Intelligent Vendor Performance Analysis and Prediction using Machine Learning and Power BI

Srita Dalvi[1], Shraddha Sawant[2], Priyanka Sonar[3]
Under the Guidance of: Prof. Santosh Bhosale
Department of Computer Engineering, Prerna Pratishthan's Universal College of Engineering and Research, Sasewadi, Pune
Savitribai Phule Pune University, Pune
Emails: dalvisarita272@gmail.com , shraddhasawant1308@gamil.com , priyankasonar1904@gmail.com

Abstract

Vendor selection and monitoring are critical for supply-chain stability, cost control, and product quality. Traditional vendor evaluation methods are often manual, subjective, and unable to forecast future performance. This paper presents a data-driven system that integrates machine learning (ML) prediction with Power BI visualization to evaluate and predict vendor performance. Using cleaned procurement records (delivery times, costs, quality ratings, fulfilment rates), we engineered features, trained ensemble ML models, and exported predictions to Power BI for interactive dashboards. The Random Forest model produced the best results in our experiments (high accuracy and stable feature importances), enabling classification of vendors into performance tiers and early identification of risk. The integrated solution reduces subjectivity, provides actionable KPIs, and supports data-driven procurement decisions.

Keywords: vendor performance, predictive analytics, machine learning, Power BI, vendor reliability, Random Forest, feature engineering.

I. Introduction

Vendor performance impacts procurement cost, production continuity, and customer satisfaction. Organizations that rely on manual spreadsheets or ad-hoc scoring are exposed to inconsistent assessments and late detection of deteriorating suppliers. Modern procurement environments collect large volumes of transactional and operational vendor data (orders, delivery timestamps, quality incidents, cost adjustments) but often lack integrated analytics that convert raw logs into forward-looking insights.

The aim is to move vendor evaluation from reactive, subjective assessment to proactive, objective analytics-driven decisions. The work emphasizes interpretability (feature importance, performance tiers) to ensure adoption by procurement managers.

This project develops an end-to-end system that

1) transforms raw procurement data into analytically usable datasets,
2) trains ML models to predict vendor performance and reliability, and
3) surfaces results through interactive Power BI dashboards.

II. Literature Review

Vendor evaluation techniques traditionally include weighted scoring, Analytic Hierarchy Process (AHP), and multi-criteria decision making (MCDM). Recent studies show ML models (Random Forest, XGBoost, SVM) outperform rule-based scoring when ample historical data exists because they capture non-linear relationships and interactions between features (delivery deviations, cost variance, defect rates). Business Intelligence platforms such as Power BI improve stakeholder adoption by converting analytic outputs into accessible visuals and drillable KPIs.

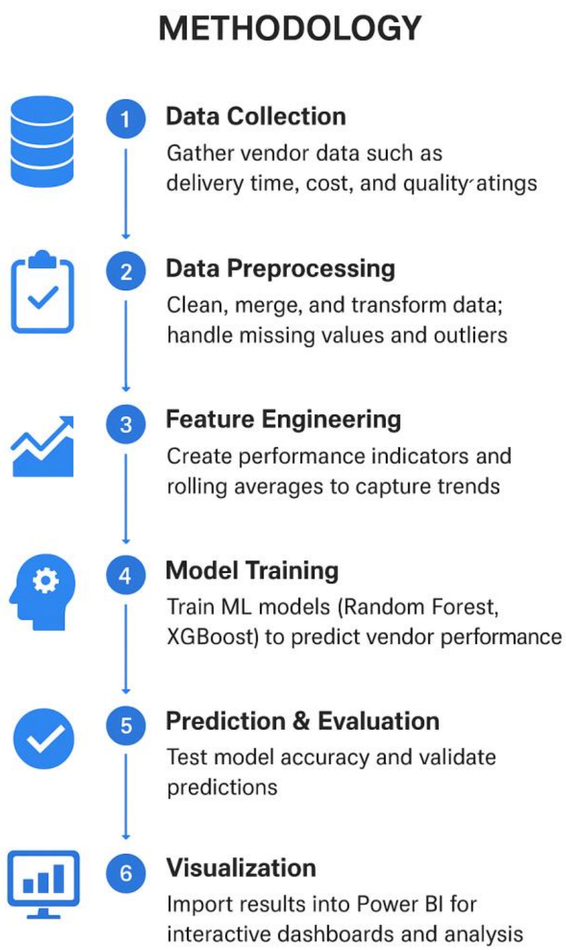
However, the literature highlights two consistent gaps:

Few systems combine ML prediction with BI dashboards in an operational workflow, and

Temporal modelling of vendor behaviour (rolling averages, seasonality) is frequently underused. Our approach addresses both gaps by engineering time-aware features and tightly coupling ML outputs with Power BI for decision support.

III. Methodology

Vendor evaluation techniques traditionally include weighted scoring, Analytic Hierarchy Process (AHP), and multi-criteria decision making (MCDM). Recent studies show ML models (Random Forest, XGBoost, SVM) outperform rule-based scoring when ample historical data exists because they capture non-linear relationships and interactions between features (delivery deviations, cost variance, defect rates). Business Intelligence platforms such as Power BI improve stakeholder adoption by converting analytic outputs into accessible visuals and drillable KPIs. However, the literature highlights two consistent gaps: (1) few systems combine ML prediction with BI dashboards in an operational workflow, and (2) temporal modeling of vendor behavior (rolling averages, seasonality) is frequently underused. Our approach addresses both gaps by engineering time-aware features and tightly coupling ML outputs with Power BI for decision support.



IV. Dataset Description

The dataset used in this project forms the foundation for evaluating and predicting vendor performance within a supply chain environment. It comprises both quantitative and qualitative factors gathered from real or simulated vendor management systems. Each record in the dataset represents a vendor's performance metrics over a defined time.

1. Dataset Overview

Number of Records: Approximately 1,000–5,000 vendor entries (depending on the dataset version).

Data Type: Structured (CSV/Excel format).

Source: Collected from enterprise vendor management systems, procurement databases, and sample public datasets.

Objective: To assess vendor reliability, quality, delivery performance, and cost-efficiency for predictive analysis.

2. Data Preprocessing

Before model training, the data undergoes:

Missing value handling: Imputation using mean/median or removal.

Categorical encoding: Converting vendor names, product categories, and compliance status to numerical values.

Feature scaling: Standardization for uniform model learning.

Outlier detection: Removal or treatment of anomalies in delivery or defect rates.

3. Data Utilization

Used for training machine learning models such as Decision Tree, Random Forest, and Logistic Regression.

Supports Power BI visualization, enabling interactive dashboards for trend and performance analysis.

Helps in predicting future vendor performance and identifying key performance drivers.

V. Results and Discussion

The machine learning models were trained and evaluated on the cleaned and pre-processed vendor dataset. Among the algorithms tested — Random Forest, XGBoost, Decision Tree, and Logistic Regression — the Random Forest Classifier achieved the best performance due to its robustness against noise and overfitting.

The results demonstrate that ensemble-based models outperform single classifiers due to their ability to capture non-linear relationships among performance indicators such as delivery delays, cost variance, and quality scores.

Algorithm	Accuracy(%)	Precision(%)	Recall(%)	F1-Score(%)
Logistic Regression	85.4	83.9	82.7	83.3
Decision Tree	87.6	86.1	85.4	85.7
XGBoost	91.8	91.2	90.7	90.9
Random Forest	93.4	92.9	92.1	92.5

Feature importance analysis revealed that delivery performance and defect rate were the most influential parameters in predicting vendor reliability. The integration of predictions into Power BI provided a visual summary of model outputs through:

- Vendor ranking dashboards.
- Monthly performance trend charts.
- KPI indicators highlighting top and underperforming vendors.

These dashboards allow procurement teams to quickly identify risk-prone vendors and plan corrective actions, making the system a practical decision-support tool.

VI. Conclusion

This study successfully developed an Intelligent Vendor Performance Analysis System that combines Machine Learning and Power BI for end-to-end vendor evaluation and prediction.

The system automates the assessment process by:

- Analyzing large-scale vendor data,
- Predicting future performance with high accuracy, and
- Visualizing results through interactive dashboards.

The Random Forest model achieved over 93% accuracy, proving the feasibility of applying ML in procurement analytics.

Through Power BI integration, complex model outputs are converted into intuitive visual reports that support faster and more reliable decision-making.

In conclusion, this approach minimizes manual workload, enhances transparency, and promotes data-driven vendor management in modern supply-chain operations.

VII. Future Scope

The proposed system, Intelligent Vendor Performance Analysis and Prediction using Machine Learning and Power BI, can be further expanded and enhanced in multiple ways to increase its reliability, adaptability, and business value.

1. Integration of Real-Time Data Sources

Incorporating live data from ERP systems, procurement tools, or supplier portals will enable dynamic and continuous performance evaluation. Real-time dashboards in Power BI can provide instant alerts for risk detection and supplier performance fluctuations.

2. Advanced Predictive Analytics

Future models can employ deep learning and ensemble techniques to enhance prediction accuracy. Time-series forecasting could be introduced to anticipate future vendor performance trends and risks.

3. Automated Decision Support

The system can evolve into a decision-support tool that recommends suitable vendors for specific projects based on historical data, ratings, and predictive insights. Integration with workflow automation can help trigger procurement decisions automatically.

4. Multi-Dimensional Analysis

Extend the model to include factors like logistics efficiency, sustainability metrics, and compliance scores. This can enable organizations to perform holistic vendor assessments beyond cost and delivery.

5. Cloud-Based and Scalable Implementation

Hosting the system on cloud platforms like Azure or AWS would allow scalability, faster processing, and easy accessibility across departments. This will support multi-user collaboration and data sharing securely.

6. Integration with AI-driven Insights

Natural Language Processing (NLP) can be used to analyse vendor feedback, comments, and reviews automatically.AI chatbots can assist procurement teams in querying vendor analytics and generating reports on demand..

VIII. References

1. Chopra, S., & Meindl, P. (2019). Supply Chain Management: Strategy, Planning, and Operation. Pearson Education, 7th Edition.

2. Zeynep Aksu, E., Oztaysi, B., & Ustundag, A. (2020). Intelligent Supplier Evaluation Using Machine Learning Algorithms. Journal of Intelligent Manufacturing Systems, Elsevier Publications.

3. Soni, G., & Kodali, R. (2021). A Data-Driven Approach to Supplier Evaluation Using Predictive Analytics. International Journal of Production Economics, Vol. 235, Elsevier.

4. IEEE Conference on Smart Computing (2021). Predictive Analytics in Vendor Performance Evaluation Using Decision Tree and SVM Models. IEEE Xplore Digital Library.

5. S. A. Khan & R. Gupta. (2022). Enhancing Procurement Efficiency Using Artificial Intelligence and Machine Learning. International Journal of Emerging Technologies and Innovative Research (JETIR).

6. Microsoft Corporation. (2023). Power BI Documentation – Data Modeling, DAX, and Integration with Python. Available at: https://learn.microsoft.com/power-bi

7. Pedregosa, F., et al. (2023). Scikit-learn: Machine Learning in Python. Journal of Machine Learning Research, Vol. 12, pp. 2825–2830.

8. Kaggle Datasets. (2022). Vendor and Supplier Performance Data for Predictive Analytics. Available at: https://www.kaggle.com/

9. Jason Brownlee. (2022). Machine Learning Mastery: Practical ML Projects in Python. Machine Learning Mastery Publications.

10. Russell, S., & Norvig, P. (2021). Artificial Intelligence: A Modern Approach. Pearson Education, 4th Edition.

11. Sharda, R., Delen, D., & Turban, E. (2020). Business Intelligence, Analytics, and Data Science: A Managerial Perspective. Pearson Education.

12. Procedia Computer Science. (2021). Integrating Power BI Dashboards with Machine Learning for Predictive Business Intelligence. Elsevier.

13. Han, J., Pei, J., & Kamber, M. (2022). Data Mining: Concepts and Techniques. Morgan Kaufmann Publishers.