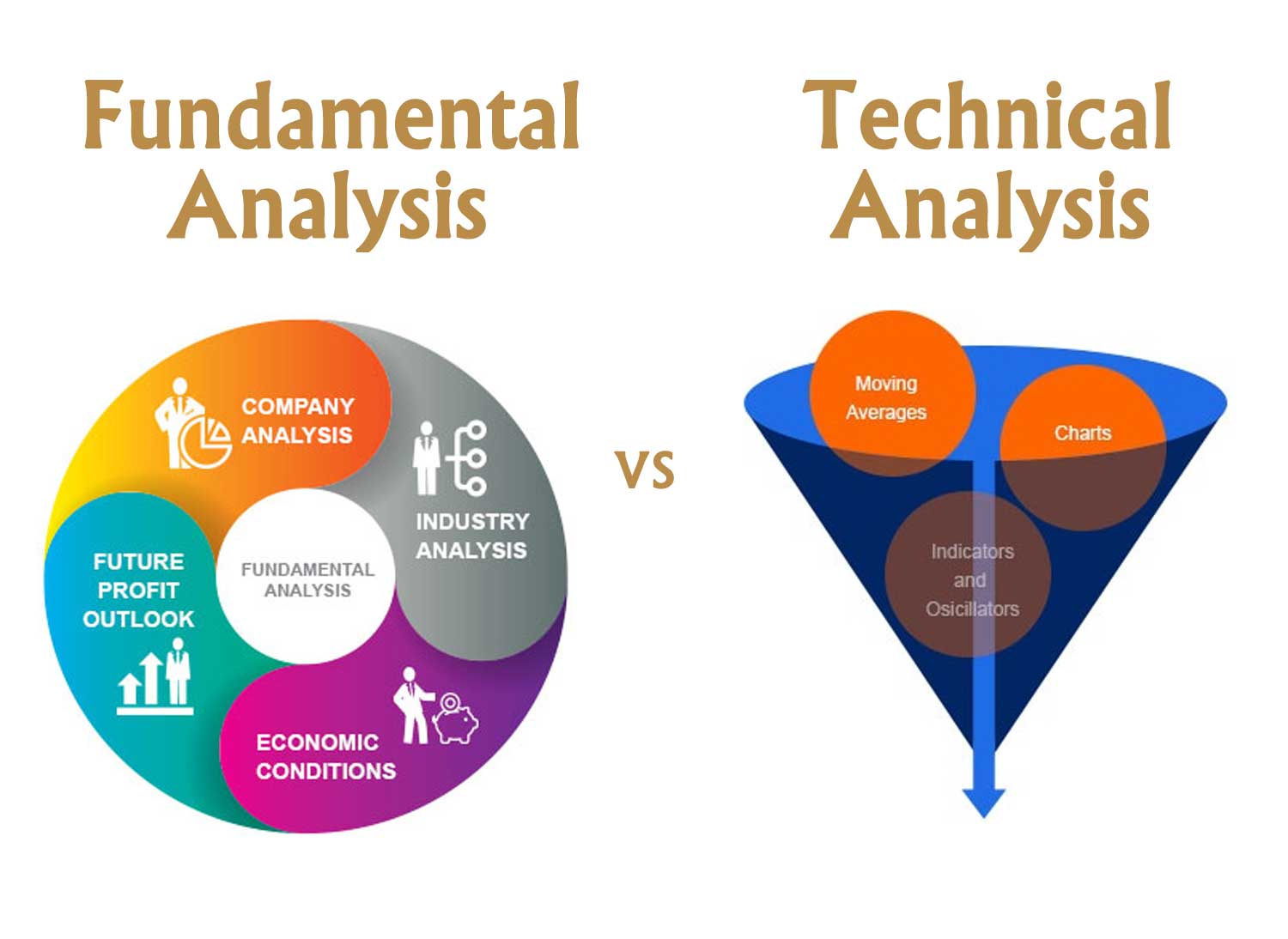
**Abstract:**

Making predictions about the stock market is challenging, since there are many factors that influence how the market behaves. For my project I’m predicting the Istanbul Stock Exchange (ISE) using the data from the UCI machine learning repository. There are several features in the dataset such as a, b, c.., and the primary attribute is the ISE feature, which is used to predict stock values using the machine learning models. The goal of this project is to compare how well two model forecast the ISE index: Long Short-Term Memory (LSTM) and Autoregressive Integrated Moving Average (ARIMA). Finding the model with the highest accuracy while utilizing the least amount of data input is one of the goals. By creating the model architecture, assess the model performance and figure out the best model for short-term stock prediction by studying the relevant literature on predicting stock using LSTM and ARIMA. The understanding of the stock exchange trends and the effectiveness of models in financial forecasting will be enhanced by the result of this study.

**Introduction:**

**Background**

Forecasting the stock market’s future using the past data and other influencing factors is referred as stock market prediction. The two main categories of analyzing stock market are fundamental analysis and technical analysis (figure1). By evaluating the company’s present financial performance and business environment, fundamental analysis projects the organization’s potential profitability. On the other hand, technical analysis involves analyzing charts i.e. is stock’s historical prices and utilizing statistical data to find market trends. In this project we will concentrate on the technical analysis to forecast the stock market.



**Figure 1: Fundamental Analysis and Technical Analysis comparison**

The ability to make better decisions and control risk is one of the main advantages that offers the investors, financial analysts, and traders to reduce risk and make smarter judgements. Because of several factors that may impact market behavior in addition to the regular fluctuations in prices, predicting stock market values is a difficult and complex undertaking. The applications of machine learning and statistical models in financial forecasting has been developing with an aim of improving the accuracy and confidence of investing predictions.

One important financial market that represents Turkeys economy performance and investment opportunities is the Istanbul Stock Exchange (ISE). Both domestic and foreign investors can benefit from understanding and forecasting the ISE attribute. A dataset from the UCI machine learning repository, which has multiple financial variables, is used in this work. The above features are essential in developing predictive models and fully understanding the connections between the targeted variable and other financial metrics.

**Problem Statement**

This study’s main goal is to use the historical financial data to anticipate the Istanbul Stock Exchange (ISE) with accuracy. Choosing the right model for prediction is important because of the complexity of financial markets. Even though they are widely used, traditional time series models like ARIMA might not be able to properly represent complex dependencies on time and non-linear relationships. On the other hand, advanced deep learning models such as LSTM network have shown progress in processing sequential data; however, they require sufficient training data and appropriate tuning. The objective of this research is to evaluate and compare the ISE indices forecasting performance of the LSTM and ARIMA models to identify which model requires less data to provide forecast that are more accurate.

**Justification of the study**

This study is justified by the potential benefits of improving investment forecasting accuracy. Accurate models may enhance financial decision-making, investment strategies and risk minimization. By evaluating the performance of LSTM and ARIMA models, identifying the benefits and drawbacks of each model, this project contributes to the ongoing research in financial projections. The findings will be useful to the investors, financial analysts and researchers who are interested in applying machine learning and statistical models in the financial market. And in cases where data is very limited, understanding which model performs best with less data might be extremely beneficial.

**Research Questions**

The research aims to provide explanation to the following questions:

1. When predicting the stock market with limited amounts of data input, which model LSTM or ARIMA provides the highest accuracy?
2. By studying market movements, daily returns, moving averages and correlations between stocks, how can we forecast future stock behavior?

**Aims and Objectives**

The aim of this research is to find out whether model like LSTM or ARIMA predicts the Istanbul Stock Exchange (ISE) index more accurate. The objective is to:

1. Explain how well the ARIMA and LSTM model predicts the ISE index.
2. Determine which model works best with minimum data input.
3. Describes some efficient financial forecasting techniques.

**Literature Review**

Talati, D., Patel, M., & Patel B. (2022) suggest that because LSTM can manage the sequential and nonlinear structure of financial data and the model is good at predicting the stock prices. The author believes that feature selection and data preparation can improve the model’s performance. The author has trained the model on three different datasets such as “Infosys (1996 to 2022), Microsoft (1986 to 2022) and TCS (2002 to 2022)”.

Huang, W. (2023), concludes that LSTM model can learn from the sequential patterns of stock data and include long term dependencies, as they are quite predictive at predicting the prices of the stock. In the study the author compared the model’s performance with the existing models, gives a 35.18% improvement in predicting the stock market. The author uses the historical stock price data of “Google stocks “.

**Methodology**

**Overview**

The goal of this project is to use machine learning techniques to predict the Istanbul Stock Exchange. This will be executed in few steps (Figure 2). Collecting and reading the data is the first step, and the comes data preprocessing, to handle the missing values and identify the features from the dataset. And then data is split into training and testing and trained using LSTM and ARIMA models. Subsequently we will test the models, and in the end, we’ll assess model performance by tuning the parameters to see which model forecast the future stocks more accurately.

* Training
* Testing
* Validation
* Clean data (missing values, handle outliers).
* Normalize data.
* Feature Selection
* Install “**ucimlrepo”**
* Load fetch” unci repo(id=247)
* Here Id is Dataset for Uci repo.

|  |
| --- |
| * Time series (ARIMA) * LSTM (Long Short-Term Memory) * ANN (Artificial Neural Network) * Update Model architecture. * Update parameters to get better accuracy.   **Dataset**: Istanbul Stock Exchange  End   * SVM (Support Vector Machine) * Time series (ARIMA) * LSTM (Long Short-Term Memory) * ANN (Artificial Neural Network) * Random Forest   Predict the model.  Apply model (Test Dataset)  Predict the future stock market.  Optimize model (Fine-tune model)  Optimize model (Fine-tune model)  Define the ML model architecture.  Train the Model &  Validattion  Split data (Train & Test)  Pre-processing Data  Data Collection  Load & Read Data  Start |

**Figure 2: Data management plan flow chart**

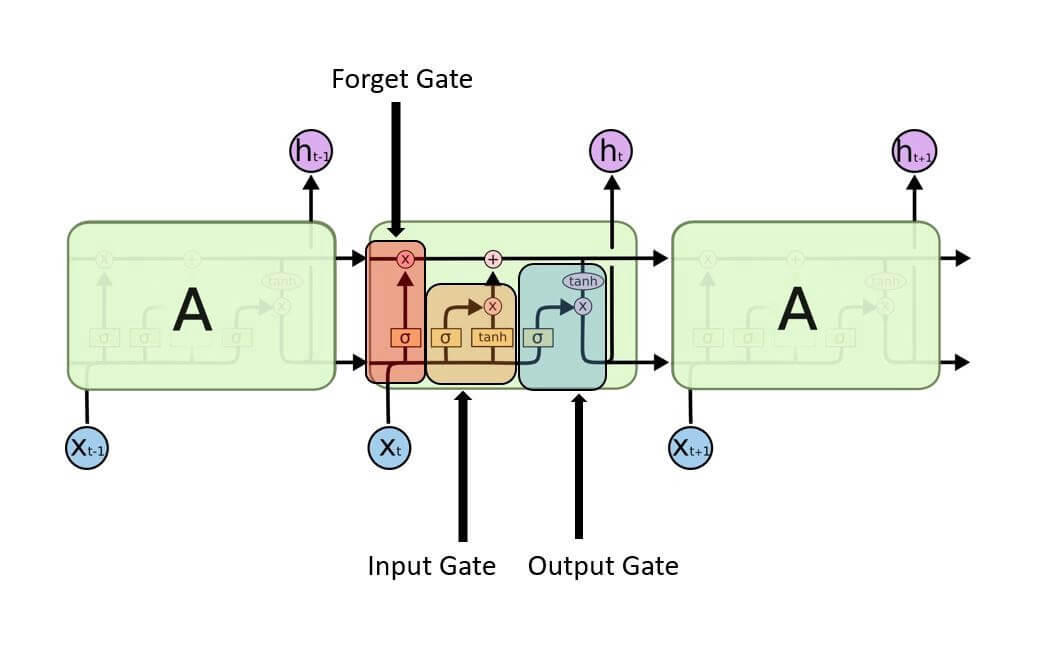
**Long Short-Term Memory (LSTM):**

Long Short-Term Memory networks are a type of Recurrent Neural Network (RNN) that process sequential data and collect long-term dependencies. When handling time-series data, like stock prices, which are generally non-stationary and show trends and changing circumstances, LSTM prove to be particularly useful.

**Why use LSTM for predicting stock market?**

With the recent breakthroughs that have been happening in data science, it is found that for almost all these sequence prediction problems, long short-Term Memory networks have been observed as the most effective solution (Talati, D., Patel, M., & Patel B. (2022) IV methodology). The basic idea of traditional statistical models such as linear regression and Autoregressive Integrated Moving Average (ARIMA) is that the data is stationary, that its statistical features such as variance and mean don’t change over time. Stock prices, on the other hand, show seasonality, patterns, and trends despite not being stationary. LSTMs are ideally suited for stock market prediction because they can manage this non-linear correlation within the data.

To handle new information, a sigmoid function is used to modify the existing data, and that’s how the neural networks operate rather than classifying information to be important or not. As a result, the data has been updated completely. LSTM, on the other hand, uses a mechanism to transmit the data known as cell states while executing minimal addition and multiplication changes to the data.



**Figure 3: Long Short-Term Memory Network**

**Essential elements of LSTM:**

* **Memory Capability**: Because LSTMs can remember information over time, they are perfect for capturing long-term dependencies in sequential data.
* **Gated mechanism**: LSTMs use gates to manage information flow, enabling them to store important data and eliminate unnecessary information.

**Functioning of LSTM:**

* **Cell state**: The sequence’s cell state serves as a memory that retains data throughout several time steps. Gates are highly controlled structures that allow information to be inserted or deleted from the cell state.
  + **Gates**:
    - Input gate: This gate selects the new data to add to the cell state.
    - Forget gate: This gate selects which data from the cell state should be removed.
    - Output gate: The output gate selects which portion of cell state to output.

**Data Collection**

**Data Preprocessing**

**Data Feature Selection**

Reference:

1. Stock Market Prediction Using LSTM Technique, <https://www.ijraset.com/research-paper/stock-market-prediction-using-lstm-technique>
2. Enhancing Stock Market Prediction Through LSTM Modeling and Analysis, <https://eudl.eu/pdf/10.4108/eai.2-6-2023.2334692>