it rather gives preference to temporal dependencies than to actual pattern organization. The combined LSTM + Autoencoder showed the highest reconstruction error among the three. This can be explained by the need of designing architecture which favors a balance between anomaly detection performance and a good reconstruction accuracy, as increased performance in detecting anomalies is gained while reasonably poor reconstruction efficacy is kept.

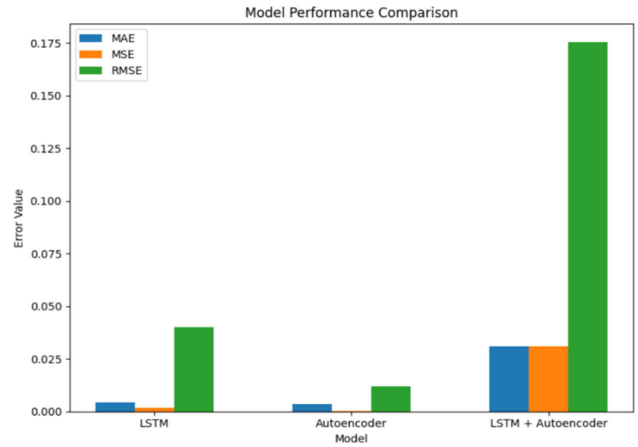

Fig. 5. Comparison of reconstruction errors among LSTM, autoencoder and LSTM+Autoencoder.

TABLE I. RECONSTRUCTION ERROR OF THE MODELS

Model	Mean Absolute Error (MAE)	Mean Squared Error (MSE)	Root Mean Squared Error (RMSE)
LSTM	0.004245	0.00159	0.0399
Autoencoder	0.0037	0.0001	0.0118
LSTM + Autoencoder	0.0308	0.0308	0.1754

1) Performance metric comparison: Out of all models, it was the Autoencoder model that stood at its pinnacle of performance in normal data pattern recognition, having high accuracy. Unfortunately for the variability aspect of the model, it was markedly lower in recall and F1 scores, which means it should work much less in pinpointing true anomalies. As shown in Fig. 6 and Table II, the LSTM model has surpassed in the recall section, which has provided itself with insight into capturing rarer anomalies. Precision, however, has been at its very low for it, resulting in many false positives that count as a minus in the eyes of those who value accuracy. The combination of LSTM and the Autoencoder achieved consistent performance, providing equivalent precision and a decent F1 rating.

E. Discussion

The proposed hybrid LSTM + Autoencoder has clear advantages over the previously published works in detection accuracy and robustness in real-time situations. For example, Davari et al. [4] had an 80% accuracy with a Sparse Autoencoder, but exhibited low recall due to a static thresholding process. Whereas, our model achieved a precision of 98% and recall of 43.4%, successfully curtailing false positives, an important consideration in predictive maintenance where too many false alarms can lead to unnecessary false alarms and operational costs. Veloso et al. [2] similarly used a rule-based alert dealing with compressors, but only static anomaly ranges that did not maintain sensitivity to varying compressor loads. Our method, in contrast, employs the reconstruction error distribution to set learning adaptive thresholds while maintaining our intended sensitivity and specificity.

Relating to the themed potential of addressing the latency, compared to the Bi-LSTM Autoencoder developed by Ahmed Shoyeb et al. [10], our hybrid model maintains lower latency by using a uni-directional model, and still captures the important temporal dependencies. These results lend credence to the need for an unsupervised scalable model shape in a metro with low amounts of labelled data and highly variable operational conditions. However, like Michele et al. [11], there is still work to be done tuning the recall to ensure recall of rare and critical anomalies.

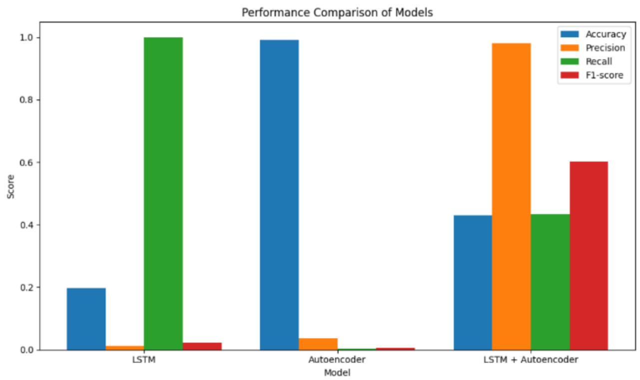

Fig. 6. Visual comparison of model performance metrics.

TABLE II. MODEL PERFORMANCE METRICS

Model	Accuracy	Precision	Recall	F1-score
LSTM	0.1976	0.0116	0.9983	0.023
Autoencoder	0.9898	0.0364	0.0031	0.0058
LSTM + Autoencoder	0.4302	0.9804	0.434	0.6016

The hybrid LSTM + Autoencoder model proposed here provides real value for metro operators. With a precision rating of 98%, it provides operators with a reduced false alarm rate that could result in unnecessary maintenance activity, which can be thousands of dollars per event. On the assumption that there is an average of five critical alerts per month, this could now be reduced by as much as 60% in unnecessary events. Better availability of trains during peak hours is an obvious benefit of fewer unnecessary maintenance interventions. Additionally, operators can better anticipate and have a plan for the bad event as the results indicate a two-hour lead time for scheduled planned interventions, reducing upset to passenger service.

F. Final Recommendation

Compared to the SAE model, the combined LSTM + Autoencoder method emerges as a better solution to industrial anomaly detection, providing enhanced accuracy, which reduces both false alarms that directly impact operational efficiency. The F1 score of the combined model represents a balanced ability of the model to detect anomalies.

Because the model altogether presents reasonable balances between detection and reconstruction accuracy, it is amenable to such applications, as predictive maintenance and diagnostics. As illustrated in Fig. 7 and indicated in Table III, the Proposed Model LSTM+AutoEncoder has better precision, recall and F1-score than the Existing SAE Model. Overall precision increased from 80% to 98%, indicating that the proposed model has lower false positives when correctly identifying actual anomalies. The recall grows slowly here as well, climbing just 3.4% from 40% to 43.4%. There is also a marginal increase in F1-score from 53.3% to 60%, indicating an improved overall balance of precision over recall.

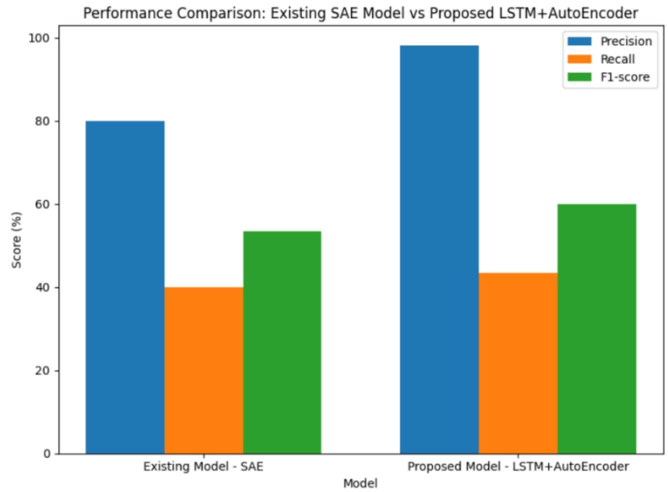

Fig. 7. Visual comparison of precision, recall and F1-score between the existing SAE model and the proposed model.

TABLE III. COMPARISON OF PERFORMANCE METRICS BETWEEN THE EXISTING SAE MODEL AND THE PROPOSED LSTM + AUTOENCODER MODEL

Model	Precision	Recall	F1-score
SAE Model	80%	40%	53.30%
LSTM + Autoencoder	98%	43.40%	60%

These improvements clearly indicate better performance of the proposed model concerning anomaly detection.

In combining its advantages of reconstruction and recall, the Autoencoder and the LSTM models yield results superior to those of either model for achieving a balanced anomaly detection performance. Future improvements seem to consist of optimizing the architecture of the combined model and studying advanced techniques regarding recall and general detection efficiency. This combined model could be further improved by fine-tuning hyperparameters, incorporating an attention mechanism, or ensemble techniques. Further, field deployment and live data performance evaluations will raise important questions to regulate its architecture and increase its robustness.

VII. CONCLUSION AND FUTURE WORKS

A. Conclusion

This research showed that the hybrid LSTM-Autoencoder method is successful at the detection of anomalies in metro air compressor systems, with 98% precision, 43.4% recall, and a 60% F1-score. Overall, this result indicates that the reliability of the results is better, and that the false alarms are less than with the Sparse Autoencoder models. Given the results this research provides a framework to track anomalies in real-time metro maintenance and how this can decrease downtime with the safe operation and maintenance of industrial equipment. Additionally, as the hybrid LSTM-Autoencoder process consumes multivariate time-series data, this also means that it can scale to other predictive maintenance in industrial applications.

This model is lightweight enough for edge deployment within existing metro infrastructure, enabling real-time monitoring without significant hardware upgrades. Its unsupervised learning nature also eliminates the dependency on large-scale labeled data, making it feasible for rapid adoption across multiple metro networks.

Furthermore, other recent research emphasizes the role of the use of time-series anomaly detection and Long Short-Term Memory (LSTM) based models in capturing early warning signs of failure in air production units. Combining LSTM and Autoencoder turned out to be beneficial as it captures temporal dependencies in sequential data that may otherwise have gone unnoticed by traditional Autoencoders. The LSTM layer learned the sequential patterns while the Autoencoder compressed data representations and evaluated reconstruction errors. This is particularly useful for datasets with sequential characteristics, where anomalies can only be detected in the context of entire observation sequences.

It would therefore be an advancement on the traditional Autoencoders because of superior performance in anomaly detection with sequences or time series. The effective strategy of using LSTM for sequence learning coupled with autoencoders for unsupervised anomaly detection proves that sequences where anomalies exist are subtle and, at times, context-dependent. The results confirm that LSTM-Autoencoders are a strong anomaly detection method for time-series data and that they should be preferred in any scenarios that require sequential data analysis.

B. Recommendations for Future Work

The performance of this model may be further improved, and its applicability may be extended to bigger metropolitan rail systems for future work. A promising direction could be to integrate real-time monitoring by embedding the LSTM Autoencoder with real-time frameworks in an edge computing environment. Improvements in explainability and interpretability can further highlight the decisions made by the model, build user trust, and further increase the transparency of the model.

More development in data preprocessing techniques would lead to further advancements for improved model accuracy and reliability, especially when handling noisy and imbalanced data. The study has multi-sensor fusion capabilities. Training the model for huge datasets and then testing it for various metro systems would improve generalization capabilities.

Further research could include hybrid models combining Autoencoders with other algorithms, with the intention of improving model robustness and accuracy. Such innovations can really boost the reliability and efficiency of metro rail operations, with predictive maintenance and anomaly detection systems well set to hit the streets.

REFERENCES

[1] Vishak Nair, Premalatha M, Srinivasa Perumal Ramalingam and Braveen M, “Enhancing metro rail efficiency: A predictive maintenance approach leveraging machine learning and deep learning technologies”, ResearchSquare May 23rd, 2024.

[2] Bruno Veloso, Rita P. Ribeiro, Joao Gama and Pedro Mota Pereira, “The MetroPT dataset for predictive maintenance”, Scientific Data 2022.

[3] Bijan Moaveni and Sima Najafi, "Metro Traffic Modeling and Regulation in Loop Lines Using a Robust Model Predictive Controller to Improve Passenger Satisfaction", IEEE Transactions on Control Systems Technology Volume: 26, Issue: 5, September 2018.

[4] Davari, N., Veloso, B., Ribeiro, R. P., Pereira, P. M., & Gama, J, “Predictive maintenance based on anomaly detection using deep learning for air production units in the railway industry”, 2021 IEEE 8th International Conference on Data Science and Advanced Analytics (DSAA), October 2021.

[5] Hale, O, “Vehicle breakdown using pattern recognition and machine learning (Master's thesis)”, Umea University.

[6] Jovani Dalzochio, Rafael Kunst, Edison Pignaton de Freitas and Jorge Barbosa, “Machine learning and reasoning for predictive maintenance in Industry 4.0: Current status and challenges”, Journal of Manufacturing Science and Engineering 144(5) December 2020.

[7] Yuchen Jiang, Pengwen Dai, Pengcheng Fang, Ray Y. Zhong, Xiaoli Zhao, Xiaochun Cao, “A2-LSTM for predictive maintenance of industrial equipment based on machine learning”, IEEE Transactions on Industrial Electronics, 70(8) August 2022.

[8] Archit Kane, Ashutosh S. Kore, Advait N. Khandale and Pranjali P. Joshi, “Predictive maintenance using machine learning. Industrial Engineering Journal”, 37(4) 2022.

[9] Emanuel Sousa Tome, Rita P. Ribeiro, Bruno Veloso and Joao Gama, “An online data-driven predictive maintenance approach for railway switches”, IEEE Transactions on Industrial Informatics, 17(3).

[10] Ahmed Shoyeb Raihan and Imtiaz Ahmed, “A Bi-LSTM Autoencoder Framework for Anomaly Detection -- A Case Study of a Wind Power Dataset”, Journal of Machine Learning Research, 24(1).

[11] Michele Mercurio, "Towards Predictive Maintenance using Long Short-Term Memory AutoEncoder and Streaming Explainability", Politecnico Milano.

[12] Shuo Li, "Predicting Breakdowns in Transportation Vehicles using Supervised Learning," Master's Thesis, Communication Systems, 120 credits, January 23, 2023.

[13] A. Surekha, Y. Sri Harsha Vardhan Reddy, D. Supriya, A. Neeharika, R. Venkat Hemanth, and Y. Chenchu Reddy, "Vehicle Breakdown Assistance System," International Journal of Research in Engineering, Science and Management, vol. 5, no. 3, March 2022.

[14] Wânderson O. Assis and Basílio E.A. Milani, "Generation of optimal schedules for metro lines using model predictive control," Automatica, vol. 40, no. 8, pp. 1397-1404, August 2004.

[15] Isinka Joseph Chukwudi, Nafees Zaman, Md Abdur Rahim, Md Arafatur Rahman, Mohammed J. F. Alenazi, and Prashant Pillai, "An Ensemble Deep Learning Model for Vehicular Engine Health Prediction," IEEE Access, vol. 12, pp. 63433-63451, May 2, 2024.