



Revisiting Feature Scaling in Linear Regression: An Empirical Study on Microsoft Stock Price Prediction

Farhan Mahfudz ^{1*}, Khoirunnisya ²

^{1*,2} Department of Informatics Engineering, Universitas Pamulang, South Tangerang City, Banten Province, Indonesia.

*Corresponding author: farhanhk12@gmail.com.

Received: March 30, 2026; Accepted: April 10, 2026; Published: April 20, 2026.

Abstract: Stock price prediction occupies a central position in quantitative finance, bearing directly on risk management, portfolio construction, and investment decision-making. This study evaluated the effect of feature scaling on linear regression performance in predicting Microsoft (MSFT) stock prices. A quantitative experimental design was employed, drawing on historical MSFT stock data spanning 2014 to 2024. Preprocessing involved data cleaning, outlier treatment via the Interquartile Range (IQR) method, and feature standardization through Z-score normalization. Two experimental conditions were tested: linear regression without feature scaling and linear regression with feature scaling. Model performance was assessed using Mean Absolute Error (MAE), Root Mean Squared Error (RMSE), Mean Absolute Percentage Error (MAPE), and the coefficient of determination (R^2). Both conditions produced nearly identical results — R^2 approaching 0.99, with negligible divergence across all error metrics. The evidence suggests that feature scaling does not meaningfully alter the predictive behavior of linear regression. For simple linear models operating without regularization, scaling appears to be an unnecessary preprocessing step, a finding that warrants more deliberate evaluation of preprocessing decisions in machine learning pipelines.

Keywords: Linear Regression; Feature Scaling; Stock Price Prediction.

1. Introduction

Investment has long been regarded as a mechanism for securing long-term financial stability, particularly through deliberate asset allocation across sectors with strong growth trajectories. The technology industry has attracted sustained investor attention — not merely because of its pace of innovation, but because of its demonstrated capacity to generate scalable, compounding returns over time (Alka *et al.*, 2025; Khatoun & Velidandi, 2025; Moro-Visconti *et al.*, 2020). Digital transformation has reshaped business operations at scale, and technology stocks have become a dominant constituent of modern investment portfolios as a result. Within this landscape, stocks represent fractional ownership in a company and remain among the most widely traded financial instruments globally. As a class of financial asset, equity investment demands rigorous, data-grounded analysis — intuition alone is insufficient, and the cost of poor forecasting can be substantial (Bang, 2024; Shah, 2025). Assets are broadly divided into real assets — property, commodities, precious metals — and financial assets, which include equities, bonds, and deposits (Tian *et al.*, 2020; Tiwari *et al.*, 2025). Among financial instruments, stocks remain particularly attractive due to their return potential, though that potential comes paired with volatility that resists easy modeling.

Predictive modeling in quantitative finance has evolved considerably over the past two decades. Traditional time-series approaches — ARIMA, GARCH — gave way to machine learning and deep learning

architectures capable of capturing nonlinear dynamics and complex market behavior (Goel & Som, 2023; Sheng & Ma, 2022; Tjøstheim, 2025). Forecasting accuracy has improved, though the inherently noisy character of financial markets continues to constrain even the most sophisticated models. Against this backdrop, linear regression persists as a relevant baseline. Its interpretability, computational efficiency, and theoretical grounding make it a standard point of reference in machine learning benchmarking (Munkhdalai *et al.*, 2022; Panovski & Zaharia, 2020). Its limitations in capturing nonlinear relationships are well-documented, but that does not diminish its utility as a diagnostic and comparative tool. Preprocessing techniques such as feature scaling are routinely applied to improve model behavior — particularly for algorithms sensitive to differences in feature magnitude. Z-score normalization and related standardization methods are widely assumed to improve model stability, accelerate convergence, and sharpen predictive accuracy (Cândido *et al.*, 2021; Shi *et al.*, 2022; Yang & Zhou, 2024). That assumption, however, is not universally supported. Ordinary Least Squares (OLS) linear regression, from a theoretical standpoint, adjusts coefficient estimates in proportion to feature scale — which implies that scaling may not change prediction outcomes in any meaningful way. The empirical record on this point is mixed, and the debate has not been resolved.

This study focuses on Microsoft Corporation (MSFT), a company with an extensive, well-documented stock history and one of the largest market capitalizations in the world. The volatility observed in major technology equities — Microsoft, Apple, and Tesla among them — makes accurate prediction genuinely difficult (Chang *et al.*, 2024; Cristescu *et al.*, 2023). The present research revisited the role of feature scaling in linear regression by testing its effect empirically on MSFT stock price prediction. The goal was not to produce a novel forecasting architecture, but to ask a more pointed question: does scaling actually matter here, and under what conditions might that answer change? By addressing this, the study seeks to contribute to a more grounded understanding of preprocessing decisions in machine learning pipelines and provide evidence-based guidance for both researchers and practitioners working in financial forecasting.

2. Related Work

2.1 Stock Price Prediction in Quantitative Finance

Stock price prediction has occupied a central position in quantitative finance research for decades, and its relevance has only grown as financial markets become more interconnected and data-rich. Accurate forecasting bears directly on risk management, portfolio construction, and strategic investment planning — not as abstract goals, but as measurable outcomes. Recent work confirms that well-designed predictive models can meaningfully improve portfolio performance metrics such as Value-at-Risk (VaR) and Sharpe ratio (Aburto *et al.*, 2023; Bagheri *et al.*, 2023). The problem, however, is not straightforward. Stock prices respond to a wide range of forces — macroeconomic conditions, investor sentiment, geopolitical disruptions, and sudden market shocks — making prediction a genuinely complex and dynamic task (Ahmed *et al.*, 2022; Goel & Som, 2023). Beyond price-level forecasting, volatility estimation has received equal attention as a component of financial modeling. GARCH models and state-space approaches have been widely applied to quantify market uncertainty and characterize risk exposure (Azman *et al.*, 2022). Taken together, these studies make clear that stock prediction is not reducible to estimating a single number — it requires understanding the underlying dynamics that drive price behavior over time.

2.2 Traditional Time-Series Models vs Machine Learning Approaches

For much of the twentieth century, stock price prediction relied on ARIMA and GARCH as the dominant analytical tools. Both are well-suited to capturing linear dependencies and short-term temporal patterns, and their theoretical properties are well understood (Goel & Som, 2023). Their limitations, however, become apparent under conditions of sustained volatility or when the data-generating process is nonlinear — which, in financial markets, is more often the rule than the exception. The shift toward machine learning and deep learning over the past decade has addressed some of these limitations. LSTM networks, CNNs, and hybrid ARIMA-LSTM architectures have all demonstrated performance gains over traditional approaches, particularly in longer forecasting horizons where temporal dependencies are more complex (Fang *et al.*, 2022; Sheng & Ma, 2022; Zhang & Hua, 2025). That said, traditional models have not been rendered obsolete. They continue to serve as baseline references — a function that is easy to underestimate but methodologically important. Without a reliable baseline, it is difficult to assess whether the added complexity of a deep learning model is actually justified by the data.

2.3 Linear Regression as a Baseline Model in Machine Learning

Linear regression, implemented via Ordinary Least Squares, estimates the relationship between dependent and independent variables by minimizing the sum of squared residuals. It is computationally inexpensive, easy to interpret, and theoretically well-grounded — properties that have made it a standard

benchmark in machine learning for decades (Munkhdalai *et al.*, 2022; Panovski & Zaharia, 2020). Several studies confirm its value as a reference point: it provides a transparent account of feature contributions and sets a performance floor against which more complex models must be judged (Gyaneshwar *et al.*, 2023). Its limitations are real and well-documented — linear regression cannot capture nonlinear relationships or complex feature interactions, both of which are common in financial data (DeMiguel *et al.*, 2021). Yet dismissing it entirely would be premature. On structured tabular datasets with relatively simple patterns, linear models remain competitive, particularly when regularization techniques such as Ridge or Lasso are applied (Jiang *et al.*, 2024; Li, 2021). The implication is not that linear regression is sufficient for all tasks, but that its continued use as a baseline is methodologically sound and practically defensible.

2.4 Feature Scaling in Machine Learning Models

Feature scaling normalizes the range of independent variables so that all features contribute proportionally during model training. The two most commonly applied methods are Min-Max normalization and Z-score standardization, both of which transform features into comparable scales while preserving their relative distributions (Cândido *et al.*, 2021). Scaling is particularly consequential for algorithms that rely on distance calculations or gradient-based optimization — in these settings, features with larger numerical ranges can dominate the learning process, biasing coefficient estimates and slowing convergence (Bassam *et al.*, 2022; Shi *et al.*, 2022). For this reason, scaling has become a near-universal step in machine learning pipelines, often applied as a default rather than a deliberate choice. Its importance is most pronounced in regularized models. In Lasso and Ridge regression, the penalty term operates directly on coefficient magnitudes, meaning that unscaled features will be penalized unevenly — a problem that can produce biased or unstable estimates (Yang & Zhou, 2024). The practical takeaway is that scaling is not universally necessary, but its necessity is model-dependent, and the distinction matters more than practitioners sometimes acknowledge.

2.5 The Debate on Feature Scaling in Linear Regression

The question of whether feature scaling is necessary for linear regression has not been settled, and the disagreement runs deeper than it might initially appear. Theoretically, OLS does not require scaling for accurate coefficient estimation — the model adjusts parameter weights in proportion to input scale, leaving prediction outcomes unchanged by linear transformations of the feature space (Cândido *et al.*, 2021). That is the clean answer from first principles. In practice, however, the picture is less tidy. When iterative optimization is involved, or when regularization is incorporated, scaling can improve numerical stability, accelerate convergence, and reduce sensitivity to poorly conditioned feature matrices (Shi *et al.*, 2022; Sperber *et al.*, 2023). Scaling also facilitates fairer comparison of feature importance, which matters when coefficient magnitudes are used to draw substantive conclusions. Empirical studies have produced conflicting results: some report negligible effects of scaling on linear regression performance, while others identify meaningful improvements in specific contexts — particularly in high-dimensional or heterogeneous datasets (Yang & Zhou, 2024). The inconsistency across studies suggests that the effect of scaling is not fixed but context-dependent, and that blanket recommendations in either direction are likely to mislead.

2.6 Research Gap and Contribution

Most research on stock price prediction has concentrated on model architecture — comparing LSTM against ARIMA, testing ensemble methods against single estimators, or evaluating hybrid frameworks against their components. Far less attention has been directed at the preprocessing decisions that precede model training. Feature scaling, in particular, tends to be applied as a default assumption rather than an empirically tested choice. The question of whether scaling actually changes outcomes in simple linear models has rarely been asked directly, and when it has, the evidence is thin. This study addressed that gap by testing the effect of Z-score normalization on linear regression performance using Microsoft stock data from 2014 to 2024. The objective was not to produce a new forecasting method, but to generate evidence-based guidance on a preprocessing decision that is routinely made without scrutiny. The findings are expected to contribute to both the theoretical understanding of OLS behavior under scaling and the practical question of when preprocessing steps are worth applying at all.

3. Methodology

This study employed a quantitative experimental design to assess the effect of feature scaling on linear regression performance in stock price prediction. The dataset consists of historical MSFT stock price data sourced from Investing.com, covering January 2014 through December 2024 — a ten-year window chosen to capture multiple market cycles, including periods of sustained growth, sharp correction, and post-pandemic recovery. Each record contains six financial attributes: opening price, closing price, daily high, daily low, trading

volume, and percentage change. All attributes were initially stored in non-numeric formats and required conversion to floating-point values before any statistical processing could proceed — a step that is easy to overlook but consequential for downstream model behavior. Preprocessing followed a structured sequence. The dataset was first examined for missing values and duplicate records, then subjected to outlier detection using the Interquartile Range (IQR) method. Identified outliers were replaced with the median value of the respective feature, a strategy that preserves record count while reducing the influence of extreme values on coefficient estimation. Feature scaling via Z-score normalization was subsequently applied to the scaled experimental condition, transforming each feature to a mean of zero and a standard deviation of one. Whether this transformation produces any measurable difference in model performance is the central empirical question the study was designed to answer. Following preprocessing, a linear regression model was constructed using the Scikit-learn library in Python. The dataset was partitioned into training (80%) and testing (20%) subsets — 2,213 records for training and 554 for testing — to evaluate generalization on unseen data. Two experimental conditions were tested under identical settings: linear regression on raw, unscaled data and linear regression on Z-score normalized data. Holding all other variables constant ensured that any observed performance differences could be attributed to scaling alone. Model performance was assessed using four metrics: Mean Absolute Error (MAE), Root Mean Squared Error (RMSE), Mean Absolute Percentage Error (MAPE), and the coefficient of determination (R^2). Results from both conditions were then compared to determine whether scaling produced any meaningful change in predictive accuracy.

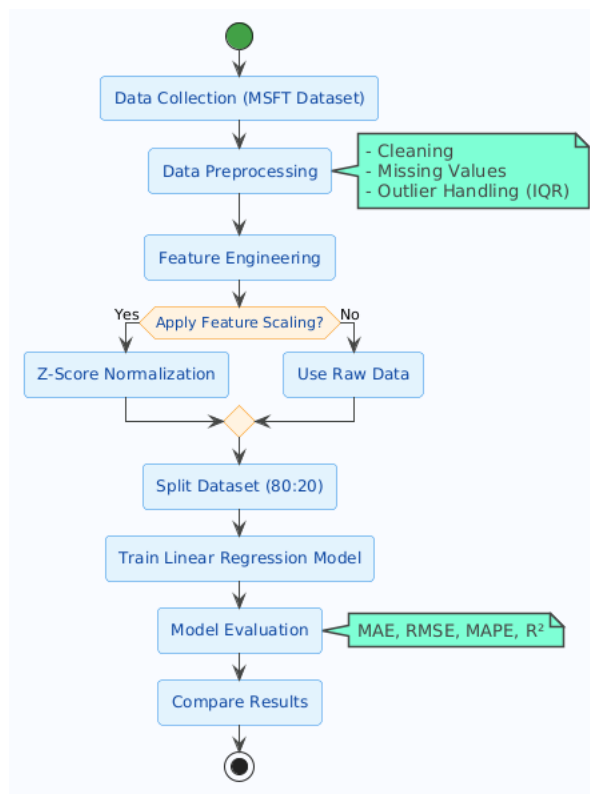


Figure 1. Systematic Research Methodology Flow

Figure 1 traces the full research workflow, from data collection through final evaluation. The process begins with loading the MSFT dataset and converting all attributes to numerical formats, followed by missing value checks, IQR-based outlier detection, and median replacement. At this point the workflow branches: one path proceeds with raw data, the other applies Z-score normalization. Both paths then follow the same sequence — train-test splitting, model training, and prediction generation on the test set. The final stage brings the two paths together for comparative evaluation across all four performance metrics. This branching structure is the methodological core of the study, designed to isolate the effect of scaling as cleanly as possible.

4. Result and Discussion

4.1 Result

4.1.1 Dataset Characteristics and Initial Exploration

The dataset comprised 2,767 records of MSFT historical stock prices spanning January 2014 to December 2024. Each record contained six financial attributes: opening price, closing price, daily high, daily low, trading volume, and percentage change. All attributes were initially stored in non-numeric formats and required conversion to floating-point values before any statistical or model-based processing could proceed. Exploratory data analysis (EDA) was then conducted to characterize the distribution and internal structure of the dataset. Descriptive statistics revealed wide distributional variation across price-based features, with the mean closing price differing substantially from both the minimum and maximum values — a pattern consistent with the long upward trend in MSFT prices over the decade studied. No missing values were detected across any attribute, which eliminated the need for imputation and strengthened the reliability of subsequent analysis. Correlation analysis produced a result that is worth examining carefully. Opening price, daily high, and daily low were all near-perfectly correlated with closing price, with coefficients approaching 1.0. Trading volume and percentage change, by contrast, showed weak or negligible correlations with the target variable. This pattern has a direct implication for model behavior: the prediction task is effectively being driven by three features that are structurally near-identical to the target. Whether the model is learning a meaningful predictive relationship or simply recovering a near-identity mapping is a question the results will need to address.

Table 1. Descriptive Statistics of Dataset

Feature	Min	Max	Mean	Std Dev
Open	34.73	467.00	174.20	124.25
High	35.88	468.35	175.86	125.35
Low	34.63	464.46	172.48	123.05
Close	34.98	467.56	174.25	124.25
Volume	7.16	202.53	29.31	13.72
Change (%)	-14.74	14.22	0.10	1.67

Figure 2 traces the data exploration sequence: dataset loading and format conversion, missing value checks, statistical summary generation, and correlation analysis. Each step was completed before any modeling decisions were made, ensuring that the dataset was adequately characterized and that no silent data quality issues carried forward into the experimental phase.

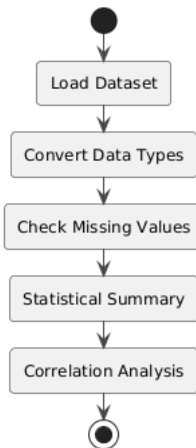


Figure 2. Data Exploration Workflow

4.1.2 Data Preprocessing and Outlier Handling

No missing values or duplicate records were identified in the dataset, indicating strong data integrity from the outset. Outlier detection using the IQR method, however, identified significant anomalies in two features: volume and percentage change. Left unaddressed, these extreme values could distort coefficient estimation and inflate error metrics — particularly in a linear model, where a single influential observation can pull the fitted line in ways that do not reflect the broader data pattern. Outliers were replaced with the median value of the respective feature. Median replacement was preferred over deletion because it preserves the original record count and is less sensitive to the distributional assumptions that underpin mean-based imputation. After treatment, the number of outliers in both features was substantially reduced, as shown in Table 2. Feature scaling via Z-score normalization was subsequently applied to the scaled experimental

condition, transforming each feature to a mean of zero and a standard deviation of one. Whether this transformation produces any measurable difference in predictive performance is the question the experimental comparison was designed to answer.

Table 2. Outlier Detection Results

Feature	Outliers Before	Outliers After
Volume	153	70
Change (%)	142	70

Figure 3 illustrates the preprocessing pipeline: raw data intake, cleaning, IQR-based outlier detection, median replacement, and Z-score normalization. Each step was applied consistently across both experimental conditions, with the exception of the scaling step itself, which was applied only in the scaled condition. This design ensures that any observed performance differences can be attributed to scaling alone.

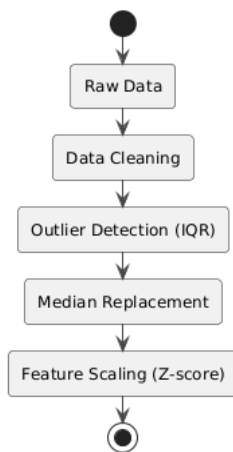


Figure 3. Data Preprocessing Pipeline

4.1.3 Model Development and Training

The dataset was partitioned into training and testing subsets using an 80:20 ratio — 2,213 records for training and 554 for testing. This split provides a reasonable balance between training data volume and evaluation coverage, and is consistent with standard practice in supervised machine learning. A linear regression model was then constructed using the Scikit-learn library in Python, trained on the training subset via the least squares method. Given the near-perfect correlations among price-based features, the model was well-positioned to identify strong linear relationships in the data. Two versions of the model were trained under identical conditions: one on raw, unscaled data and one on Z-score normalized data. Holding all other variables constant was essential — without this control, any observed differences in performance could not be cleanly attributed to scaling.

Table 3. Dataset Splitting

Dataset Type	Records	Percentage
Training	2,213	80%
Testing	554	20%

Figure 4 outlines the training workflow: dataset splitting, model fitting on the training subset, and prediction generation on the test set. Both experimental conditions followed the same sequence, differing only in whether the input features were scaled prior to training.

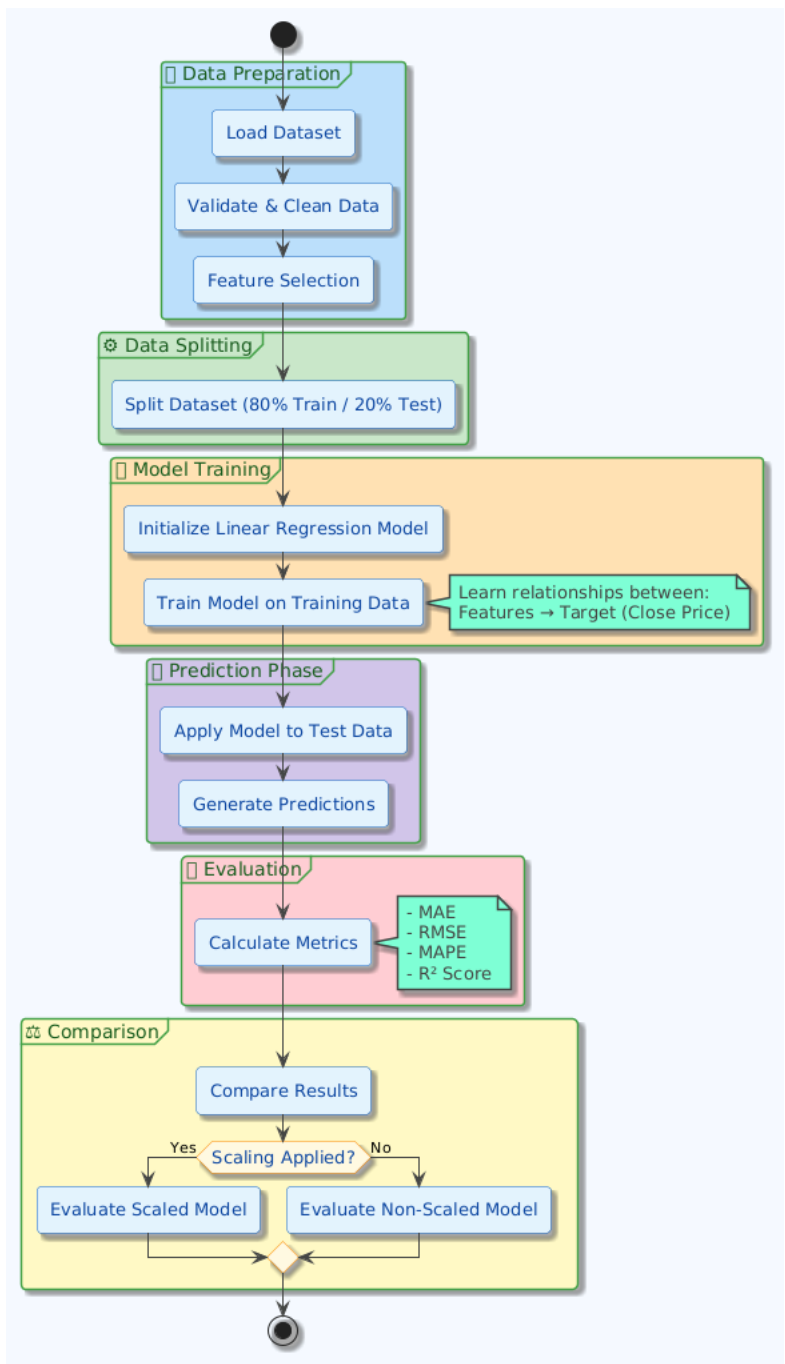


Figure 4. Model Training Process

4.1.4 Model Evaluation and Performance Comparison

Both models achieved R^2 values of approximately 0.99, indicating that the fitted models account for nearly all variance in the test data. MAE, RMSE, and MAPE were virtually identical across the two conditions. The MAPE of 0.50% in both cases confirms that average prediction error was negligible in percentage terms. The difference in MAE between the unscaled and scaled models — 70.74 versus 74.79 — is small enough to be practically meaningless, and RMSE was identical at 103.79 across both conditions. What the results do not show is equally telling: there is no evidence that scaling improved numerical stability, reduced error, or altered the model's behavior in any detectable way. For this dataset and this model, the scaling step produced no measurable benefit.

Table 4. Model Performance Comparison

Metric	Without Scaling	With Scaling
RMSE	103.79	103.79
MAE	70.74	74.79
MAPE	0.50%	0.50%
R^2	0.99	0.99

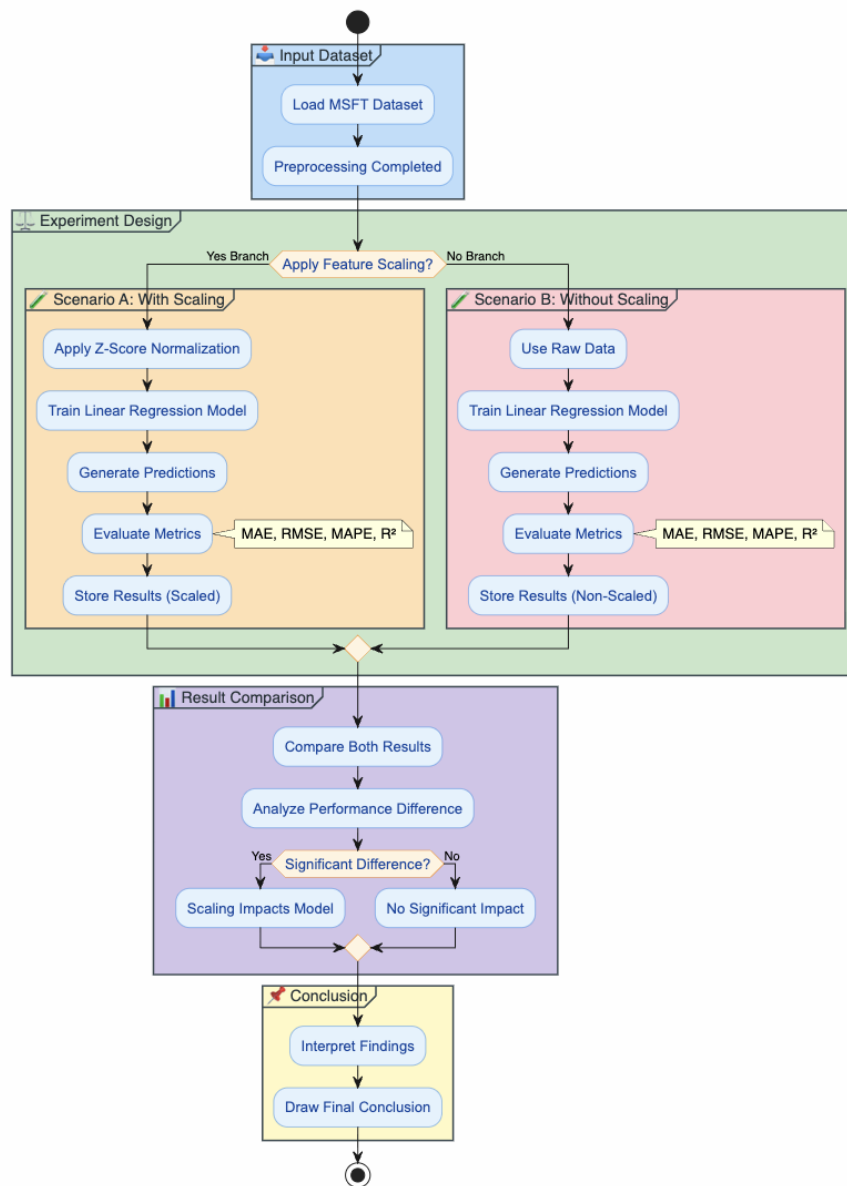


Figure 5. Model Comparison Flow

Figure 5 illustrates the comparative evaluation structure: both models — raw and scaled — were assessed using the same four metrics under identical test conditions. The near-identical results across all metrics confirm that scaling did not alter predictive performance in any meaningful direction, which is precisely the finding the experimental design was built to detect.

4.1.5 Summary of Findings

Feature scaling did not affect linear regression performance in any meaningful way across this dataset. That result is consistent with the theoretical expectation: OLS adjusts coefficient magnitudes in proportion to feature scale, leaving the underlying predictive relationship unchanged regardless of whether features are standardized. The R^2 of 0.99 warrants careful interpretation, however. A near-perfect fit may reflect something more mundane than genuine predictive power — specifically, the fact that opening price, daily high, and daily low are structurally near-identical to closing price by construction. The model may be recovering an algebraic near-identity rather than learning a generalizable relationship. Variance Inflation Factor (VIF) analysis would be a necessary next step to assess the severity of multicollinearity and determine whether the coefficient estimates carry any interpretive weight. Preprocessing steps, this study suggests, should be evaluated against the specific properties of the model and data in use — not applied as a reflex.

4.2 Discussion

The finding that feature scaling produced no meaningful change in linear regression performance aligns with the theoretical foundation of OLS, which adjusts parameter magnitudes proportionally to input scale

(Cândido *et al.*, 2021). Prediction outcomes are invariant to linear transformations of the feature space — a mathematical property, not an empirical accident. The near-identical evaluation metrics across both experimental conditions are exactly what theory predicts, and the data bear that out. That said, dismissing scaling entirely would be a mistake — and the boundary conditions matter here. For gradient-based optimization methods, scaling accelerates convergence and reduces the risk of numerical instability (Shi *et al.*, 2022; Bassam *et al.*, 2022). In regularized regression — Ridge, Lasso, elastic net — scaling is not optional: the penalty term operates on raw coefficient magnitudes, and unscaled features will be penalized unevenly, producing biased estimates (Dani *et al.*, 2025; Habib & Okayli, 2024; Hassan & Baghban, 2025; Ko *et al.*, 2025). The negligible effect observed here is specific to a simple, unregularized linear model. Generalizing this finding beyond that boundary would be unwarranted.

The R^2 of approximately 0.99 is striking, but it deserves scrutiny rather than satisfaction. Opening price, daily high, and daily low are structurally correlated with closing price by construction — they are all derived from the same underlying trading session. A model trained on these features may be capturing little more than a near-identity mapping, which would explain the near-perfect fit without implying genuine predictive generalizability. As Gyaneshwar *et al.* (2023) and Munkhdalai *et al.* (2022) both note, multicollinearity of this kind can inflate apparent model performance while undermining the interpretability of individual coefficients. VIF analysis is the appropriate diagnostic, and its absence is a limitation of the current study that future work should address.

From a broader perspective, the results reinforce the continued relevance of linear regression as a baseline model. Despite the widespread adoption of LSTM, ensemble methods, and other architectures, linear models retain value precisely because they are interpretable and computationally inexpensive (Bian *et al.*, 2023; Sakib *et al.*, 2024; Shome *et al.*, 2024). On structured tabular data with strong linear relationships, they can perform competitively with far more complex alternatives (Hollmann *et al.*, 2025; Khosravi *et al.*, 2023; Nagaraju *et al.*, 2023). Complexity is not always a virtue — and this study offers a modest but concrete illustration of that point. The broader implication for preprocessing practice is this: scaling should be applied selectively, not reflexively. The assumption that it universally improves model performance is not supported by the evidence here, and the literature is inconsistent on the point (Cândido *et al.*, 2021; Sperber *et al.*, 2023). Preprocessing decisions should be grounded in the specific properties of the model and dataset in question — not in habit or convention. Future work should extend this analysis to regularized models, alternative asset classes, and multiple market conditions, with statistical significance testing to determine whether observed performance differences are substantively meaningful or merely numerical noise.

5. Conclusion and Recommendations

Feature scaling, applied as Z-score normalization, did not meaningfully affect the predictive performance of linear regression on Microsoft stock price data. Both the scaled and unscaled models produced nearly identical MAE, RMSE, MAPE, and R^2 values — with R^2 approaching 0.99 in both conditions. This outcome is consistent with the theoretical properties of OLS: linear regression is scale-invariant in terms of prediction accuracy when implemented without regularization, as the model adjusts coefficient magnitudes proportionally to input scale. The high R^2 obtained across both conditions points to strong linear relationships among the input features, though it simultaneously raises the possibility of multicollinearity — a concern that warrants diagnostic follow-up via Variance Inflation Factor (VIF) analysis before any strong interpretive claims are made about individual coefficients.

The practical implication is straightforward, if somewhat inconvenient for those who treat scaling as a default step: preprocessing techniques should be evaluated against the specific model and dataset in use, not applied out of habit. For simple linear models without regularization, scaling adds no measurable value. For regularized models — Ridge, Lasso, elastic net — or those relying on gradient-based optimization, it remains important and should not be skipped. The line between these two cases is not always obvious in advance, which is precisely why empirical testing matters. Future research should extend this analysis across a wider range of algorithms, asset classes, and market conditions, incorporating statistical significance testing to determine whether performance differences are substantively meaningful rather than incidental. Expanding the comparison to include regularized linear models and nonlinear baselines would be a natural and productive next step.

References

- Aburto, L., Romero-Romero, R., Linfati, R., & Escobar, J. W. (2023). An approach for a multi-period portfolio selection problem by considering transaction costs and prediction on the stock market. *Complexity*, 2023, 1–15. <https://doi.org/10.1155/2023/3056411>
- Ahmed, W. S., Mehmood, A., Sheikh, T., & Bachaya, A. (2022). Unveiling the linkages between emerging stock market indices and cryptocurrencies. *Asian Academy of Management Journal*, 27(2). <https://doi.org/10.21315/aamj2022.27.2.9>
- Alka, T. A., Raman, R., & Suresh, M. (2025). Critical success factors for successful technology innovation development in sustainable energy enterprises. *Scientific Reports*, 15(1), Article 14138. <https://doi.org/10.1038/s41598-025-98725-2>
- Azman, S., Pathmanathan, D., & Thavaneswaran, A. (2022). Forecasting the volatility of cryptocurrencies in the presence of COVID-19 with the state space model and Kalman filter. *Mathematics*, 10(17), Article 3190. <https://doi.org/10.3390/math10173190>
- Bagheri, F., Recupero, D. R., & Sirnes, E. (2023). Leveraging return prediction approaches for improved value-at-risk estimation. *Data*, 8(8), Article 133. <https://doi.org/10.3390/data8080133>
- Bang, C. G. (2024). *Data-driven decision-making for business*. Routledge. <https://doi.org/10.4324/9781003457787>
- Bassam, A. M., Phillips, A. B., Turnock, S. R., & Wilson, P. A. (2022). Ship speed prediction based on machine learning for efficient shipping operation. *Ocean Engineering*, 245, Article 110449. <https://doi.org/10.1016/j.oceaneng.2021.110449>
- Bian, L., Qin, X., Zhang, C., Guo, P., & Wu, H. (2023). Application, interpretability and prediction of machine learning method combined with LSTM and LightGBM: A case study for runoff simulation in an arid area. *Journal of Hydrology*, 625, Article 130091. <https://doi.org/10.1016/j.jhydrol.2023.130091>
- Cândido, J., Haesen, J., Aniche, M., & van Deursen, A. (2021). An exploratory study of log placement recommendation in an enterprise system. *Proceedings of the 18th International Conference on Mining Software Repositories (MSR)*, 143–154. <https://doi.org/10.1109/msr52588.2021.00027>
- Chang, V., Xu, Q. A., Chidozie, A., & Wang, H. (2024). Predicting economic trends and stock market prices with deep learning and advanced machine learning techniques. *Electronics*, 13(17). <https://doi.org/10.3390/electronics13173396>
- Cristescu, M. P., Mara, D. A., Nerişanu, R. A., Cuda, L. C., & Maniu, I. (2023). Analyzing the impact of financial news sentiments on stock prices: A wavelet correlation. *Mathematics*, 11(23). <https://doi.org/10.3390/math11234830>
- Dani, Y., Belouaggadia, N., & Jammoukh, M. (2025). Predicting CO2 emissions in Morocco: Exploring the use of ridge regression with data preprocessing and feature impact analysis. *Environmental Science and Pollution Research*, 32(45), 25618–25642. <https://doi.org/10.1007/s11356-025-37156-y>
- DeMiguel, V., Gil-Bazo, J., Nogales, F. J., & Santos, A. A. P. (2021). *Can machine learning help to select portfolios of mutual funds?* SSRN. <https://doi.org/10.2139/ssrn.3768753>
- Fang, F., Ventre, C., Basios, M., Kanthan, L., Martínez-Rego, D., Wu, F., & Li, L. (2022). Cryptocurrency trading: A comprehensive survey. *Financial Innovation*, 8(1). <https://doi.org/10.1186/s40854-021-00321-6>
- Goel, H., & Som, B. K. (2023). Stock market prediction, COVID-19 pandemic and neural networks: An SCG algorithm application. *Economia*, 24(1), 134–146. <https://doi.org/10.1108/econ-07-2022-0101>

- Gyaneshwar, A., Mishra, A., Chadha, U., Vincent, P. M. D. R., Rajinikanth, V., Ganapathy, G. P., & Srinivasan, K. (2023). A contemporary review on deep learning models for drought prediction. *Sustainability*, *15*(7), Article 6160. <https://doi.org/10.3390/su15076160>
- Habib, M., & Okayli, M. (2024). Evaluating the sensitivity of machine learning models to data preprocessing technique in concrete compressive strength estimation. *Arabian Journal for Science and Engineering*, *49*(10), 13709–13727. <https://doi.org/10.1007/s13369-024-08776-2>
- Hassan, R., & Baghban, A. (2025). Pioneering machine learning techniques to estimate thermal conductivity of carbon-based phase change materials: A comprehensive modeling framework. *Case Studies in Thermal Engineering*, *73*, Article 106648. <https://doi.org/10.1016/j.csite.2025.106648>
- Hollmann, N., Müller, S., Purucker, L., Krishnakumar, A., Körfer, M., Hoo, S. B., Schirrmeister, R. T., & Hutter, F. (2025). Accurate predictions on small data with a tabular foundation model. *Nature*, *637*(8045), 319–326. <https://doi.org/10.1038/s41586-024-08328-6>
- Jiang, H., Dong, Y., & Wang, J. (2024). Electricity price forecasting using quantile regression averaging with nonconvex regularization. *Journal of Forecasting*, *43*(6), 1859–1879. <https://doi.org/10.1002/for.3103>
- Khatoon, U. T., & Velidandi, A. (2025). An overview on the role of government initiatives in nanotechnology innovation for sustainable economic development and research progress. *Sustainability*, *17*(3). <https://doi.org/10.3390/su17031250>
- Khosravi, B., Weston, A. D., Nugen, F., Mickley, J. P., Maradit Kremers, H., Wyles, C. C., Carter, R. E., & Taunton, M. J. (2023). Demystifying statistics and machine learning in analysis of structured tabular data. *The Journal of Arthroplasty*, *38*(10), 1943–1947. <https://doi.org/10.1016/j.arth.2023.08.045>
- Ko, K., Al Doulat, A., & Ku, S. (2025). Ridge, LASSO, and elastic net regression: Fathers' psychological well-being in familial contexts. *Journal of Social and Personal Relationships*. Advance online publication. <https://doi.org/10.1177/02654075251412474>
- Li, Q. (2021). The use of artificial intelligence combined with cloud computing in the design of education information management platform. *International Journal of Emerging Technologies in Learning (IJET)*, *16*(5), 32. <https://doi.org/10.3991/ijet.v16i05.20309>
- Moro-Visconti, R., Rambaud, S. C., & Pascual, J. L. (2020). Sustainability in FinTechs: An explanation through business model scalability and market valuation. *Sustainability*, *12*(24). <https://doi.org/10.3390/su122410316>
- Munkhdalai, L., Munkhdalai, T., Pham, V.-H., Hong, J., Ryu, K. H., & Theera-Umpon, N. (2022). Neural network-augmented locally adaptive linear regression model for tabular data. *Sustainability*, *14*(22), Article 15273. <https://doi.org/10.3390/su142215273>
- Nagaraju, T. V., Mantena, S., Azab, M., Alisha, S. S., El Hachem, C., Adamu, M., & Rama Murthy, P. S. (2023). Prediction of high strength ternary blended concrete containing different silica proportions using machine learning approaches. *Results in Engineering*, *17*, Article 100973. <https://doi.org/10.1016/j.rineng.2023.100973>
- Panovski, D., & Zaharia, T. (2020). Long and short-term bus arrival time prediction with traffic density matrix. *IEEE Access*, *8*, 226267–226284. <https://doi.org/10.1109/access.2020.3044173>
- Sakib, M., Mustajab, S., & Alam, M. (2024). Ensemble deep learning techniques for time series analysis: A comprehensive review, applications, open issues, challenges, and future directions. *Cluster Computing*, *28*(1), 73. <https://doi.org/10.1007/s10586-024-04684-0>
- Shah, K. N. (2025). Data-driven automation and AI/ML: Revolutionizing financial decision-making. In *AI and automation in financial services* (pp. 249–271). Springer. https://doi.org/10.1007/978-3-031-92916-8_13

- Sheng, Y., & Ma, D. (2022). Stock index spot–futures arbitrage prediction using machine learning models. *Entropy*, 24(10), Article 1462. <https://doi.org/10.3390/e24101462>
- Shi, X., Jiang, D., Qian, W., & Liang, Y. (2022). Application of the Gaussian process regression method based on a combined kernel function in engine performance prediction. *ACS Omega*, 7(45), 41732–41743. <https://doi.org/10.1021/acsomega.2c05952>
- Shome, A., Mukherjee, G., Chatterjee, A., & Tudu, B. (2024). Study of different regression methods, models and application in deep learning paradigm. In *Deep learning concepts in operations research* (pp. 130–152). CRC Press. <https://doi.org/10.1201/9781003433309-13>
- Sperber, C., Gallucci, L., Mirman, D., Arnold, M., & Umarova, R. M. (2023). Stroke lesion size: Still a useful biomarker for stroke severity and outcome in times of high-dimensional models. *NeuroImage: Clinical*, 40, Article 103511. <https://doi.org/10.1016/j.nicl.2023.103511>
- Tian, Y., Lu, Z., Adriaens, P., Minchin, R. E., Caithness, A., & Woo, J. (2020). Finance infrastructure through blockchain-based tokenization. *Frontiers of Engineering Management*, 7(4), 485–499. <https://doi.org/10.1007/s42524-020-0140-2>
- Tiwari, A. K., Abdullah, M., Sarker, P. K., & Abakah, E. J. A. (2025). Real-world asset tokens and commodities: Static and dynamic linkages. *China Accounting and Finance Review*, 27(5), 759–788. <https://doi.org/10.1108/CAFR-05-2024-0054>
- Tjøstheim, D. (2025). Selected topics in time series forecasting: Statistical models vs. machine learning. *Entropy*, 27(3), Article 279. <https://doi.org/10.3390/e27030279>
- Yang, N., & Zhou, W. (2024). Feature selection for explaining yellowfin tuna catch per unit effort using least absolute shrinkage and selection operator regression. *Fishes*, 9(6), Article 204. <https://doi.org/10.3390/fishes9060204>
- Zhang, L., & Hua, L. (2025). Major issues in high-frequency financial data analysis: A survey of solutions. *Mathematics*, 13(3), Article 347. <https://doi.org/10.3390/math13030347>.