FINANCIAL SCIENCES
NAUKI O FINANSACH

Financial Sciences. Nauki o Finansach
Year 2026, Vol. 31, No. 1
ISSN 2449-9811
journals.uew.pl/fins

Determinants of FinTech Stock Returns: Evidence from the Warsaw Stock Exchange Using Machine Learning

Katarzyna Perez

Poznan University of Economics and Business, Poland

e-mail: katarzyna.perez@ue.poznan.pl

ORCID: 0000-0003-3331-8456

Katarzyna Czaplińska

Poznan University of Economics and Business, Poland

e-mail: czaplinska.kat@gmail.com

ORCID: 0009-0006-5685-3434

Quote as: Perez, K., & Czaplińska, K. (2026). Determinants of FinTech Stock Returns: Evidence from the Warsaw Stock Exchange Using Machine Learning. *Financial Sciences*, *31*(1), 46-59.

DOI: 10.15611/fins.2026.1.04

JEL: G17, G12, C45

Abstract

Aim: The aim of this study was to identify the key determinants of stock returns of FinTech companies listed on the Warsaw Stock Exchange and to evaluate the usefulness of machine learning methods in predicting the direction of FinTech stock returns in this emerging European market.

Methodology: The study was based on a sample of 12 FinTech companies observed over the period 2011–2025. The analysis employed financial indicators describing economic and financial condition of companies, including profitability, growth, leverage, and liquidity measures. A Random Forest classification model was applied, supported by correlation analysis, data preprocessing techniques and logistic regression used for additional robustness.

Findings: The results indicated that traditional correlation methods do not reveal significant relationships between financial variables and stock returns, suggesting a complex and nonlinear structure of the analysed market. The machine learning model achieved moderate predictive performance. Feature importance analysis identified operating profit margin, revenue growth, and return on equity as the most influential variables.

Implications: The findings highlight the limitations of traditional econometric approaches and support the application of machine learning methods in financial market analysis. For practitioners, the results suggest that operational performance and growth indicators are key factors in evaluating FinTech companies.

Originality/value: The article contributes to the literature by combining machine learning techniques with the analysis of stock return determinants in the FinTech sector within an emerging market context, which remains underexplored.

Keywords: FinTech, stock returns, machine learning, Random Forest, Warsaw Stock Exchange

1. Introduction

The financial services sector has undergone significant transformation in recent years due to rapid technological progress and the growing role of digital innovation. Financial technologies, commonly referred to as FinTech, combine financial services with modern digital solutions such as artificial intelligence, big data analytics, cloud computing, and blockchain technologies. As a result, FinTech companies increasingly influence the structure and functioning of financial markets, challenging traditional financial institutions and reshaping the ways financial services are delivered and consumed (Kou & Lu, 2025; Lee & Shin, 2018).

The rapid development of the FinTech sector has attracted increasing attention from both practitioners and researchers. FinTech companies are often characterised by strong growth potential, high levels of innovation, and the ability to scale their business models quickly, while at the same time they operate in a highly dynamic and uncertain environment shaped by technological change, regulatory developments, and evolving consumer expectations. These characteristics make the valuation and prediction of the market performance of FinTech companies particularly challenging.

In parallel with the growth of the FinTech sector, financial research has increasingly adopted advanced analytical techniques, including machine learning methods. Compared with traditional econometric models, machine learning algorithms are capable of identifying complex nonlinear relationships and interactions between financial variables, which may improve the accuracy of stock return predictions. Previous studies demonstrated that machine learning approaches can significantly enhance forecasting performance in empirical asset pricing and financial market prediction (Gu et al., 2020). Consequently, these methods are increasingly used in areas such as stock return prediction, risk assessment, and financial decision-making.

Despite the growing literature on financial technologies and the expanding use of machine learning in finance, empirical studies focusing specifically on the stock market performance of FinTech companies remain relatively limited. Earlier research mainly examined the role of FinTech in financial intermediation, innovation in financial services, and the interactions between FinTech companies and traditional financial institutions (Carlini et al., 2022). Only a limited number of studies have analysed the determinants of their stock returns, and even fewer applied machine learning techniques in this context (Gil-Corbacho et al., 2023; Meher et al., 2024).

This research gap is particularly visible in emerging European capital markets. In the case of Poland, existing studies primarily focused on the development of the FinTech ecosystem, regulatory frameworks, and the structure of the financial innovation market. However, empirical analyses investigating the determinants of stock returns of FinTech companies listed on the Warsaw Stock Exchange remain scarce.

Therefore the aim of this study was to identify the determinants of stock returns of selected FinTech companies listed on the Warsaw Stock Exchange and to evaluate the usefulness of machine learning methods in predicting the direction of stock returns. To achieve this objective, the authors employed the Random Forest classification model, supported by correlation analysis, time-ordered validation, and comparison with logistic regression as a benchmark model.

The empirical results indicated that selected financial indicators play a meaningful role in predicting the direction of stock returns, although their predictive importance varies across variables. The applied machine learning model demonstrated moderate but meaningful predictive performance, suggesting that financial indicators contain some useful information about the direction of FinTech stock returns, although their explanatory power remains limited. At the same time, the results suggested that relationships between financial indicators and stock returns in the FinTech sector may be complex and nonlinear, which supports the application of advanced analytical methods.

The contribution of this study is twofold. First, the article extends the literature on stock return determinants by focusing on FinTech companies operating in the Polish capital market, which remains relatively underexplored in academic research. Second, it contributes to the growing body of research on the application of machine learning methods in finance by demonstrating the usefulness of the Random Forest algorithm in predicting stock return directions in the FinTech sector.

The remainder of the article is structured as follows. Section 2 presents the literature review on the development of the FinTech sector and the application of machine learning methods in financial market analysis. Section 3 describes the research methodology and the dataset used in the study. Section 4 presents the empirical results. Finally, Section 5 discusses the findings and provides concluding remarks.

2. Literature Review

2.1. FinTech as a Transforming Segment of the Financial System

Financial technologies (FinTech) have become an integral part of the modern financial system, significantly transforming the way financial services are created, distributed, and consumed. The concept of FinTech refers to the use of advanced digital technologies to improve or automate financial services and processes (Mention, 2019; Puschmann, 2017). The development of FinTech has been driven by technological progress, increased data availability, and changing consumer expectations, leading to the emergence of innovative business models that challenge traditional financial institutions (Lee & Shin, 2018).

The literature stresses that FinTech should not be understood solely as a technological phenomenon, but rather as a structural transformation of the financial sector. FinTech companies operate at the intersection of finance and technology, often offering more flexible, user-oriented, and cost-efficient solutions compared to traditional banks (Thakor, 2020). At the same time, the growing importance of FinTech has led to increased competition, collaboration, and integration between FinTech companies and incumbent financial institutions.

Recent studies also highlighted the systemic implications of FinTech development, including its impact on financial stability, market structure, and investment behaviour. For example, Carlini et al. (2022) showed that investment in the FinTech sector may influence stock returns and risk profiles of financial institutions, indicating that FinTech is becoming an important factor in financial markets. Similarly, Kou and Lu (2025) underlined that the rapid expansion of FinTech is closely linked to the adoption of artificial intelligence and data-driven decision-making processes. The growing importance of FinTech is also associated with its potential to improve the efficiency of financial intermediation by reducing costs and the increasing accessibility of financial services (Philippon, 2020).

2.2. Determinants of Stock Returns

The identification of stock return determinants has been a central topic in financial research for decades. Traditional approaches are based on asset pricing models such as the Capital Asset Pricing Model (CAPM) and multifactor models, including the Fama–French framework. These models explain stock returns through systematic risk factors, company characteristics, and macroeconomic variables.

However empirical evidence suggests that the relationships between financial variables and stock returns are often complex, unstable, and nonlinear. This is particularly relevant in the case of innovative and high-growth sectors such as FinTech, where traditional financial indicators may not fully capture company performance and market expectations. Recent studies also indicated that additional factors, such as investor sentiment, technological innovation, and intangible assets, may play a significant role in explaining stock returns (Chen et al., 2023).

Research comparing FinTech companies with traditional financial institutions suggests that the determinants of their stock performance may differ significantly. Gil-Corbacho et al. (2023) showed that FinTech companies exhibit distinct risk-return characteristics, which may require alternative analytical approaches. These findings support the need to explore new methods and models tailored to the unique features of the FinTech sector.

2.3. Machine Learning in Financial Market Prediction

In response to the limitations of traditional econometric models, financial research has increasingly turned to machine learning techniques. Machine learning algorithms, such as decision trees, random forests, support vector machines, and neural networks, are capable of processing large datasets and capturing complex nonlinear relationships between variables.

A growing body of literature demonstrates the effectiveness of machine learning in predicting stock returns and other financial variables. Henrique et al. (2019) provided a comprehensive review of machine learning applications in financial market prediction, highlighting their advantages in terms of flexibility and predictive accuracy. Gu, Kelly, and Xiu (2020) showed that machine learning methods can significantly outperform traditional regression-based models in empirical asset pricing by capturing nonlinear interactions and high-dimensional patterns in financial data. Similar conclusions were reported by Krauss et al. (2017), who demonstrated that machine learning models such as random forests and neural networks can outperform traditional approaches in stock return prediction tasks.

More recent studies confirmed that machine learning has become an important approach in empirical asset pricing and stock market prediction. Bagnara (2024) reviewed the asset pricing literature using machine learning and suggested that recent studies can be grouped into several methodological streams, including regularisation methods, dimension reduction, regression trees and random forests, neural networks, and comparative model analyses. This supports the use of tree-based models such as Random Forest, in studies that aimed to capture nonlinear relationships between financial indicators and stock returns. Ayyildiz and Iskenderoglu (2024), comparing several machine learning algorithms in stock market prediction, showed that predictive performance depends on the market and the selected algorithm, which justifies the comparison of Random Forest with a simpler benchmark model including logistic regression.

Recent research also increasingly emphasised interpretability. Zografopoulos et al. (2025) demonstrated that interpretable machine learning can be used not only to forecast industry returns but also to identify the relative contribution of individual predictors. Finally, Giantsidi and Tarantola (2025), in their review of recent deep learning applications in financial forecasting, underlined the growing importance of robustness, model evaluation, and interpretability in financial prediction models.

There are two conclusions arriving from the above. Firstly, the studies mentioned indicated that machine learning is useful in financial market analysis, but its results should be interpreted carefully and supported by appropriate validation procedures. Secondly, machine learning methods, including random forests, are particularly relevant in the context of emerging and innovation-driven sectors such as FinTech, where traditional analytical approaches may be insufficient.

2.4. FinTech Development Around the World

Although the global FinTech sector has been widely analysed, its development varies significantly across regions due to differences in institutional, regulatory, and economic conditions (Lee & Shin, 2018; Thakor, 2020). In Europe the growth of FinTech has been supported by regulatory initiatives, digital transformation strategies, and increased access to funding, although the level of adoption and market maturity differs across countries (Cumming et al., 2023).

The global scale of the sector confirms its growing importance. In 2024 the number of FinTech companies worldwide exceeded 29,000, with major hubs located in the United States, China, and Europe. At the same time market forecasts indicate that global FinTech revenues may surpass USD 18 trillion by 2028, reflecting continued expansion driven by technological innovation and increasing consumer adoption ((Statista, 2026)). In Europe the FinTech market is dominated by the United Kingdom, Germany, and France, which exhibit the highest levels of user adoption and revenue generation. For example, in 2023 the United Kingdom alone accounted for over 100 million FinTech users and generated more than USD 8 billion in revenue, highlighting its leading role in the European FinTech landscape (Statista, 2026b) .

In Central and Eastern Europe, including Poland, the FinTech sector has developed dynamically in recent years. Poland is increasingly recognised as one of the leading FinTech markets in the region, with a growing number of companies, a diversified structure of services, and a supportive institutional environment. According to recent industry reports, the number of FinTech companies in Poland has more than doubled since 2018, reaching approximately 383 entities in 2025, reflecting strong market expansion and increasing demand for digital financial solutions (Cashless.pl, 2025; FinTech Poland, 2023). The sector is dominated by payment services, although other segments such as InsurTech, blockchain, and asset management are also developing rapidly. Moreover, the financial condition of the FinTech sector has improved in recent years, with a declining share of loss-making companies and increasing profitability across the sector (FinTech Poland, 2023).

Despite the dynamic growth of the FinTech sector in Poland and across Europe, important gaps remain in the academic literature. Earlier studies primarily focused on ecosystem development, regulatory frameworks, and market trends, emphasising institutional conditions and the innovation environment (Iwanicz-Drozdowska et al., 2023; Kliber et al., 2020). In contrast, empirical research examining the behaviour of FinTech companies listed on the stock exchange is scarce – in particular, there is a lack of studies analysing the determinants of stock returns of FinTech companies using quantitative and data-driven approaches, especially in emerging European markets. This indicates that, while the sector is well described from an institutional and developmental perspective, its capital market dimension remains underexplored (Anielak, 2019).

The review of the existing literature further suggests that three major research streams have developed relatively independently. The first focuses on the growth and transformation of the FinTech sector, the second examines the determinants of stock returns within the framework of asset pricing theory, and the third explores the application of machine learning methods in financial market prediction, yet the intersection of these areas remains insufficiently explored. In particular there is a lack of empirical studies that simultaneously analyse the determinants of stock returns of FinTech companies using machine learning techniques.

This research gap is especially pronounced in the context of emerging European markets. In Poland the existing literature concentrates mainly on institutional and descriptive aspects of FinTech development, while quantitative analyses of listed FinTech firms remain limited. Hence a clear need for research integrating financial analysis with machine learning methods in order to better understand the behaviour of FinTech companies on the capital market.

3. Methodology

The aim of the study was to identify and analyse the factors determining stock returns of FinTech companies listed on the Warsaw Stock Exchange (WSE) over the period 2011–2025. To achieve this objective, machine learning techniques were applied in order to assess the impact of selected financial indicators on stock return dynamics. The article addresses the following research question: What factors significantly determine the stock returns of companies operating in the FinTech sector?

The research procedure consisted of three main stages: (1) data collection and preparation, (2) model development and analysis, (3) interpretation of results.

In the first stage, data were collected and prepared for study. The authors utilized financial data and stock price information for 12 FinTech companies listed on the Warsaw Stock Exchange. The data covered the period from 2011 to 2025 and were obtained from publicly available sources, including stooq.pl and the Biznes Radar database. The selection of companies was based on their inclusion in the Polish FinTech Map (Cashless.pl) and their listing on the WSE. Entities operating as part of larger capital groups (e.g. Allegro Pay) as well as non-independent financial services (e.g. Żappka Pay) were excluded from the sample.

The selected companies represent various segments of the FinTech sector and differ in terms of financial performance and business models. This diversity allowed for a broader analysis of the determinants of stock returns.

The prepared dataset was used to conduct empirical analysis and develop the predictive model. The dependent variable was the direction of stock return, defined as a binary classification target indicating whether the stock return increased or did not increase. The explanatory variables included return on equity (ROE), operating profit margin, Debt/EBITDA ratio, quarterly changes in revenue (Δ revenue QoQ), quarterly changes in EBIT (Δ EBIT QoQ), logarithm of market capitalisation, and current ratio. These variables reflect three main dimensions of financial condition assessment: profitability, debt and liquidity as well as the dynamics of financial results and the scale of business activity.

The selection of the explanatory variables was based on financial theories such as fundamental analysis, trade-off theory, and the Fama-French three-factor model (1993). The inclusion of profitability and growth variables (ROE, operating profit margin, Δ revenue QoQ, Δ EBIT QoQ) was derived from fundamental analysis, according to which a company's financial performance and growth potential influence its market valuation and stock returns. The inclusion of the Debt/EBITDA ratio was based on the trade-off theory, according to which the level of corporate debt influences financial risk and firm value. The current ratio was included as a measure of short-term liquidity and financial stability, while the logarithm of market capitalisation was included in line with the Fama-French model, which highlights the importance of company size in explaining stock returns.

The data cleaning procedure included the identification of missing values, removal of observations with missing dependent variables, and winsorisation of financial variables at the 1st and 99th percentiles in order to limit the impact of extreme observations. Next, the variables were standardised before model estimation. To explore relationships between variables, both Pearson's correlation and Kendall's Tau correlation coefficients were calculated. Pearson's correlation measures linear dependence, whereas Kendall's Tau captures monotonic relationships and is more robust to non--normal distributions and outliers. All data processing and model development procedures were conducted using Python. The analysis was performed with the use of standard scientific libraries, including Pandas and NumPy for data manipulation, Scikit-learn for machine learning modelling, and Matplotlib and Seaborn for data visualisation.

Next, a machine learning approach was employed to identify the key determinants of stock returns. A Random Forest classification model was selected due to its ability to capture nonlinear relationships, handle complex interactions between variables, and perform well in relatively small and medium-sized datasets. The model also enables the assessment of variable importance, which facilitates the interpretation of the influence of individual predictors on the dependent variable.

The dataset was split into training and test sets using a time-ordered approach, with an 80/20 cutoff. The first 80% of observations was used for the model estimation, while the remaining 20% served for the evaluation. The use of a sequential split, rather than random sampling, was motivated by the nature of the financial data, characterised by temporal dependencies. This approach helps to avoid look-ahead bias and better reflects real-world conditions of predictive model use, where forecasts are generated based on historical information.

To further assess the robustness and temporal stability of the predictive results, additional validation procedures were conducted using Walk-Forward Validation and Time Series Cross-Validation. These approaches are particularly appropriate for financial time series because they preserve the chronological order of observations and reduce the risk of look-ahead bias. In contrast to random cross-validation, which may lead to overly optimistic results in time-dependent data, time-aware validation procedures better reflect real-world forecasting conditions, where predictions are made using only information available up to a given point in time.

For additional robustness, the Random Forest classification model was compared with logistic regression, which served as a simpler benchmark model. This comparison made it possible to assess whether the nonlinear tree-based model provided additional predictive value relative to a standard linear classification approach. Both models accounted for class imbalance through the use of balanced class weights.

In the final stage of the study, the variable importance measure provided by the Random Forest algorithm was applied. This method evaluates the contribution of each variable by measuring the reduction in prediction error attributed to that variable across all trees in the model. Variables with higher importance scores are considered to have a stronger impact on the prediction of stock return direction.

4. Results

The empirical analysis began with a descriptive overview of the dataset. Table 1 presents summary statistics for the variables used in the study, including mean, median, standard deviation, minimum and maximum values as well as skewness (before and after winsorisation) and kurtosis.

Analysis of the descriptive statistics revealed that the financial variables of FinTech companies exhibited strong asymmetry and the presence of extreme values, especially in the case of ROE, operating margin, and debt ratios, which required winsorisation before modelling.

A variance analysis of the financial variables of the enterprises surveyed was also performed, as presented in Table 2.

Table 1. Descriptive statistics

Features	Mean	Median	St. dev.	min	max	skewness	Skewness after winsorisation	kurtosis
ROE	-0.18	0.05	2.52	-53.13	8.81	-15.80	-3.92	309.75
Operating profit margin	-1.22	0.03	7.61	-96.00	0.34	-9.45	-7.79	97.93
Net debt/EBITDA	1.93	1.81	39.71	-901.00	375.17	-16.62	1.04	418.46
Delta Revenue QoQ	1.46	0.03	29.41	-1.79	749.67	25.41	5.03	647.47
Delta EBIT QoQ	-1.66	-0.01	35.64	-825.49	137.90	-19.37	-1.86	436.30
Log (Market cap)	16.71	16.86	3.70	0.00	22.54	-2.51	-2.53	10.01
Current ratio	9.61	1.79	76.03	0.00	1889.00	23.01	5.04	566.37
Rate of return (t+1)	0.08	-0.01	0.46	-0.94	5.51	4.61	4.61	38.17

Source: own elaboration.

Table 2. Descriptive statistics

Companies	ROE variance	Operating profit margin variance	Net debt/EBITDA variance	Delta Revenue QoQ variance	Delta EBIT QoQ variance	Current ratio variance	Rate of return (t + 1) variance
AILLERON	0.01	0.01	2.00	0.74	30.89	2.30	0.06
ASSECO POLAND	0.00	0.00	0.35	0.03	0.02	0.06	0.02
COMPERIA	0.15	0.05	9.43	0.12	205.79	1.79	0.09
DATAWALK	74.14	8.94	28.25	10.41	12 612.18	70.52	0.40
FINTECH	0.16	405.12	267.90	10 788.79	91.94	2 769.03	0.30
HONEYPAYMENT	0.00	254.64	422.91	0.82	1 226.10	94 394.45	0.17
IGORIA TRADE	0.18	50.09	75.49	3.96	14.87	1.72	0.38
KBJ	0.01	0.00	5.03	0.27	24.05	3.99	0.17
MINERAL	0.20	0.00	15.82	0.45	348.58	0.71	0.09
MPAY	0.03	0.11	14 466.08	0.91	1 024.57	0.89	0.70
OPTEAM	0.63	0.00	116.31	0.28	64.63	3.82	0.09
SYGNITY	0.48	0.02	3 112.36	0.23	133.47	0.13	0.09

Source: own elaboration.

The analysed variables showed high volatility, which reflected the diversity of FinTech companies operating on the WSE. In particular, variables related to profitability, e.g. return on equity and operating profit margin, exhibited significant variability, indicating differences in financial performance across the sector. Similarly, growth indicators, including quarterly changes in revenue and EBIT, demonstrated high volatility, which is typical for innovative and rapidly developing sectors such as FinTech. The presence of variability in the dataset supports the use of machine learning methods, which are well-suited for capturing complex and nonlinear relationships in heterogeneous data.

The next step was the examination of the relationships between the variables using correlation analysis. The Pearson correlation matrix (Figure 1) indicated that there were no strong linear relationships between the analysed variables.

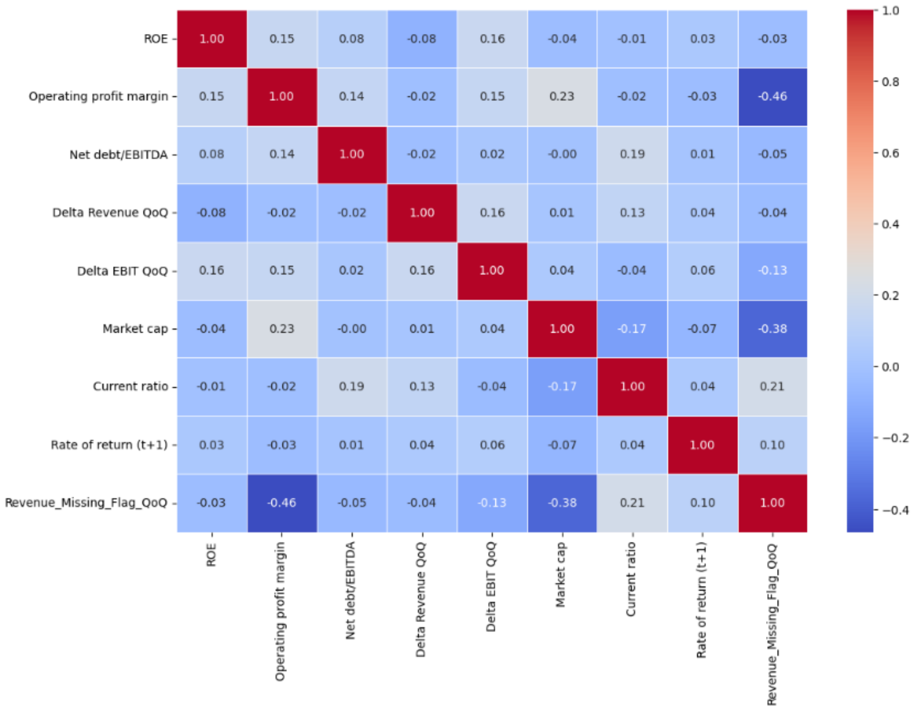

Fig. 1. The Pearson Correlation Matrix

Source: own elaboration.

To further explore potential dependencies, Kendall's Tau nonparametric correlation was calculated (Figure 2). The results confirmed the absence of strong monotonic relationships between the variables.

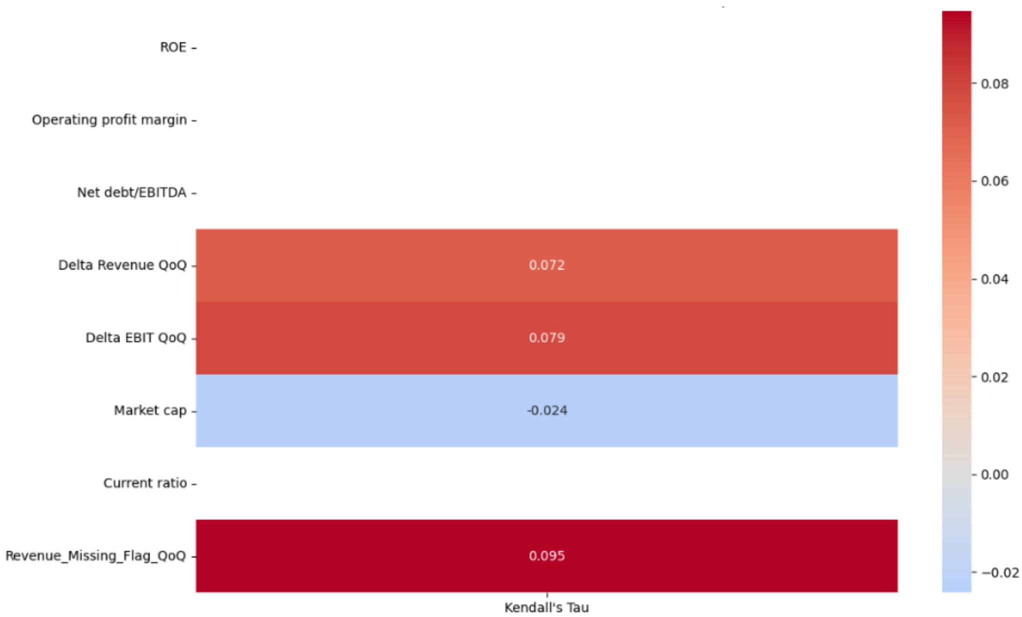

Fig. 2. Kendall's Tau Nonparametric Correlation Matrix

Source: own elaboration.

The lack of statistically significant linear and monotonic correlations suggests that the relationships between the analysed financial variables and future stock returns were complex and potentially nonlinear. This finding is consistent with the nature of financial markets, which are often characterised as complex adaptive systems.

Given these results, a Random Forest classification model was developed to predict the direction of future stock returns (increase vs. non-increase). The model achieved an accuracy of 0.61 and a ROC-AUC of 0.65. These values indicated a moderate predictive performance, suggesting that the model was able to distinguish between positive and non-positive returns better than random classification, although with a non-negligible level of classification error. The results support the presence of nonlinear relationships between financial indicators and stock returns which are not captured by traditional correlation measures.

The robustness analysis based on Walk-Forward Validation and Time Series Cross-Validation confirms that the predictive signal was not limited to a single train-test split. Across different temporal splits, the Random Forest model achieved a mean Accuracy of 0.5514, with a standard deviation of 0.0514, and a mean ROC-AUC of 0.5937, with a standard deviation of 0.0312. These results indicated that the model generally performed above the random benchmark, although its predictive power remained moderate. The relatively low standard deviations suggested that the model performance was reasonably stable across different time windows, while the moderate average values confirmed that firm-level financial indicators explain only part of the variation in stock return direction.

The comparison with logistic regression provided additional evidence on the predictive performance of the Random Forest model. On the single time-ordered train-test split, the Random Forest classifier achieved higher accuracy than logistic regression, 0.6119 compared to 0.5821, and a higher ROC-AUC, 0.6165 compared to 0.5643. The Random Forest model also achieved higher recall for the positive class, reaching 0.72 against 0.50 for logistic regression, indicating a stronger ability to identify periods of positive stock returns. However, the advantage of Random Forest should be interpreted as moderate rather than substantial. This suggests that the nonlinear model captures some additional information compared with the linear benchmark, but the overall predictive potential of company-level financial variables remains limited.

To further investigate the determinants of stock returns, a feature importance analysis based on the Random Forest model was conducted (Figure 3).

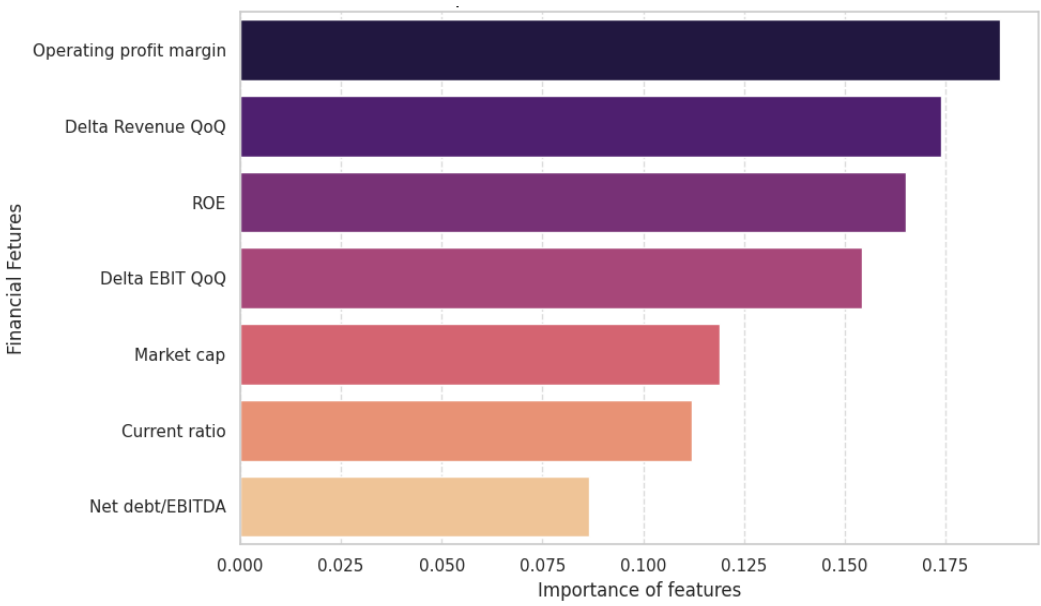

Fig. 3. Random Forest Classification - Feature Importance Analysis on Return Rate

Source: own elaboration.

The results indicated that operating profit margin was the most important variable influencing stock returns, followed by quarterly revenue growth (Δ revenue QoQ) and return on equity (ROE). These findings suggested that measures of operational performance and growth dynamics played a key role in explaining the direction of stock returns in the FinTech sector.

The results may reflect structural characteristics commonly associated with the Polish FinTech sector, including the importance of operational efficiency, scalability, and the ability to maintain profitability in a competitive and regulated market environment as well as technology-driven business models in shaping investor perception and market valuation. The dominant role of operating margin indicates that investors do not focus mainly on revenue growth dynamics, but primarily on whether this growth translates into sustainable operating profitability. The second key factor, quarterly revenue growth, shows that the market responds strongly to signals of expansion and adoption of digital financial solutions. However, revenue growth appears to be valued positively mainly when accompanied by stable or improving operating profitability, highlighting the importance of growth quality rather than scale alone.

The significant role of ROE further suggests that investors consider the efficiency of equity capital utilisation when evaluating FinTech companies. In contrast, the relatively low importance of the Debt/EBITDA ratio stresses that leverage-related factors play a less significant role in explaining short-term stock return dynamics. These findings are consistent with the characteristics frequently attributed to the FinTech sector in the literature, where operational efficiency, scalability, and the ability to sustain profitability play a central role in company valuation and investor perception.

The results reveal that financial indicators contained some predictive information regarding stock return direction, although their explanatory power remained limited. In practice this means that the model was able to capture certain patterns and regularities in the data, but not with sufficient strength to deliver highly reliable forecasts in every case. The obtained performance was better than random classification, indicating that financial variables do provide useful signals. However, the moderate level of accuracy showed that these signals were incomplete and should not be interpreted as a fully robust basis for prediction. The model is better suited as a supporting analytical tool than as a standalone decision-making mechanism, and its outputs should be complemented with additional financial or macroeconomic information.

5. Discussion and Conclusion

The results of this study provide important insights into the determinants of stock returns of FinTech companies listed on the Warsaw Stock Exchange. The analysis demonstrates that relationships between financial indicators and future stock returns were not captured by traditional correlation measures. Both Pearson's and Kendall's Tau correlations indicated a lack of statistically significant linear or monotonic relationships, suggesting that the informational structure of the analysed market was more complex than assumed by classical econometric approaches.

This finding is consistent with previous research emphasising that financial markets, particularly those related to innovative sectors, exhibit nonlinear and dynamic characteristics (Gu et al., 2020; Henrique et al., 2019). The inability of traditional correlation analysis to detect meaningful relationships supports the argument that more advanced analytical methods are required when analysing sectors such as FinTech.

The application of the Random Forest classification model provided further evidence supporting this perspective. The model achieved an accuracy of 61% and a ROC-AUC of 0.65, indicating moderate but meaningful predictive performance. These results reveal that the model was able to distinguish between positive and non-positive returns better than random classification, although its predictive power remained limited, which is broadly consistent with the literature showing that machine learning models can improve financial prediction when nonlinearities and interactions between variables are present. However, the comparison with logistic regression indicated that the advantage of Random

Forest was moderate rather than substantial, suggesting that predictive performance depends not only on the selected algorithm, but also on the structure of the dataset, the market context, and the nature of the explanatory variables. This conclusion is in line with recent evidence showing that the effectiveness of machine learning methods in stock market prediction is context-dependent (Ayyildiz & Iskenderoglu, 2024).

The feature importance analysis revealed that operating profit margin, revenue growth (Δ revenue QoQ), and return on equity (ROE) were the most influential variables in predicting stock return direction. This result is in line with previous studies suggesting that profitability and growth indicators play a significant role in asset pricing, particularly in sectors characterised by innovation and high growth potential (Fama & French, 2015; Gil-Corbacho et al., 2023). Moreover, the relatively lower importance of leverage (Debt/EBITDA) implies that capital structure may be less relevant for short--term return prediction in the FinTech sector.

This interpretation is consistent with the recent shift in machine learning-based financial forecasting toward more interpretable models. Rather than treating predictive models solely as black-box tools, recent studies stressed that feature importance measures can help identify economically meaningful predictors of stock returns and improve the interpretability of machine learning results (Bagnara, 2024; Zografopoulos et al., 2025). The applied feature importance analysis made it possible to link the predictive performance of the Random Forest model with specific financial characteristics of FinTech companies, particularly operational efficiency, growth dynamics, and capital profitability.

The dominant role of operating profit margin indicates that operational efficiency and the ability to generate sustainable core earnings are particularly important for investors. This finding may reflect the specific characteristics of the FinTech sector, where business models are often based on scalability and technological innovation. In such environments, the market may place greater emphasis on operational performance than on traditional balance-sheet indicators.

The moderate predictive performance of the model also suggests that stock returns in the FinTech sector are influenced by a broader set of factors beyond company-level financial variables. This observation is consistent with the literature on financial markets as complex adaptive systems, where prices are affected not only by fundamentals but also by external factors such as macroeconomic conditions, regulatory changes, and investor sentiment (Thakor, 2020).

The research results contribute to the existing literature in two main ways. First, they extend research on stock return determinants by focusing specifically on FinTech companies in an emerging European market. Second, they provide empirical evidence supporting the usefulness of machine learning methods in analysing financial data characterised by nonlinearity and complexity. The findings highlight that machine learning models should be treated as complementary tools rather than as fully predictive solutions.

The literature still lacks comparable studies that analyse FinTech companies using machine learning methods, particularly in the context of emerging markets, which may be partly explained by the relatively short history of the FinTech sector and limited data availability. Many FinTech companies operate outside public markets or within larger corporate structures, which restricts access to consistent financial data. Furthermore, machine learning models require large and well-structured datasets, which may not always be available for this sector.

These limitations should not discourage further research. The FinTech sector in Poland demonstrates strong growth potential, supported by increasing demand for digital financial services and technological innovation (FinTech Poland, 2023). As the sector continues to develop, the availability of data and research opportunities is likely to improve. Empirical studies on return prediction may therefore provide valuable insights for both academics and market participants, particularly in supporting investment decision-making.

The study was subject to several limitations. First, the sample included only 12 companies listed on the Warsaw Stock Exchange, which limits the generalisability of the results. Second, the model relied exclusively on company-level financial variables and did not incorporate macroeconomic or market--wide factors, which may influence stock returns.

Moreover, the relatively moderate predictive performance of the model (ROC-AUC = 0.65) indicated that financial variables explain stock return dynamics only to a limited extent. This is consistent with the nature of short-term stock returns, which are often driven by external factors such as market sentiment, macroeconomic shocks, and regulatory changes. The relatively low feature importance values suggest that no single variable dominates the prediction process, further supporting the view that stock return formation is influenced by multiple interacting factors.

Future research could improve the model by incorporating additional variables, including technical indicators, macroeconomic variables, market sentiment measures, and alternative data sources. Recent reviews of machine learning and deep learning applications in financial forecasting emphasise that multi-source feature integration may improve model robustness and predictive performance, particularly in noisy and nonlinear financial environments (Giantsidi & Tarantola, 2025). Further research should continue to pay attention to model interpretability, standardised validation procedures, and the risk of overfitting, especially when working with relatively small samples.

References

Anielak, K. (2019). Fintech as a Source of Financial Innovations on the Polish Financial Services Market. *Zeszyty Naukowe SGGW w Warszawie – Problemy Rolnictwa Światowego*, *19*(1), 162–171. https://doi.org/10.22630/PRS.2019.19.1.15

Ayyildiz, N., & Iskenderoglu, O. (2024). How effective is machine learning in stock market predictions? *Heliyon*, *10*(2), e24123. https://doi.org/10.1016/J.HELIYON.2024.E24123

Bagnara, M. (2024). Asset Pricing and Machine Learning: A critical review. *Journal of Economic Surveys*, *38*(1), 27–56. https://doi.org/10.1111/JOES.12532

Carlini, F., Del Gaudio, B. L., Porzio, C., & Previtali, D. (2022). Banks, FinTech and stock returns. *Finance Research Letters*, *45*, 102252. https://doi.org/10.1016/J.FRL.2021.102252

Cashless.pl. (2025). *Map of Polish FinTech 2025*.

Chen, R., Huang, J., Jin, C., Yang, Y., & Chen, B. (2023). Multidimensional attention to Fintech, trading behavior and stock returns. *International Review of Economics and Finance*, *83*, 373–382. https://doi.org/10.1016/j.iref.2022.09.007

Cumming, D., Johan, S., & Reardon, R. (2023). Global fintech trends and their impact on international business: a review. *Multinational Business Review*, *31*(3), 413–436. https://doi.org/10.1108/MBR-05-2023-0077

Fama, E. F., & French, K. R. (1993). Common risk factors in the returns on stocks and bonds. *Journal of Financial Economics*, *33*(1), 3–56. https://doi.org/10.1016/0304-405X(93)90023-5

Fama, E. F., & French, K. R. (2015). A five-factor asset pricing model. *Journal of Financial Economics*, *116*(1), 1–22. https://doi.org/10.1016/j.jfineco.2014.10.010

FinTech Poland. (2023). *How to do FinTech in Poland*.

Giantsidi, S., & Tarantola, C. (2025). Deep learning for financial forecasting: A review of recent trends. *International Review of Economics & Finance*, *104*, 104719. https://doi.org/10.1016/J.IREF.2025.104719

Gil-Corbacho, A., Miralles-Quirós, M. M., & Miralles-Quirós, J. L. (2023). Stock market performance differences between fintech and traditional financial firms. *Studies of Applied Economics*, *41*(1). https://doi.org/10.25115/SAE.V41I1.9078

Gu, S., Kelly, B., & Xiu, D. (2020). Empirical Asset Pricing via Machine Learning. *Review of Financial Studies*, *33*(5), 2223–2273. https://doi.org/10.1093/rfs/hhaa009

Henrique, B. M., Sobreiro, V. A., & Kimura, H. (2019). Literature review: Machine learning techniques applied to financial market prediction. *Expert Systems with Applications*, *124*, 226–251. https://doi.org/10.1016/j.eswa.2019.01.012

Iwanicz-Drozdowska, M., Cichowicz, E., Cicirko, M., Kawiński, M., & Nowak, A. K. (2023). New technologies in the financial industry: case of Poland. *Economics and Business Review*, *9*(3), 98–123. https://doi.org/10.18559/EBR.2023.3.926

Kliber, A., Będowska-Sójka, B., Rutkowska, A., Świerczyńska, K., Zdunkiewicz, W. (2020). *FinTechs in Poland: Insights, Trends and Perspectives, Technical Report.* https://doi.org/10.13140/RG.2.2.31841.53605

Kou, G., & Lu, Y. (2025). FinTech: A literature review of emerging financial technologies and applications. *Financial Innovation*, *11*(1). https://doi.org/10.1186/S40854-024-00668-6

Krauss, C., Do, X. A., & Huck, N. (2017). Deep neural networks, gradient-boosted trees, random forests: Statistical arbitrage on the S&P 500. *European Journal of Operational Research*, *259*(2), 689–702. https://doi.org/10.1016/J.EJOR.2016.10.031

Lee, I., & Shin, Y. J. (2018). Fintech: Ecosystem, business models, investment decisions, and challenges. *Business Horizons*, *61*(1), 35–46. https://doi.org/10.1016/J.BUSHOR.2017.09.003

Meher, B. K., Singh, M., Birau, R., & Anand, A. (2024). Forecasting stock prices of fintech companies of India using random forest with high-frequency data. *Journal of Open Innovation: Technology, Market, and Complexity*, *10*(1), 100180. https://doi.org/10.1016/J.JOITMC.2023.100180

Mention, A. L. (2019). The Future of Fintech. *Research-Technology Management*, *62*(4), 59–63. https://doi.org/10.1080/08956308.2019.1613123

Philippon, T. (2020). *On fintech and financial inclusion* (841; BIS Working Papers).

Puschmann, T. (2017). Fintech. *Business and Information Systems Engineering*, *59*(1), 69–76. https://doi.org/10.1007/s12599-017-0464-6

Statista. (2026a, January 9). *Number of fintechs worldwide by region 2008-2025*. https://www.statista.com/statistics/893954/number-fintech-startups-by-region/

Statista. (2026b, June 4). *Fintech in Europe – statistics & facts*. https://www.statista.com/topics/3397/fintech-market-in-europe

Thakor, A. V. (2020). Fintech and banking: What do we know? *Journal of Financial Intermediation*, *41*, 100833. https://doi.org/10.1016/J.JFI.2019.100833

Zografopoulos, L., Iannino, M. C., Psaradellis, I., & Sermpinis, G. (2025). Industry return prediction via interpretable deep learning. *European Journal of Operational Research*, *321*(1), 257–268. https://doi.org/10.1016/J.EJOR.2024.08.032