# Concise Report on Stock Price Prediction System

## Overview

This report outlines the methodologies and technologies used in developing a comprehensive Python-based system for predicting stock prices. The system leverages data from the Alpha Vantage API and performs advanced predictive modeling using machine learning algorithms. The project also incorporates sentiment analysis from news headlines to provide contextual insights that potentially influence stock prices.

## System Components

1. **Data Fetching and Management**:
   - **API Interaction**: Utilizes the Alpha Vantage API for real-time financial data. Data requests respect API rate limits through a custom-built rate limiter.
   - **Caching Mechanism**: Implements data caching to optimize API usage and limit data fetches, using pickle for serialization and timestamping for freshness validation.
2. **Predictive Modeling**:
   - **Data Preprocessing**: Scales numerical features using `StandardScaler` to normalize data, ensuring consistent input for machine learning models.
   - **Model Training**: Employs `GradientBoostingRegressor`, a robust algorithm suitable for non-linear trends observed in stock price movements.
   - **Future Price Prediction**: Uses `Cubic Hermite Spline` interpolation to predict future stock prices based on trends derived from historical data.
3. **Sentiment Analysis**:
   - **News Data Fetching**: Gathers relevant news headlines using the News API, keyed by stock symbols.
   - **Sentiment Evaluation**: Analyzes headlines using a pre-trained BERT model (`bert-base-multilingual-uncased-sentiment`) to assess sentiment, which is then correlated with stock data.
4. **Visualization**:
   - **Trend Visualization**: Generates both static (Matplotlib) and interactive (Plotly) visualizations showcasing stock trends and volume changes.
   - **Sentiment Integration**: Visualizes sentiment trends alongside stock price data, providing a holistic view of potential market influences.

## Methodologies

- **Machine Learning**: The system trains a gradient boosting model on historical stock data, learning patterns and anomalies that affect stock prices.
- **Interpolation**: For predicting future stock prices, the model outputs are used as a basis for cubic Hermite spline interpolation, which provides a smoothed trend prediction over a specified future period.
- **Sentiment Analysis**: By incorporating sentiment analysis, the system adds a layer of qualitative analysis, considering how public perception and news can impact stock behavior.

## API Management

- **Rate Limiting**: Ensures compliance with API constraints by calculating the necessary delay between consecutive API calls, thus avoiding over-polling and potential bans.
- **Data Freshness**: The caching system includes a freshness mechanism, where data older than a specified threshold is fetched anew, ensuring that the analysis is based on the latest available data.

## Conclusion

The developed system is a robust tool for financial analysis and stock price prediction, combining advanced data science techniques with practical API management and visualization. By integrating quantitative stock data with qualitative news sentiment, it provides a comprehensive tool for traders and analysts to assess potential stock movements and make informed decisions. The modular design and extensive documentation ensure that the system can be easily adapted and extended to include additional data sources or modeling techniques.