International Journal of Research Publications in Engineering, Technology and Management (IJRPETM)

|www.ijrpetm.com | ISSN: 2454-7875 | editor@ijrpetm.com |A Bimonthly, Peer Reviewed & Scholarly Journal|

||Volume 7, Issue 6, November – December 2024||

DOI:10.15662/IJRPETM.2024.0706026

AI-Augmented Big Data Frameworks for Smart Urban Infrastructure Maintenance Prediction

Mallesham Goli

Independent Researcher, India

mallesham.goli.research01@gmail.com

ABSTRACT: Innovations in artificial intelligence (AI) and predictive analytics are driving the convergence of big data, sensor networks, the Internet of Things, and cyber-physical systems into environments known as smart cities. By augmenting conventional city information systems with predictive modeling capabilities that exploit the wealth of high-frequency operational data generated by sensor-equipped urban infrastructure assets, it is possible to implement proactive preventive maintenance strategies that promise to enhance infrastructure robustness, safety, performance, sustainability, and cost-effectiveness. Such strategies, termed predictive maintenance, exploit predictive models of asset health (measured via condition indicators), typically learned using supervised learning techniques, and create actionable information to optimise maintenance scheduling and better allocate limited maintenance budgets.

Despite considerable interest in predictive maintenance research, most efforts have focused on specific domains (e.g., buildings, roads, bridges, tunnels) in isolation. Thus, a big-data framework architecture that provides for the integrated prediction of asset health indicators across different urban infrastructure systems and their underlying sub-systems is developed. In the proposed framework, an architecture for the ingestion and integration of diverse high-dimensional data types is presented, together with a scalable ensemble-learning-based approach for feature engineering that exploits the multi-variate time-series nature of the data generated by real-world monitoring systems. By providing a generic architecture for the predictive maintenance of infrastructure assets, the framework lays the foundations for effective predictive maintenance strategies in smart cities.

KEYWORDS: Smart cities; Big data; Artificial intelligence; Urban infrastructure; Predictive maintenance.

I. INTRODUCTION

Smart urban infrastructure generates vast quantities of operational data, which can be structured and analyzed to predict the failure of critical components, allowing maintenance to be performed just-in-time rather than just-in-case. When combined with advanced supervised learning techniques, such predictive models yield a high level of accuracy. Sensor data streams from similar assets in related locations can, however, exhibit fundamental differences, so that transfer learning techniques may be required to tune predictive models for specific locations. All smart cities face a similar problem—continuous operation of highly critical assets that represent billions of dollars in investment and yet remain functional to a large degree beyond their original design life. Sensor data from these assets can be used to build predictive models.

In a move towards smarter urban infrastructure systems, cities are deploying sensors in an attempt to monitor the health of their assets. They are seeking the most effective methods to aggregate and fuse these data to glean useful insights from them at various levels of granularity. Smart cities are hence building big data solution platforms to ingest, store, catalogue, and distribute all sensor data from the range of infrastructure assets—hospitals, roads, drainage, buildings, and 12 others. The cumulative sensor data generated by such systems can be used in combination with the advances in augmented big data frameworks and artificial intelligence to provide predictive maintenance capabilities that reduce operational risk and maintenance costs through a predictive maintenance capability.

1.1. Research design

The development of a multi-layered architecture for big data systems designed specifically for a smart urban environment is explored in a comprehensive manner, identifying appropriate combinations of the constituent components and integrating them into a cohesive architecture. Each layer is built on considerations of the MapReduce geographic distribution and support for decentralized data acquisition, social and sensor networks driven by citizen adoption, and integration of technologies that make the construed architecture interoperable with national and

|www.ijrpetm.com | ISSN: 2454-7875 | editor@ijrpetm.com |A Bimonthly, Peer Reviewed & Scholarly Journal|

||Volume 7, Issue 6, November – December 2024||

DOI:10.15662/IJRPETM.2024.0706026

international infrastructures for data and services. The underlying science of a big data architecture supports the analytical processes enabling predictive maintenance of urban infrastructure: a big-data-supported framework for predictive maintenance integrates Machine Learning classifiers, ensemble techniques, and deep neural networks with feature-engineering approaches that create meaningful higher-level descriptors of infrastructure asset condition and health, providing sound bases for predictive maintenance models.

The dynamic nature of urban infrastructure condition and usage creates great challenges for its maintenance, usually reactive and determined by budget restrictions rather than condition assessments. A framework for predictive models considers transportation networks (bridges, pavements, tunnels), buildings, and utility networks, with radar, vision, and acoustic sensors enabling technology-supported analysis of conditions. Models for individual data sources, training and optimization of ensemble classifiers and deep neural networks, and case studies for predicting infrastructure health based on a combination of sources support the appositeness of the framework and its predictive-behaviour perspective.

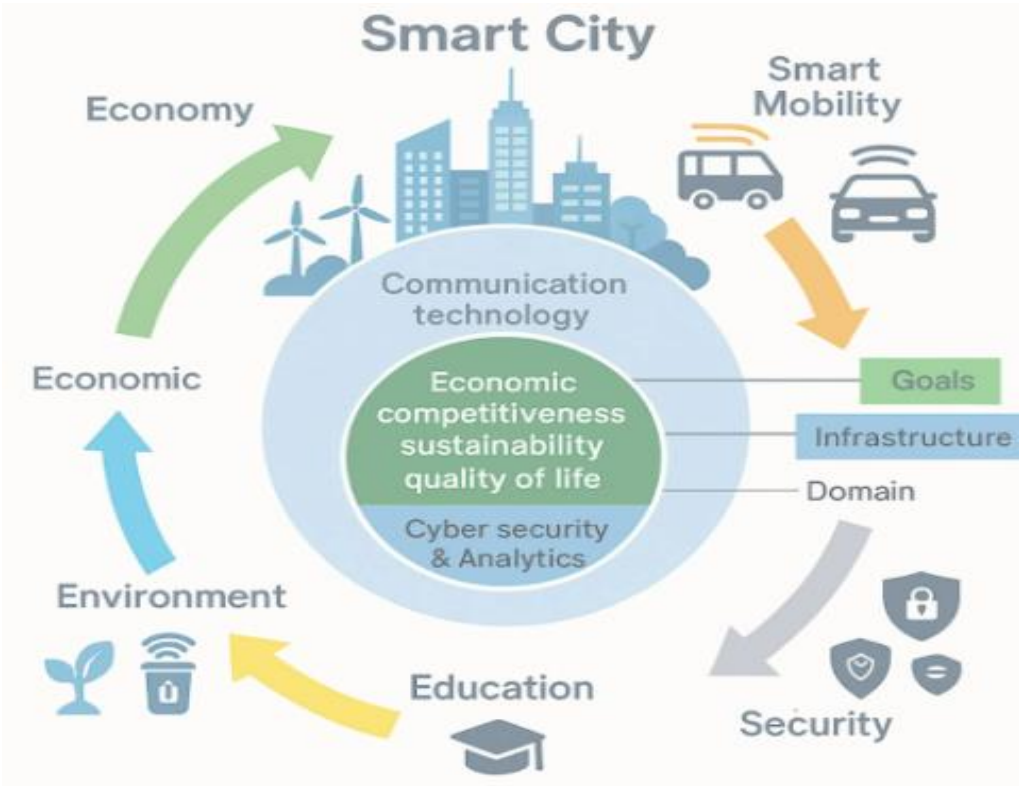


Fig 1: Leveraging artificial intelligence to enable sustainable urban development

1.2. Background and Significance

Big Data frameworks have changed the way data from urban environments is analyzed and leveraged to deliver new services to cities and their citizens. AI outcomes containing information on individual organizations, services, or real estate contribute to the Big Data landscape and create pressure for the construction sector, infrastructure service providers, and operators of urban transportation networks. Degradation is a common cause of emergency maintenance. Maintenance based on actual conditions is common in large industrial sectors such as aerospace, energy generation, and automotive. Common objective is to reduce maintenance or repair costs. The risk of failure at a critical time is low, so monitoring the risk of failure is crucial to creating services for this. Traffic assists diagnostic capabilities through intelligent transportation systems equipped with neural networks. Remaining useful life determination is achieved through LSTM networks. Predicting maintenance using a variety of data sources provides significant savings over traditional preventive maintenance by 20%. Predictive maintenance for telecommunications quality is based on Big Data and ML algorithms, which predict possible interruptions. These tools allow providing telephony, Internet and other services with a reliability index below certain levels.

International Journal of Research Publications in Engineering, Technology and Management (IJRPETM)

|www.ijrpetm.com | ISSN: 2454-7875 | editor@ijrpetm.com |A Bimonthly, Peer Reviewed & Scholarly Journal|

||Volume 7, Issue 6, November – December 2024||

DOI:10.15662/IJRPETM.2024.0706026

In this scenario, an AI-augmented Big Data framework is needed to provide predictive maintenance services. It integrates data from digital twins, sensors, SCADA systems, mobile applications, social networks, and external sources. Predictive models determine the health of infrastructure assets and consequently the risk of failure. The models reduce monitoring costs and help to plan maintenance actions and resource allocation. The architecture supports the continuous learning of the predictive models, ensuring quality in the information generated. The quality of the Information has a direct impact on these models.

Equation 1: Integrated data-fusion equation

Let

- S_t= sensor measurements at time t
- I_t= inspection data at time t
- O_t= operational/environmental data at time t
- M_t= maintenance/history data at time t

We want one combined state/input vector.

Step 1: Represent each data source as a vector

$$S_t = [s_{1,t}, s_{2,t}, \dots, s_{p,t}]^\top$$
$$I_t = [i_{1,t}, i_{2,t}, \dots, i_{q,t}]^\top$$
$$O_t = [o_{1,t}, o_{2,t}, \dots, o_{r,t}]^\top$$
$$M_t = [m_{1,t}, m_{2,t}, \dots, m_{u,t}]^\top$$

Step 2: Normalize each source

Because the sources have different units and scales, define standardized variables:

$$\tilde{s}_{k,t} = \frac{s_{k,t} - \mu_{s_k}}{\sigma_{s_k}}, \tilde{\imath}_{k,t} = \frac{i_{k,t} - \mu_{i_k}}{\sigma_{i_k}}$$

and similarly for $O_t$and M_t.

So the normalized vectors are

$$\tilde{S}_t, \tilde{I}_t, \tilde{O}_t, \tilde{M}_t$$

Step 3: Concatenate them

The simplest fused representation is just stacking all components:

$$X_t = [\tilde{S}_t^\top, \tilde{I}_t^\top, \tilde{O}_t^\top,$$

II. THEORETICAL FOUNDATIONS

The course of digitalization in the last decade has generated various research initiatives on the use of Big Data for the creation of digital twins of the urban environment and the associated data layers. The search for intelligent digital services in the urban environment is served by the concept of Urban Informatics, which looks at the demand side of urban data and how cloud-based Artificial Intelligence services can be served from Urban Big Data platforms. The development of data services for Intelligent Predictive Maintenance proposes Data Science models to predict the future state of Urban Infrastructure Assets within their respective Hierarchical Lifecycles through the aggregation of Urban Big Data, based on condition, time and usage information, with an architecture designed for urban infrastructure such as Buildings, Roads, Bridges, Airports, Underground Pipelines, Water Distribution Networks, Airports and Air Traffic, Waste Collection and Urban Mobility.

A specific Artificial Intelligence for Predictive Maintenance (AIPM) framework deepens research on the supply side of Urban Informatics, aimed at predicting the level of health of the assets that logically make up a city. AIPM describes formal and systematic methods of Data Science for the definition of predictive models of Urban Infrastructure Asset Health, and identify a corresponding architecture for Predictive Service oriented Urban Big Data Cloud Platforms. Big Data Ingestion and Integration are associated with Feature Engineering methods for the prediction of the Health state of Urban Infrastructure Assets involved in Transport Networks, in Buildings & Utilities, as well as in Airports and Air Traffic. The assessment of Urban Infrastructure through Health Gauge Models focusses on Performance Metrics for predictive models, thru Benchmarking & Case Studies, in relation to Predictive Maintenance Prives of Transport Networks, of Building & Utilities, as well as of Airports & Air Traffic.

|www.ijrpetm.com | ISSN: 2454-7875 | editor@ijrpetm.com |A Bimonthly, Peer Reviewed & Scholarly Journal|

||Volume 7, Issue 6, November – December 2024||

DOI:10.15662/IJRPETM.2024.0706026

2.1. Big Data Architectures for Urban Informatics

Existing Big Data architectures for urban informatics typically comprise six functional components: data ingestion, data storage, data management, data analytics and model development, data visualization, and data dissemination. Data ingestion initializes and triggers the data-handling cycle, directing urban data fusion in appropriate ways at appropriate times. A Big Data warehouse comprises storage tiers, often including hot, warm, and cold storage nodes, as in Hadoop, GlusterFS, Swift, and other cluster-based storage approaches. The next functional component, data management, deals with low-level data scenarios involving traditional data management methods—for instance, database operations (e.g., creating and populating tables)—and technologies such as relational, SQL-style, or NoSQL-based databases.

Artificial intelligence (AI) and machine learning (ML) support the analytical processes that enable knowledge discovery from complex, large-scale, uncertain, and real-time urban data. These processes include transformation operations, mapping (relating input to output data), data classification (associating data with predefined classes), learning (training predictive ML models), detection (identifying data points that correspond to rare patterns), forecasting (predicting future data values), and recommendation (making suggestions based on learned representations). These increasingly popular AI-based algorithms also require extensive use of data transformation techniques—often including feature extraction for time series and hypertext entities, image preprocessing, and network-to-graph conversion—to derive health indicators for predictive maintenance of infrastructure assets across Smart Cities.

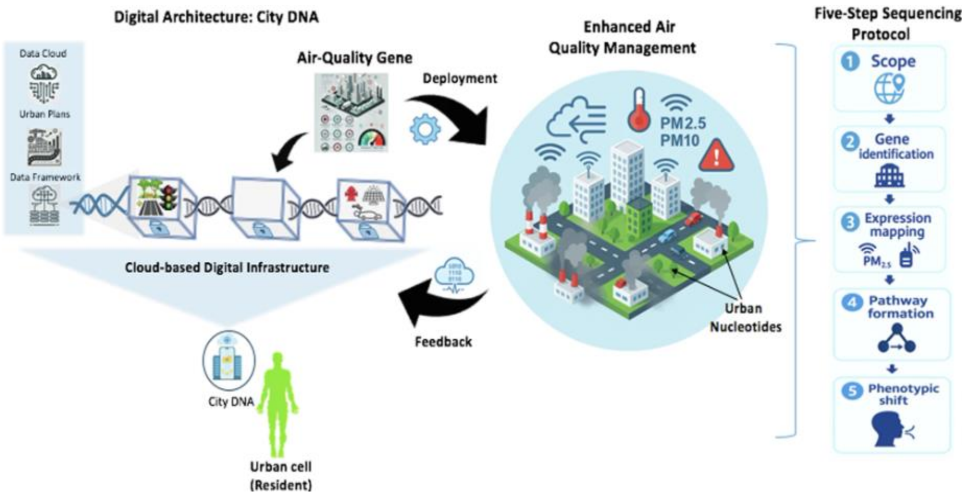


Fig 2: Big Data Architectures for Urban Informatics

2.2. Artificial Intelligence for Predictive Maintenance

Predictive maintenance in transportation networks, buildings, utilities, and other elements of smart urban infrastructure can be embedded into the cognitive service layer of these AI-augmented big data frameworks. AI is mature enough in particular domains to transfer knowledge from historical to new situations. Many of the typical applications of predictive maintenance, including supervised learning, reinforcement learning, or even structured knowledge representation and reasoning, can be directly executed in the context of urban informatics architecture. With good feature engineering, effective and accurate predictive models may also be automatically or semi-automatically generated, as long as data quality is sufficiently high. Many of these applications have been developed extensively in other areas, including industrial plants or robotics; therefore, the questions that remain mainly concern their scalability and interoperability across multiple assets, urban infrastructures, and cities.

Predictive maintenance enhances the resilience of smart urban infrastructures. The remaining lifetime of critical infrastructure assets can be estimated at large scale, warning functions can be implemented to alert infrastructure managers of potential problems, and service and maintenance scheduling can be optimized. A rich literature has been developed on the predictive maintenance of transport networks, energy and water utilities, buildings, and even urban systems in their totality. A special subdomain focuses on the different elements of physical coping capacity, whether clustering tools for event detection or models that quantify direct and indirect consequences and impacts on urban dynamics and economy. Applied predictive maintenance services cover particular domains, including road pavements,

International Journal of Research Publications in Engineering, Technology and Management (IJRPETM)

|www.ijrpetm.com | ISSN: 2454-7875 | editor@ijrpetm.com |A Bimonthly, Peer Reviewed & Scholarly Journal|

||Volume 7, Issue 6, November – December 2024||

DOI:10.15662/IJRPETM.2024.0706026

tunnels, railways, bridges, dams and levees, underground infrastructure, buildings, wind farms, sea ports, and urban systems as a whole.

Equation 2: Health indicator / health index equation
Let X_t be the fused input from Equation 1.
We define a scalar health index H_t such that:

- $H_t \approx 1$: healthy
- $H_t \approx 0$: failed or severely degraded

A standard derivation is a weighted linear score followed by a bounded transformation.

Step 1: Linear degradation score
Assume each feature contributes to degradation with weight w_j:

$$z_t = w_0 + \sum_{j=1}^{d} w_j \, x_{j,t}$$

where d is the number of features in X_t.

Step 2: Convert score to bounded health value
We want health in $(0, 1)$, so use the logistic function:

$$H_t = \frac{1}{1 + e^{-z_t}}$$

Substitute z_t:

$$H_t = \frac{1}{1 + \exp\left(-w_0 - \sum_{j=1}^{d} w_j \, x_{j,t}\right)}$$

III. FRAMEWORK ARCHITECTURE

In a big data environment, utility asset health indicators and predictive models need to be created from high-frequency data acquired from a variety of sources. Data ingestion and integration ingest and harmonise data streams from multiple sources, covering different distinct periods and locations while providing a level of temporal and spatial disaggregation suitable for feature extraction. The zone and temporal scales of feature generation then match those of predictive modelling.

Three distinct types of predictive models are considered. The first predicts the remaining useful life of a utility infrastructure asset and governs the maintenance decision of just-in-time, condition-based maintenance. The second predicts the immediate remaining capacity of a transport infrastructure facility and governs the preconditions for permitting loading. The third identifies the upcoming risk of failure or loss of function for an urban building or utility system, enabling prioritisation for urban resilience. Predictive models are built, deployed, and used to provide alarm and warning messages at different time horizons.

3.1. Data Ingestion and Integration

Big-data systems typically rely on process pipelines as the main mechanism for ingesting data. Their most basic form consists of a sequence of processing stages, each of which performs operations on a batch of data that has been pulled from a preceding stage. Data to be processed is stored in a distributed file system or object store, and at each stage computation consists of drawing a subset of this data into memory, performing some computation on it, and writing the output back to external storage. Operational simplicity and fault-tolerant distributed execution are attractive properties of process pipelines, but their limitations—the absence of support for low-latency operations and for continuous streams—have given rise to more versatile ingestion frameworks.

In the context of predictive maintenance, a wide variety of data sources can be ingested and integrated. Sensors installed on infrastructure assets provide a stream of data on their real-time condition, operation, and environmental factors. Transport service providers typically maintain and publish historical datasets of service disruptions along with their causes, such as major weather events (flooding, snow, etc.), planned works, and accidents. Public authorities

International Journal of Research Publications in Engineering, Technology and Management (IJRPETM)

|www.ijrpetm.com | ISSN: 2454-7875 | editor@ijrpetm.com |A Bimonthly, Peer Reviewed & Scholarly Journal|

||Volume 7, Issue 6, November – December 2024||

DOI:10.15662/IJRPETM.2024.0706026

control datasets on accidents, transit times, and unavailability of connection lines, among others. Similar datasets are available for utility networks and buildings. Data on operational costs and service level agreements can be obtained as well. The availability of this wealth of information allows the modeling of asset health status and long-term predictions of remaining useful life and residual life.

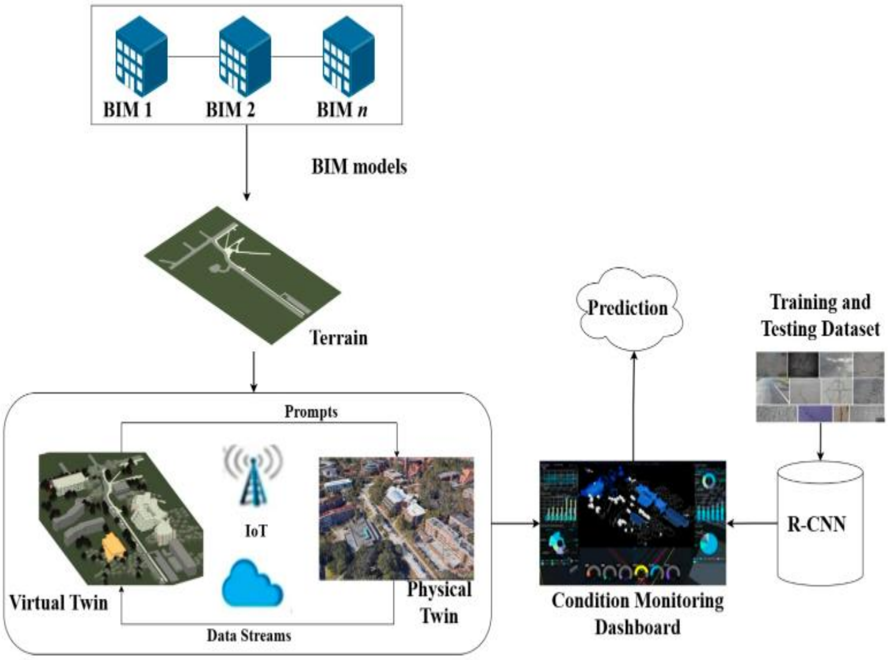


Fig 3: Cognitive Predictive Maintenance of Urban Assets Using City Information Modeling

3.2. Feature Engineering for Infrastructure Asset Health

As various types of big data are collected from the environment and urban infrastructure, healthy state indicators for predictive maintenance applications should be engineered by identifying the relevant data sources and implementing feature engineering techniques in the context of supervised or unsupervised machine learning. Indicative health states are often represented by either condition indicators or health indices, by using labelled data, e.g. in a transportation network, recorded BMS traffic data can be related to current pavement conditions by its surface distresses or cracks using supervised machine learning techniques such as Random Forest and/or Neural Network classification. On the other hand, asymptomatic sensory data, such as the datastream collected by pH temperature sensors in underground drainage networks with initial no-history of defect (not-default), can be related to the actual cracks deterioration process using Survival Analysis and other non-discriminative approaches to derive an overall degradation pattern. When these health states, based on physical performance of the infrastructure assets, are defined through predictive maintenance algorithms, these are integrated in real systems to assess their deviation levels and recognised as one of the eight designations suitable for use in hybrid predictive maintenance models, aimed at predicting imminent failures.Failure prediction and condition-based approaches have been successfully hybridised with survival analysis, to reduce the number of system-calibration cycles and increase its relevance in a real-world scenario.

Towards a more systematic approach to predictive maintenance feature engineering, the literature has undertaken a special focus on building and utility systems, relying on the advanced capabilities provided by the ITS for real-Time Quality Monitoring of high-dimensional data in order to generative condition indicators for the networked smart buildings and hybrid-pipe drainage systems.Built on the existing background for building and utility systems, simplifying assumptions have been made on the remaining urban infrastructure domains, i: (i) data-driven quality-attestment and -gated modelling and control of high-dimensional sensing data, and (ii) infrastructure systems are integrated into the ITS but hygiene quality monitoring is not performed, allowing the direct integration of smart-drainage-system status information in infrastructure predictive-maintenance feature engineering.

|www.ijrpetm.com | ISSN: 2454-7875 | editor@ijrpetm.com |A Bimonthly, Peer Reviewed & Scholarly Journal|

||Volume 7, Issue 6, November – December 2024||

DOI:10.15662/IJRPETM.2024.0706026

Equation 3: Remaining Useful Life (RUL) equation
Define:

- T_f= failure time
- t= current time
- $RUL(t)$= time left before failure

Step 1: Definition of RUL
By definition,

$$RUL(t) = T_f - t$$

That is the most basic equation.

Step 2: Connect failure time to health threshold
Suppose failure occurs when health drops to a critical threshold H_{crit}.
If current health is H_t, and health degrades approximately linearly:

$$H(\tau) = H_t - \lambda(\tau - t)$$

where $\lambda > 0$is the degradation rate.
At failure time $\tau = T_f$,

$$H(T_f) = H_{\text{crit}}$$

Substitute:

$$H_{\text{crit}} = H_t - \lambda(T_f - t)$$

Rearrange:

$$\lambda(T_f - t) = H_t - H_{\text{crit}}$$
$$T_f - t = \frac{H_t - H_{\text{crit}}}{\lambda}$$

Since $RUL(t) = T_f - t$,

IV. APPLICATIONS IN URBAN INFRASTRUCTURE

For smart cities, AI-augmented predictive maintenance is primarily applicable to two categories of built infrastructure: transportation networks and supporting buildings and utilities. While the rationale for predictive maintenance is evident for both, differences in maintained system availability give rise to different applications.

Urban transportation networks—roads, bridges, tunnels, railways, subways, airports, and seaports—are critical components of smart cities, enabling the safe and timely transport of people and goods in a cost-efficient manner. Investing in their maintenance has been found to be economically superior to underinvestment (with unnecessary repairs) or overinvestment (with excessive closure costs). Because travel is a necessity for city dwellers, maintenance activities must be perfected to be invisible to users. Consequently, at least part of the maintenance function must be predictive, using simulated or analytical models augmented by actual performance data. Predictive models, usually based on machine learning, indicate potential failures in sufficient temporal advance to plan the necessary maintenance work optimally. Predictive models applied to asset health are a special case of predictive maintenance. Such models, based on data from sensors, inspections, and maintenance history, infer the remaining useful life of a specific asset—be it a road, tunnel, bridge, or railway—in sufficient advance to plan the necessary maintenance work without unnecessary closings.

While smart cities are being designed and built from the ground up, the residential and commercial buildings providing the city's ecology still require maintenance. The purpose of this maintenance is twofold: to avoid accidents and to ensure that the buildings continue to serve their functions. Built externally, the various critical services—electricity, telecommunications, water, drainage, sewage, and solid waste disposal—also require maintenance. The consumption of these services and the inconvenience posed by their interruption register the effectiveness of their predictive maintenance.

|www.ijrpetm.com | ISSN: 2454-7875 | editor@ijrpetm.com |A Bimonthly, Peer Reviewed & Scholarly Journal|

||Volume 7, Issue 6, November – December 2024||

DOI:10.15662/IJRPETM.2024.0706026

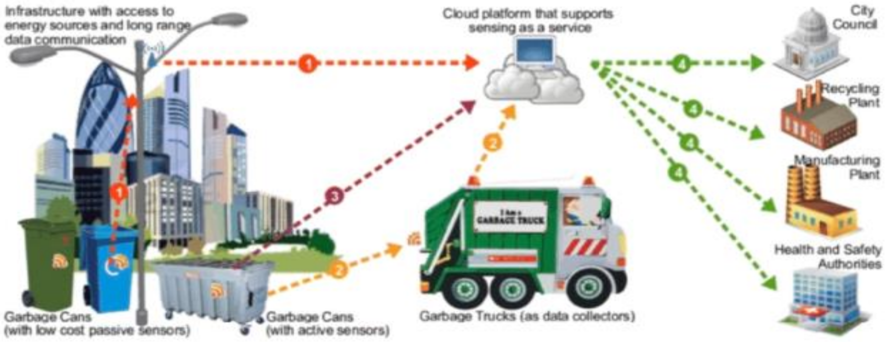


Fig 4: IoT for Smart Cities

4.1. Transportation Networks

Traffic-related infrastructure accounts for the highest proportion of big data generation among all groups of urban assets by a considerable margin due to the continuous monitoring of dynamic parameters, including status, consumption levels, occupancy, and even the structural health of critical components. Predictive maintenance models for urban transportation networks apply a variety of techniques on heterogeneous datasets, including decision trees trained on publicly available traffic data to estimate bus breakdowns, support vector machines on Internet of Things (IoT) big data to predict metro incidents, generalized additive models to assess the probability of subway tunnel flooding, recurrent neural networks on sensor data to forecast the failure of airplane landing gears, and deep learning on social media to classify disturbance impacts of the urban traffic condition. Regression models are also employed on sensor data to detect abnormal traffic levels for use in demand estimations, while an ensemble-based location-monitoring strategy on GPS data has been proposed for taxi positioning in urban areas where the signal is lost.

Smart traffic management has been further advanced with the integration of traffic data and vehicular videos. Traffic incidents such as lane blockage, accidents, and vehicles showing abnormal behaviors have been detected with extreme-gathering-based approaches, utilizing functional and structural properties of the urban traffic network. The construction of incident-response-aware policy sets for variational traffic flows has led to the use of traffic video data to enhance the detection of traffic incidents and to semantically classify the state of urban traffic by capturing unseen implicit relationships between traffic state and the operation of traffic control devices in the visual. Probabilistic models of incident detection, response, and clearance have provided a systematic decision-making framework for intelligent vehicular scheduling during incidents.

Equation 4: Failure-risk / warning probability equation

Let $Y_t \in \{0,1\}$, where

- $Y_t = 1$: failure occurs within a warning horizon Δ
- $Y_t = 0$: no failure within Δ

We want

$$P(Y_t = 1 \mid X_t)$$

Step 1: Use odds of failure

Assume the log-odds of failure is linear in the features:

$$\log\left(\frac{P(Y_t = 1 \mid X_t)}{1 - P(Y_t = 1 \mid X_t)}\right) = \beta_0 + \beta^\top X_t$$

Step 2: Exponentiate both sides

$$\frac{P(Y_t = 1 \mid X_t)}{1 - P(Y_t = 1 \mid X_t)} = e^{\beta_0 + \beta^\top X_t}$$

Let

$$p_t = P(Y_t = 1 \mid X_t)$$

Then

|www.ijrpetm.com | ISSN: 2454-7875 | editor@ijrpetm.com |A Bimonthly, Peer Reviewed & Scholarly Journal|

||Volume 7, Issue 6, November – December 2024||

DOI:10.15662/IJRPETM.2024.0706026

$$\frac{p_t}{1-p_t} = e^{\beta_0+\beta^\top X_t}$$

Step 3: Solve for p_t

$$p_t = (1-p_t)e^{\beta_0+\beta^\top X_t}$$
$$p_t = e^{\beta_0+\beta^\top X_t} - p_t e^{\beta_0+\beta^\top X_t}$$
$$p_t\left(1+e^{\beta_0+\beta^\top X_t}\right) = e^{\beta_0+\beta^\top X_t}$$
$$p_t = \frac{e^{\beta_0+\beta^\top X_t}}{1+e^{\beta_0+\beta^\top X_t}}$$

Equivalent form:

$$p_t = \frac{1}{1+e^{-(\beta_0+\beta^\top X_t)}}$$

4.2. Buildings and Utilities

Maintenance of buildings in urban environments faces the conflicting demands for ensuring safety and aesthetic appeal while promoting resilience and minimizing cost and energy use. The desire to diffuse knowledge of the structure's condition across the party entitled to carry out remedial actions necessitates making data on material deterioration available to a large number of stakeholders. Buildings also represent significant energy costs and carbon emissions, which are projected to continue to increase unless maintenance of the buildings is improved. AI techniques are enabling new maintenance practices based on building energy simulations, predictive control, police and fire departments' reports, social media postings, and occupant behavior and satisfaction feedback, and produce operational energy use feedback.

Modern utilities increasingly use a sensor-enabled wide-area monitoring system to detect, characterize, and locate developing events of physical damages using the information about currents at the sensors. Such information when combined with GIS data can support predictive maintenance of power transmission and subtransmission lines. However, sensors can malfunction and send false alarms that harm system reliability. AI classifiers have been developed that exploit the latent and higher-order correlations between data from multiple sensors to detect and identify fault conditions. The labelling of historical sensor data can be expensive when it is done by human experts and possibly unreliable when it is done by non-expert users or caused by non-fault conditions. The classification model can be improved by transferring knowledge from similar yet labelled classification problems, even with little labelled source domain data.

V. EVALUATION AND VALIDATION

The predicted health of infrastructure assets can be leveraged in several predictive maintenance applications. For a given asset class, a wide variety of possible predictive models can be trained relating the predicted health to the predicted time to failure. It is therefore essential to always evaluate the expected accuracy of a model before deploying it; this can involve splitting the dataset into training, validation, and target subsets, or employing k-fold cross-validation. Yet another approach is benchmarking against a set of models of totally different structures, e.g. the mean, a random predictor, or simple regression models based on the largest features. In addition, for real-time warning scenarios, different cumulative/density risk functions can be selected based on the underlying cost-benefit framework. Finally, for case studies such as within-group or across-group assessments, transfer learning methods can be adapted to urban infrastructure by making use of the temporal dimension, so as not to be constrained by the usually scarce availability of labelled data.

To enable dedicated AI-augmented big-data frameworks for predictive maintenance in urban infrastructure, it is crucial not only to develop supporting architectures, but also to ensure that training, applying, and evaluating the predictive models for long-term asset health and short-term time-to-failure forecasts are a simple plug-and-play exercise. Yet predictive maintenance on urban infrastructure remains an unexploited field of application for such frameworks. For these asset categories, several highly structured frameworks – for data ingestion and integration, predictive health assessment, predictive time-to-failure models, and final application of the whole structure – can be readily defined and adapted, thus paving the way for urban-infrastructure deployment of dedicated AI solutions supporting long-term predictive maintenance.

|www.ijrpetm.com | ISSN: 2454-7875 | editor@ijrpetm.com |A Bimonthly, Peer Reviewed & Scholarly Journal|

||Volume 7, Issue 6, November – December 2024||

DOI:10.15662/IJRPETM.2024.0706026

5.1. Performance Metrics for Predictive Models

Well-designed predictive models enable stakeholders in urban infrastructure to proactively anticipate asset health conditions and schedule maintenance actions in a timely manner. A model's accuracy and reliability are crucial since inappropriate maintenance scheduling may exacerbate an asset's already compromised capability and increase budget overruns. Examining the model performance metrics specifically in the context of predictive maintenance reveals that different measures are relevant depending on the maintenance strategy employed. Consequently, measures to evaluate predictive models for predictive maintenance, focusing primarily on predictive maintenance based on appropriate condition thresholds of monitored health parameters, can be viewed in three groups: classification performance measures, regression performance measures, and specialized predictive maintenance performance measures.

Classification performance measures are directly relevant to predictive maintenance strategies based on a binary decision that is performed when the health indicator enters the warning condition. Examples of measures in this category include accuracy, precision, recall, F1 score, as well as the area under the receiver operating characteristic curve and the area under the precision-recall curve. Regression performance measures apply when the decision on corrective actions is based on an explicit condition policy that defines specific values (e.g., "0") for maintenance actions. Examples include mean absolute error and root mean square error. Finally, specialized measures assess the actual overhead costs due to maintenance actions, allowing the evaluation of the models in a more realistic context. One example measure is the total expected cost (expressed in monetary units) incurred within a certain forecasting horizon.

5.2. Benchmarking and Case Studies

Rigorous evaluation and testing are fundamental to assess the effectiveness and reliability of framework capabilities. For predictive frameworks, performance is benchmarked using an array of model types and configurations based on learning methods (supervised, semi-supervised, weakly supervised and unsupervised), learning algorithms (decision trees, rule-based and ensemble), types of features (statistical, spectral, and model-based) and types and combinations of external data, to obtain models that best characterize the phenomena at hand, supported by an appropriate selection of performance metrics. In terms of applied case studies, the AI-augmented big data PF architecture is applied to urban road-maintenance management in the context of transportation network predictive maintenance and to the adaptive monitoring and risk-based inspection of building-integrity conditions in the context of building predictive maintenance.

Predictive models are developed to forecast road-pavement condition at specific locations and times in a network. Motor-vehicle-volume data from extensive secondary sources are possible covariates for the predicted time series of road-condition indicators. Sensor data from a mobile-phone-water-level survey serve to calibrate a parallel hydrological model. Indicators from the Bayesian Network of Bayesian Inference for Road Condition+ Bayesian Model Fusion model collectively formulate a clear heat map of predictive risk for a road segment over a four-day storm period. The proposed framework and supportive models and data sources enable responsive management by establishing a temporal prediction capability, by identifying critical locations and times, and by allowing risk-adaptive interventions.

Equation 5: Predictive-model training loss equation

Let

- $y_i \in \{0,1\}$: true label for sample i
- $\hat{p}_i = P(Y_i = 1 \mid X_i)$: predicted failure probability
- N: number of training samples

Step 1: Likelihood of one sample

For Bernoulli data:

$$P(y_i \mid X_i) = \hat{p}_i^{y_i}(1-\hat{p}_i)^{1-y_i}$$

Step 2: Likelihood of all samples

Assuming independent samples:

$$L = \prod_{i=1}^{N} \hat{p}_i^{y_i}(1-\hat{p}_i)^{1-y_i}$$

Step 3: Take log-likelihood

|www.ijrpetm.com | ISSN: 2454-7875 | editor@ijrpetm.com |A Bimonthly, Peer Reviewed & Scholarly Journal|

||Volume 7, Issue 6, November – December 2024||

DOI:10.15662/IJRPETM.2024.0706026

$$\log L = \sum_{i=1}^{N}[y_i \log \hat{p}_i + (1 - y_i)\log(1 - \hat{p}_i)]$$

Step 4: Convert to minimization objective
Machine learning usually minimizes negative average log-likelihood:

$$\mathcal{L} = -\frac{1}{N}\log L$$

Therefore,

$$\mathcal{L} = -\frac{1}{N}\sum_{i=1}^{N}[y_i \log \hat{p}_i + (1 - y_i)\log(1 - \hat{p}_i)]$$

VI. DEPLOYMENT CONSIDERATIONS

Prescriptive analytics and improved decision support for urban infrastructure depend on the proposed feature engineering, predictive maintenance, and health estimation of infrastructure assets. These components are embedded into a scalable and interoperable framework that can operate across urban domains, thus permitting composite predictions over networks of different asset types. Prescriptive analytics based on innovative feature definitions is assessed relative to commonly employed asset-repair algorithms, spanning data-rich domains where infrastructure condition is well known, such as custom bikeshare data applications, and headroom for automated intervention on decision-support systems for simpler, data-scarce domains, such as hospital networks, hydrocarbon supply, and utility outage mapping problems. Deployment considerations relate to real-time feasibility of prediction-model operations, the capacity of the integrated-framework architectural blueprint to support healthengineered descriptions of urban-assets and networks, and high-volume prediction-model development across domains. Scalability and interoperability are examined via the spatial-on-spatial operation of a city-scale data-epidemiological health model for cycling infrastructure.

The predicted asset-health measures can be exploited to suggest prescriptive maintenance solutions and greater reliability of damage-pattern transfer when training-and-test datasets are sampled to minimise the potential impact of the transfer-learning challenge. These data-analytic building blocks serve catchment areas of disparate sizes, unique infrastructures with different functional roles, and data volumes for which prediction-model construction rely on disparate training-and-test-sampling conditions. By contrast, beyond-catchment network-health estimation is performed using portable data-epidemiological machine-learning structures for which the health state must remain invariant across applied datasets.

6.1. Scalability and Interoperability

The predictive maintenance framework architecture has been designed specifically to support the demands of fast decision-making in operation-critical processes and to be deployed on a large scale across urban-scale infrastructure systems. These requirements are addressed through low-latency automated integration of the underlying AI-model Inference Engine – including its transdisciplinary feature-engineering and predictive-model development constituting submodules – into a big-data analytics architecture based on the Lambda architecture.

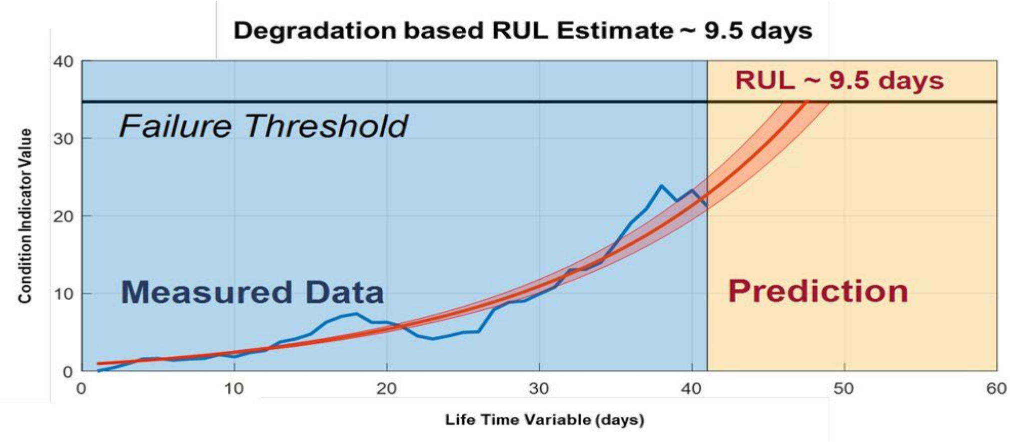
Signal ingestion is decomposed into three separate data paths: (1) real-time spatial–temporal event streams, (2) batch-processing external big data stores residing out-of-band of the urban system, and (3) precisely-timed sensor data streams supported by GPS, RF-ID or Sensor Web technologies. Typically, interconnected urban infrastructure systems create a globally-scaled data-sea showing strong spatial correlation, i.e., physical laws of nature allow signals to be freely propagated through space, thus forming an implicit data-sharing backbone. Signals flowing in one city can be shared as training/test data for predictive-maintenance task within other cities, even if these signals/conditions are unseen by the other cities. An architectural module is dedicated to enable such smooth data sharing across signals by leveraging Transfer Learning methods.

International Journal of Research Publications in Engineering, Technology and Management (IJRPETM)

|www.ijrpetm.com | ISSN: 2454-7875 | editor@ijrpetm.com |A Bimonthly, Peer Reviewed & Scholarly Journal|

||Volume 7, Issue 6, November – December 2024||

DOI:10.15662/IJRPETM.2024.0706026



6.2. Privacy, Security, and Compliance

Given the breadth of data types and sources in a Predictive Maintenance Big Data Framework, strong legal protections are required against unauthorized access, data breaches, and discrepancies. The GDPR in Europe and the Health Insurance Portability and Accountability Act in the United States are two noteworthy examples of laws intended to protect personally identifiable information (PII), health information, and accounts with financial institutions in transit, and in stasis. When such regulations apply, citizens are made aware of the extent of data collection in order to provide consent. Research suggests that transparency with respect to data collection builds trust and thus engenders compliance.

Governments and private corporations deviate from this best practice as they claim legalized exemption from, or ability to grant unsupervised immunity against, transparency provisions when the public interest is at stake. By necessity, urban predictive maintenance frameworks are involuntary participatory frameworks for data collection. Individuals, organizations, and governance bodies directly affected have little say in the balance of risks and rewards associated with the Big Data Against Humanity database. Unfair, unjust, or illegal use of the dataset poses substantial long-term reputational risks. Data owners, contributors, and consumers cannot be forced to negotiate Big Data contributions based on the contributions of others without leaving that portion of the database vulnerable to opportunistic exploitation.

VII. CHALLENGES AND FUTURE DIRECTIONS

A smart city is evolving into a massive ecological, hydrological, and meteorological system in which infrastructure development and the actual situation of cities and provinces are in a dynamic adaptive evolutionary relationship. The original watering and humidifying process of regions, water network systems, and others in cities and provinces has greatly dampened urban heat islandeffect and seasonal change. With the rapid growth of modern society, the intelligent management of urban public utilities playing crucial role in transportation networks, buildings, energy systems, underground pipes, etc. is wider used. Domestic and foreign studies on the intelligent management of urban public facilities are gradually deepening, with big data, the Internet of Things, artificial intelligence, and other technologies being used to enhance the intelligent management level of urban public facilities.

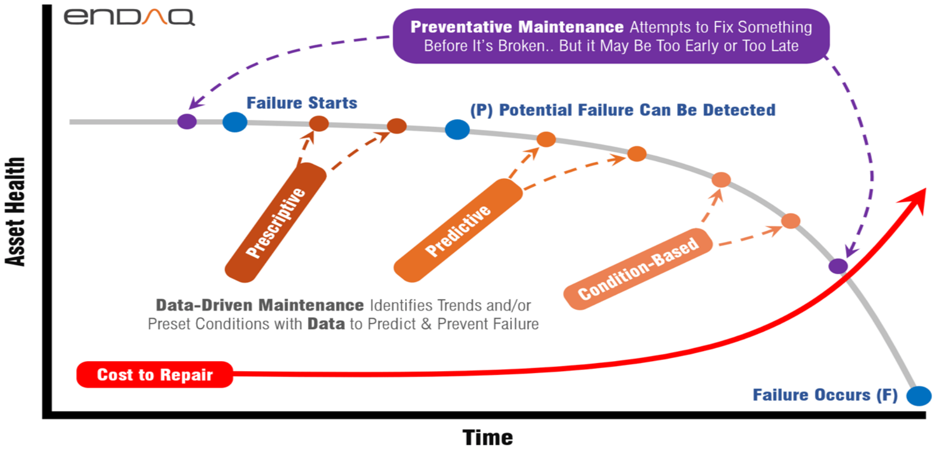
Inevitably, the application of these modern methods still faces many obstacles. Following aspects should be given high attention in future research, such as data sensing and collection decisions driving the algorithm application, sensor network failure, abnormal data detection and compensation and privacy security issues. The demand for automatic decision and evaluation under big data background through transfer learning in different space and time dimensions are important development directions in the near future.

International Journal of Research Publications in Engineering, Technology and Management (IJRPETM)

|www.ijrpetm.com | ISSN: 2454-7875 | editor@ijrpetm.com |A Bimonthly, Peer Reviewed & Scholarly Journal|

||Volume 7, Issue 6, November – December 2024||

DOI:10.15662/IJRPETM.2024.0706026



7.1. Data Quality and SensorReliability

The widespread adoption of AI technologies for predictive maintenance in smart urban environments relies on continuous monitoring via large numbers of sensors integrated into civil infrastructure. There are growing concerns that these data-generating sources introduce new challenges in terms of quality and reliability. Sensor-generated data are often unstructured and noisy, with uncertainties related not only to the measurement process and underlying environmental conditions but also to the sensors' own state of health on production, transport and installation, placement through correspondent level of redundancy. Real-world learning curve studies have shown that the transfer from the experimental to the operational phase of automated monitoring systems are inherently tricky. Moreover, sensor information must be qualitatively consistent and quanti-tatively representative to make automated predictive maintenance decisions reliable. The ongoing overexploitation of sensors for data generation raises additional issues regarding their availability and reliability. In practice, instrumentation reliability has not been critically monitored during the data generation phase. Although the aforementioned issues regarding the attendant data reliability, relevance and redundancy may have already been considered in prior work, they have not yet attained the attention necessary for widespread operational application of AI technologies. Therefore, the inscription of such phenomena in the predictive maintenance modelling framework constitutes a considerable change of paradigm: transfer learning techniques can now be adopted directly at the raw data level. Transfer from source to target domain becomes transfer among the various sensors negatively affecting the classifier sor neural net or predictive model employed.

7.2. Transfer Learning and Adaptation

In dynamic environments, models are often trained in one context yet deployed in another. In predictive maintenance of urban infrastructure, these scenarios are common; labeled data are available for one set of assets and potentially adapted to another, commonly used but rarely realized. A small set of the target assets need to be extensively labeled with ground-truth data to transfer the predictive model successfully to the target domain. In recent years, big data and AI architectures have initiated research in urban informatics and predictive maintenance of urban infrastructure. However, building scalable solutions is still a challenge due to largescale data ingestion from different sources. The AI engine enables model deployment, training, evaluation, and operation, and predictive maintenance modeling is distributed in different geographic locations and periods. Scalability issues have been addressed with distribution of model training, validation, and evaluation. Transfer learning and adaptation are becoming commonplace for supervised-learning solutions in supervised contexts, where the model is trained in one context (the source) and is used in another (the target). In predictive maintenance of urban infrastructure, this situation is common: labeled data are available for one set of assets and the model is then transferred to other assets of a different category for which maintenance history is either not available or remains short. A small set of target assets need to be extensively labeled with ground-truth data with time-consuming expert knowledge to transfer the model from one category to another using either transfer learning with source-trained parameters or transfer adaptation using different label mappings using direct optimal transport. This approach adapts a well-trained model on a source domain for maintenance while avoiding further labeling in a related yet different target domain.

Layer	Function	Technologies Mentioned	Role in Predictive Maintenance
Data Ingestion	Collect sensor & external data	IoT, APIs, pipelines	Real-time monitoring
Data Storage	Store large datasets	Hadoop, GlusterFS	Historical + real-time storage
Data Management	Organize data	SQL / NoSQL DBs	Data structuring

|www.ijrpetm.com | ISSN: 2454-7875 | editor@ijrpetm.com |A Bimonthly, Peer Reviewed & Scholarly Journal|

||Volume 7, Issue 6, November – December 2024||

DOI:10.15662/IJRPETM.2024.0706026

Layer	Function	Technologies Mentioned	Role in Predictive Maintenance
Data Analytics	Model building	ML, AI	Failure prediction
Visualization	Present insights	Dashboards	Decision support
Dissemination	Share results	Cloud platforms	Stakeholder communication

Table: Big Data Architecture Layers

VIII. CONCLUSIONS

Research on the use of big data techniques to ensure predictive maintenance of smart urban infrastructure has increased in recent years. However, a comprehensive framework architecture that addresses the different aspects of these applications has remained lacking. Urban data sources are highly heterogeneous, leading to several challenges in the data ingestion and integration process. Moreover, feature engineering for infrastructure asset health is often neglected or not performed satisfactorily, even when a predictive maintenance solution is proposed. The combination of missing or poor-quality features slows down model training or leads to unreliable predictions. Additionally, the inherent presence of historical data in urban systems supports AI techniques, which could be applied more broadly to analyze other data sources directly. To fill these gaps and address the limitations, the proposed framework architecture integrates the different components and stages of the process and guides the development of predictive maintenance solutions. The first impression of operating such a comprehensive AI-augmented big data framework reveals how well it captures the many different levels of complexity. The presented AI algorithms cover a wide spectrum of needs, offering valuable predictive information about the future of the infrastructure asset health. The availability of enriched data pipelines is instrumental to the operationalization of predictive maintenance solutions, facilitating predictive model training and improving predictive performance. Combining textual and visual data sources for the same infrastructure asset could foster knowledge transfer for detecting anomalies in less monitored assets.

REFERENCES

[1] Kummari, D. N., & Burugulla, J. K. R. (2023). Decision Support Systems for Government Auditing: The Role of AI in Ensuring Transparency and Compliance. International Journal of Finance (IJFIN)-ABDC Journal Quality List, 36(6), 493-532.
[2] Garapati, R. S. (2022). Web-Centric Cloud Framework for Real-Time Monitoring and Risk Prediction in Clinical Trials Using Machine Learning. Current Research in Public Health, 2, 1346.
[3] Mangalampalli, B. M. Intelligent Data Profiling for Healthcare Data Lakes Using AI-Enhanced Analytics.
[4] Kolla, S. H. (2023). Deep Learning–Driven Retrieval-Augmented Generation for Enterprise ITSM Automation: A Governance-Aligned Large Language Model Architecture. Journal of Computational Analysis and Applications, 31(4).
[5] Mangalampalli, B. M. Generative AI Applications In Healthcare Data Mart Design And Optimization.
[6] Mangala, N. (2022). Real-Time Data Quality Monitoring and Gating Frameworks in Cloud-Based Data Pipelines. International Journal of Research and Applied Innovations, 5(6), 8197-8219.
[7] Nagubandi, A. R. (2023). Advanced Multi-Agent AI Systems for Autonomous Reconciliation Across Enterprise Multi-Counterparty Derivatives, Collateral, and Accounting Platforms. International Journal of Finance (IJFIN)-ABDC Journal Quality List, 36(6), 653-674.
[8] Pamisetty, V., & Amistapuram, K. Smart Decision Support Systems For Dynamic Tax Policy Optimization Using Reinforcement Learning.
[9] Singireddy, J. (2023). Finance 4.0: Predictive analytics for financial risk management using AI. European Journal of Analytics and Artificial Intelligence (EJAAI) p-ISSN, 3050-9556.
[10] Mangala, N. (2021). Optimizing Large-Scale ETL Pipelines Using Medallion Architecture on Azure Data Lake. Journal of Artificial Intelligence and Big Data, 1(1), 1-20. https://doi.org/10.31586/jaibd.2021.1361
[11] Valiki, D., & Segireddy, A. R. (2023). Deep Learning Architectures Deployed on Cloud Platforms for Dynamic Financial Risk Evaluation and Market Prediction. American International Journal of Computer Science and Technology, 5(5), 12-24.
[12] Nandan, B. P. (2024). Semiconductor Process Innovation: Leveraging Big Data for Real-Time Decision-Making. Journal of Computational Analysis and Applications (JoCAAA), 33(08), 4038-4053.
[13] Singireddy, J. (2024). Deep Learning Architectures for Automated Fraud Detection in Payroll and Financial Management Services: Towards Safer Small Business Transactions. Journal of Artificial Intelligence and Big Data Disciplines, 1(1), 75-85.

International Journal of Research Publications in Engineering, Technology and Management (IJRPETM)

|www.ijrpetm.com | ISSN: 2454-7875 | editor@ijrpetm.com |A Bimonthly, Peer Reviewed & Scholarly Journal|

||Volume 7, Issue 6, November – December 2024||

DOI:10.15662/IJRPETM.2024.0706026

[14] Yandamuri, U. S. AI-Driven Decision Support Systems for Operational Optimization in Hospitality Technology
[15] Sheelam, G. K. (2024). Deep Learning-Based Protocol Stack Optimization in High-Density 5G Environments. European Advanced Journal for Science & Engineering (EAJSE)-p-ISSN, 3050-9696.
[16] Pamisetty, A., Adusupalli, B., Mashetty, S., & Singreddy, S. (2024). Redefining Financial Risk Strategies: The Integration of Smart Automation, Secure Access Systems, and Predictive Intelligence in Insurance, Lending, and Asset Management. Sneha, Redefining Financial Risk Strategies: The Integration of Smart Automation, Secure Access Systems, and Predictive Intelligence in Insurance, Lending, and Asset Management (December 05, 2024).
[17] Kolla, S. H. (2024). RETRIEVAL-AUGMENTED GENERATION WITH SMALL LLMS FOR KNOWLEDGE-DRIVEN DECISION AUTOMATION IN ENTERPRISE SERVICE PLATFORMS. Turkish Journal of Computer and Mathematics Education (TURCOMAT), 15(3), 476-486.
[18] Singreddy, S. (2024). Applying deep learning to mobile home and flood insurance risk evaluation. Available at SSRN 5238946.
[19] Garapati, R. S. (2023). Optimizing Energy Consumption in Smart Build-ings Through Web-Integrated AI and Cloud-Driven Control Systems.
[20] Inala, R. (2022). Engineering Data Products for Investment Analytics: The Role of Product Master Data and Scalable Big Data Solutions. International Journal of Scientific Research and Modern Technology, 155-171.
[21] Singireddy, S. (2024). The Integration of AI and Machine Learning in Transforming Underwriting and Risk Assessment Across Personal and Commercial Insurance Lines. Journal of Computational Analy- sis and Applications(JoCAAA), 33(08), 3966-3991.
[22] Kummari, D. N. (2023). AI-powered demand forecasting for automotive components: A multi-supplier data fusion approach. European Advanced Journal for Emerging Technologies (EAJET)-p-ISSN, 3050-9734.
[23] Inala, R. (2023). Big Data Architectures for Modernizing Customer Master Systems in Group Insurance and Retirement lanning. Educational Administration: Theory and Practice, 29 (4), 5493–5505
[24] Nandan, B. P. (2024). Revolutionizing Semiconductor Chip Design through Generative AI and Reinforcement Learning: A Novel Approach to Mask Patterning and Resolution Enhancement. International Journal of Medical Toxicology and Legal Medicine, 27(5), 759-772.
[25] Kolla, T. (2024). AI-Powered Data Catalog Systems For Healthcare Data Discovery And Governance. South Eastern European Journal of Public Health, 2296–2311. https://doi.org/10.70135/seejph.vi.7077
[26] Inala, R., & Somu, B. (2024). Agentic AI in Retail Banking: Redefining Customer Service and Financial Decision-Making. Journal of Artificial Intelligence and Big Data Disciplines, 1(1).
[27] Recharla, M. (2024). Advances in Therapeutic Strategies for Alzheimer's Disease: Bridging Basic Research and Clinical Applications. American Online Journal of Science and Engineering (AOJSE)(ISSN: 3067-1140), 2(1).
[28] Segireddy, A. R. (2024). Machine Learning-Driven Anomaly Detection in CI/CD Pipelines for Financial Applications. Journal of Computational Analysis and Applications, 33(8).
[29] Amistapuram, K. (2024). Smart Decision Support Systems For Dynamic Tax Policy Optimization Using Reinforcement Learning. Available at SSRN 6143426
[30] Kolla, S. K. (2024). Federated Machine Learning On Big Healthcare Data For Privacy-Preserving Analytics. The Review of Diabetic Studies, 175-190.
[31] Singireddy, J. (2024). AI-Driven Payroll Systems: Ensuring Compliance and Reducing Human Error. American Data Science Journal for Advanced Computations (ADSJAC) ISSN, 3067-4166.
[32] Yandamuri, U. S. (2023). An Intelligent Analytics Framework Combining Big Data and Machine Learning for Business Forecasting. International Journal Of Finance, 36(6), 682-706.
[33] Pamisetty, A. (2024). Leveraging Agentic AI and Cloud Infrastructure for Predictive Logistics in National Food Supply Chains. Available at SSRN 5262994.
[34] Deep Learning-Driven Optimization of ISO 20022 Protocol Stacks for Secure Cross-Border Messaging. (2024). MSW Management Journal, 34(2), 1545-1554.
[35] Pamisetty, A. (2024). Leveraging Big Data Engineering for Predictive Analytics in Wholesale Product Logistics. Available at SSRN 5231473.
[36] Kolla, S. K. (2023). Explainable AI and ML Models for Transparent Clinical Decision Support. Journal for ReAttach Therapy and Developmental Diversities, 6, 2444-2460.
[37] Nagabhyru, K. C. (2023). Accelerating Digital Transformation with AI Driven Data Engineering: Industry Case Studies from Cloud and IoT Domains. Educational Administration: Theory and Practice, 29(4), 5898-5910.
[38] Aitha, A. R. (2023). CloudBased Microservices Architecture for Seamless Insurance Policy Administration. International Journal of Finance (IJFIN)-ABDC Journal Quality List, 36(6), 607-632.
[39] Sheelam, G. K. (2024). Towards autonomic wireless systems: integrating agentic AI with advanced semiconductor technologies in telecommunications. Am. Online J. Sci. Eng., 3(4), 234-256.

[40] Yandamuri, U. S. (2022). Big Data Pipelines for Cross-Domain Decision Support: A Cloud-Centric Approach. International Journal of Scientific Research and Modern Technology (IJSRMT).
[41] Amistapuram, K. (2024). Federated Learning for Cross-Carrier Insurance Fraud Detection: Secure Multi-Institutional Collaboration. Journal of Computational Analysis and Applications (JoCAAA), 33(08), 6727-6738.
[42] Reddy Segireddy, A. (2024). Federated Cloud Approaches for Multi-Regional Payment Messaging Systems. Turkish Journal of Computer and Mathematics Education (TURCOMAT), 15(2), 442-450.
[43] Aitha, A. R. (2023). Cloud-Native Big Data AI/ML Framework for Risk Intelligence and Fraud Control in Banking and Insurance Ecosystems. Available at SSRN 6157967.
[44] Kolla, T. (2023). Predictive ETL Failure Detection in Healthcare Data Pipelines Using Anomaly Detection Algorithms. International Journal of Medical Toxicology & Legal Medicine.
[45] Gottimukkala, V. R. R. (2022). Licensing Innovation in the Financial Messaging Ecosystem: Business Models and Global Compliance Impact. International Journal of Scientific Research and Modern Technology, 1(12), 177-186.
[46] Pamisetty, V. (2024). AI-Driven Decision Support for Taxation and Unclaimed Property Management: Enhancing Efficiency through Big Data and Cloud Integration. Available at SSRN 5250776.
[47] Mahesh Recharla, "Integrated Genomic and Neurobiological Pathway Mapping for Early Detection of Alzheimer's Disease," International Journal of Advanced Research in Computer and Communication Engineering (IJARCCE), DOI: 10.17148/IJARCCE.2023.12122.
[48] Nagabhyru, K. C. (2023). From Data Silos to Knowledge Graphs: Architecting CrossEnterprise AI Solutions for Scalability and Trust. Available at SSRN 5697663. [91]Meda, R. (2024). Predictive Maintenance of Spray Equipment Using Machine Learning in Paint Application Services. European Data Science Journal (EDSJ) p-ISSN, 3050-9572.
[49] Sheelam, G. K., & Koppolu, H. K. R. (2024). From Transistors to Intelligence: Semiconductor Architectures Empowering Agentic AI in 5G and Beyond. Journal of Computational Analy- sis and Applications(JoCAAA), 33(08), 4518-4537.
[50] Bandi, V. D. V. K. (2024). Automated Feature Engineering Systems in Large-Scale Healthcare Data Environments. Journal of Neonatal Surgery, 13.
[51] Mangalampalli, B. M. (2024). AI-Enhanced Data Governance: Automating Compliance In Healthcare Analytics Platforms. The Review of Diabetic Studies, 191-204.
[52] Meda, R. (2024). Enhancing Paint Formula Innovation Using Generative AI and Historical Data Analytics. American Advanced Journal for Emerging Disciplinaries (AAJED) ISSN, 3067-4190.
[53] Bandi, V. D. V. K. (2024). AI-Driven Predictive Risk Modeling Architectures for Financial Systems. International Journal Of Finance, 37(3), 54-7.
[54] Kolla, S. K. (2023). Big Data–Driven Machine Learning Frameworks for Clinical Risk Prediction. International Journal of Medical Toxicology and Legal Medicine, 26(3), 44-59.
[55] Pamisetty, V. (2024). Transforming taxation systems through predictive analytics and AI-driven compliance monitoring tools. Am Data Sci J Adv Comput, 3, 55-68.
[56] Garapati, R. S. (2022). AI-Augmented Virtual Health Assistant: A Web-Based Solution for Personalized Medication Management and Patient Engagement. Available at SSRN 5639650.
[57] Nagabhyru, K. C. (2024). Data Engineering in the Age of Large Language Models: Transforming Data Access, Curation, and Enterprise Interpretation. Computer Fraud and Security.
[58] Aitha, A. R. (2022). Cloud Native ETL Pipelines for Real Time Claims Processing in Large Scale Insurers. Available at SSRN 5532601.
[59] Meda, R. (2023). Data Engineering Architectures for Scalable AI in Paint Manufacturing Operations. European data science journal.
[60] Gottimukkala, V. R. R. (2023). Privacy-Preserving Machine Learning Models for Transaction Monitoring in Global Banking Networks. International Journal of Finance (IJFIN)-ABDC Journal Quality List, 36(6), 633-652.
[61] Davuluri, P. N. Integrating Artificial Intelligence into Event-Driven Financial Crime Compliance Platforms.
[62] Kolla, S. H. (2022). Knowledge Retrieval Systems for Enterprise Service Environments. International Journal of Intelligent Systems and Applications in Engineering, 10, 495-506.
[63] Bandi, V. D. V. K. (2024). Intelligent Data Platforms For Personalized Retail Analytics At Scale. Metallurgical and Materials Engineering, 30 (4), 1011–1027.
[64] Mangala, N. (2022). Implementing Databricks Unity Catalog For Centralized Data Governance In Multi-Business-Unitenterprises. Journal of International Crisis and Risk Communication Research , 101–122. https://doi.org/10.63278/jicrcr.vi.3738.
[65] Davuluri, P. N. AI-Augmented Sanctions Screening: Enhancing Accuracy and Latency in Real Time Compliance Systems.
[66] Meda, R. (2024). Agentic AI in Multi-Tiered Paint Supply Chains: A Case Study on Efficiency and Responsiveness. Journal of Compu-tational Analysis and Applications (JoCAAA), 33(08), 3994-4015.