# Prediction of Tencent Share Price Based on ARIMA, SVR and LSTM

Xinyu Yang
*School of Economics, Jinan University, Guangzhou, China*

Keywords: Stock Price Prediction, ARIMA, SVR, LSTM, Tencent.

Abstract: As a matter of fact, the predicting stock price can not only be a hard chicken to pluck but also it is an exciting research area of financial markets for the researchers' due high return profit that can be made upon successful prediction accuracy, which matters based on informed investment decision especially in contemporary high volatility market. With this in mind, in this article, the stock price of Tencent Holdings is predicted using 1,203 parameter sets by comparing three models, i.e., ARIMA (Autoregressive Integrated Moving Average), Support Vector Regression (SVR), and Long Short-Term Memory Networks (LSTM). According to the analysis, the results reveal that SVR is very well suited for short-term predictions due to its ability to capture market volatility. In contrast, LSTM can handle long term patterns. The ARIMA model, on the other hand, hamstrung to make sense of non-linear data by its linear foundation. While they differ in performance, it is a useful guidance to investors choosing between models for specific needs.

## 1 INTRODUCTION

Predicting stock prices is one of the most important research topics in financial markets, using various mathematical and statistical methods. The early research in this regard was based on the Random Walk Theory (RWT), according to which stock prices are a completely random process, and future movements of share price depends only upon past movement of market information not past prices (Fama, 1995). While that theory bolstered the efficient market hypothesis, it was later oppugned because it did not consider the regularities in stock prices.

Recent trends in finance and statistics have upheld Autoregressive Integrated Moving Average (ARIMA) model as one of the paramount methods for stock price forecasting. ARIMA is actually a model that Autoregressive (AR) and Moving Average (MA) components to support non-stationary data. It models future values as linear combinations of past data and errors thus making it suitable for short-term projections (Ariyo et al., 2014). As one can see in the case of Weng, he used a ARIMA model to predict stock worth with 939 closing prices from Construction Bank and already acquired good results for short-range prescience (Weng, 2023). Whilst ARIMA is one of the most accurate time series models for the stock market, it fails to provide an interpretable feature that clearly indicates when you should invest in long term stocks because investing based off seasonality or general bullish/bearish trends will lead to huge errors due to autocorrelation.

Besides traditional statistical models, Support Vector Regression (SVR), as a machine learning method is promising in stock price prediction. By the nature of its algorithm itself: linear and non-linear problems can both be addressed accurately. SVR, considering the principle of structural risk minimization, can provide strong generalization ability for financial time series forecasting task. According to Li, it achieves an annualized return of 14.75% based on the multifactor stock selection model constructed by SVR with a trading adjustment frequency 30 days ago in and out (Li, 2022).

Long Short-Term Memory (LSTM) can capture long-term dependencies in time series data, making them very successful for financial modelling where the traditional Recurrent Neural Network (RNN) could suffer from vanishing gradient problem (Bhandari et al., 2022). For example, the study carried out by Yang and Wang utilized a deep LSTM network to forecast stock indices from 30 global markets (Yang & Wang, 2019). LSTM offered the best forecasting performance over other methods (e.g., SVR, RNN and ARIMA) through various maturities. In addition, hybrid models which combines the traditional statistical methods with deep learning like



Yang, X.
Prediction of Tencent Share Price Based on ARIMA, SVR and LSTM.
DOI: 10.5220/0013213800004568
In *Proceedings of the 1st International Conference on E-commerce and Artificial Intelligence (ECAI 2024)*, pages 229-233
ISBN: 978-989-758-726-9

ARIMA-LSTM robustly provides better results than the solo counterparts (Abdulrahman et al., 2020).

This paper focuses on predicting the stock price of Tencent Holdings, a subject that has been explored in some studies. In a study conducted by Shi and Zhuang, they compared different soft computing techniques for prediction of the fusion defect. ANN showed more accurate performance to predict the fusion output among all other models used in this research (Shi & Zhuang, 2019). Wang et al. introduced an Event Attention Network (EAN) to predict short-term stock price trends of companies like Tencent using social media and news data (Wang et al., 2019). Lu et al. compared stock prices of two internet companies in China: Tencent and Alibaba using Capital Asset Pricing Model (CAPM), Dividend Discount Mode (DDM) and Fama-French Three-Factor Model (FF3F), and Tencent showed a larger proportion of expected returns (Lu et al., 2021). Zhou proposed an LSTM model combined with multidimensional input and sentiment analysis to improve the predictability of Tencent's stock price (Zhou, 2021).

This study aims to apply the 1,203 parameters into ARIMA, SVR and LSTM to find out which model is predictive power for forecasting Tencent's stock price movements. The one problem with ARIMA is it handles short-term linear trends whereas long term could be non-linear during explained period. SVR helps in short term predictions because of the kernel trick that SVRs use by properly taking care of non-linear relations. Given that LSTM are able to capture long-term dependencies in time series, this makes them a very good option for performing long term forecasts. This paper performs a systematic comparison of this series predictive power for the price trajectory. The subsequent sections describe the data and methods used, a comprehensive analysis on model performance as well as provide practical implications for investment decisions. Finally, this study concluded comparison of all these models that which model works best for stock price prediction.

## 2 DATA AND METHOD

The data used in this study was sourced from investing. com providing 34,909 data of Tencent Holdings Limited daily since its first listing on HKEx from June 17, 2004 to September 9, 2024. Entries have date, close, open, High and Low prices of the day in HKD, traded volume in million and range of fluctuation (%). The closing price is used as a dependent variable to predict changes in stock prices at the next trading day, whereas opening, high, low and the others are independent variables that reflect the direction on how this dynamic may evolve. All computation for this study was performed on an environment with TensorFlow 2.9.0, Python 3.8, CUDA11, 80GB RAM, AMD EPYC 7642 and RTX 3090 via the cloud computing platform AutoDL. To facilitate model training and testing, the dataset was divided into two parts: the first 3,990 days of data were used for model training, and the subsequent 997 days were reserved for testing and evaluating predictive performance. As ARIMA is a univariate model, only the date and closing price were used for its training. On the other hand, SVR and LSTM employed all variables. Additionally, the data were normalized before training SVR and LSTM models to ensure efficient training and accurate predictions.

This paper analyzes and forecasts Tencent's stock price using three different forecasting models, namely ARIMA, SVR, and LSTM. Each model uses different methods to find the optimal parameters and quantifies the prediction effect of the model through the evaluation indexes such as Coefficient of Determination ($R^2$), Mean Squared Error (MSE), Mean Absolute Error (MAE) and Mean Percentage Absolute Error (MPAE). In order to select an appropriate ARIMA model, an Augmented Dickey-Fuller Test (ADF) was first performed on the closing price data to determine the smoothness of the data and the order of difference. The test results indicated that the closing price data was non-stationary and required first order differencing. Subsequently, the auto_arima function was used to automatically select the optimal model order from 147 parameter combinations based on the Akaike Information Criterion (AIC). The final optimal ARIMA model obtained is ARIMA(5,1,3), i.e. p=5, d=1, q=3. For the SVR model, this paper optimizes the model parameters by hyperparameter grid search to find the optimal parameter combination from 32 different parameter combinations. In the process of parameter tuning, 5-fold cross-validation is used and negative_mean_squared_error is used as the scoring criterion. The final optimal parameter combination obtained is penalty parameter (C) = 100, ε-insensitive loss function (epsilon) = 0.01, and kernel function is linear. The hyperparameters of the LSTM model were tuned by Keras Tuner, traversing 1024 different parameter combinations. The final optimal LSTM model consists of two layers of LSTM, the first layer has 100 neurons and returns sequences, and the second layer has 100 neurons and does not return sequences, both with a dropout rate of 0.2. The model uses Adam's optimizer, with a learning_rate of 0.01, and has been trained with 500 epochs to achieve stable predictions. Stable prediction results were achieved. The loss function is Mean Square Error (MSE) and the batch size is 32.

# 3 RESULTS AND DISCUSSION

## 3.1 Model Performances

In this study, the authors employed three distinct models, i.e., ARIMA, SVR, and LSTM, to forecast the closing price of Tencent Holdings' stock. The predictive efficacy of each model was evaluated using a set of performance metrics, including MSE, MAE, $R^2$, and MPAE. As illustrated in Table 1, the subsequent section presents a comprehensive performance and comparative analysis of each model.

Table 1: Model Prediction Performance Comparison Table.

| Model | MSE | MAE | $R^2$ | MPAE |
|---|---|---|---|---|
| ARIMA | 40491.03 | 179.15 | -2.0378 | 53.29% |
| SVR | 14.83 | 2.85 | 0.9989 | 0.70% |
| LSTM | 196.63 | 9.13 | 0.9848 | 2.23% |

The ARIMA model employs time series differencing to process non-stationary data, exhibiting efficacy in capturing short-term fluctuations. However, the model demonstrates suboptimal performance in long-term forecasting. The ARIMA model yielded a MSE of 40491.03, a MAE of 179.15, a $R^2$ of -2.04, and a MPAE of 53.29%. The negative $R^2$ value indicates that the ARIMA model does not fit the stock price data well, and that the prediction results are significantly inaccurate. As illustrated in Figure 1, the ARIMA model is unable to accurately capture the primary trend of the stock price, particularly when the stock price exhibits significant fluctuations. In such instances, the model's prediction demonstrates a persistent deviation.

The SVR model demonstrates an exceptional capacity for forecasting Tencent's stock price movements. The model exhibits a MSE of 14.83, a MAE of 2.85, a high $R^2$ of 0.9989, and a MPAE of 0.70%.These results proved that the SVR model can well fit the actual data of stock price and achieve high prediction accuracy, especially for those nonlinear time series with complex patterns. Figure 2 suggests that it can perfectly reproduce the trend of real stock price fluctuation and short-term predictive performance, as a SVR model.

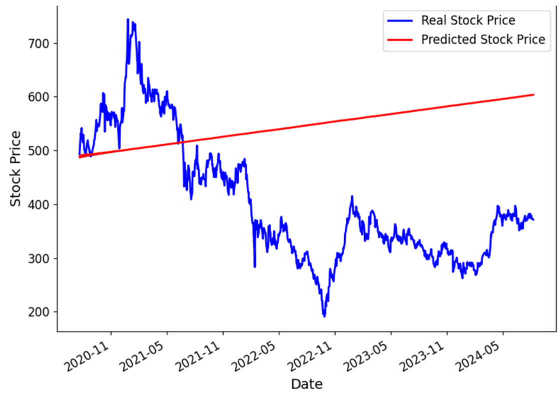


Figure 1: Tencent Stock Price Prediction Using ARIMA (Photo/Picture credit: Original).

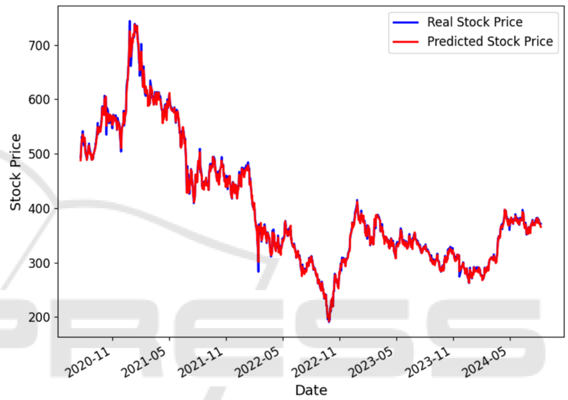


Figure 2: Tencent Stock Price Prediction Using SVR (Photo/Picture credit: Original).

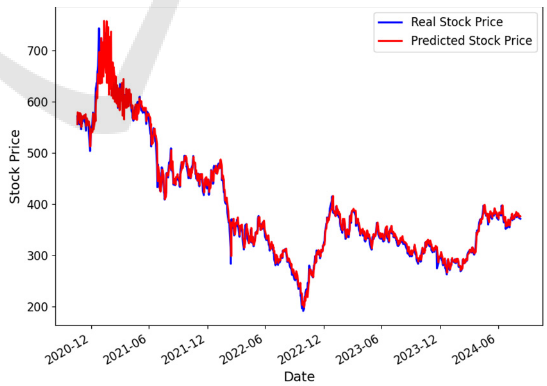


Figure 3: Tencent Stock Price Prediction Using LSTM (Photo/Picture credit: Original).

Since there is high time-series with the dataset, LSTM model outperforms in processing complex signals likely capturing long-term dependencies and nonlinear features. Model has MSE: 196.63, MAE: 9.13, $R^2$ : 0.9848 and MPAE: 2.23%. Even though the

LSTM model performs lower than SVR, it shows unique power to catch stock price changes over long-time horizons. Figure 3 illustrates that the LSTM model is able to follow long-term trends of stock prices relatively well. More importantly, it can perform well to keep a reasonably accurate level of prediction especially when facing long-term stock price fluctuation.

The SVR model has the best fitting result and robustness in short-term forecasting among all models, while LSTM can better capture long term trends. On the other hand, a ARIMA model that only uses both autoregressive and difference terms of time series show relative weak performance in the complex stock market.

## 3.2 Explanation and Implications

While the LSTM model showed under-perfromace on short term compared to SVR, it had a significant ability in capturing long range dependencies. The specific memory structure of the LSTM model allows for capturing long-term trends and patterns in time series data particularly relevant to various strategies. The LSTM model will help investors to understand the long-term trajectory of its stock price, given that it is a technology giant and has good scope for growth in future as well, so they can design their investment portfolio accordingly. This by integrating the short-term trend and long term factoring, investor could shift their investment style according to market dynamics focusing more towards a balance of profit between game than from a view with only eye on the ball.

## 3.3 Limitations and Prospects

This research uses three models in this paper to systematically analyze and forecast Tencent stock. However, the research of this study with ARIMA, SVR and LSTM models sftill has some limitation. In other words, the ARIMA model is only suitable for univariate time series prediction and it cannot include external factors that affect stock prices like macroeconomic data or company financials. The ARIMA model competence is limited updating data set with changing market context. Second, despite the improved nonlinear processing by SVR and LSTM models, these more advanced methods also suffer from higher computational demand. The LSTM model is a time and computationally exhaustive process, both for training the network —that involves hyperparameter tuning; background search over every possible value. In addition, the dataset can be divided into three cases with different lengths to better detect the short-, medium- and long-term predictions of the three models.

A more limitation is the curtailment of data set. This study uses daily stock prices that are based on historical data of the market, which can reflect market behavior in hindsight but does not completely represent all complex and uncertain future dynamics. Moreover, training the model does not have include those other external factors that one cannot control with most weight as news/policy/economy etc may fluctuate the stock prices to another front which is difficult for the model to treat accurately.

With this in mind, future research could advance these findings at several levels. Next, future work may implement similar multivariate time series models with exogenous features (e.g., macroeconomic metrics and industrial trends) to enhance the precisions of stock price predictions. Finally, the combination with further advanced deep learning such as attention mechanisms and graph neural networks together with nonlinear models like SVR or LSTM might be another way to improve performance. It might therefore increase the prediction performance of the model. Additionally, by incorporating new sources of data such as technical analysis and sentiment analysis which covers various market sentiments from company news reports to social media may give a more holistic view on providing inputs for making stock price predictions. Another important issue is the establishment of hybrid models that combine different model capabilities together. For example, a hybrid model of SVR short-term prediction and LSTM long-term prediction could be established to grasp the market volatility in the short term by ignoring trend alternation for a longer period. This will provide investors a better ability to analyze the market and make more data-driven investment decisions in one of the fastest moving markets.

# 4 CONCLUSIONS

To sum up, this study systematically examines the predictability of Tencent Holdings stock price using ARIMA, SVR and LSTM models. The results demonstrate that SVR is good at predicting the short-term stock price and recording change in market, while LSTM has a better ability to catching long-term patterns and dependence. While on the other hand, even though ARIMA model has great uses for stationary time series but in a lot of classical and complex financial time series problems it provides very poor results due to its linearity nature. Besides, it fails to consider multivariate information and external market factors in this paper which may result in prediction accuracy is not high. Further research may combine multiple models as well as incorporate more externals data sources, such market sentiment to enhance the prediction. The results of this paper add to investors in the stock market gives a reference on improving short-term or long-short strategies according to different forecasting models.

## REFERENCES

Abdulrahman, U. F. I., Ussiph, N., Hayfron-Acquah, B., 2020. *A hybrid arima-LSTM model for stock price prediction.* International Journal of Computer Engineering and Information Technology, 12(8), 48-51.

Ariyo, A. A., Adewumi, A. O., Ayo, C. K., 2014. S*tock price prediction using the ARIMA model.* 2014 UKSim-AMSS 16th international conference on computer modelling and simulation, 106-112.

Bhandari, H. N., Rimal, B., Pokhrel, N. R., Rimal, R., Dahal, K. R., Khatri, R. K., 2022. *Predicting stock market index using LSTM.* Machine Learning with Applications, 9, 100320.

Fama, E. F., 1995. *Random walks in stock market prices.* Financial Analysts Journal, 51(1), 75-80.

Li, H. Y., 2022. *Research on quantitative stock selection based on SVR.* Industrial Innovation Research, 21, 127-129.

Lu, S., Li, X., Qi, Y., Zheng, Z., 2021. C*omparison of Different Asset Pricing Models Based on Alibaba and Tencent Stocks.* 2021 3rd International Conference on Economic Management and Cultural Industry (ICEMCI 2021) 3118-3124.

Shi, C., Zhuang, X., 2019. *A study concerning soft computing approaches for stock price forecasting.* Axioms, 8(4), 116.

Sun, L., Yuan, H., 2023. *Research on the tencent company stock price based on ARIMA model.* Advances in Economics, Management and Political Sciences, 64(1), 263–268.

Wang, Y., Li, Q., Huang, Z., Li, J., 2019. *EAN: Event attention network for stock price trend prediction based on sentimental embedding.* Proceedings of the 10th ACM conference on web science 311-320.

Weng, Z., 2023. *Analysis and prediction of stock prices based on ARIMA model: A case study of China Construction Bank*. Modern Information Technology, 14, 137-141.

Yang, Q, Wang, C., 2019. *Research on global stock index prediction based on deep learning LSTM neural network*. Statistical Research, 3, 65-77.

Zhou, S., 2021. *A Stock Prediction Method Based on LSTM.* Emerging Trends in Intelligent and Interactive Systems and Applications: Proceedings of the 5th International Conference on Intelligent, Interactive Systems and Applications (IISA2020), 186-193.