Media as a Signal: Evaluating the Contribution of Economic News Sentiment to Predicting EUR/HUF Movements

By

Zita Gréta Szekér

Submitted to Central European University - Private University
Undergraduate Studies

In partial fulfillment of the requirements for the degree of Bachelor of Science and Bachelor of Arts in Data Science and Society

Supervisor: Petra Kralj Novak

Vienna, Austria

2025

Copyright Notice

[1] Icon by Font Awesome: https://fontawesome.com/

Author’s Declaration

I, the undersigned, **Zita Gréta Szekér**, candidate for the BA and BSc in Data Science and Society declare herewith that the present thesis titled “*Media as a Signal: Evaluating the Contribution of Economic News Sentiment to Predicting EUR/HUF Movements*” is exclusively my own work, based on my research and only such external information as properly credited in notes and bibliography. I declare that no unidentified and illegitimate use was made of the work of others, and no part of the thesis infringes on any person’s or institution’s copyright.

I also declare that no part of the thesis has been submitted in this form to any other institution of higher education for an academic degree.

Vienna, 26 May 2025

Zita Gréta Szekér

Abstract

This thesis tests whether the inclusion of media sentiment features enhances the predictive accuracy of euro-forint exchange rate movements. The analysis examines the role of economic news sentiment in the context of smaller and less liquid currency markets with the case of Hungary. A multilingual, data-driven framework is developed that extracts sentiment from economic news articles from three media sources: Portfolio.hu, Euronews, and Reuters. Sentiment is extracted from headlines and lead sentences with pre-trained transformer-based models. These sentiment indicators are combined with economic variables and used as inputs to an XGBoost classification model, which is tested with a sliding-window evaluation over the years 2014-2024. The results do not reveal a statistically significant improvement in predictive power from incorporating sentiment features, challenging the strength of media sentiment as a signal in this setting. This result can be an indication of information efficiency, timing mismatch, or structural properties of Hungary's foreign exchange market. However, feature importance analysis does reveal that sentiment, particularly domestic publications, consistently contributes to model decisions, illustrating the potential for such variables as monitoring tools. In general, the thesis provides a reproducible approach to sentiment-based prediction in smaller markets and contributes to the growing literature on alternative data in financial prediction.

Acknowledgements

I would like to thank my supervisor, Petra Kralj Novak, for her continuous support and guiding suggestions throughout the semester. I also wish to thank my family and friends for their encouragement and help with proofreading. Furthermore, I acknowledge the use of AI software (ChatGPT 4.0) in this thesis for occasional support with grammar and spelling checks.

Contents

Abstract iii

Acknowledgments iv

1 Introduction 1

2 Theoretical Background and Literature Review 3

3 Data and Sentiment Index Construction 7

3.1 Data Collection 7

3.1.1 Economic Data 7

3.1.2 News Data 8

3.2 Data Preprocessing 9

3.3 Sentiment Extraction and Aggregation 10

3.3.1 Models Used 10

3.3.2 Article-level Sentiment Score 11

3.3.3 Daily Aggregation 12

4 Analysis and Methodology 13

4.1 Exploratory Data Analysis 13

4.1.1 Target Variable 13

4.1.2 Economic Indicators 14

4.1.3 News Sentiment Indicators 16

4.2 XGBoost Modelling 18

4.2.1 Model Selection and Configuration 18
4.2.2 Experimental Setup 19
4.2.3 Evaluation and Comparison 19

5 Results **21**
5.1 Statistical Evaluation 21
5.2 Feature Importance 22
5.3 Robustness Checks 23

6 Discussion **25**
6.1 Interpretation of Results 25
6.2 Analytical Limitations 27
6.3 Contribution and Future Research 28

7 Conclusion **29**

Bibliography **31**

A Appendix **34**
A.1 Tables 34
A.2 Figures 36

List of Figures

4.1 EUR/HUF exchange rate from 2014 to 2024 and the distribution of its daily percentage changes, with class cutoff thresholds indicated 14
4.2 Correlation matrix of final selected economic variables 15
4.3 Daily article counts by source over time, showing variation in volume and temporal coverage across the three outlets. 17
4.4 Distribution of average daily sentiment scores by source, reflecting differences in score dispersion 17
4.5 Monthly average sentiment scores from Portfolio.hu, Reuters, and Euronews between 2014 and 2024 18

5.1 Distribution of accuracy scores for both models and histogram of pairwise accuracy differences (news model minus economic model) 22
5.2 Feature importance ranking for economic and sentiment-enhanced models, with sentiment features from Portfolio.hu, Euronews, and Reuters highlighted 23

A.1 Yearly distribution of target classes (Up, Stable, Down) based on daily EUR/HUF percentage changes 36
A.2 Proportions of predicted sentiment labels across sources, with “Neutral” being the dominant class in all three cases 37
A.3 Robustness check #1: Feature importance and accuracy distributions for the model with reweighted sentiment scores using a 40:60 headline-to-body ratio . 37
A.4 Robustness check #2: Feature importance and accuracy distributions for the model with reweighted sentiment scores using a 60:40 headline-to-body ratio . 37

A.5 Robustness check #3: Accuracy distributions for economic and news-enhanced models under binary classification setting . . . 38
A.6 Robustness check #3: Mean absolute SHAP values for binary classification models, comparing economic-only and news-enhanced models with sentiment features highlighted . . . 38

List of Tables

A.1 Description of economic variables used in the model . . . 35
A.2 English keywords used for article filtering . . . 35
A.3 Hungarian keywords used for article filtering . . . 36

List of Abbreviations

API	Application Programming Interface
ECB	European Central Bank
EU	European Union
EUR	Euro currency abbreviation
EUR/HUF	Euro-forint currency exchange rate, usually expressed in daily closing prices
FX	Foreign Exchange
HUF	Hungarian Forint
LSEG	London Stock Exchange Group
MNB	Magyar Nemzeti Bank (Hungarian Central Bank)
SHAP	SHapley Additive exPlanations
USD	United States dollar currency abbreviation
USD/HUF	Dollar-forint currency exchange rate
XGBoost	Extreme Gradient Boosting algorithm

Chapter 1

Introduction

Economic reporting is more than just event documentation; it shapes market interpretations of the events themselves. Tone, framing, and repetition in media coverage can have a substantial impact on investor sentiment and contribute to the short-term movements of the market. In the era of Big Data, advancements in natural language processing have made it possible to systematically extract sentiment from large collections of news articles. In sentiment analysis, *sentiment* refers to the emotional tone expressed in text, typically categorized as positive, negative, or neutral (Mao et al., 2024). These techniques create new opportunities for the application of media indicators as predictive variables in financial models, especially in cases where conventional economic indicators have had low efficacy, such as in the prediction of short-term exchange rates. Previous research, such as that of Uhl (2017), illustrates the valuable informational contribution of news sentiment for major exchange rates such as EUR/USD (euro-dollar), indicating that sentiment-based approaches result in improved performance metrics. However, there is limited evidence of such a connection for smaller currencies like the Hungarian Forint. Frömmel et al. (2011) emphasize the special features of the Hungarian foreign exchange market, highlighting its smaller size, lower liquidity, and economic ties to the European Union. The authors also point out that in a transition economy such as Hungary, there is a greater amount of private information in the market. These attributes might influence the way news sentiment interacts with economic indicators, thereby forming an imperative setting for research. The research question logically follows from this hypothesis: Does the inclusion of news sentiment features result in significantly higher predictive accuracy in the case of the EUR/HUF (euro-forint) exchange rate? This thesis addresses this question by focusing on economic news from three relevant media outlets and evaluating the predictive and explanatory value of sentiment features using both accuracy-based metrics and model-derived feature importance. The relevance of news sentiment is assessed by comparing a model trained on economic indicators alone with the same model augmented by sentiment features. Using a structured, data-driven approach tailored to the FX (foreign exchange) prediction task, the findings show that adding news sentiment features does not result in a statistically significant increase in accuracy in the context of the EUR/HUF exchange rate, even though the sentiment variables carry useful information for the models. These results add to recent work on alternative data in financial forecasting by highlighting that sentiment features may have limited predictive value but still offer interpretive usefulness in the context of emerging market currency prediction.

The remaining part of the paper proceeds as follows: Chapter 2 lays the theoretical and empirical foundation of the research through a literature review, highlighting the existing works that influenced the methodological design of the thesis. Chapter 3 moves on to discuss how the required datasets were sourced, assembled, and transformed to create reliable inputs for the machine learning algorithm, with an emphasis on the sentiment extraction steps. Chapter 4 describes the final input data and XGBoost (extreme gradient boosting) modelling approach in detail. It begins with a statistical overview of the input variables, followed by a presentation of the training and evaluation strategy. Building on this, Chapter 5 presents the findings of the research and supports the conclusions with robustness checks. Chapter 6 interprets the results and discusses their limitations. The paper concludes with a summary of the main findings and their implications.

Chapter 2

Theoretical Background and Literature Review

This chapter considers the theoretical fundamentals and empirical evidence related to the research question of this thesis. It begins with a brief overview of academic literature on exchange rate determination and information processing in financial markets. It next examines empirical studies on how news and media relate to financial markets more broadly. Finally, the chapter narrows its focus to research directly related to FX prediction using text-based market sentiment proxies, with particular attention to machine learning methodologies.

The foundations of exchange rate predictability theory are closely related to the influential work of Meese and Rogoff (1983), who highlighted the inherent difficulty of the task by showing that out-of-sample forecasts based on three structural macroeconomic models failed to outperform a naïve random walk. Subsequent literature responded to this challenge by exploring alternative modelling approaches, among which Taylor rule-based models have shown promise (Rossi, 2013). The Taylor rule, in its original form, serves as a monetary policy guideline for central banks setting short-term interest rates (Rossi, 2013). Molodtsova and Papell (2009) incorporated assumptions derived from purchasing power parity[1] into the Taylor rule, achieving forecast performance that exceeded the random walk benchmark under certain conditions. Other frequent predictors include commodity prices, portfolio balances and differen-

[1] Purchasing Power Parities (PPP) are exchange rates adjusted for differences in price levels between countries, allowing for more accurate cross-country comparisons of real economic values (OECD, 2024).

tials in output, money or productivity, with limited predictive power (Rossi, 2013). Short-run forecasting variants consider high-frequency data or forward-looking indicators, such as assets expressing financial market expectations, due to the frequency limitations of many macroeconomic variables (Dal Bianco et al., 2012). However, the persistent difficulty in outperforming the random walk baseline can also be interpreted through the lens of the Fama's Efficient Market Hypothesis (1970), which posits that asset prices fully reflect all available information. If FX rates are informationally efficient in this sense, it would imply that publicly observable economic indicators and fundamentals are already incorporated into prices, leaving little room for systematic prediction improvements. These works influenced the economic variable selection criteria described in Section 3.3.1.

While traditional exchange rate determination models rely heavily on macroeconomic fundamentals, empirical evidence suggests that financial markets also react strongly to information conveyed through news and the media. New pieces of information, particularly macroeconomic, firm-specific, and market news, are incorporated into investors' valuations and expectations about future market conditions and often trigger behavioral reactions such as overreaction or underreaction. These responses can lead to aggregate-level shifts in financial markets (Shantha and Ram, 2019). For currency markets specifically, the effect of macroeconomic news arrivals persists over the subsequent days, as shown by Evans and Lyons (2005), contradicting the "instantaneous" adjustment implied by the semi-strong form of market efficiency (Fama, 1970). Supporting this, Larsen and Thorsrud (2019) find that aggregated news topics extracted from newspaper text can predict asset prices and economic outcomes, suggesting that releases contain valuable forward-looking information not immediately incorporated into markets. Many researchers have expanded this idea from the pure content of news to the sentiment it conveys, recognising that emotional tone can influence investor behavior. Among a growing number of similar studies, Fazlija and Harder (2022) demonstrate promising results by using sentiment scores extracted from Bloomberg and Reuters news to forecast the direction of Standard & Poor's 500 stock index. This area of literature increasingly focuses on combining different types of machine learning approaches with various text-based sentiment indices to enhance predictive performance. These developments motivate the application of similar sentiment-based techniques to foreign exchange markets, which remain relatively underexplored compared to equity markets.

Studies that combine sentiment analysis and machine learning for foreign exchange rate prediction provide both methodological inspiration and proof of concept for this thesis. As an early empirical foundation, Evans and Lyons (2008) highlight the contribution of news sentiment to FX prediction, establishing that macroeconomic announcements explain more than 30% of the daily exchange rate change. However, the authors also note that directional effects of scheduled macroeconomic announcements are often difficult to distinguish from other factors at the daily frequency. Nevertheless, Uhl (2017) shows that news sentiment can indeed predict movements in the dollar-euro exchange rate and outperforms traditional momentum strategies. The study uses Reuters NewsAnalytics data and filtering techniques, which influenced the decision to use Reuters data in this thesis. In a related approach, Nassirtoussi et al. (2015) propose a multi-layer text-mining algorithm to predict intraday directional movements in the euro-dollar currency pair using financial news headlines. Their model integrates semantic abstraction, sentiment weighting, and targeted feature reduction, achieving up to 83.3% accuracy. The study demonstrates the predictive value of sentiment found in the article headlines, which is an important consideration in this thesis. More recently, Olaiyapo (2024) applies both lexicon-based and machine learning sentiment analysis to news and social media data to generate trading signals for major USD currency pairs. These trading signals achieved a 12% profit over the testing period, indicating the potential of the approach for informed trading decisions. Similarly, Semiromi et al. (2020) use news articles at the time of events in the economy calendar to predict intraday directional movements in four major currencies. They extract news sentiments using a dictionary-based approach and include text-mining indices, such as term frequency. The authors show the predictive power of these factors through experiments with three machine learning models. The consistent outperformance of the extreme gradient boosting (XGBoost) model influenced the model selection decision in this thesis.

Surprisingly, such sentiment-based forecasting applications within FX markets are still targeted toward major currency pairs, with rare references to smaller or emerging market currencies. This thesis begins to address this gap by exploring sentiment-based forecasting of the euro-forint exchange rate, a currency pair that presents an intriguing case study due to Hungary's status as a member of the European Union but with its own national currency.

Chapter 3

Data and Sentiment Index Construction

This chapter establishes the foundation of the modelling dataset by describing how the economic and news data were processed into structured inputs. It begins by explaining how relevant economic and news data were obtained and selected for use as model inputs. Next, it outlines the key data transformation decisions made to prepare the data for analysis. The final section describes how sentiment scores were derived and aggregated from the news articles.

3.1 Data Collection

This section outlines the data sources, extraction steps, and variable selection criteria for the two input categories: macroeconomic indicators and economic news publications. All data cover the period from 2014 to 2024 and are collected at a daily frequency.

3.1.1 Economic Data

The macroeconomic variables used in this thesis were sourced from the LSEG (London Stock Exchange Group) Workspace database (n.d.), with indicators selected based on their economic relevance and availability for timely prediction of EUR/HUF exchange rate movements. Given the complex nature of foreign exchange rate prediction, strict rules had to be applied to reduce the dimensionality of the input space. Therefore, the variable selection was guided by four principles. First, the variables must be closely tied to either the euro or forint, or broadly capture investor expectations and global financial conditions that could influence exchange rate dynamics. Second, variables were required to be available on a daily frequency. Third, they had to be made publicly accessible by the end of each trading day to ensure their usability in next-day forecasts. As the predictive timeline of this thesis is from the end of one day to the next, it is crucial to restrict inputs to information that would have been public knowledge by the close of the same day. This inevitably excludes otherwise relevant variables from the analysis that are published retrospectively on a monthly or quarterly basis, such as inflation or gross domestic product. Fourth, given the short-term nature of this prediction timeline, variables relating to short time horizons are preferred over longer ones. To illustrate, longer-term yields such as 2-year or 10-year bonds were excluded in favor of short-term rates like overnight, 1-week, or 1-month instruments, which better reflect immediate market conditions.

Building on the above priorities, the variables of interest were selected based on the work of Plankadaras et al. (2015), applied to the context of the forint and euro. The final set includes daily series on short-term interest rates (e.g., Hungary's base rate, ECB deposit facility rate), exchange rates (e.g. USD/HUF), equity indices (e.g. BUX), commodity prices (e.g. oil, gold), and volatility measures (e.g. HUF option-implied volatility, intraday Forint volatility). Each variable was chosen for its theoretical or empirical relevance to exchange rate dynamics, and its availability by the end of the day. More detailed variable descriptions are provided in Table A.1 of the appendix. The time series data for each of these factors was extracted using the LSEG Workspace API (Application Programming Interface) and consists of end-of-day values as reported at the close of each trading day.

3.1.2 News Data

The collection of economic news articles was centred around three primary sources: Portfolio.hu, Euronews, and Reuters. By focusing on just these three outlets, the analysis remains manageable from a complexity perspective while still reflecting economic news from domestic, European, and global sources relevant to the EUR/HUF rate. The local economic narrative is reflected by Portfolio.hu, which is among the ten most-read Hungarian-language newspapers in the country (Portfolio Csoport, 2018), standing out from its competitors due to its emphasis on business, economic, and financial news. Euronews, the leading international news channel in Europe (Euronews, n.d.), provides a comprehensive overview of political and economic developments across the continent, making it a suitable source for capturing regional dynamics relevant to both the euro and forint movements. Reuters contributes a global dimension through its extensive coverage of international financial markets (Reuters, n.d).

Article retrieval from these sources was conducted using the Event Registry API (n.d.) by applying a keyword-based filtering strategy tailored to economic content. This approach was chosen to ensure that only articles relevant to economic topics were captured during the extraction process, thereby reducing the need for extensive post-processing.[1] To ensure comprehensive coverage across the long time span, articles were extracted month by month. This setup accommodated the API's usage limits while allowing for manual checks to ensure all relevant content was included. The keywords list used for filtering was assembled using the topic modelling results found in the paper of Larsen and Thorsrud (2019), extended by the "*tokens of interest*" mentioned in the work of Barbaglia et al. (2023) and additional forint-specific terms. This list was applied in Hungarian for Portfolio.hu and in English for Euronews articles. For Reuters, a simplified filter was used to capture internationally significant news, selecting only articles with mentions of "*Hungary*" or "*forint*" in the headline, due to the high volume and broad scope of global coverage. The complete keyword lists in both languages are provided in the appendix, tables A.2 and A.3. Using these filters in the API queries resulted in data sets containing over 60 thousand articles from Portfolio.hu, above 50 thousand from Euronews, and around 3000 from Reuters over the years between 2014 and 2024.

3.2 Data Preprocessing

This section outlines the preparatory steps applied to the raw data, including cleaning, transformation, and the construction of target and derived variables. For all macroeconomic data, the dates were filtered to cover only trading days, as the majority of variables were only published these days. An additional derived feature was created to capture intraday volatility in

[1] The Event Registry API operates on a usage quota system, which also influenced the choice to query data in smaller, controlled batches.

the EUR/HUF rate, defined as the difference between the daily high and low prices divided by the opening price. This measure serves as a proxy for market uncertainty and relative daily price range. For the remaining economic variables (excluding intraday volatility), day-to-day percentage changes were calculated based on closing values to capture short-term market movements and to standardise the scale across features. The most important transformation is the creation of the categorical target variable from the EUR/HUF rate. After computing the daily percentage change, the values were shifted by a day so that each row contains the economic information available on a given day and the corresponding exchange rate movement on the following day. These values were then converted into a three-class categorical variable with labels "*Up*", "*Stable*", and "*Down*" by dividing the full distribution of percentage changes from 2014 to 2024 into three equally sized bins. Further details about the target variable are shown in section 4.1.1. For the news data, the main preparatory step required before sentiment classification was removing any irrelevant content, such as author names, advertisement fragments, and disclaimers. In case of both input types, the topic of missing values has to be addressed. In both datasets, missing values were examined and confirmed to reflect genuine data gaps rather than extraction errors. Since no systematic patterns were found, these values were retained and handled internally by the selected machine learning model, which supports missing-value processing.

3.3 Sentiment Extraction and Aggregation

This section explains the process of constructing daily news sentiment scores from the raw articles for all three news outlets. It first describes the predefined sentiment classifiers used, then explains how individual articles were scored, and finally details how these scores were aggregated at the daily level.

3.3.1 Models Used

Two pre-trained transformer-based sentiment classifiers were used throughout the analysis: one for the English language news and another for Hungarian. Transformer-based sentiment analysis models such as BERT (Bidirectional Encoder Representations from Transformers) and its derivatives (e.g. RoBERTa, DistilBERT) are widely favoured in recent literature due to their ability to capture semantic nuance and context more effectively than traditional lexicon-based methods (Bashiri and Naderi, 2024). This analysis utilised a pre-trained DistilRoBERTa-based model (Romero, 2021), fine-tuned specifically on financial sentences to capture domain-specific sentiments in English-language text. As no equivalent model was available for Hungarian text, a general-purpose XLM-RoBERTa-based model (Laki and Yang, 2023) was applied to the Portfolio.hu articles. The English-language model outputs one of three sentiment classes (positive, neutral, negative), while the Hungarian-language model extends this structure to five classes by adding *very positive* and *very negative*. These models provide the sentiment labels that serve as the input for per-article sentiment assessment.

3.3.2 Article-level Sentiment Score

The sentiment of each article was calculated by classifying and weighting its most informative text segments: headlines and leading sentences. Headlines play a crucial role in conveying the core message of news stories and are often the primary source of information for market participants forming expectations (Peramunetilleke and Wong, 2001). The decision to exclude the full length of articles is based on the journalistic convention of presenting key information first, followed by less relevant background information, referred to as "*lead bias*" by Zhu et al. (2021). These authors focus on the first three sentences to improve zero-shot news summarisation models. Applying these ideas to the current analysis, sentiment classifiers were run on each article's headline and first three sentences. As the classifiers return sentiment categories, these must be converted to numerical values and normalised to a -1 to 1 scale to achieve interpretable aggregations, which was done for both the 3-class and 5-class models. The final article-level score was computed as a weighted average, assigning 50% to the headline and 50% to the mean sentiment of the first three sentences. This approach reflects the assumption that headlines carry as much informational weight as the initial content of the article.

3.3.3 Daily Aggregation

Building on the article-level sentiment scores, daily sentiment indices are computed by averaging across all articles published by the given news outlet on a particular day. Days without any publication from the source are treated as missing values. The final step before merging the resulting data with economic variables involves aligning calendar dates with trading days. This inevitably requires discarding articles published over the weekend, as no corresponding economic data is available, and prevents artificially inflating sentiment scores on Fridays or Mondays by shifting weekend content to adjacent trading days. The resulting daily sentiment scores were then joined to the input economic dataset. This process produced three variables representing daily news sentiment scores, one for each source: Portfolio.hu, Euronews, and Reuters.

Chapter 4

Analysis and Methodology

This thesis applies a systematic, data-driven methodology to investigate whether incorporating news sentiment improves the prediction of daily shifts in the EUR/HUF exchange rate. The chapter opens with a statistical description of the input variables, highlighting their key characteristics and relationships. It then presents the machine learning modelling strategy, including the training setup, evaluation metrics, and a comparative analysis of models with and without sentiment features, providing the framework for the outcomes discussed in the next chapter.

4.1 Exploratory Data Analysis

This section offers a systematic examination of the target series, economic variables, and sentiment indicators, forming the basis for subsequent experimental phase and model specification.

4.1.1 Target Variable

Building on the preprocessing steps detailed in section 3.2, the target variable used in this analysis is a three-class categorical variable describing the daily directional change in EUR/HUF closing prices. The left side of figure 4.1 displays the temporal changes in the exchange rate between 2014 and 2024, showing a general increasing trend with significant volatility over time. Although the histogram on the right side of Figure 4.1 may appear roughly bell-shaped, the distribution deviates from normality due to heavy tails, as confirmed by a Shapiro-Wilk test and an excess kurtosis of 3.6. Most daily percentage changes fall between –2% and 2%, and the distribution shows a slight right skew, consistent with the upward trend in the exchange rate. The vertical lines indicate the cutoffs used to define the three classes: *Down*, *Stable*, and *Up*, set at –0.138% and 0.159% to divide the distribution into equally sized categories.

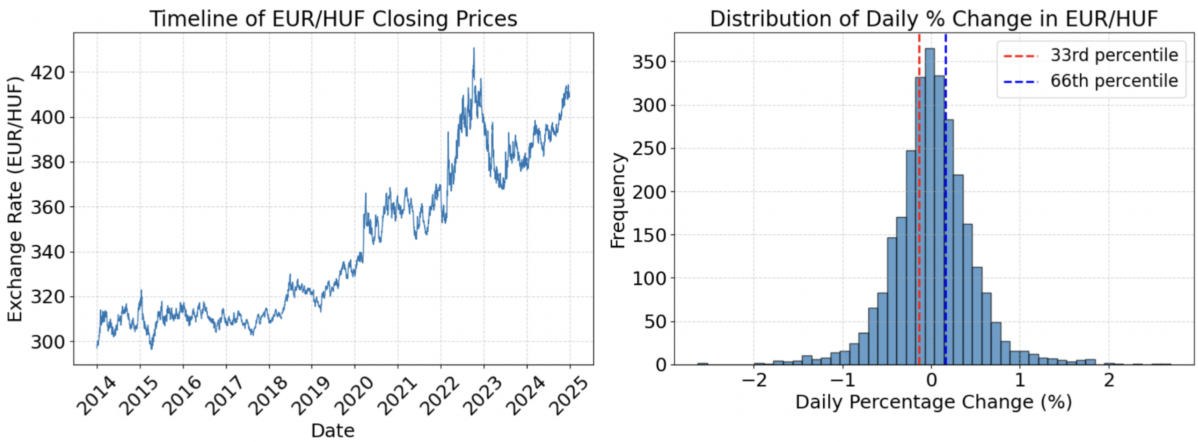

Figure 4.1: EUR/HUF exchange rate from 2014 to 2024 and the distribution of its daily percentage changes, with class cutoff thresholds indicated

The discretisation of the target variable was based on the full distribution of percentage changes to avoid potential data leakage during the modelling phase, where subsamples of the timeline are used in a sliding-window approach. While this ensures consistent class boundaries across all training and test periods, the relative frequency of the *Down*, *Stable*, and *Up* classes shifts over time. These temporal imbalances may affect model performance in later windows by introducing uneven class representation. The yearly class distribution is shown in Figure A.2 in the Appendix.

4.1.2 Economic Indicators

The economic component of the input data comprises 13 variables observed over 2,870 trading days, corresponding to 11 years of data from 2014 to 2024, with approximately 250 trading days per year. These features capture indicators of general investor behavior as well as developments specific to the euro and the Hungarian forint. Descriptions of each variable are provided in Table A.1.

The feature space has been reduced from the original 17 variables, as many of these indicators have correlations at near 1, and including them would not add information to the model. Therefore, the overnight BUBOR rates, HUFONIA rates, and 3-month benchmark rates were excluded in favor of the Hungarian Central Bank's base rate.[1] Similarly, the 1-month EURIBOR rate has been dropped due to strong overlap with the ECB deposit rate.[2] The correlations between the final set of variables are shown in Figure 4.2. Several features still exhibit high correlations around 0.85, which is expected given the interrelated nature of macro-financial indicators. The observed multicollinearity in the input space influenced the model selection, as the chosen machine learning algorithm should be robust to highly correlated features to ensure stable and interpretable predictions.

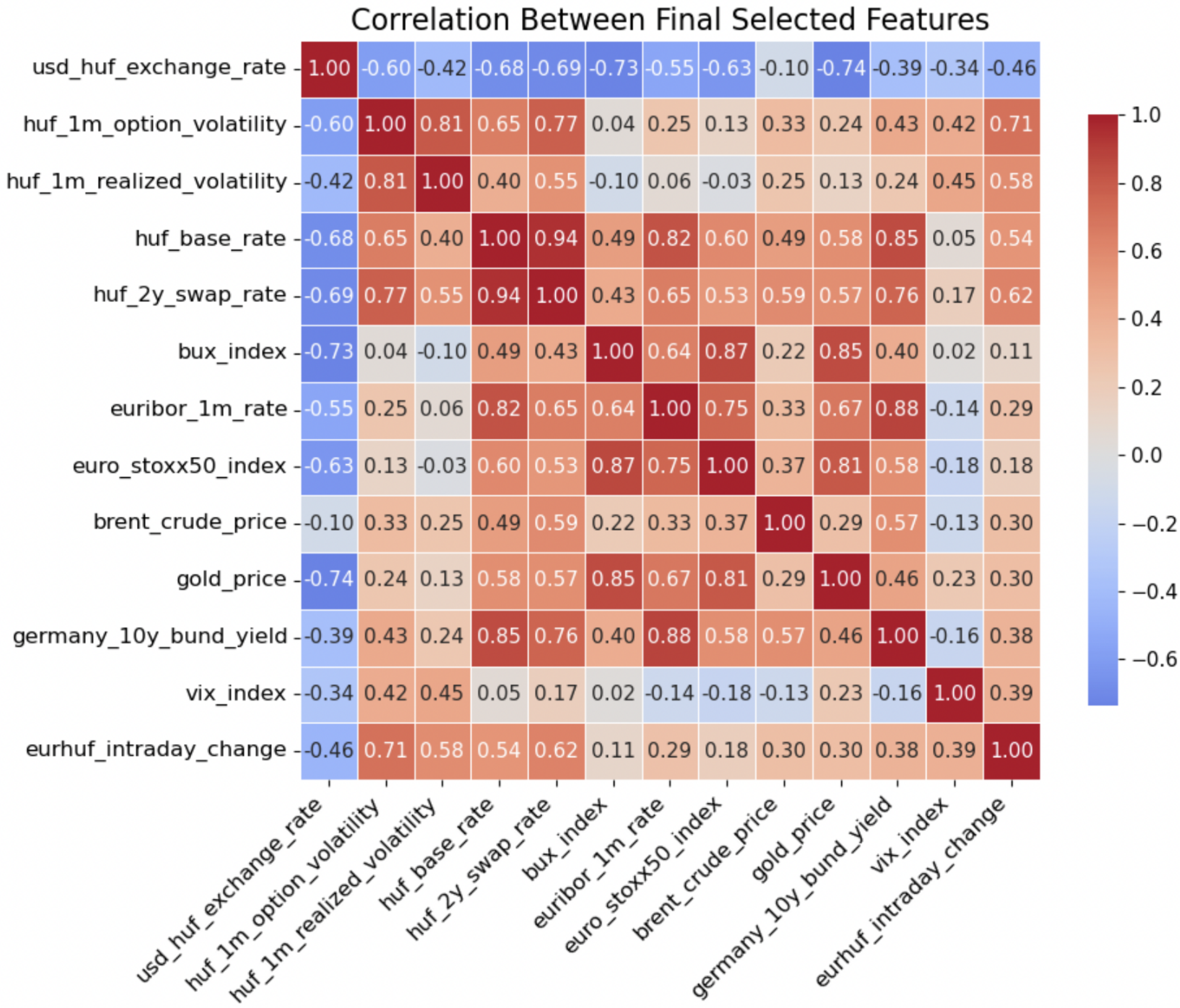

Figure 4.2: Correlation matrix of final selected economic variables

[1]BUBOR (Budapest Interbank Offered Rate) and HUFONIA (Hungarian Forint Overnight Index Average) are interbank lending rates reflecting short-term liquidity in the Hungarian market. The base rate is the main policy rate set by the Hungarian Central Bank (MNB) to guide monetary conditions.

[2]EURIBOR (Euro Interbank Offered Rate) is a benchmark interest rate at which eurozone banks lend to one another.

The final set of 13 economic features provides a high-frequency overview of daily market conditions and monetary developments related to both the euro and the forint. Expressing most variables as daily percentage changes allows the dataset to capture short-term fluctuations in the financial indicators. This economic input forms the baseline for the prediction task and serves as the benchmark against which the added predictive value of news sentiment will be assessed in the modelling phase.

4.1.3 News Sentiment Indicators

The economic news sentiment component consists of three average daily sentiment variables, one from each of the three considered sources. These variables capture variations in volume, sentiment distribution, and temporal patterns across sources between 2014 and 2024.

The underlying volume and frequency of economic articles varies between news sources, as shown in Figure 4.3. While the 61,749 articles on Portfolio.hu are concentrated in recent years, the comparable volume of 54,036 articles from Euronews shows a steady increase over time, after a sharp spike in 2017. In contrast, Reuters contributed only 3,069 articles over the entire period, with sparse and irregular coverage that was particularly limited before 2015 and early 2024. The number of articles directly influences the reliability and stability of daily sentiment scores, as days with sparse coverage can produce more volatile or less representative sentiment values.

Figure 4.4 presents the distribution of the average daily sentiment scores in the three different news sources. The large variability in Reuters is explained by its low volume and frequency of articles. Euronews scores are centered around zero and slightly skewed toward positive values. Its excess kurtosis of 6.87 indicates a heavy-tailed distribution, with occasional extreme sentiment values. Portfolio.hu displays a lower standard deviation and a distribution centered around mildly negative values. It exhibits pronounced excess kurtosis (3.91), indicating frequent central values alongside extreme outliers. This pattern results from the 5-class structure of its sentiment classifier and the absence of the extreme "*very positive*" and "*very negative*" categories. Since the normalised scale spans –1 to 1, the missing outer classes compress the range of possible scores, as illustrated in Figure A.2 in the appendix.

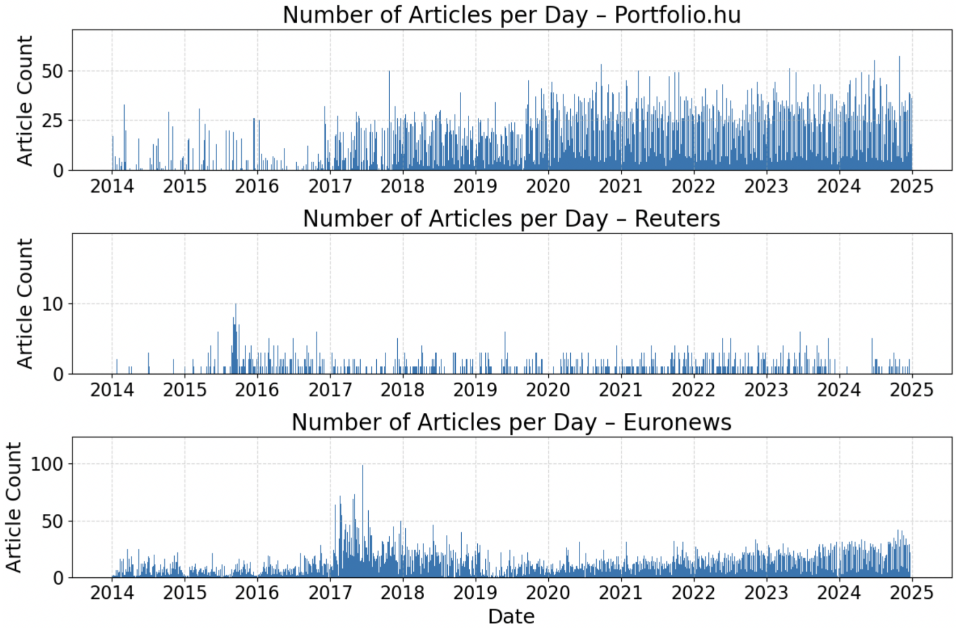

Figure 4.3: Daily article counts by source over time, showing variation in volume and temporal coverage across the three outlets.

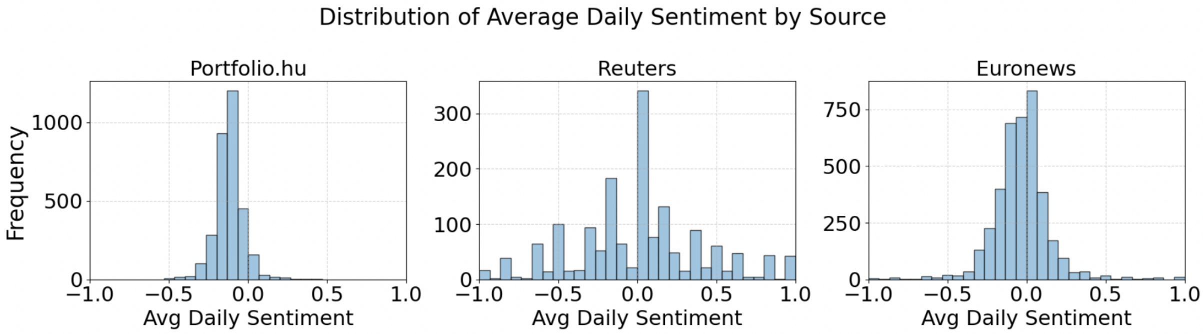

Figure 4.4: Distribution of average daily sentiment scores by source, reflecting differences in score dispersion

Figure 4.5 confirms these observations by displaying the monthly average sentiment scores per source. Portfolio.hu consistently trends negative, Euronews remains near neutral with a slight positive lean, while Reuters shows greater volatility due to lower coverage and includes several missing segments caused by months with no articles. Differences in coverage and sentiment across sources should be considered when interpreting their role in the modelling analysis. Figure 4.5 also highlights the rare comovement of the sentiment indicators, which is supported by their correlation coefficients of near 0.

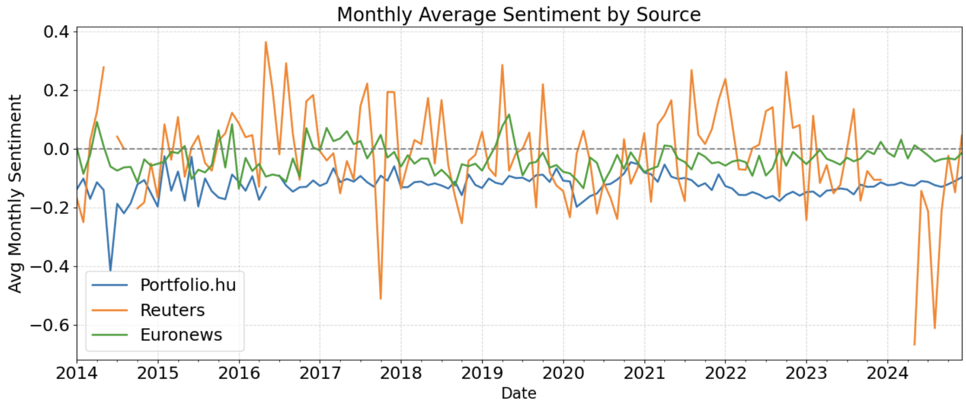

Figure 4.5: Monthly average sentiment scores from Portfolio.hu, Reuters, and Euronews between 2014 and 2024

4.2 XGBoost Modelling

This subsection concludes the methodology section by outlining a machine learning framework to address the research question: whether incorporating news sentiment significantly improves the prediction of EUR/HUF movement direction, beyond what can be achieved using economic variables alone. The underlying idea in this subsection is to compare the predictive performance of two models: one using only economic variables and another incorporating news sentiment, by constructing a distribution of accuracy scores from a series of sliding-window evaluations conducted across the 2014 to 2024 period. This involves three main steps: first, tuning the parameters of a general model to apply consistently across all window operations; second, running the sliding-window experiments and extracting accuracy and feature importance measures; and third, comparing the results between the two models.

4.2.1 Model Selection and Configuration

The XGBoost model used across all training windows was selected through hyperparameter tuning with the RandomizedSearchCV algorithm from the sklearn library (Pedregosa et al., 2011). XGBoost is a powerful tree-based machine learning model, widely adopted due to its scalability, built-in regularisation, and strong performance for a wide range of tasks (Chen and Guestrin, 2016). It is particularly suitable for this analysis, since it is robust to multicollinearity in the input space, as demonstrated in the work of Shen et al. (2024). To prevent potential data leakage[3], the baseline model was tuned using data from 2014 to 2015, which was then excluded from subsequent analysis. As the aim is to measure the added predictive power of sentiments, the parameter selection was done on the economic inputs only to isolate this effect. The model was optimised using a randomised search of 50 parameter combinations with five-fold time series cross-validation, yielding a best configuration with a cross-validated accuracy of 36.6%. This is slightly higher than the majority-class baseline accuracy of 34.6% for this time frame. The chosen model includes a maximum depth of 7, 120 estimators, a learning rate of 0.25, and subsample and column sampling rates of 0.7 and 0.9, respectively. Early stopping conditions, as well as L1 and L2 regulisation parameters were included to prevent overfitting the models[4].

4.2.2 Experimental Setup

Using the optimal hyperparameters from the randomised search, sliding window evaluations were performed by training on a 3-year period, testing on the subsequent 3 months, and shifting each window forward by one month in each iteration. The process began with the first training window from 2017 to 2019 and continued until the final testing window at the end of 2024. For each iteration, the model's accuracy on the test set and the SHAP (SHapley Additive exPlanations) feature importance scores were stored (Lundberg and Lee, 2017). The procedure was repeated for both input configurations (with and without the news sentiment variables), resulting in 71 accuracy scores for each setup.

4.2.3 Evaluation and Comparison

The two models, with and without news sentiment components, are compared on the basis of the accuracy scores and feature importance values obtained from the sliding-window experiments. Since both models were trained and tested on the same time intervals, the resulting accuracy scores form paired observations, allowing for direct statistical comparison. For such paired data, the appropriate statistical test is either a paired t-test or a Wilcoxon signed-rank

[3] Data leakage refers to the unintentional use of information from outside the training set during model development, which can lead to overly optimistic performance estimates.

[4] Early stopping halts training when performance on a validation set stops improving, while L1 and L2 regularisation penalize overly complex models to reduce the risk of overfitting.

test, depending on whether the assumptions of the parametric t-test were satisfied.

Regarding feature importance, SHAP values are averaged across target classes, test samples, and sliding windows to produce mean absolute feature importance scores. These values reveal which features consistently contribute most to the model's predictions and how the inclusion of news sentiment variables alters the relative importance of the input features.

Chapter 5

Results

The analysis reveals that even though some sentiment indices show up among the most important features, adding news sentiments to the model does not result in a statistically significant accuracy improvement, indicating that their informational value may not translate into measurable predictive gains. This chapter presents the conclusions from the final model comparison using statistical testing and feature importance analysis, and supports the findings through robustness checks.

5.1 Statistical Evaluation

Having obtained the paired observations of 71 accuracy scores for each model (with and without sentiment components) according to the experimental design, the initial step in the statistical evaluation is to determine whether the assumptions of a parametric test are met in the data. The accuracy differences between the models violate the assumptions of a paired t-test on two fronts. First, the Shapiro-Wilk test result indicates that the differences are not normally distributed (Shapiro and Wilk, 1965). Second, the observations are not independent due to the temporal overlap in training data across sliding windows. As a result, the one-sided Wilcoxon signed-rank test is used, as it is a nonparametric method suitable for paired data without assumptions about the distribution of the underlying data (Wilcoxon, 1992). It is conducted as a one-sided test, as the aim is to determine whether there is a statistically significant improvement in model accuracy when adding additional features, rather than just detecting a general difference. The null hypothesis in this case is that the median difference between news-augmented accuracy and economic-only accuracy is zero, meaning that including news data does not improve predictive accuracy. The alternative hypothesis is that the median difference is greater than zero, reflecting an increase in predictive accuracy when including news indices. To reject the null hypothesis and find evidence supporting the alternative one, the p-value of the test has to be lower than 0.05. However, the test results show a p-value of 0.7, which implies no evidence of a significant accuracy improvement when including economic media sentiment. This finding is supported by Figure 5.1, where the left panel shows the distributions of accuracy scores for both models, revealing substantial overlap and no visible shift in favour of the news-augmented model. The right panel illustrates the distribution of pairwise accuracy differences, centered closely around zero with no clear skew toward positive values, reinforcing the finding that adding news sentiment does not necessarily improve model performance.

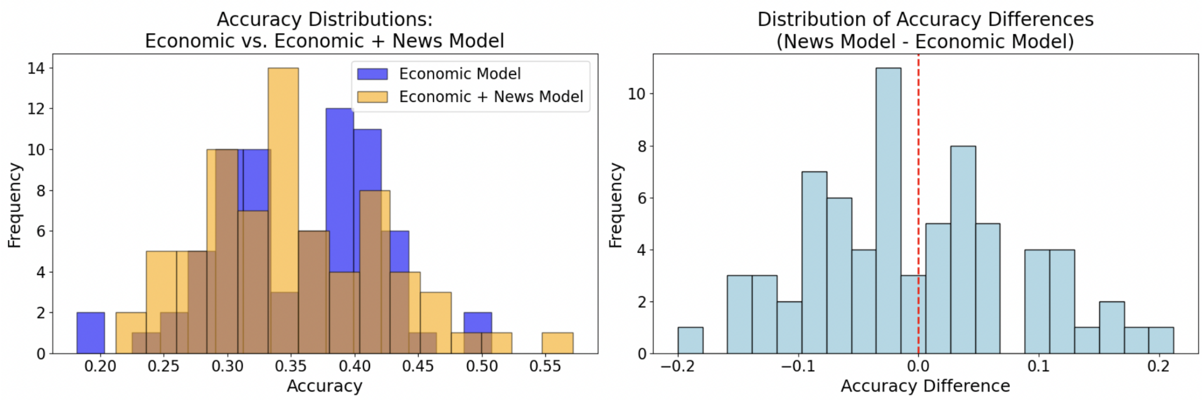

Figure 5.1: Distribution of accuracy scores for both models and histogram of pairwise accuracy differences (news model minus economic model)

5.2 Feature Importance

Only looking at feature importance differences between the two models may create a misleading impression about the significance of sentiment indices. When observing the aggregated SHAP values for the two models shown in Figure 5.2, the importance of sentiment scores from Portfolio.hu ranks second among the other features. This means that the model frequently relies on this feature in its decision-making process. The significance of this variable indicates the relevance and timeliness of domestic economic news from Portfolio.hu, which captures

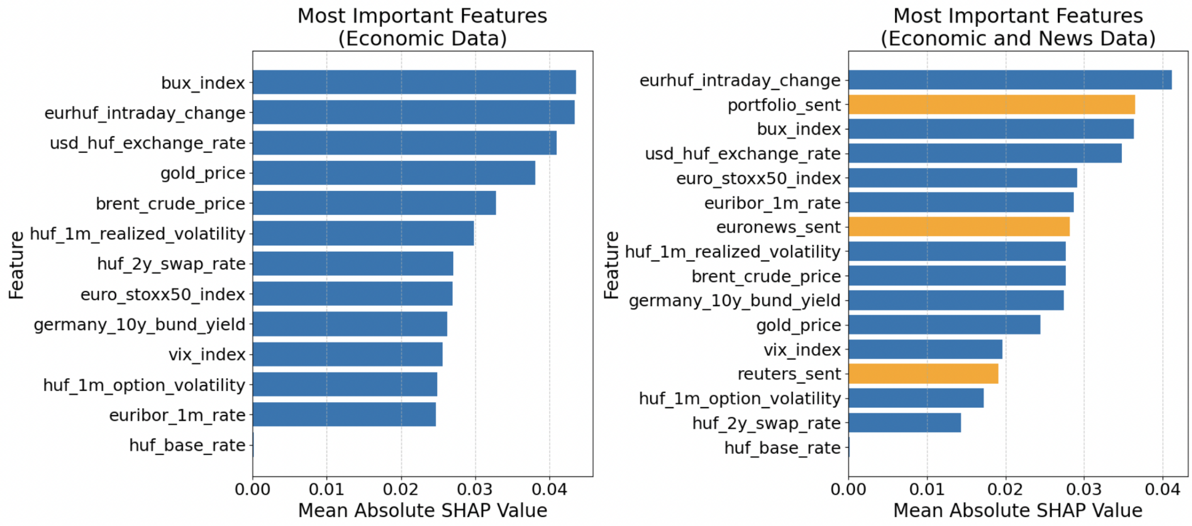

Figure 5.2: Feature importance ranking for economic and sentiment-enhanced models, with sentiment features from Portfolio.hu, Euronews, and Reuters highlighted

local market sentiment more directly than macroeconomic indicators alone. Sentiment scores extracted from Euronews also appear moderately influential, while Reuters sentiment ranks among the least important features. This likely reflects the lower volume of Hungary-specific content published by Reuters compared to domestic or regional sources. It should be noted that these SHAP values are averaged across many model runs with different data, so the exact values are difficult to interpret. Since they have no fixed scale, they should be understood in terms of ranking, showing which features are more important than others, not by how much. Therefore, the reasonable conclusion from the feature importance differences is that while the models rely on some sentiment indices, the information these variables provide is not sufficient to produce a significant accuracy improvement between the models.

5.3 Robustness Checks

The validity of the findings is reinforced through robustness checks targeting two modelling assumptions. These checks examine the sensitivity of the results to changes in specific analytical choices. The first robustness check tests the assumption that daily aggregated sentiment scores should assign equal weight to headlines and the leading three sentences. This is achieved by reweighting the aggregation in both directions: first by assigning 60% to headlines and 40% to the article body, and then by reversing the weights to 40% for headlines and 60% for the body. The model's performance and feature importance are then re-evaluated under both scenarios. In both cases, the overall conclusion of no significant accuracy increase holds, as the p-values remain above 0.05, with only minor changes in the feature importance rankings. When headlines are given majority weight, the Wilcoxon test p-value decreases slightly to 0.46, the importance of sentiments from Portfolio.hu decreases, and Reuters improves its ranking by one spot. Conversely, when the leading three sentences are emphasized in the weighting, the p-value of the statistical test increases to 0.768 and Portfolio.hu still loses its initial second place in the feature importance ranking. These results reveal that the findings are robust to moderate variations in how sentiment scores are aggregated. The feature importances and sentiment score distributions for both cases are shown in Figures A.3 and A.4 in the appendix. The second robustness check involves simplifying the prediction task by transforming the target variable from a three-class to a binary classification problem. The two classes were assigned based on an equal split of the full distribution of percentage changes. This robustness check also required updating the baseline XGBoost model parameters to accommodate a binary target. Re-evaluating the resulting models reveals that there is still no significant accuracy improvement when including media sentiment indicators. This is because the Wilcoxon test p-value increased to 0.8, which does not allow for rejection of the null hypothesis. The feature selection ranking was more sensitive to this change, as Euronews sentiment scores lost their importance dramatically and Portfolio.hu scores moved down the ranking by two spots. This is illustrated in Figure A.6 in the appendix, along with the corresponding sentiment score distribution in Figure A.5. Overall, the robustness checks reveal that feature importance scores are more sensitive to small adjustments in the methodological assumptions, while the conclusion of no significant accuracy difference remains stable.

Chapter 6

Discussion

This chapter critically evaluates the empirical findings by considering their broader context, theoretical relevance, and methodological boundaries. It begins with an interpretation of the model results and their alignment with existing literature, followed by a discussion of key limitations and the implications they hold for related future research.

6.1 Interpretation of Results

Before relating the conclusions in this thesis to existing research, it is important to first discuss the context in which these results should be interpreted and explore possible reasons for the absence of statistically significant findings. As presented in section 5.1, while adding sentiment variables did not lead to statistically significant improvements in predictive accuracy, they still appeared as relevant features, with findings remaining stable across multiple robustness checks. These outcomes should be understood within the specific context of the EUR/HUF exchange rate, Hungary's position as an EU member state with its own national currency, and the three news sources analysed, as these factors likely shaped both the sentiment signals and the model's performance. Generalisations should therefore remain within these boundaries. Within this scope, several factors explain why the addition of sentiment variables did not lead to measurable improvements in predictive accuracy. Firstly, the limited feasibility of FX rate prediction might be reflected in these results, as argued by the Efficient Market Hypothesis and supported by the influential work of Meese and Rogoff (1983). This means that public news sentiments, especially from dominant news sources like the ones used in this analysis, may already be priced in. Secondly, closely connected to this, there could be an analytical mismatch between the reaction of FX rates to news and the timing used in this thesis. Using daily sentiment aggregates may lag or miss intraday price formations, so the predictive value of sentiment may be reduced if markets react before the daily average is captured. This lag could lead to underestimating the informational content of news in fast-moving FX markets. Thirdly, certain non-economic factors such as political developments or geopolitical risks may influence the EUR/HUF exchange rate but are not systematically captured by the sentiment indicators derived from purely economic news. Fourthly, the presence of private information in transitional economies like Hungary, as noted by Frömmel et al. (Frömmel et al., 2011), may weaken the observable relationship between public news sentiment and exchange rate movements. Despite the lack of significant increase in accuracy, the feature importance rankings still suggest that the sentiment scores contained useful informational content from the model's perspective. The frequent reliance on Portfolio.hu sentiment suggests that domestic, Hungarian-language news carries informational value and may reflect investor attention to local developments. Euronews showed only moderate importance, likely due to its broader regional focus, while the low importance of Reuters may be explained by its infrequent coverage, similar to the limited predictive value of the HUF base rate, which is adjusted only occasionally. Connecting these results to related literature, the comparison is challenging due to the differences in scope, especially in terms of currency pairs, the economic background of the countries involved, and news coverage used. Much of the existing literature, such as Uhl (2017) and Nassirtoussi et al. (2015), focuses on improving the performance of single models and reports positive effects of news sentiment indices without formal hypothesis testing. Nonetheless, the observed relevance of Portfolio.hu sentiment supports findings in the literature that highlight the potential value of domain-specific and localised news sentiment in shaping market expectations.

6.2 Analytical Limitations

The design of the empirical framework reflects careful methodological choices, yet it also includes structural limitations in data coverage, sentiment representation, and modelling assumptions that affect both performance and interpretability. Starting with the data extraction process, the focus on three media outlets, restriction to economic news and keywords-based article collection may introduce selection bias and reduce the representativeness of the derived sentiment components. The uneven source coverage and low article volume in the case of Reuters also impacts the reliability of the derived features. When constructing the sentiment variables, the dependence on two pre-trained classifiers introduces a degree of opacity to the scoring process, as these models function as black boxes with limited interpretability and provide the user with little information about potential structural biases. The limitations coming from the classifiers are particularly important to mention for Hungarian articles, as no domain-specific financial sentiment model was available in that language. Focusing sentiment extraction to article headlines and the first three sentences, while based on empirical justification, may have excluded relevant context or narrative shifts found deeper in the text. Aggregating article-level sentiment scores on a daily basis might miss fast-moving intraday effects or underestimate effects of influential news such as surprising central bank announcements. Furthermore, while disregarding weekend articles is statistically justified, the Efficient Market Hypothesis suggests that relevant information may still be priced in at the market open, potentially reducing the model's ability to capture timely sentiment effects. As for the modelling process, the discretisation of the target variable is purely data-driven, aiming to create balanced class distributions, but it does so without reference to macroeconomic theory, which may limit the financial relevance of the resulting categories. The temporal overlap between the windows used in the experimental setup introduces dependence between the observations, only allowing the use of a non-parametric statistical test instead of a stronger parametric one.

6.3 Contribution and Future Research

The findings of this thesis highlight the potential informational value of domestic, local-language news sentiment when predicting exchange rate movements in smaller markets such as Hungary, especially in the context of its economic integration with the European Union. In contrast to most existing research that focuses on major currency pairs, this analysis contributes to bridging a gap in the literature by addressing a less-explored, externally sensitive market. The results introduce skepticism about the predictive power of news sentiment in fast-moving FX markets, at least within the framework and time resolution used here. Still, domestic news sentiment may hold diagnostic or monitoring value for analysts focused on the Hungarian market. The model's frequent reliance on Portfolio.hu sentiment in feature importance rankings further underlines the relevance of local, language-specific media in reflecting market-relevant information, even when overall predictive gains are limited. Methodologically, this thesis introduces a multilingual sentiment scoring framework using three news sources with different geographic orientations, paired with statistical testing and SHAP-based feature importance analysis. Unlike many previous studies that evaluate a single fitted model, this thesis employed a sliding-window evaluation and formal statistical testing to assess the generalisable effect of sentiment features. This design can serve as a replicable structure for future work on low-visibility currencies or illiquid asset classes. Given the focus on EUR/HUF and the specific media sources selected, the findings remain context-dependent, but the underlying approach can be adapted to other emerging market settings. Future research could address current limitations by incorporating intraday or event-specific sentiment effects, expanding to alternative media sources such as social media or political content, and applying econometric or causal inference techniques such as Granger causality to better understand the interaction between sentiment and exchange rate dynamics.

Chapter 7

Conclusion

This thesis set out to examine whether adding media sentiment features results in significantly better accuracy for EUR/HUF movement prediction, with a broader aim of understanding the dynamics between news sentiment features and FX prediction in case of smaller, less liquid currency markets. This goal was achieved through a framework that combines automated sentiment extraction from multilingual news articles with a focus on context-rich text segments such as headlines and leading sentences, applying a machine learning framework tailored to time-sensitive financial prediction. The empirical analysis relied on a sliding-window evaluation and feature importance metrics to ensure that results were robust and interpretable over time. The news data was narrowed down to three main news outlets with differing geographical coverage, which necessarily restricted the scope of the sentiment input. The main finding of this thesis reflects scepticism regarding the power of media sentiment as a signal, as the statistical test showed no measurable increase in the accuracy levels when including media sentiment in this context. This result can be viewed as a confirmation of the efficient market hypothesis, a result of analytical decisions regarding the timeline of prediction, or a consequence of structural characteristics of the EUR/HUF market, such as the influence of private information and the exclusion of non-economic drivers not captured by the sentiment indicators. Nevertheless, in terms of feature importance, these sentiment features, especially the one derived from local language publications, frequently contributed to the model's decisions, suggesting their value as a monitoring variable for analysts and traders. These results were strengthened through robustness checks on certain analytical decisions. It is important to mention that the used methodology is subject to limitations, such as the restricted scope of news sources, the use of general-purpose sentiment classifiers in the case of Hungarian texts, and the reliance on daily aggregation, which may overlook intraday effects. This motivates future research that expands the current framework by incorporating higher-frequency sentiment signals, broadening the range of media inputs, and applying causal inference methods to more precisely isolate the effects of sentiment on exchange rate movements. Overall, this study contributes to a more nuanced understanding of the role of news sentiment in currency forecasting for emerging markets, offering both methodological advances and practical insights for analysts, researchers, and financial decision-makers.

Bibliography

Barbaglia, Luca, Sergio Consoli, and Sebastiano Manzan (2023). "Forecasting with Economic News". en. In: *Journal of Business & Economic Statistics* 41.3, pp. 708–719. ISSN: 0735-0015, 1537-2707. DOI: 10.1080/07350015.2022.2060988.

Bashiri, Hadis and Hassan Naderi (2024). "Comprehensive review and comparative analysis of transformer models in sentiment analysis". en. In: *Knowledge and Information Systems* 66.12, pp. 7305–7361. ISSN: 0219-3116. DOI: 10.1007/s10115-024-02214-3.

Chen, Tianqi and Carlos Guestrin (2016). "XGBoost: A Scalable Tree Boosting System". en. In: *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. San Francisco California USA: ACM, pp. 785–794. ISBN: 978-1-4503-4232-2. DOI: 10.1145/2939672.2939785.

Dal Bianco, Marcos, Maximo Camacho, and Gabriel Perez Quiros (2012). "Short-run forecasting of the euro-dollar exchange rate with economic fundamentals". In: *Journal of International Money and Finance* 31.2, pp. 377–396. ISSN: 0261-5606. DOI: 10.1016/j.jimonfin.2011.11.018.

Euronews (n.d.). *About Euronews*. Accessed May 4, 2025. URL: https://www.euronews.com/about.

Evans, Martin D. D. and Richard K. Lyons (2005). "Do currency markets absorb news quickly?" In: *Journal of International Money and Finance*. Exchange Rate Economics 24.2, pp. 197–217. ISSN: 0261-5606. DOI: 10.1016/j.jimonfin.2004.12.004.

— (2008). "How is macro news transmitted to exchange rates?" In: *Journal of Financial Economics* 88.1, pp. 26–50.

Event Registry (n.d.). *Event Registry: Global News Intelligence Platform*. Accessed May 4, 2025. URL: https://eventregistry.org/.

Fama, Eugene F. (1970). "Efficient Capital Markets: A Review of Theory and Empirical Work". In: *The Journal of Finance* 25.2. Publisher: [American Finance Association, Wiley], pp. 383–417. ISSN: 0022-1082. DOI: 10.2307/2325486.

Fazlija, Bledar and Pedro Harder (2022). "Using Financial News Sentiment for Stock Price Direction Prediction". en. In: *Mathematics* 10.13. Number: 13 Publisher: Multidisciplinary Digital Publishing Institute, p. 2156. ISSN: 2227-7390. DOI: 10.3390/math10132156.

Frömmel, Michael, Norbert Kiss M., and Klára Pintér (2011). "Macroeconomic announcements, communication and order flow on the Hungarian foreign exchange market". en. In: *International Journal of Finance & Economics* 16.2, pp. 172–188. ISSN: 1099-1158. DOI: 10.1002/ijfe.424.

Laki, László János and Zijian Győző Yang (2023). "Sentiment Analysis with Neural Models for Hungarian". In: *Acta Polytechnica Hungarica* 20.5, pp. 109–128. DOI: 10.12700/APH.20.5.2023.5.8.

Larsen, Vegard H. and Leif A. Thorsrud (2019). "The value of news for economic developments". In: *Journal of Econometrics*. Annals Issue in Honor of John Geweke "Complexity

and Big Data in Economics and Finance: Recent Developments from a Bayesian Perspective” 210.1, pp. 203–218. ISSN: 0304-4076. DOI: 10.1016/j.jeconom.2018.11.013.
LSEG Refinitiv (n.d.). *Workspace*. Accessed May 4, 2025. URL: https://www.refinitiv.com/en/products/refinitiv-eikon.
Lundberg, Scott M and Su-In Lee (2017). “A Unified Approach to Interpreting Model Predictions”. In: *Advances in Neural Information Processing Systems 30*. Ed. by I. Guyon et al. Curran Associates, Inc., pp. 4765–4774. URL: http://papers.nips.cc/paper/7062-a-unified-approach-to-interpreting-model-predictions.pdf.
Mao, Yanying, Qun Liu, and Yu Zhang (2024). “Sentiment analysis methods, applications, and challenges: A systematic literature review”. In: *Journal of King Saud University - Computer and Information Sciences* 36.4, p. 102048. ISSN: 1319-1578. DOI: https://doi.org/10.1016/j.jksuci.2024.102048.
Meese, Richard A. and Kenneth Rogoff (1983). “Empirical exchange rate models of the seventies”. en. In: *Journal of International Economics* 14.1-2, pp. 3–24. ISSN: 00221996. DOI: 10.1016/0022-1996(83)90017-X.
Molodtsova, Tanya and David H. Papell (2009). “Out-of-sample exchange rate predictability with Taylor rule fundamentals”. In: *Journal of International Economics* 77.2, pp. 167–180. ISSN: 0022-1996. DOI: 10.1016/j.jinteco.2008.11.001.
Nassirtoussi, Arman Khadjeh et al. (2015). “Text mining of news-headlines for FOREX market prediction: A Multi-layer Dimension Reduction Algorithm with semantics and sentiment”. In: *Expert Systems with Applications* 42.1, pp. 306–324. ISSN: 0957-4174. DOI: https://doi.org/10.1016/j.eswa.2014.08.004.
OECD (2024). *Purchasing Power Parities (PPP) - OECD Data*. Accessed: 2024-05-13. URL: https://www.oecd.org/en/data/datasets/purchasing-power-parities.html.
Olaiyapo, O. F. (2024). “Applying News and Media Sentiment Analysis for Generating Forex Trading Signals”. en. In: *Review of Business and Economics Studies* 11.4. Number: 4, pp. 84–94. ISSN: 2311-0279. DOI: 10.26794/2308-944X-2023-11-4-84-94.
Pedregosa, Fabian et al. (2011). “Scikit-learn: Machine learning in Python”. In: *the Journal of machine Learning research* 12, pp. 2825–2830.
Peramunetilleke, Desh and Raymond K Wong (2001). “Currency Exchange Rate Forecasting from News Headlines”. en. In: *Australian Computer Science Communications* 24.2, pp. 131–139.
Plakandaras, Vasilios, Theophilos Papadimitriou, and Periklis Gogas (2015). “Forecasting Daily and Monthly Exchange Rates with Machine Learning Techniques”. en. In: *Journal of Forecasting* 34.7. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/for.2354, pp. 560–573. ISSN: 1099-131X. DOI: 10.1002/for.2354.
Portfolio Csoport (2018). *TOP15: A legolvasottabb oldalak között a Portfolio*. Accessed May 4, 2025. URL: https://portfoliocsoport.hu/hirek/2018/04/10/top15-a-legolvasottabb-oldalak-kozott-a-portfolio-62.
Reuters (n.d). *About Reuters*. https://reutersagency.com/about/. Accessed: 2025-05-23.
Romero, Manuel (2021). *distilroberta-finetuned-financial-news-sentiment-analysis*. Hugging Face model repository. Developed by Manuel Romero, username mrm8488. Accessed May 4, 2025. URL: https://huggingface.co/mrm8488/distilroberta-finetuned-financial-news-sentiment-analysis.
Rossi, Barbara (2013). “Exchange Rate Predictability”. en. In: *Journal of Economic Literature* 51.4, pp. 1063–1119. ISSN: 0022-0515. DOI: 10.1257/jel.51.4.1063.
Semiromi, Hamed Naderi, Stefan Lessmann, and Wiebke Peters (2020). “News will tell: Forecasting foreign exchange rates based on news story events in the economy calendar”. In:

The North American Journal of Economics and Finance 52, p. 101181. ISSN: 1062-9408. DOI: https://doi.org/10.1016/j.najef.2020.101181.

Shantha, Gowri B. and Vedantam Seetha Ram (2019). “Influence of news on rational decision making by financial market investors”. English. In: *Investment Management & Financial Innovations* 16.3. Num Pages: 142-156 Place: Sumy, Ukraine Publisher: Business Perspectives Ltd. Section: Articles, pp. 142–156. ISSN: 18104967. DOI: 10.21511/imfi.16(3).2019.14.

Shapiro, Samuel Sanford and Martin B Wilk (1965). “An analysis of variance test for normality (complete samples)”. In: *Biometrika* 52.3-4, pp. 591–611.

Shen, Xinyu et al. (2024). *Harnessing XGBoost for Robust Biomarker Selection of Obsessive-Compulsive Disorder (OCD) from Adolescent Brain Cognitive Development (ABCD) data*. arXiv:2407.00028 [q-bio] version: 1. DOI: 10.48550/arXiv.2407.00028.

Uhl, Matthias W. (2017). “Emotions Matter: Sentiment and Momentum in Foreign Exchange”. In: *Journal of Behavioral Finance* 18.3, pp. 249–257. ISSN: 1542-7560. DOI: 10.1080/15427560.2017.1332061.

Wilcoxon, Frank (1992). “Individual comparisons by ranking methods”. In: *Breakthroughs in statistics: Methodology and distribution*. Springer, pp. 196–202.

Zhu, Chenguang et al. (Apr. 2021). *Leveraging Lead Bias for Zero-shot Abstractive News Summarization*. arXiv:1912.11602 [cs]. DOI: 10.48550/arXiv.1912.11602.

Appendix A

Appendix

A.1 Tables

Variable Name	Description
eur_huf_exchange_rate	Target variable in this analysis. Closing EUR/HUF exchange rate prices.
usd_huf_exchange_rate	Closing US Dollar vs Hungarian Forint exchange rate prices.
huf_1m_option_volatility	Implied volatility from 1-month at-the-money options on the USD/HUF exchange rate. It captures investor expectations of short-term exchange rate fluctuations and is a forward-looking measure of market uncertainty.
huf_1m_realized_volatility	Historical volatility of the HUF exchange rate over a 1-month time frame, calculated from past price movements. A backward-looking measure of market uncertainty.
huf_base_rate	Hungarian central bank's (MNB) base rate, the main tool for managing inflation and short-term interest rates.
overnight_bubor_rate	Budapest Interbank Offered Rate for one-day loans. Reflects short-term liquidity conditions and expectations of monetary policy.
hufonia_rate	Daily average of unsecured overnight interbank lending rates. Focuses on the interbank market for the Hungarian Forint.
huf_3m_benchmark_rate	Yield on Hungarian 3-month treasury bills. A short-term interest rate reflecting monetary expectations.
huf_2y_swap_rate	Fixed rate exchanged for a floating rate in 2-year HUF interest rate swaps. Reflects medium-term expectations.
bux_index	Benchmark stock index of the Budapest Stock Exchange. Captures domestic equity performance and investor sentiment.
ecb_deposit_rate	Interest rate on overnight bank deposits at the European Central Bank. Indicates Eurozone monetary policy stance.

Continued on next page

Variable Name	Description
euro_stoxx50_index	Index of 50 major Eurozone companies. Proxy for economic outlook and investor sentiment in the region.
brent_crude_price	Global oil benchmark. Reflects inflation expectations and production costs relevant for macroeconomic conditions.
gold_price	Global safe-haven asset. Serves as a hedge against financial uncertainty and inflation.
germany_10y_bund_yield	Long-term Eurozone benchmark rate. Reflects investor expectations about inflation and economic growth.
eurhuf_intraday_change	Relative intraday volatility measure for EUR/HUF, computed as high–low range over the opening price.
vix_index	Volatility index measuring implied equity market risk. Serves as a global proxy for investor risk aversion.

Table A.1: Description of economic variables used in the model

Keyword List for Article Selection (terms with * were stemmed)			
ECB	Stock*	Central bank	Yield*
Pension	Economy	Treasury	Currency
Monetary*	Expenditure	Energy price	Production*
Bond	Loan*	Financial crisis	GDP
Output	Money supply	Unemployment*	Base rate
Deficit	Growth	Industry*	Manufacturing*
Invest*	Exchange rate	Interest rate	Labor market
Export	Import	Demand*	Consumption*
Credit*	Inflation*	Tax*	Real estate
Budget	Crisis	Volatility	Savings*
Debt	Employment*		

Table A.2: English keywords used for article filtering

*Note: Terms marked with * were applied in stemmed form (e.g., producti*, invest*) to match multiple lexical variations during article extraction.*

Hungarian Keyword List for Article Selection (terms with * were stemmed)			
termel*	ipar	gyárt*	tőzsde
részvény	befektet*	árfolyam	GDP
export	import	gazdaság	forint
kibocsát*	kereslet	fogyaszt*	kamat
hitel	kölcsön*	állampapír	adósság
hozam	kockázat	valuta	deviza
jegybank	MNB	bank	kincstár
pénzkínál*	alapkamat	monetáris	tartalék
munkaerő	munkanélkül	bér	nyugdíj
adó	költségvet*	kiadás	hiány
válság	ingatlan	energiaár	infláció

Table A.3: Hungarian keywords used for article filtering

*Note: Terms marked with * were applied in stemmed form (e.g., termel*, befektet*) to match multiple lexical variations during article extraction.*

A.2 Figures

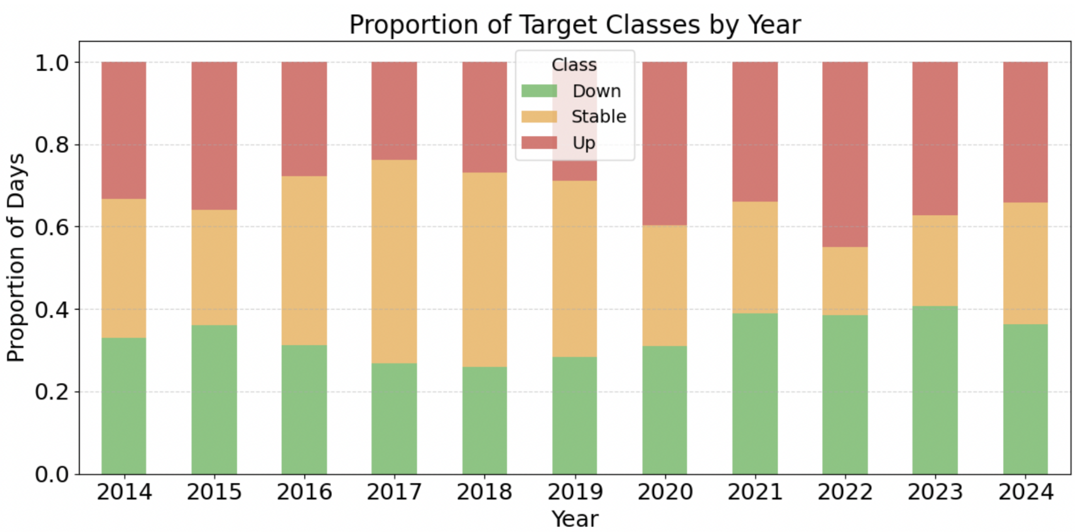

Figure A.1: Yearly distribution of target classes (Up, Stable, Down) based on daily EUR/HUF percentage changes

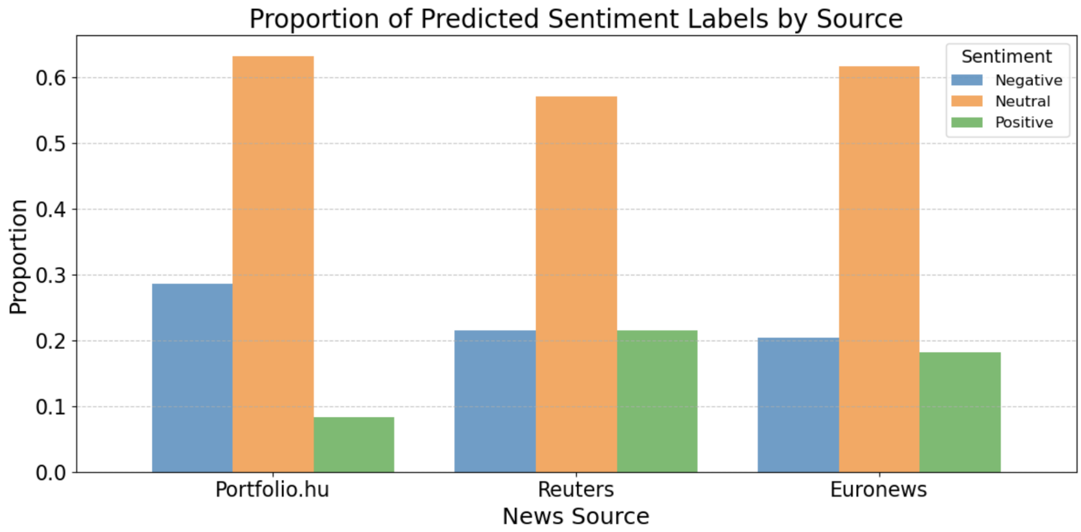

Figure A.2: Proportions of predicted sentiment labels across sources, with "Neutral" being the dominant class in all three cases

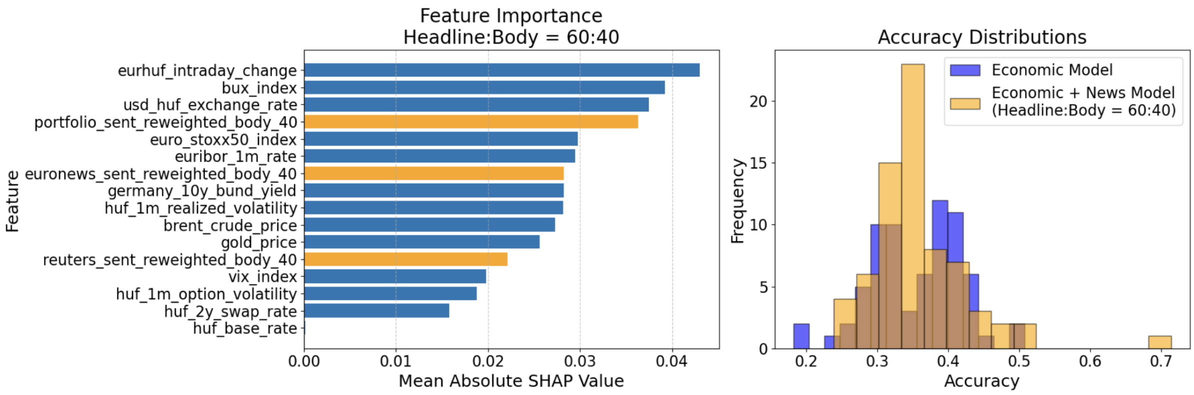

Figure A.3: Robustness check #1: Feature importance and accuracy distributions for the model with reweighted sentiment scores using a 40:60 headline-to-body ratio

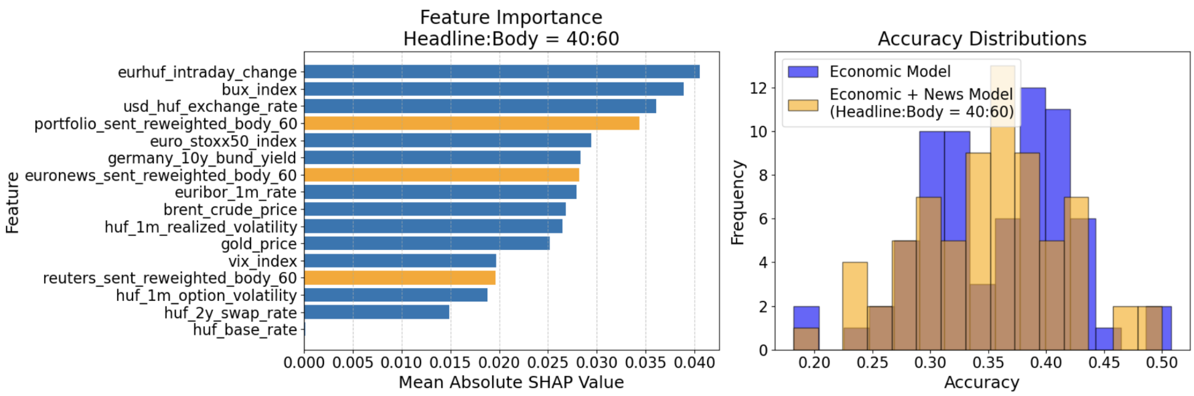

Figure A.4: Robustness check #2: Feature importance and accuracy distributions for the model with reweighted sentiment scores using a 60:40 headline-to-body ratio

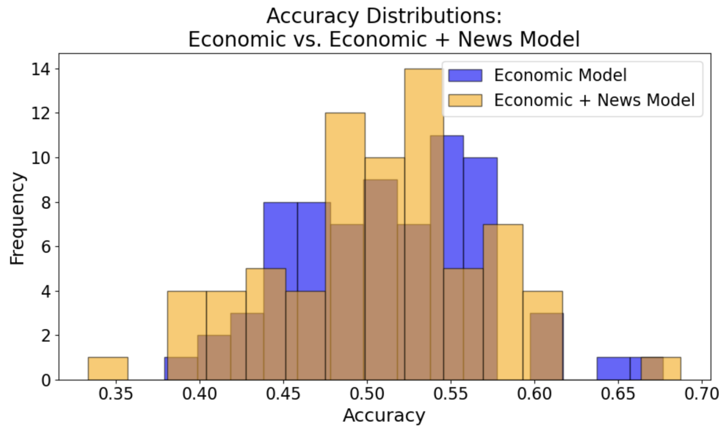

Figure A.5: Robustness check #3: Accuracy distributions for economic and news-enhanced models under binary classification setting

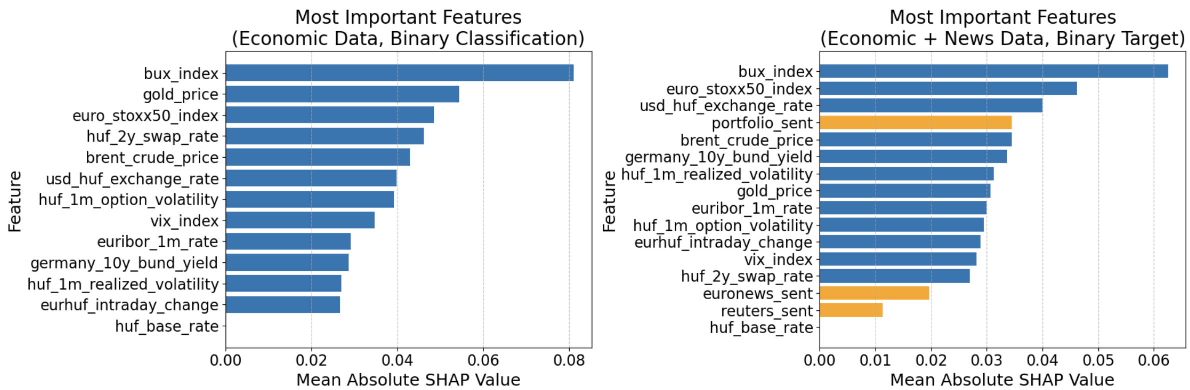

Figure A.6: Robustness check #3: Mean absolute SHAP values for binary classification models, comparing economic-only and news-enhanced models with sentiment features highlighted