Prediction of Market Value Based on Changes in Capital and Asset Structure Using Long Short-Term Memory Network

Kacper Sobczyk
ksobczyk@student.agh.edu.pl

May 2024

Abstract

The study examines the effectiveness of Long Short-Term Memory (LSTM) networks in identifying patterns within financial data to improve market value predictions. The model is trained on historical data from the Polish stock exchange (GPW) covering the period from 1998 to 2022 on a quarterly basis and evaluated on its predictive accuracy using Mean Absolute Error (MAE), Root Mean Squared Error (RMSE), and R-squared (R^2). Experiments summary, limitations, and suggestions for future research are also discussed.

1 Introduction

1.1 Motivation

The valuation of companies and their stocks is a critical aspect of financial markets, influencing investment decisions, corporate strategies, and economic policies. Accurate prediction of market value is essential for investors seeking to maximize returns and minimize risks. Traditional methods for predicting market value often rely on linear models and fundamental analysis, which may not fully capture the complexities of financial data. With the advent of advanced computational techniques and the availability of large datasets, machine learning approaches, particularly Long Short-Term Memory (LSTM) networks, have shown promise in enhancing predictive accuracy by identifying intricate patterns within the data.

Long Short-Term Memory networks, a type of recurrent neural network (RNN), have been successfully applied to various domains, including natural language processing, speech recognition, and now, financial forecasting. The strength of LSTMs lies in their ability to capture temporal dependencies in sequential data, making them well-suited for analyzing time series and structured financial information. This paper explores the application of LSTMs to predict market value based on changes in capital and asset structure, leveraging historical data from the Polish stock exchange. By transforming financial metrics into formats that can be processed by LSTMs, we aim to enhance the accuracy and reliability of market value predictions.

The motivation for this research stems from the limitations of traditional financial analysis methods and the potential of machine learning techniques to address these challenges. Traditional models often assume linear relationships and may overlook complex interactions within financial data. Furthermore, these models might not fully exploit the wealth of historical data available, leading to suboptimal predictive performance.

Machine learning, and LSTMs in particular, offer a new paradigm for financial forecasting. LSTMs can learn from large datasets and identify non-linear relationships, making them ideal for capturing the dynamic and multifaceted nature of financial markets. The goal of this study is to develop a robust LSTM model that can predict market value with higher accuracy than traditional methods, thereby providing investors and analysts with a powerful tool for decision-making.

This research also aims to contribute to the growing body of literature on the application of deep learning in finance. By focusing on the Polish stock exchange, this study provides insights into an emerging market that has received less attention in the context of machine learning applications compared to more developed markets. The findings of this research could have implications for both academic research and practical applications, highlighting the potential of LSTMs to revolutionize financial forecasting and investment strategies.

1.2 Theoretical Background

The capital structure of a company has been a central topic in modern finance theory for over four decades, reflecting its crucial role in corporate financial management and market dynamics [1]. The complexity and versatility of capital make the analysis of a company's capital structure essential for understanding its financial health and strategic direction. Capital, at its core, is an economic value with the potential for growth, leading to the creation of 'added value' such as profit or interest. This abstract notion of capital transforms into tangible financial resources or goods used in the production process, aiming to stimulate further development [2].

The conscious shaping and assessment of a company's capital structure involve analyzing the relationships and proportions between different types of financial resources, including equity and debt. The capital structure primarily refers to the configuration of a company's sources of financing, where equity represents ownership interest and debt constitutes borrowed funds subject to interest payments. Effective management of capital structure is vital as it influences a company's ability to generate added value, maintain financial stability, and pursue long-term growth [2].

Equity capital is pivotal in the functioning of a company, providing the necessary financial foundation for initiating business operations and safeguarding against potential losses. It also serves as a source of funding for ongoing business activities. Although equity financing is considered the safest option due to its non-repayment obligation, it often involves higher costs compared to debt financing. This higher cost is attributed to the opportunity cost of equity holders' funds and the potential dilution of ownership [9].

Conversely, debt capital arises from obligations to creditors and is characterized by lower acquisition costs. Debt financing can be beneficial due to tax advantages, as interest payments are tax-deductible, reducing the overall tax burden on the company. However, increased reliance on debt elevates financial risk and may adversely impact the company's creditworthiness. The challenge for financial managers lies in balancing the benefits of debt, such as cost advantages and financial leverage, against the associated risks [8].

The determinants of a company's capital structure are multifaceted, encompassing both internal factors related to the firm's operations and external factors from the macroeconomic environment. Internal factors, or microeconomic factors, include company size, industry characteristics, cost of capital, financial liquidity, risk tolerance, market position, asset structure, profitability, dividend policy, and investment strategy. These factors directly influence managerial decisions regarding the optimal mix of equity and debt [8].

External factors, or macroeconomic factors, encompass the broader economic context in which the company operates. These include the level of economic development, inflation rates, interest rates, fiscal policies, legal regulations, efficiency of the banking sector, technological advancements, and the development of financial markets. These macroeconomic determinants shape the external financing environment, impacting a company's ability to access capital and influencing its capital structure decisions [8].

Both microeconomic and macroeconomic factors significantly affect the choice of capital structure and, consequently, the market value of the enterprise. By thoroughly analyzing and adapting to these determinants, companies can optimize their financial strategies and enhance their market value.

1.3 Related Work

The market value of an enterprise reflects the price at which it can be bought or sold under open and competitive market conditions, assuming rational behavior and adequate information on both sides of the transaction [11]. This definition underscores the importance of transparency and information availability in determining fair market value. Various financial indicators are crucial for assessing an enterprise's value. Traditional indicators, such as return on equity, sales growth, fixed asset investments, and changes in working capital, provide insights into a company's financial performance. Additionally, competitive advantages and strategic positioning play a significant role in valuation [7].

Maximizing enterprise value is a fundamental objective in contemporary economic approaches, particularly for public companies listed on stock exchanges. For these companies, market value is often equated with market capitalization, which is the product of the total number of shares and their current market price. This metric provides a clear and quantifiable measure of a company's value in the financial market, essential for investors and shareholders evaluating the company's investment potential [10].

Understanding market value involves comparing a company's share prices to benchmark indices representing similar enterprises within the same industry. This comparative analysis helps eliminate the influence of general market trends, allowing a focus on the individual performance and value of the enterprise [4].

Over the past few decades, the financial sector has experienced significant transformation due to advancements in information technology. The implementation of machine learning (ML) and deep learning (DL) algorithms has revolutionized data analysis and investment strategy optimization. These technologies enable the processing and interpretation of vast amounts of financial data, offering greater accuracy and speed in predicting market trends and making investment decisions [5].

Machine learning algorithms, including support vector machines (SVM), decision trees, and random forests, have found widespread applications in finance. These methods are used for various tasks, such as regression analysis, real estate price forecasting, credit card fraud detection, and stock price prediction. ML algorithms excel in identifying patterns within data and improving predictive accuracy by minimizing errors [3].

Deep learning, a subset of machine learning, has garnered special interest for its ability to handle complex financial problems. Deep learning models are inspired by the human brain's neural networks, where information is transmitted between neurons through synapses. Similarly, in artificial neural networks, data moves between nodes (analogous to neurons) via connections whose strengths (weights) are adjusted during the learning process [6].

Neural networks' ability to learn and adapt by adjusting their weights based on training data allows them to recognize patterns, predict outcomes, and make decisions with remarkable accuracy. This capability makes neural networks particularly suitable for financial forecasting, where they can uncover trends and insights not readily apparent through traditional methods. Neural networks' applications in finance include image recognition for fraud detection, natural language processing for sentiment analysis, and complex time-series forecasting for stock prices and market trends [6].

In practice, the iterative adjustment of weights in neural networks enables these models to evolve and improve their predictive capabilities over time. By learning from historical data, neural networks can adapt to changing market conditions and provide valuable insights for financial decision-making. This adaptability and learning ability distinguish neural networks from traditional statistical methods, offering a powerful tool for financial analysis and forecasting [6].

2 Problem Formulation and Proposed Solution

2.1 Proposed Approach

This approach involves several key steps, including data collection and preprocessing, feature extraction, model architecture design, training, and evaluation. The aim is to leverage LSTMs to capture complex patterns in financial data that traditional models may miss.

2.1.1 Data Collection and Preprocessing

The data for this study was obtained from two primary sources:

- **Company Financial Records**: The dataset includes Excel files from 350 companies listed on the Polish stock exchange (GPW), covering the period from 1998 to 2022 on a quarterly basis. These files contain comprehensive information about the companies' financial statuses at the end of each period. For this study, all data directly related to capital and asset structure, such as equity, debt, short-term and long-term liabilities, industry sector, and number of shares during the period, were extracted. Additionally, indirectly related economic indicators that involve capital or assets were considered.
- **Historical Market Data from Stooq.pl**: This source provides historical market data. By matching the dates and companies, the average market value of each company's stock on the balance sheet date was aligned.

The data preprocessing steps include:

- **Data Cleaning**: Handling missing values, outliers, and inconsistencies in the dataset to ensure the quality of data.
- **Feature Engineering**: Creating new features that may better capture the underlying financial dynamics. This includes adding columns that show the difference between quarters, two quarters, and four quarters to better represent the essence of the study. Additionally, ratios like debt-to-equity and return on assets were created. One-hot encoding was also employed for categorical variables to convert them into a format suitable for machine learning algorithms.
- **Normalization**: Using techniques such as `StandardScaler` from sklearn to standardize the features, ensuring that they have a mean of zero and a standard deviation of one.
- **Dataset Preparation**: Two datasets were prepared for the experiments. The first dataset includes only the specific values directly related to the study. The second dataset encompasses all related factors, including indirectly related economic indicators, where capital and asset structure play a role. Additionally, two corresponding Data Frames were created to facilitate the experiments.

2.1.2 Model Architecture Design

The architecture of the LSTM model used in this study is detailed in the following code snippet:

```python
class LSTMModel(nn.Module):
    def __init__(self, input_size, hidden_size,
       num_layers):
        super(LSTMModel, self).__init__()
        self.hidden_size = hidden_size
        self.num_layers = num_layers
        self.lstm = nn.LSTM(input_size, hidden_size,
           num_layers, batch_first=True)
        self.fc = nn.Linear(hidden_size, 1)

    def forward(self, x):
        h0 = torch.zeros(self.num_layers, x.size(0),
           self.hidden_size).to(x.device)
        c0 = torch.zeros(self.num_layers, x.size(0),
           self.hidden_size).to(x.device)
        out, _ = self.lstm(x, (h0, c0))
        out = self.fc(out[:, -1, :])
        return out
```

Listing 1: LSTM Model Definition

The `LSTMModel` class defines the architecture of our LSTM network. The model consists of an LSTM layer followed by a fully connected (linear) layer. The key components of the model are:

- **Input Size**: This parameter determines the number of input features for each time step in the input sequence.
- **Hidden Size**: This parameter defines the number of features in the hidden state of the LSTM.
- **Number of Layers**: This parameter sets the number of stacked LSTM layers in the network.
- **LSTM Layer**: The core of the model, which processes the input sequence and learns the temporal dependencies.
- **Fully Connected Layer**: This layer maps the output of the LSTM layer to the final prediction.

In the `forward` method, the hidden state (`h0`) and cell state (`c0`) are initialized to zeros and are passed along with the input sequence (`x`) to the LSTM layer. The LSTM processes the input sequence and outputs the final hidden state, which is then passed through the fully connected layer to produce the final prediction.

LSTM networks are particularly suitable for financial data due to their ability to handle the sequential nature of time series data. Financial markets are influenced by a variety of factors over time, and capturing these temporal dependencies is crucial for accurate forecasting. LSTMs are designed to remember information for long periods and are robust against issues like vanishing gradients, which often affect traditional RNNs. This makes LSTM networks a powerful tool for modeling complex temporal patterns in financial data.

2.1.3 Training and Evaluation

To train the LSTM model, the dataset is split into training and testing sets using the `train_test_split` function from `scikit-learn`. This ensures that the model is evaluated on a separate subset of the data that it has not seen during training, providing a more accurate assessment of its performance. The training process involves the following key steps:

- **Optimization**: Optimizer Adam is utilized to minimize the loss function. This optimizer adjusts the weights of the neural network to reduce the difference between the predicted and actual values.
- **Loss Function**: A suitable loss function like Mean Squared Error (MSE) is implemented to quantify the prediction errors. The choice of loss function is crucial as it guides the optimizer in updating the model parameters.
- **Evaluation Metrics**: The model performance is assessed using metrics such as Mean Absolute Error (MAE), Root Mean Squared Error (RMSE), and R-squared (`r2_score`). These metrics provide different perspectives on the accuracy and reliability of the predictions.

Mean Absolute Error (MAE): The Mean Absolute Error (MAE) measures the average absolute difference between the predicted and actual values. It provides a straightforward interpretation of the average error magnitude, making it an intuitive metric for assessing model performance. The formula for MAE is:

$$\text{MAE} = \frac{1}{n}\sum_{i=1}^{n}|y_i - \hat{y}_i|$$

where y_i represents the actual value, $\hat{y}_i$ represents the predicted value, and n is the number of observations.

Root Mean Squared Error (RMSE): The Root Mean Squared Error (RMSE) is the square root of the Mean Squared Error (MSE) and provides an estimate of the standard deviation of the prediction errors. RMSE is more sensitive to large errors than MAE, making it a useful metric when large deviations are particularly undesirable. The formula for RMSE is:

$$\text{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}$$

where y_i represents the actual value, $\hat{y}_i$ represents the predicted value, and n is the number of observations.

R-squared ($\mathbf{R^2}$): The R-squared (R^2) metric represents the proportion of the variance in the dependent variable that is predictable from the independent variables. It provides a measure of how well the model captures the variability of the target variable. An R^2 value closer to 1 indicates a better fit. The formula for R^2 is:

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{n}(y_i - \bar{y})^2}$$

where y_i represents the actual value, $\hat{y}_i$ represents the predicted value, and $\bar{y}$ is the mean of the actual values.

To prevent overfitting, an early stopping mechanism is implemented. This technique halts the training process when the validation loss stops improving for a specified number of epochs (patience). By monitoring the validation loss and stopping early when there is no significant improvement, the model's generalization capability is maintained. The early stopping mechanism is integrated into the training loop as follows:

```
if test_loss < best_test_loss:
    best_test_loss = test_loss
    patience_counter = 0
else:
    patience_counter += 1
    if patience_counter >= patience:
        print("Early stopping triggered")
        break
```

By following this training and evaluation approach, we aim to develop a robust LSTM model capable of accurately predicting market value based on changes in capital and asset structure.

2.2 Libraries

The implementation of the Long Short-Term Memory (LSTM) network for market value prediction leverages several Python libraries that are essential for data manipulation, model building, and evaluation. Below is a detailed description of the libraries and their roles in this study:

- **PyTorch**: PyTorch is an open-source machine learning library widely used for deep learning applications. It provides a flexible platform for building and training neural networks. This study utilizes several submodules of PyTorch:
  - `torch`: The core library providing data structures for multi-dimensional tensors and a wide range of mathematical operations on them.

  - `torch.nn`: This submodule provides modules and classes for building neural network layers, including the Long Short-Term Memory (LSTM) layers and fully connected layers used in the model.
  - `torch.optim`: This submodule includes optimization algorithms such as Adam, which is used for updating the network's weights based on the computed gradients.
  - `torch.utils.data`: This submodule provides utilities for handling data, including the `DataLoader` and `TensorDataset` classes, which facilitate efficient data batching and shuffling during training.

- **pandas**: Pandas is a powerful data manipulation and analysis library that provides data structures like DataFrames, which are essential for handling and preprocessing structured data. In this study, pandas is used for loading and cleaning the financial datasets, as well as for organizing the processed data into a suitable format for model training.
- **NumPy**: NumPy is the fundamental package for scientific computing in Python, providing support for arrays and matrices, along with a wide range of mathematical functions to operate on these data structures. NumPy is used in this study for numerical operations and for converting data into formats compatible with PyTorch.
- **scikit-learn**: Scikit-learn is a machine learning library that offers simple and efficient tools for data mining and data analysis. It is used for:
  - `train_test_split`: A function to split the dataset into training and testing sets.
  - `mean_squared_error`, `mean_absolute_error`, `r2_score`: Functions for evaluating the performance of the model using various metrics.
- **Matplotlib**: Matplotlib is a plotting library used for creating static, animated, and interactive visualizations in Python. In this study, it is used to visualize the training and testing loss over epochs.

These libraries and their respective submodules provide a comprehensive toolkit for implementing, training, evaluating, and visualizing the LSTM model, facilitating a thorough analysis of its performance in predicting market values.

2.3 Algorithms

The following subsections detail the primary functions used in the provided Long Short-Term Memory (LSTM) model implementation, including descriptions of their purposes and instructions for replicating the experiment.

```python
def set_seed(seed):
    np.random.seed(seed)
    torch.manual_seed(seed)
    if torch.cuda.is_available():
        torch.cuda.manual_seed(seed)
        torch.cuda.manual_seed_all(seed)
    torch.backends.cudnn.deterministic = True
    torch.backends.cudnn.benchmark = False
```

Listing 2: Setting the seed for reproducibility

The `set_seed` function ensures that the experiment is reproducible by setting a fixed seed for random number generators in NumPy and PyTorch. This function also configures CUDA to use deterministic algorithms to further guarantee reproducibility when using GPUs.

2.4 Preprocessing the Data

```
def preprocess_data(df):
    X = df.drop('target', axis=1)
    y = df['target']
    return X, y
```

Listing 3: Preprocessing the data

The `preprocess_data` function separates the input features (`X`) from the target variable (`y`). This function assumes that the input DataFrame (`df`) contains a column named 'target' which represents the target variable for prediction.

```
def reshape_data(X, y):
    X_train, X_test, y_train, y_test =
        train_test_split(X, y, test_size=0.2,
        random_state=42)
    X_train = torch.tensor(X_train.values, dtype=torch
        .float32).unsqueeze(1)
    X_test = torch.tensor(X_test.values, dtype=torch.
        float32).unsqueeze(1)
    y_train = torch.tensor(y_train.values, dtype=torch
        .float32).unsqueeze(1)
    y_test = torch.tensor(y_test.values, dtype=torch.
        float32).unsqueeze(1)
    return X_train, X_test, y_train, y_test
```

Listing 4: Reshaping the data

The `reshape_data` function splits the data into training and testing sets and reshapes them for input into the LSTM model. The function uses an 80-20 split ratio and ensures the data is in the appropriate tensor format for PyTorch.

```
def train_model(model, train_loader, test_loader,
   criterion, optimizer, epochs=100, patience=10):
    train_losses = []
    test_losses = []
    best_test_loss = float('inf')
    patience_counter = 0

    for epoch in range(epochs):
        model.train()
        running_loss = 0.0
        for inputs, targets in train_loader:
            optimizer.zero_grad()
            outputs = model(inputs)
            loss = criterion(outputs, targets)
            loss.backward()
            optimizer.step()
            running_loss += loss.item() * inputs.size
                (0)
        epoch_loss = running_loss / len(train_loader.
            dataset)
        train_losses.append(epoch_loss)

        model.eval()
        test_loss = 0.0
        with torch.no_grad():
            for inputs, targets in test_loader:
                outputs = model(inputs)
                loss = criterion(outputs, targets)
                test_loss += loss.item() * inputs.size
                    (0)
        test_loss = test_loss / len(test_loader.
            dataset)
        test_losses.append(test_loss)

        print(f'Epoch {epoch + 1}/{epochs}, Train Loss
            : {epoch_loss:.4f}, Test Loss: {test_loss
            :.4f}')

        if test_loss < best_test_loss:
            best_test_loss = test_loss
            patience_counter = 0
        else:
            patience_counter += 1
            if patience_counter >= patience:
                print("Early stopping triggered")
                break

    loss_df = pd.DataFrame({'Epoch': list(range(1, len
        (train_losses) + 1)), 'Train Loss':
        train_losses, 'Test Loss': test_losses})
    loss_df.to_csv("lstm_training_results.csv", index=
       False)
```
Listing 5: Training the model

The `train_model` function trains the LSTM model using the specified loss criterion and optimizer. It tracks training and testing losses over epochs, implements early stopping based on a patience parameter, and saves the loss history to a CSV file.

```
def evaluate_model(model, test_loader):
    model.eval()
    predictions, actuals = [], []
    with torch.no_grad():
        for inputs, targets in test_loader:
            outputs = model(inputs)
            predictions.extend(outputs.numpy())
            actuals.extend(targets.numpy())
    return np.array(predictions), np.array(actuals)
```
Listing 6: Evaluating the model

The `evaluate_model` function assesses the model's performance on the test set. It collects the model's predictions and the actual targets for evaluation.

```
def main():
    set_seed(42)
    mini_df = pd.read_csv('mini_df_scaled.csv')
    X, y = preprocess_data(mini_df)
    X_train, X_test, y_train, y_test = reshape_data(X,
        y)

    train_dataset = TensorDataset(X_train, y_train)
    test_dataset = TensorDataset(X_test, y_test)

    train_loader = DataLoader(train_dataset,
       batch_size=32, shuffle=True)
    test_loader = DataLoader(test_dataset, batch_size
       =32, shuffle=False)

    input_size = X_train.shape[2]
    hidden_size = 50
    num_layers = 2

    model = LSTMModel(input_size, hidden_size,
       num_layers)
    criterion = nn.MSELoss()
    optimizer = optim.Adam(model.parameters(), lr
       =0.001)
```

```python

    train_model(model, train_loader, test_loader, criterion, optimizer, epochs=50, patience=10)

    y_pred, y_true = evaluate_model(model, test_loader)

    rmse = np.sqrt(mean_squared_error(y_true, y_pred))
    mae = mean_absolute_error(y_true, y_pred)
    r2 = r2_score(y_true, y_pred)

    print(f'RMSE: {rmse}')
    print(f'MAE: {mae}')
    print(f'R2: {r2}')

if __name__ == "__main__":
    main()
```

Listing 7: Main function

The `main` function orchestrates the entire process: setting the seed, loading and preprocessing the data, defining data loaders, initializing the model, training it, and finally evaluating its performance.

3 Experimental Results

3.1 Results and Analysis

Results of experiments on two datasets: `nn_df` and `mini_df`.

The `nn_df` dataset includes all columns directly and indirectly related to capital and asset structure, in addition to feature engineering techniques applied to enhance the dataset. This comprehensive dataset aims to capture a broader range of financial indicators and their relationships.

On the other hand, the `mini_df` dataset focuses exclusively on the direct state of capital and asset structure, with feature engineering applied as well. This dataset aims to streamline the input features to those most directly relevant to the model's objective.

The performance of the LSTM model is evaluated using three metrics: Root Mean Squared Error (RMSE), Mean Absolute Error (MAE), and R-squared (R^2). The results are summarized in Table 1.

Dataset	**RMSE**	**MAE**	R^2
`nn_df`	0.2934	0.0391	0.9166
`mini_df`	0.2669	0.0428	0.9220

Table 1: Performance of the LSTM model on the `nn_df` and `mini_df` datasets.

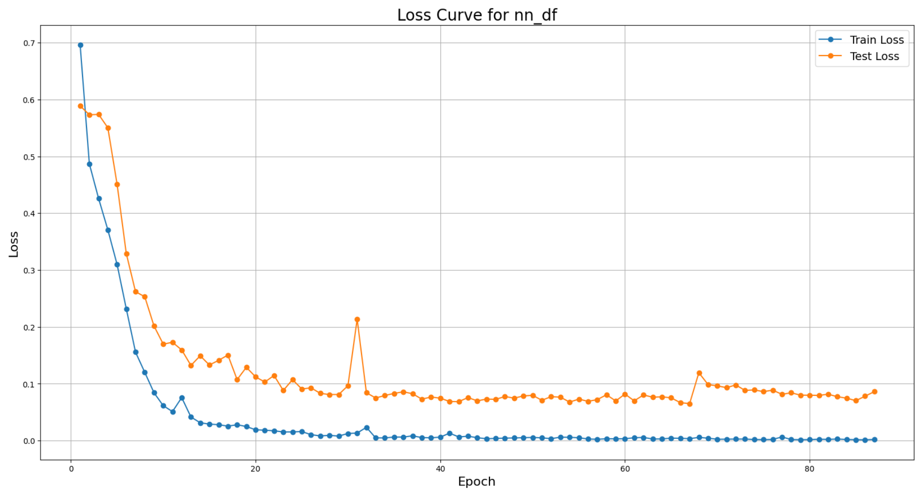

Figure 1: Loss Curve for LSTM Model on `nn_df`

nn_df Dataset:

- The training loss decreases steadily over the epochs, indicating that the model is learning and fitting the training data well.

- The test loss also decreases but shows more fluctuations compared to the training loss. This indicates that while the model is improving on the training data, it experiences some variability when evaluated on the test data.

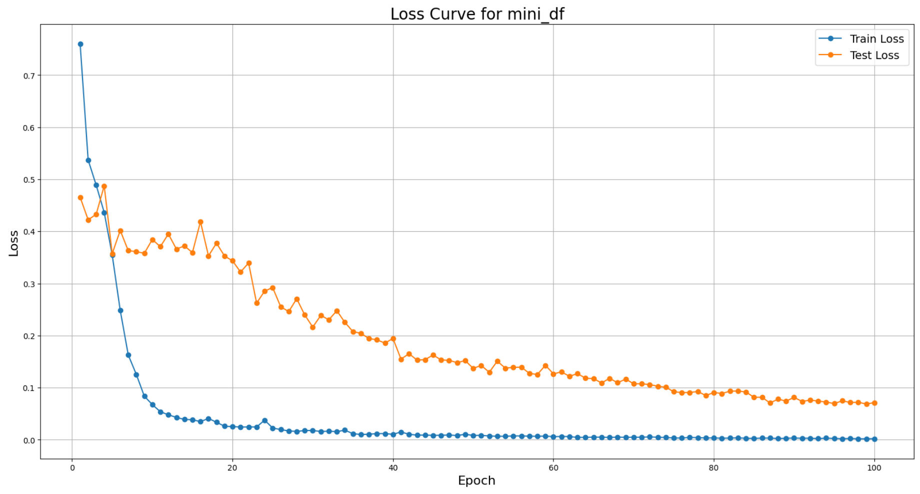

Figure 2: Loss Curve for LSTM Model on `mini_df`

mini_df Dataset:

- Similar to the `nn_df` dataset, the training loss decreases consistently, showing effective learning by the model.
- The test loss decreases and shows fewer fluctuations compared to the `nn_df` dataset, suggesting a more stable performance on the test data.
- The test loss stabilizes around the 30th epoch, with minor fluctuations, indicating that the model generalizes well without significant overfitting.

From the results, the following observations can be made:

- The model achieved a lower RMSE on the `mini_df` dataset compared to the `nn_df` dataset. Specifically, the RMSE decreased from 0.2934 to 0.2669, indicating that the model's predictions are closer to the actual values on the `mini_df` dataset.
- The MAE is slightly higher on the `mini_df` dataset (0.0428) compared to the `nn_df` dataset (0.0391). Although the difference is small, it suggests that the average absolute error of the model's predictions is slightly greater on the `mini_df` dataset.

- The R^2 value, which indicates the proportion of variance in the dependent variable that is predictable from the independent variables, is higher for the `mini_df` dataset (0.9220) than for the `nn_df` dataset (0.9166). This suggests that the model explains a greater portion of the variance in the target variable for the `mini_df` dataset.

Overall, the `mini_df` dataset yielded better performance metrics, indicating that the LSTM model may generalize better on this dataset. These results suggest that the model's ability to predict market values is more accurate and reliable when using the `mini_df` dataset.

3.2 Summary

The use of a Long Short-Term Memory (LSTM) network for predicting market value based on changes in capital and asset structure has proven to be effective. The experimental results demonstrate that the LSTM model achieved high performance, particularly on the `mini_df` dataset, where it reached a Mean Absolute Error (MAE) of 0.0428. This indicates that, on average, the model's predictions deviate by 0.0428 units from the actual values, showcasing its precision in capturing the true market values.

The results from the `nn_df` dataset also show a commendable performance, with an MAE of 0.0391. This slight difference in MAE between the two datasets suggests that the model performs slightly better with the `nn_df` dataset, which includes a broader range of financial indicators and relationships. However, the `mini_df` dataset, which focuses exclusively on the direct state of capital and asset structure, still provides a highly accurate prediction.

In addition to MAE, the R^2 value provides further insights into the model's performance. The R^2 value of 0.9220 on the `mini_df` dataset indicates that the model explains 92.20% of the variance in the target variable, highlighting its reliability in predicting market values. Similarly, an R^2 value of 0.9166 on the `nn_df` dataset demonstrates the model's strong capability to capture the variance in a broader set of financial indicators.

Despite the high performance, it is important to note that such a model might not always be sufficient in every context. Financial markets are influenced by a myriad of factors, some of which may not be fully captured by the datasets used. The complexity and dynamic nature of financial data mean that the model's performance can vary depending on the quality and relevance of the input data.

Moreover, while the MAE and R^2 values are impressive, there are scenarios where even higher accuracy may be required, particularly in high-stakes financial decision-making environments. Hence, further improvements and refinements to the model, as well as the inclusion of additional relevant features, could be necessary to achieve even better results.

In summary, the LSTM model demonstrated a robust ability to predict market values with a high degree of accuracy, particularly reflected in the MAE and

R^2 metrics. These results are promising and indicate that with continued development and adaptation, such models can become valuable tools in financial analysis and forecasting.

3.3 Limitations and Future Work

The performance of the LSTM model is highly dependent on the quality and completeness of the financial data. Any inaccuracies or missing data can significantly affect the model's predictions. Additionally, the dataset used in this study is limited to companies listed on the Polish stock exchange (GPW). Extending the dataset to include companies from different regions and sectors could provide more generalized insights. The inclusion of more diverse data could help in understanding and predicting market values in a broader context.

While both datasets included several relevant features, there may be other important variables not considered in this study. Incorporating additional features, such as macroeconomic indicators, market sentiment, and geopolitical factors, could potentially improve the model's predictive accuracy. In particular, market sentiment is a major factor influencing market value. Investor sentiment, driven by news, social media, and overall market trends, can have significant impacts on stock prices and market movements. This study did not implement sentiment analysis, which is a critical component in understanding market behavior and making informed investment decisions. Integrating sentiment analysis into the model could enhance its ability to capture short-term market fluctuations and improve prediction accuracy.

To better capture the temporal dependencies in financial data, future work could explore the integration of different neural network architectures. For example, combining LSTM networks with Convolutional Neural Networks (CNNs) could leverage the strengths of both architectures, improving the model's ability to predict market values over time by capturing both spatial and temporal patterns. Additionally, exploring other architectures such as Gated Recurrent Units (GRUs) or Transformer models, which have shown success in sequential data tasks, could provide further improvements in predictive performance by capturing more complex dependencies and interactions in the data.

Future research could focus on identifying and incorporating additional features that capture the underlying economic and market conditions. Feature selection techniques, such as Principal Component Analysis (PCA) or Recursive Feature Elimination (RFE), could be employed to identify the most relevant features. These techniques can help in reducing the dimensionality of the data, improving model efficiency, and potentially increasing predictive accuracy by focusing on the most impactful variables.

Implementing advanced regularization techniques, such as dropout or L2 regularization, can help prevent overfitting and improve the model's generalization capabilities. Dropout implementation can reduce the risk of overfitting by randomly omitting neurons during the training process, thus making the model more robust. Additionally, exploring model ensemble methods, where multiple models are combined to produce a single prediction, can further enhance performance and reliability. Model ensembles can capture a wider range of patterns and reduce the variance in predictions, leading to more stable and accurate outcomes.

References

[1] Saugata Banerjee, Almas Heshmati, Clas Wihlborg, et al. *The dynamics of capital structure.* New York University-Salomon Center-Leonard N. Stern School of Business, 1999.

[2] Sylwia Bętkowska. Determinanty struktury kapitału w przedsiębiorstwie. *Finanse, Rynki Finansowe, Ubezpieczenia*, 82(1):385–396, 2016.

[3] Sheng Chen and Hongxiang He. Stock prediction using convolutional neural network. In *IOP Conference series: materials science and engineering*, volume 435, page 012026. IOP Publishing, 2018.

[4] Aswath Damodaran. *Finanse korporacyjne: teoria i praktyka.* Wydawnictwo Helion, 2007.

[5] Jesus Cuauhtemoc Tellez Gaytan, Karamath Ateeq, Aqila Rafiuddin, Haitham M Alzoubi, Taher M Ghazal, Tariq Ahamed Ahanger, Sunita Chaudhary, and GK Viju. Ai-based prediction of capital structure: Performance comparison of ann, svm, and lr models. *Computational Intelligence and Neuroscience*, 2022, 2022.

[6] Kevin Gurney. *An introduction to neural networks.* CRC press, 2018.

[7] Adriana Kaszuba-Perz and Paweł Perz. Rola przedsiębiorcy w budowaniu wartości przedsiębiorstwa rodzinnego. *Studia Ekonomiczne*, 388:134–145, 2019.

[8] Agnieszka Kurczewska. Determinanty struktury kapitałowej przedsiębiorstwa. *Prace Naukowe Uniwersytetu Ekonomicznego we Wrocławiu*, 111:326–337, 2010.

[9] Anna Kwiecień. Struktura kapitału a stopa zwrotu–analiza przypadku. *Studia Ekonomiczne*, 222:139–152, 2015.

[10] Alfred Rappaport. *Wartość dla akcjonariuszy: poradnik menedżera i inwestora.* WIG-Press, 1999.

[11] Sara Rupacz and Izabela Jonek-Kowalska. Model szacowania wartości przedsiębiorstwa na podstawie danych publikowanych przez giełdę papierów wartościowych w warszawie. *Management & Quality/Zarządzanie i Jakość*, 4(3), 2022.