# Stock price prediction – Team 9

## Introduction:

### Background

Predicting stock prices has long been a challenging yet essential task for investors, traders, and financial institutions. Accurate predictions can help in making informed investment decisions, managing risks, and maximizing returns. This document presents an analysis of stock price prediction using machine learning and data analysis techniques.

### Objective

The primary objective of this project is to develop a predictive model for stock prices that leverages historical data, financial indicators, news sentiment analysis, and economic indicators. The aim is to provide valuable insights into future stock price movements.

### Scope

This analysis focuses on a specific set of stocks and considers various features, including historical price data, financial ratios, news sentiment, and macroeconomic factors. The project will employ machine learning algorithms to develop predictive models and evaluate their performance.

## Methodology:

### Data Collection

* Gathering historical stock price data from reliable sources.
* Collecting financial indicators such as P/E ratio, EPS, and volume.
* Scraping news articles and conducting sentiment analysis.
* Incorporating macroeconomic indicators like GDP growth and interest rates.

### Data Preprocessing

* Handling missing data and outliers.
* Normalizing or scaling numerical features.
* Encoding categorical variables.
* Exploratory Data Analysis (EDA) for data understanding.

### Feature Engineering

* Creating new features based on domain knowledge.
* Lagging features to capture temporal patterns.
* Selecting the most relevant features for modeling.

### Model Selection

* Evaluating different machine learning models (e.g., linear regression, decision trees, LSTM).
* Fine-tuning hyperparameters using cross-validation.

### Model Evaluation

* Assessing model performance using appropriate metrics.
* Backtesting the model with historical data.
* Conducting sensitivity analysis.

### Performance Metrics

* Mean Absolute Error (MAE)
* Root Mean Square Error (RMSE)
* Mean Absolute Percentage Error (MAPE)
* R-squared (R²)

## Data Sources:

### Historical Stock Price Data

* Source: Stock exchange data (e.g., NYSE, NASDAQ)
* Contains historical daily or intraday price and volume information.

### Financial Indicators

* Source: Financial reports (e.g., 10-K filings)
* Includes fundamental indicators like P/E ratio, EPS, and revenue.

### News Sentiment Analysis

* Source: News websites and APIs
* Utilizes Natural Language Processing (NLP) for sentiment analysis of news articles.

### Economic Indicators

* Source: Government reports (e.g., Bureau of Economic Analysis)
* Comprises macroeconomic data such as GDP growth, unemployment rate, and interest rates.