# Comparison of Classical ARIMA Forecasting with the Machine Learning LSTM Method: A Case Study on the DAX 50 ESG Index.

**Manuel Rosinus, M. Sc.**

**Financial Markets and Capital Flows**
**12th Annual**
**27. & 28.05.2025**

www.vsfs.cz

# Manuel Rosinus

- 38-Year-old, from **Germany**
- **Ph.D. Student** in the 3rd year at the University of Finance and Administration, VŠFS, Prague
- **Master's Degree** (2013) in Applied **Computer Science** Major in Business

Since 2020, **Program Manager** at **SAP,** responsible for customer-specific developments and transformations.

accenture

2013 – 2022 Technology Consulting (Automotive and Banking)

www.vsfs.cz

# Agenda

- Introduction & Motivation
- Theoretical Background
  Brief overview of classical vs. modern forecasting approaches
  - ARIMA as classical methods
  - LSTM Machine Learning: leverage machine learning to capture nonlinear patterns
- Data and Methods
- Results
- Discussion
- Future Outlook
- Q & A

www.vsfs.cz

# Motivation

- Compare forecasting accuracy between the classical econometric ARIMA model and the machine learning LSTM approach

- Assess implications for efficiency and predictability in ESG-oriented financial markets

A case study comparing Classical ARIMA Forecasting with Machine Learning LSTM on the DAX 50 ESG Index.

www.vsfs.cz

# Theoretical Background: Forecasting Models

**Classical Methods: ARIMA (Autoregressive Integrated Moving Average)**

- Basic concept: Uses past values and errors of the time series for forecasting. Based on linear relationships.
- Strengths: Good interpretability, established method.
- Limitations: Assumption of linearity and stationarity, often cannot capture complex non-linear patterns.

**Machine Learning: LSTM (Long Short-Term Memory)**

- A type of "learning" system, inspired by the functioning of the brain.
- Can recognize complex patterns and long-term dependencies in sequential data.
- Strengths: Ability to model non-linear relationships.
- Limitations: Explainability (Black-Box)

www.vsfs.cz

# Theoretical Background: ARIMA

**ARIMA (Autoregressive Integrated Moving Average)**

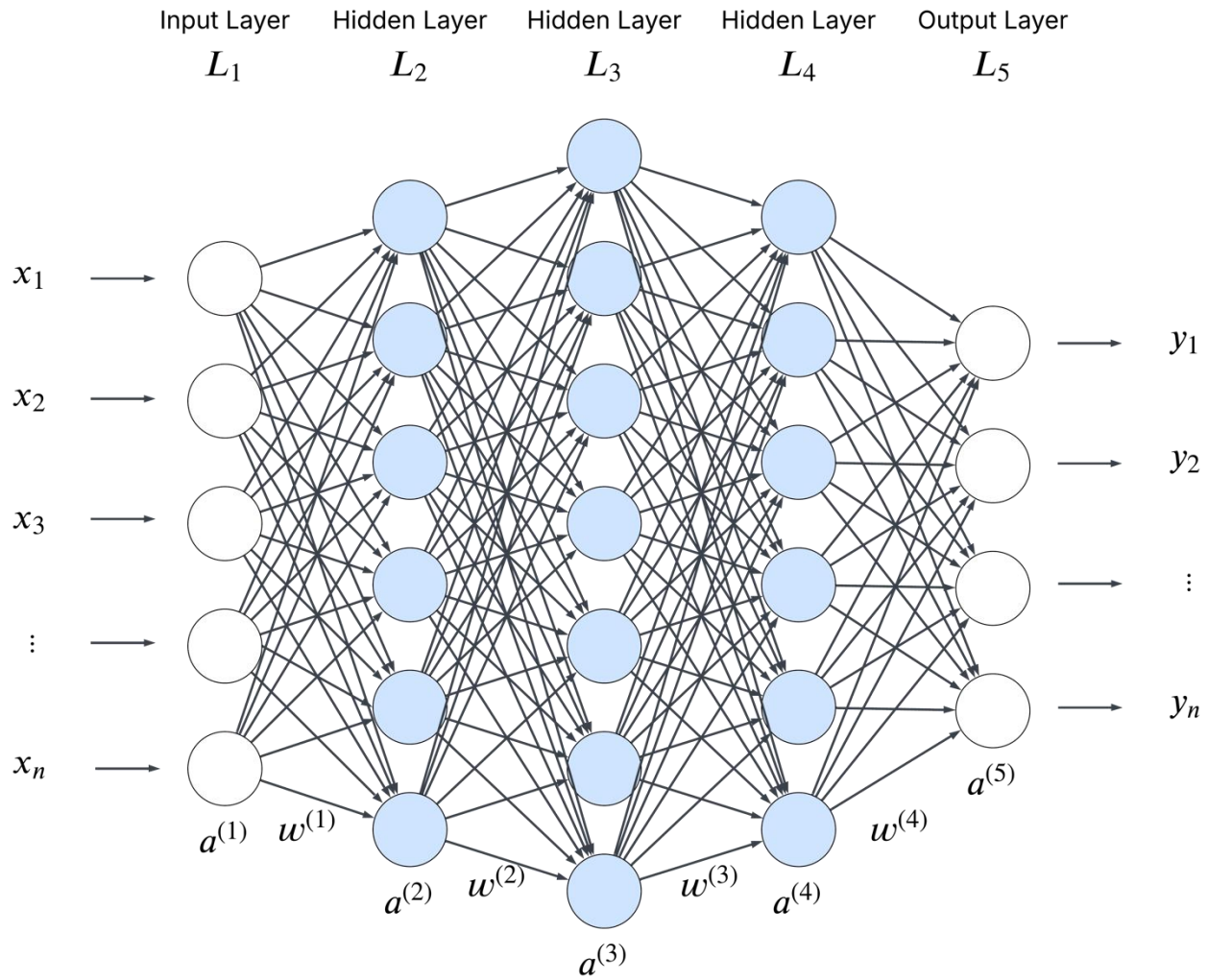
**Auto Regressive (AR):** $y_t = C + \phi_1 y_{\{t-1\}} + \phi_2 y_{\{t-2\}} + \ldots + \phi_q y_{\{t-q\}} + \epsilon_t$

**Integrated (I): Represents the differencing of raw observations to allow the time series to become stationary.**

**Moving Average (MA):** $y_t = \mu + \epsilon_t + \theta_1 \epsilon_{\{t-1\}} + \theta_2 \epsilon_{\{t-2\}} + \ldots + \theta_q \epsilon_{\{t-q\}}$

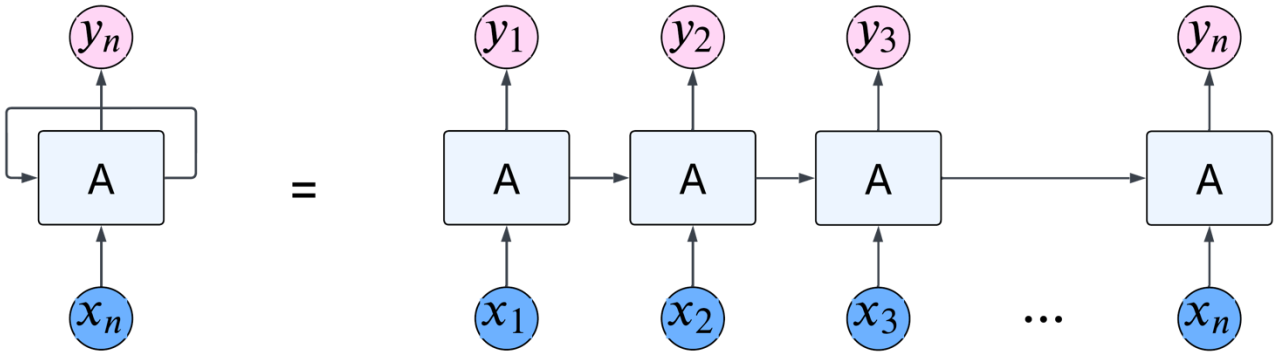
www.vsfs.cz

# Theoretical Background: LSTM: ANN




www.vsfs.cz

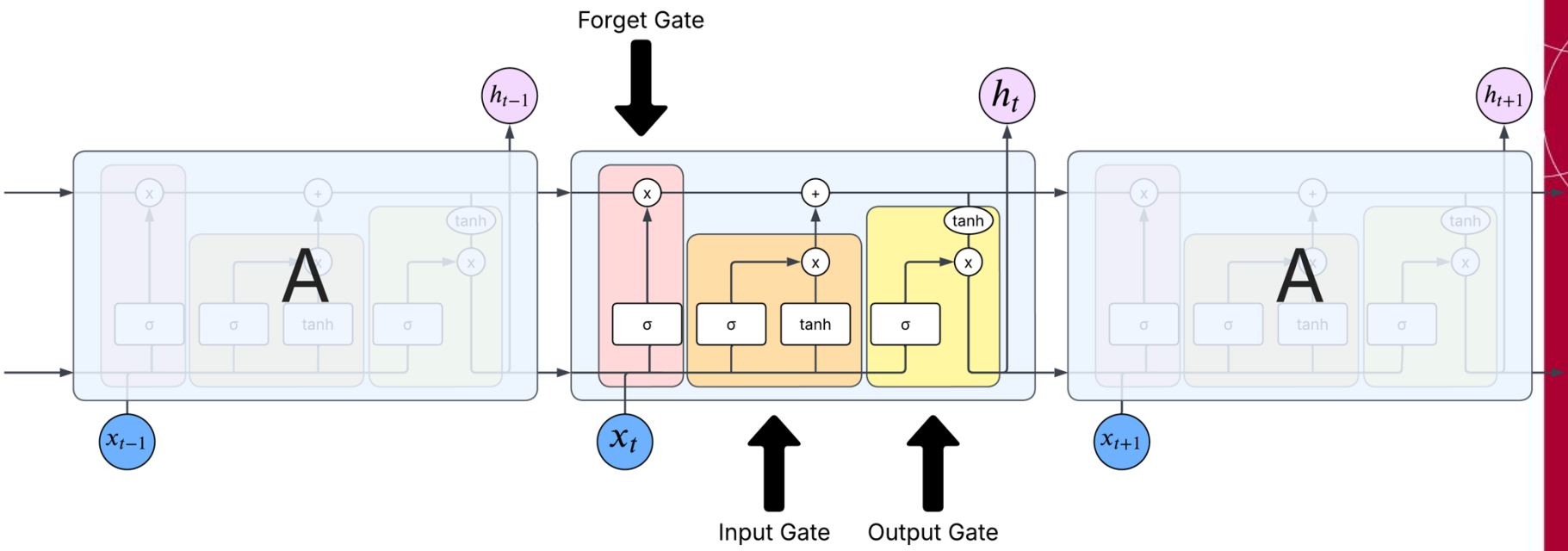
# Theoretical Background: LSTM: RNN




Unrolling the Forward Pass

www.vsfs.cz

# Theoretical Background: LSTM

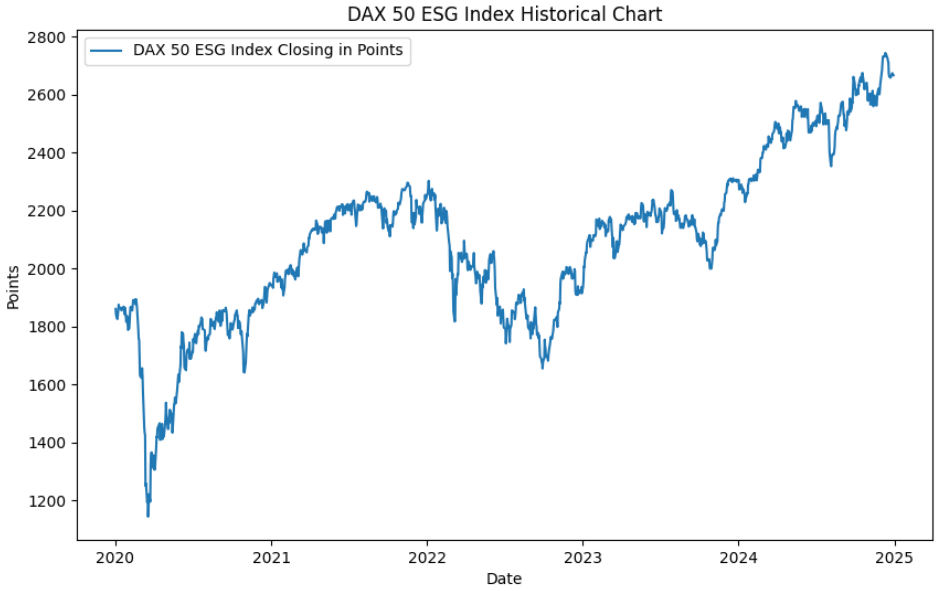



www.vsfs.cz

# Data

The DAX 50 ESG tracks the **50 largest** and most liquid German market stocks that have passed **standardized ESG screens** related to Global Standards Screening, as well as their **involvement in controversial weapons, tobacco production, thermal coal, nuclear power, and military contracting**. These stocks feature comparably good performance based on their Environmental, Social, and Governance criteria.

Observations: Daily Closing Prices from **January 1, 2020,** to **December 31, 2024**.




www.vsfs.cz

# Method (1/2)

**Method 1: Split**

The data source is split into **80 % training** and **20% test** datasets.

The training set is used to estimate the ARIMA parameters and train the LSTM neural network.

**Method 2: Expanding Window**

The models are trained and fitted on the initial 70% of the data, and the **1-step-ahead** forecast is calculated for each day t in the test set. Afterwards, the t+1 actual value is added to the model.

www.vsfs.cz

# Method (2/2)

**Evaluation Methods:**

- **Mean Absolute Error (MAE**): Average magnitude of errors. Good for datasets with outliers or non-normal errors.
- **Root Mean Squared Error (RMSE):** Measures error magnitude with emphasis on large deviations (quadratic penalty). Penalizing large errors strongly.
- **Mean Absolute Percentage Error (MAPE):** Errors as a percentage, scale-independent. Allows for error comparison across different scales.
- **Mean Directional Accuracy (MDA):** Focuses on predicting the correct direction rather than the magnitude.

**Diebold-Mariano** test was used to compare the predictive accuracy of the best-performing ARIMA specification against the final LSTM model.

www.vsfs.cz

# ARIMA on DAX 50 ESG

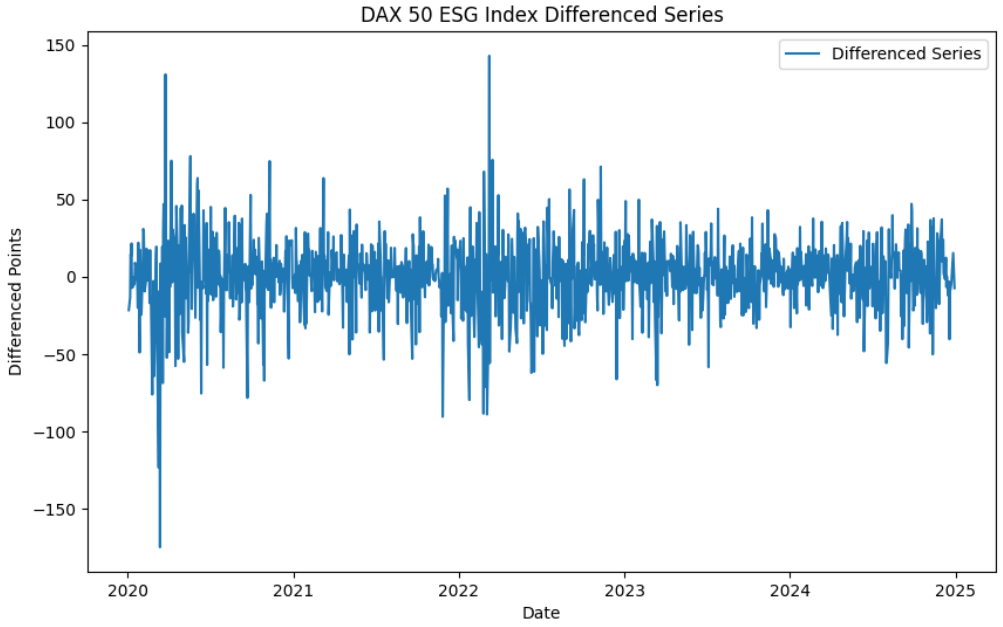
## ARIMA (p,d,q) Parameters Estimation:

- ***AR*** ***p***: the number of lag observations in the model, also known as the lag order.
- ***I*** ***d***: the number of times the raw observations are differenced; also known as the degree of differencing.
- ***MA*** ***q***: the size of the moving average window, also known as the order of the moving average.

www.vsfs.cz

# ARIMA on DAX 50 ESG

Augmented Dickey-Fuller (ADF) validates the stationarity property after first differentiation:



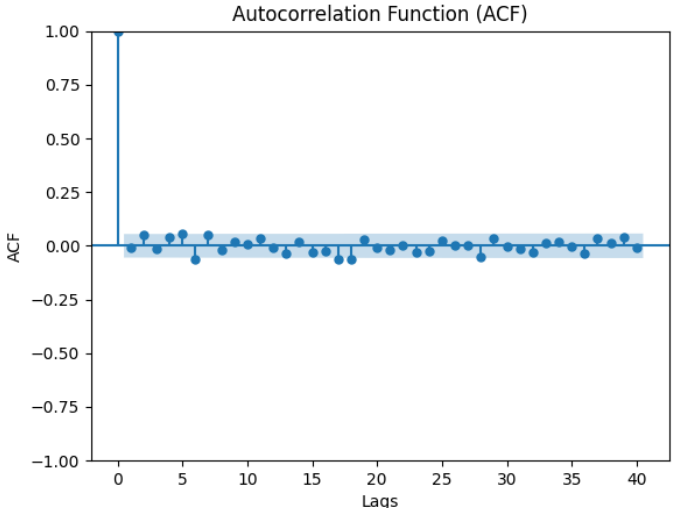

| Parameter | Value |
|---|---|
| **ADF Test Statistic:** | -14.9966 |
| **p-value:** | 0.00000 |
| **Critical value (1%):** | -3.4340 |
| **Critical value (5%):** | -2.8631 |
| **Critical value (10%):** | -2.5676 |

www.vsfs.cz

# ARIMA on DAX 50 ESG

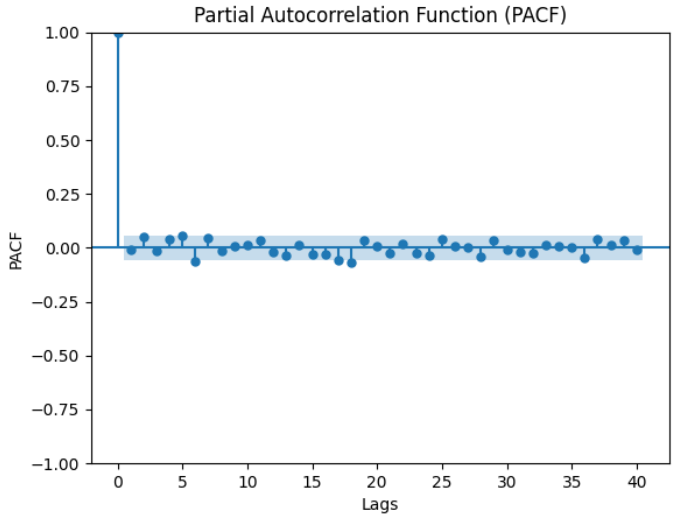
**ACF (Autocorrelation Function)**:

Helps determine *q*, the number of past forecast errors to include.




**PACF (Partial Autocorrelation Function)**:

Helps determine *p*, the number of lagged observations to consider.




www.vsfs.cz

# ARIMA on DAX 50 ESG

- ACF/PACF do not visually show any significant lags
- Models using higher ARIMA parameters are generally more prone to overfitting.
- By iterating through all combinations of p, q, and d for values < 100, the ARIMA(10,1,1) model achieved the lowest **Akaike Information Criterion (AIC)** value among the evaluated ARIMA models.
- **AIC** is a metric used in model selection to assess the trade-off between model complexity and model fit. It achieves this by penalizing models with a higher number of parameters. A lower AIC indicates that the model obtains a more balanced representation of the data, providing a more concise yet practical representation.
- As a result, the **ARIMA(5,1,1)** model was chosen as the most suitable option for forecasting in this investigation.
- Additionally, the following common ARIMA parameter combinations are also used: ARIMA(1,1,0), ARIMA(0,1,1), ARIMA(1,1,1), and ARIMA(5,1,1).

www.vsfs.cz

# LSTM on DAX 50 ESG

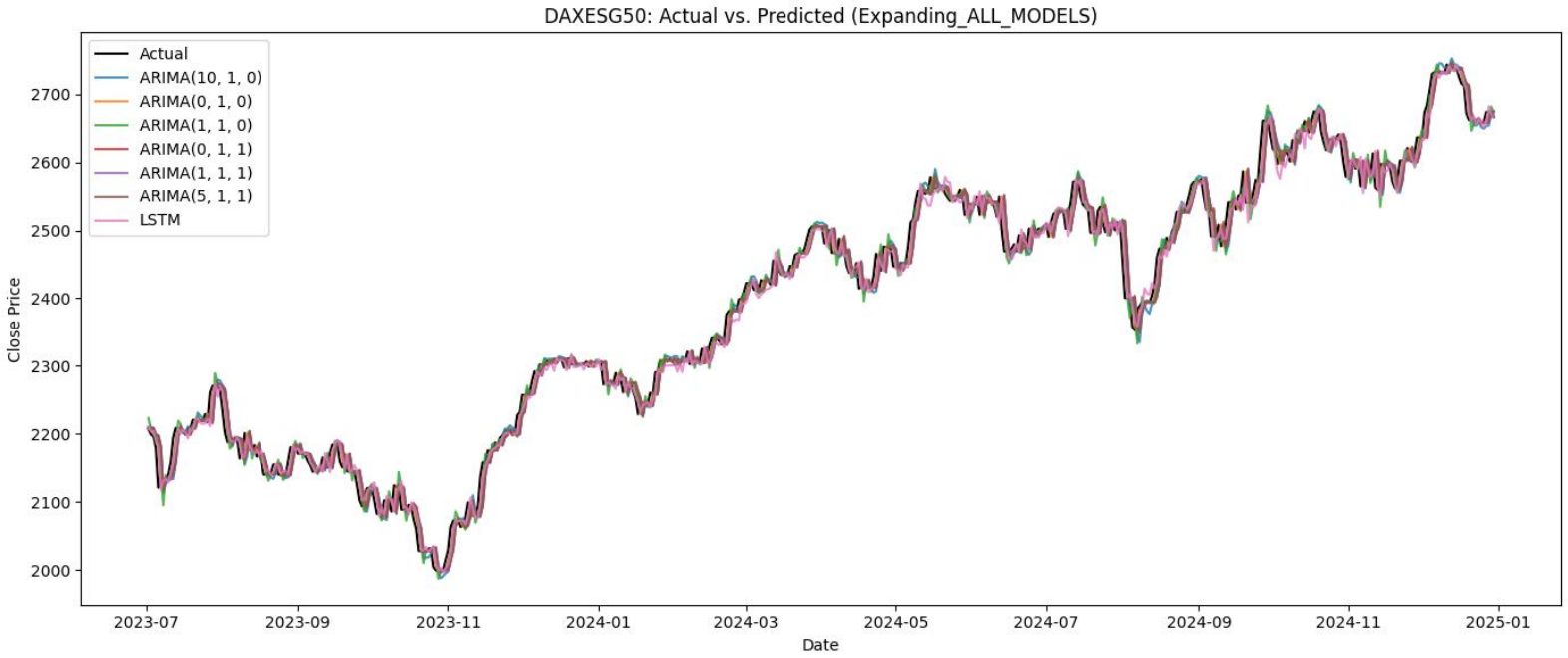
- **MinMaxScaler** to fit the full bandwidth of the closing prices within the range of 0 and 1 (on training data only)
- **Hyperparameter Tuning via** Random Search strategy (Python Keras)

The final LSTM model, consisting of an LSTM layer with 50 units followed by a dense output layer, is trained using a batch size of 32 and early stopping to prevent overfitting.

Training proceeds for up to 200 epochs, monitored via a 15% validation split, with a patience of four epochs set to halt training when improvement stalls.

www.vsfs.cz

# Results




www.vsfs.cz

# Results

| Model | Validation Method | MAE [Index Points] | RMSE [Index Points] | MAPE [%] | Directional Accuracy [%] |
|---|---|---|---|---|---|
| ARIMA(0, 1, 1) | Expanding Window | 15,4632 | 22,0198 | 0,6490 | 33,2724 |
| ARIMA(1, 1, 0) | | 14,2671 | 18,9889 | 0,5986 | 35,8318 |
| ARIMA(1, 1, 1) | | 15,6808 | 22,3392 | 0,6582 | 36,3803 |
| ARIMA(10, 1, 0) | | 16,1193 | 21,9944 | 0,6760 | **36,3803** |
| ARIMA(5, 1, 1) | | 15,8936 | 22,3449 | 0,6670 | 35,8318 |
| LSTM | | **12,6630** | **17,1242** | **0,5291** | 34,1865 |
| ARIMA(0, 1, 1) | Split Dataset | 145,9097 | 166,2172 | 5,7002 | 36,5385 |
| ARIMA(1, 1, 0) | | 201,2446 | 228,9584 | 7,8566 | 30,2198 |
| ARIMA(1, 1, 1) | | 146,1120 | 166,4455 | 5,7080 | 36,5385 |
| ARIMA(10, 1, 0) | | 155,9145 | 177,5403 | 6,0895 | 36,8132 |
| ARIMA(5, 1, 1) | | 146,9816 | 167,4292 | 5,7419 | 36,8132 |
| LSTM | | **12,3228** | **16,9290** | **0,4900** | **36,8132** |

The Diebold-Mariano test results confirmed that the superior predictive accuracy of the LSTM model compared to the ARIMA(1,1,0) was statistically significant ($p < 0.001$) under both the Split (DM statistic = 32.30) and Expanding Window (DM statistic = 4.89) frameworks.

www.vsfs.cz

# Discussion

The results indicate that the LSTM network effectively captured the underlying temporal dependencies of the DAXESG50 index data over the examined period (2020-2024), resulting in significantly more accurate point forecasts than the traditional ARIMA approach.

In this case study, it was shown in the example of the DAX ESG index as a representative of a **highly liquid marke**t that any (linear) predictable patterns in historical data, such as those considered by an ARIMA model, **can not be used to forecast the market**.

Otherwise, market participants would rapidly exploit this, leaving little to no systematic opportunity for forecasting.

**LSTM's** ability to capture **nonlinear patterns** gave it a **modest improvement**. However, the improvement is limited, which also aligns with the idea that most patterns are arbitraged away in liquid markets.

www.vsfs.cz

# Future Outlook

**Working in Progress:**

Can the Machine Learning long-term forecast be improved by incorporating ESG ratings? (Multivariate forecasting instead of univariate)

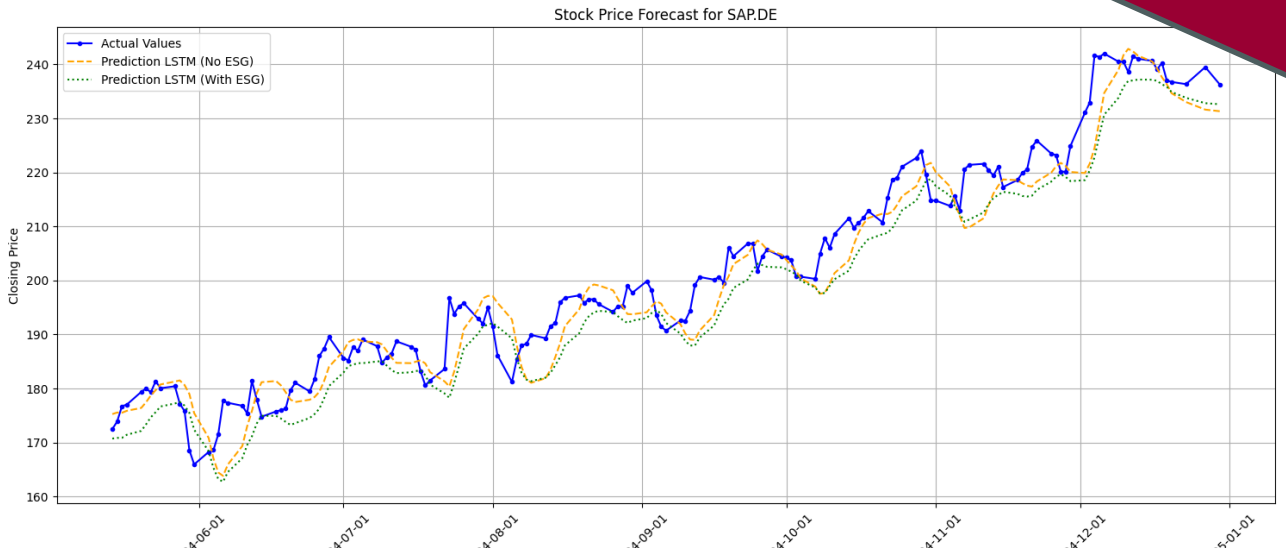
**Challenges:**

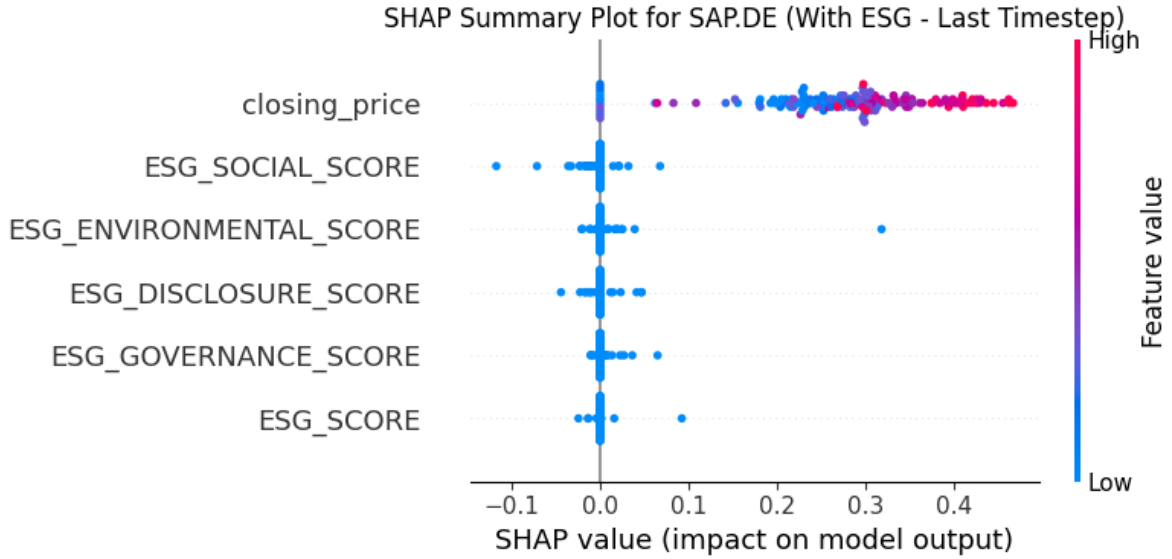
- Non-Standardized ESG Reporting
- ESG-Data availability (not available or expensive, especially historical data)
- DAX 50 ESG Index History
- ESG Data Publication Periods (yearly, and delayed)

www.vsfs.cz

# Future Outlook

preliminary results




www.vsfs.cz

# Thank you

Manuel.Rosinus@gmail.com

LinkedIn Manuel Rosinus

www.vsfs.cz

# ESG Reportings

| Regulation / Law | Legal Level | Applies To | Effective Since | Applicability to DAX ESG 50 Companies? |
|---|---|---|---|---|
| **Non-Financial Reporting Directive (NFRD)** | EU Directive | Large public-interest entities with >500 employees | January 1, 2017 | Yes; implemented in Germany via the CSR-Richtlinie-Umsetzungsgesetz |
| **Corporate Sustainability Reporting Directive (CSRD)** | EU Directive | Expands scope to large companies and listed SMEs | January 5, 2023; reporting from fiscal year 2024 | Yes; DAX ESG 50 companies are required to report under CSRD from FY 2024 |
| **European Sustainability Reporting Standards (ESRS)** | EU Standards | Companies under CSRD | Adopted July 31, 2023 | Yes; DAX ESG 50 companies must align their reports with ESRS, legally binding with CSRD. |
| **CSR-Richtlinie-Umsetzungsgesetz (CSR-RUG)** | German National Law | Transposes NFRD into German law | April 11, 2017 | Yes; mandates non-financial statements in management reports |
| **Lieferkettensorgfaltspflichtengesetz (LkSG)** | German National Law | Companies with >3,000 employees (from 2023); >1,000 employees (from 2024) | January 1, 2023 | Yes; many DAX ESG 50 companies meet the employee threshold |
| **EU Taxonomy Regulation** | EU Regulation | Financial market participants and large companies | July 1, 2021 | Yes; DAX ESG 50 companies must disclose alignment with taxonomy criteria |
| **Sustainable Finance Disclosure Regulation (SFDR)** | EU Regulation | Financial market participants | March 10, 2021 | Yes; applicable to DAX ESG 50 companies in the financial sector |
| **Global Reporting Initiative (GRI) Standards** | International Standard | Voluntary, but widely adopted | First released in 2000; GRI Standards since 2016 | Yes; many DAX ESG 50 companies use GRI for sustainability reporting |

www.vsfs.cz

# Impact of ESG Factors on Stock Performance

## Positive Impacts on Stock Performance

- **Higher ESG Ratings and Returns**: Companies with strong ESG performance often have higher stock prices and better investment performance, as they are considered more sustainable and financially stable (Kulal *et al.*, 2023).
- **Reduced Stock Price Crash Risk**: Good ESG performance can lower the likelihood of stock price crashes by reducing earnings management and corporate risk, especially in non-state companies (Luo *et al.*, 2024).
- **Lower Volatility:** ESG scores, particularly in the social dimension, are associated with reduced stock price volatility, suggesting that ESG practices can contribute to financial stability (Mechrgui and Theiri, 2024).

www.vsfs.cz

# Impact of ESG Factors on Stock Performance

## Negative or Neutral Impacts

- **ESG Rating Disagreement**: Disagreement among ESG ratings can lead to a risk premium, with stock returns positively related to such disagreements, particularly in the environmental dimension (Gibson Brandon, Krueger and Schmidt, 2021).
- **No Significant Correlation with Returns**: Certain studies have failed to establish a significant correlation between ESG performance and stock returns, suggesting that ESG factors may not be the only factor that determines of financial performance (La Torre *et al.*, 2020; *Analysis of the Impact of ESG on Financial Performance | Highlights in Science, Engineering and Technology*, no date).
- **Negative Correlation with Financial Metrics**: Evidence of a negative correlation between expected returns and ESG risk scores indicates that higher ESG risks are associated with reduced returns (Teja and Liu, 2024).

www.vsfs.cz

# Impact of ESG Factors on Stock Performance

## Mixed Findings

- **Impact on Risk Management**: ESG practices are essential to risk management, potentially reducing default risks and regulatory risks, although the correlation with stock prices can be mixed (Teja and Liu, 2024).
- **Circular Economy and ESG**: Although ESG profiles can positively impact stock performance, specific initiatives such as the circular economy do not inherently affect stock returns (Morea *et al.*, 2022).

www.vsfs.cz

# References


Kulal, A. et al. (2023) 'Impact of Environmental, Social, and Governance (ESG) Factors on Stock Prices and Investment Performance', Macro Management & Public Policies, 5(2), pp. 14–26. Available at: https://doi.org/10.30564/mmpp.v5i2.5659

Luo, W. et al. (2024) 'Can good ESG performance reduce stock price crash risk? Evidence from Chinese listed companies', Corporate Social Responsibility and Environmental Management, 31(3), pp. 1469–1492. Available at: https://doi.org/10.1002/csr.2646

Mechrgui, S. and Theiri, S. (2024) 'The effect of environmental, social, and governance (ESG) performance on the volatility of stock price returns: the moderating role of tax payment', Journal of Financial Reporting and Accounting, ahead-of-print(ahead-of-print). Available at: https://doi.org/10.1108/JFRA-03-2024-0140.

Gibson Brandon, R., Krueger, P. and Schmidt, P.S. (2021) 'ESG Rating Disagreement and Stock Returns', Financial Analysts Journal, 77(4), pp. 104–127. Available at: https://doi.org/10.1080/0015198X.2021.1963186

La Torre, M. et al. (2020) 'Does the ESG Index Affect Stock Return? Evidence from the Eurostoxx50', Sustainability, 12(16), p 6387. Available at: https://doi.org/10.3390/su12166387

Teja, K.R. and Liu, C.-M. (2024a) 'ESG Investing: A Statistically Valid Approach to Data-Driven Decision Making and the Impact of ESG Factors on Stock Returns and Risk', IEEE Access, 12, pp. 69434–69444. Available at: https://doi.org/10.1109/ACCESS.2024.3401873

Morea, D. et al. (2022) 'Circular Economy Impact Analysis on Stock Performances: An Empirical Comparison with the Euro Stoxx50® ESG Index', Sustainability, 14(2), p. 843. Available at: https://doi.org/10.3390/su14020843.


**The empirical part of this case study was done using Python and the following libraries.**

Software versions used:

Python : 3.12.7
pandas : 2.2.3
numpy : 1.26.4
matplotlib : 3.9.2
seaborn : 0.13.2
statsmodels : 0.14.4
pmdarima : 2.0.4
scikit-learn : 1.5.2
tensorflow : 2.18.0
keras : 3.9.1
keras-tuner : 1.4.7

www.vsfs.cz