Sentiment-Driven Forecasting LSTM Neural Networks for Stock Prediction-Case of China Bank Sector

Shangshang Jin
Department of Art and Science, Johns Hopkins University, Washington, D.C., United States

***Abstract*—This study explores the predictive analysis of public sentiment in China's financial market, focusing on the banking sector, through the application of machine learning techniques. Specifically, it utilizes the Baidu Index and Long Short-Term Memory (LSTM) networks. The Baidu Index, akin to China's version of Google Trends, serves as a sentiment barometer, while LSTM networks excel in analyzing sequential data, making them apt for stock price forecasting. Our model integrates sentiment indices from Baidu with historical stock data of significant Chinese banks, aiming to unveil how digital sentiment influences stock price movements. The model's forecasting prowess is rigorously evaluated using metrics such as R-squared (R^2), Root Mean Square Error (RMSE), Mean Absolute Error (MAE), Mean Absolute Percentage Error (MAPE), and confusion matrices, the latter being instrumental in assessing the model's capability in correctly predicting stock up or down movements. Our findings predominantly showcase superior prediction performance of the sentiment-based LSTM model compared to a standard LSTM model. However, effectiveness varies across different banks, indicating that sentiment integration enhances prediction capabilities, yet individual stock characteristics significantly contribute to the prediction accuracy. This inquiry not only underscores the importance of integrating public sentiment in financial forecasting models but also provides a pioneering framework for leveraging digital sentiment in financial markets. Through this endeavor, we offer a robust analytical tool for investors, policymakers, and financial institutions, aiding in better navigation through the intricate financial market dynamics, thereby potentially leading to more informed decision-making in the digital age.**

Keywords*—*Machine learning; LSTM; sentiment; forecasting; banking sector

I. INTRODUCTION

The digital epoch ushered in a plethora of tools and platforms, with search engines emerging as crucial conduits for gauging public sentiment. Shiller [1] accentuates the importance of grounding research in actual human behavior to comprehend how individuals genuinely think and act. This sentiment concurs with the dynamics of the digital realm, where search query fluctuations concerning financial entities can unveil early signs of shifting public interest or potential concerns. These subtle shifts, characteristic of the digital age, harbor the potential to impact stock market trends [2].

This study endeavors to tap into public sentiment within China's financial market, probing how sentiments harvested from digital platforms can act as a predictive lens for scrutinizing and foreseeing market dynamics. Central to this exploration is the Baidu index, a tool often drawn parallel to China's version of Google, emblematic of the digital transition. Baidu, surpassing its initial function as a search engine, has evolved into a dynamic digital nexus, reflecting the collective sentiment of its extensive user base. This discernment elevates tools like the Baidu index to a significant pedestal in contemporary financial prediction frameworks, particularly against the backdrop of digital transformation sweeping through banks and financial institutions [3].

Banks are pivotal players in China’s economic narrative, entrusted with capital provisioning, trade facilitation, and financial stability maintenance [4]. The banking sphere, headlined by colossal entities like the Industrial and Commercial Bank of China (ICBC) and the Agricultural Bank of China (ABC), commands substantial influence over China's economic fabric. In the digital era, their roles and impact have evolved, with stock performance mirroring not only the health of individual institutions but also the broader economic vitality [5]. Observing the stock trends of these institutions can significantly inform both domestic economic strategies and global financial outlooks. Traditional stock prediction methodologies, shackled by inherent constraints, often grapple with accuracy challenges, paving the path for Neural Networks—specifically, the Long Short-Term Memory (LSTM) networks—renowned for their prowess in time series predictions owing to their capability to process long-term data dependencies.

Nonetheless, the nuanced realm of stock market predictions often finds the exclusive reliance on raw financial metrics inadequate. The digital epoch grants stakeholders a platform to voice their opinions, assumptions, and concerns. Tapping into this vast sentiment reservoir, particularly reflected in search patterns on platforms like Baidu, can unveil precious insights. The fusion of sentiment analysis with LSTM models ushers in a novel frontier for stock prediction, especially within crucial sectors like China's banking domain.

This inquiry embarks on an innovative journey, integrating sentiment indices from the Baidu platform with LSTM neural networks to refine the prediction framework for China's banking stocks. Our scrutiny especially shines a light on stalwarts like Industrial and Commercial Bank of China (ICBC), Agricultural Bank of China (ABC), Bank of Communications (BoCom), CCB (China Construction Bank), China Merchants Bank (CMB), Shanghai Pudong Development Bank (SPDB), and Industrial Bank (IB). Overlaying sentiment metrics with historic stock data, we endeavor to ascertain the predictive potency of sentiments, gauged via the Baidu index, in molding future stock valuations within the banking sector.

This endeavor distinctively enriches the literature by elucidating the intertwined roles of banks in financial stability, regulatory landscapes, investor interests, and the digital transformation of the banking sector. Accurate forecasting of bank stock prices enhances our comprehension of impending financial scenarios and vulnerabilities. By leveraging sentiment analytics, we aim to foresee shifts in the continually evolving regulatory labyrinth banks traverse, equipping stakeholders with insights into potential policy alterations. Recognizing the magnetic appeal the banking sector exudes for investors, our model strives to equip them with precise forecasting tools. Additionally, as the banking sector navigates through transformative digital innovations, our sentiment-driven approach seeks to decipher public perceptions, offering insights into how fintech advancements might reshape bank stock trajectories.

Our exposition is organized as follows: Section II delves into pertinent literature, mapping the journey of sentiment-imbued stock predictions. Section III unravels our methodology, elucidating the integration of Baidu's sentiment indicators with LSTM networks. In Section IV, we disseminate our empirical findings, accentuating the merits and limitations of our model. Finally, Section V encapsulates a reflective summary and contemplates potential directions for future exploration in this domain.

II. Related Work

Research posits that investor sentiment plays a predictive role in cross-sectional stock returns, suggesting that the collective mood of investors can drive stock price fluctuations and impact their anticipated returns [6, 7, 8]. Moreover, an increasing body of literature is exploring the influence of Internet search activities on financial markets, delving into how online information-seeking behaviors may reflect or shape investor sentiment and, in turn, market dynamics.

Li et al. [9] marked one of the initial strides into understanding how financial news articles sway stock prices. Their contention was that while many had mapped the influence of news in a bag-of-words paradigm, the more profound underlying sentiment—bridging the chasm between lexical patterns and stock fluctuations—remained an undercharted territory. By implementing sentiment dictionaries, their work enriched the sentiment space, and the results affirmed that sentiment-infused models yield superior accuracy over the conventional bag-of-words counterparts.

Pivoting to the realm of digital social interactions, Guo et al. [10] embarked on an exploration capitalizing on user comments from Xueqiu, a focal social networking site in the Chinese stock market landscape. Their research sought to dynamically unravel the web of interactions between investor sentiment and the stock market. One salient revelation was that sentiment data's efficacy as a predictive tool is tethered to the degree of public attention the stock garners. This insight underscores the dynamism and contingent nature of sentiment's predictive power.

Smales in [11] delved deeper into the mechanics of how investor sentiment could cause asset prices to deviate from their fundamental equilibrium. He asserted that understanding these deviations is paramount, given their potential ramifications on capital allocation and its costs. Through a comprehensive analysis spanning a quarter-century, he demonstrated that sentiment indicators, exemplified by the Volatility Index (VIX), forge a robust bond with stock returns. Further granularity revealed that some stocks, especially those in tech or smaller capitalization brackets, display heightened susceptibility to sentiment fluctuations. Intriguingly, recessional periods amplify this sentiment-driven volatility.

Bringing sophisticated machine learning into the fold, Xing et al. [12] championed the sentiment-aware volatility forecasting (SAVING) model. Their venture sought to elucidate the bidirectional dance between asset prices and market sentiment. This synthesis of deep learning and sentiment analysis manifested in results that eclipsed traditional forecasting tools and rival neural network architectures in terms of precision.

Lastly, the ubiquity of social media platforms in shaping stock predictions has witnessed academic spotlight. Swathi et al. [13] echoed the notion that platforms like Twitter, given their massive user engagement, have metamorphosed into sentiment barometers. By fusing the Teaching and Learning Based Optimization (TLBO) technique with Long Short-Term Memory (LSTM) frameworks, they demonstrated an unprecedented accuracy in forecasting stock price movements based on social media sentiment.

III. Methodology

This section elucidates the systematic approach employed to forecast the stock prices of significant Chinese banks, leveraging sentiment analysis and LSTM Neural Networks. We initiate with an introduction to LSTM and its aptness for time series forecasting and subsequently incorporate sentiment data derived from the Baidu Index. Acknowledging the latent impact of sentiment on stock market behavior, the Baidu Index is introduced as a lagged variable. Employing grid search, the model's hyperparameters are optimized. Evaluation metrics such as R^2, RMSE, MAE, and MAPE are then utilized to determine the model's predictive accuracy.

A. *Long Short-Term Memory (LSTM) Networks*

LSTM (Long Short-Term Memory) networks, proposed by Hochreiter and Schmidhuber [14], are a specialized subset of Recurrent Neural Networks (RNN), meticulously crafted to grasp long-term dependencies within sequential datasets. Their intrinsic capability to retain patterns across extended sequences renders them particularly proficient for tasks such as stock price prediction.

Refer to Fig. 1 for a visualization of the LSTM model, structured as follows:

- 1 Input Layer: Accepts the sequence of stock prices and sentiment indices.

- 2 Hidden Layers: Composed of LSTM units that can remember patterns in data over long periods.
- 1 Output Layer: Produces the predicted stock price.

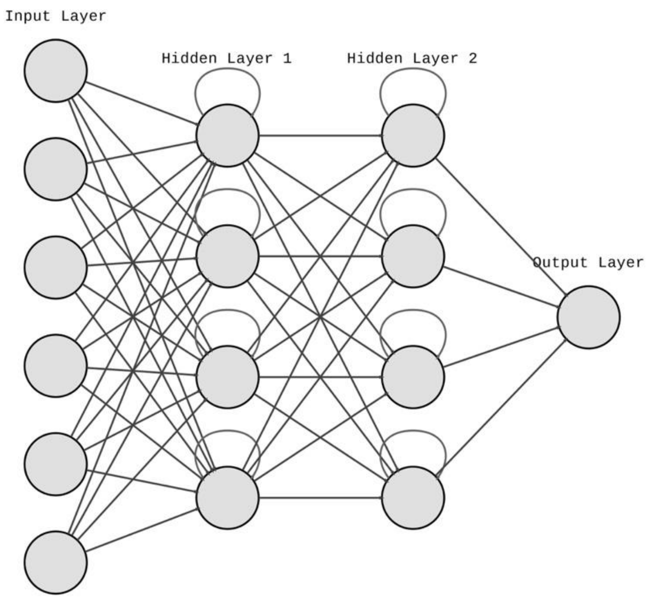

Fig. 1. LSTM structure.

Mathematically, the operations of an LSTM unit are encapsulated as:

$$\begin{aligned} f_t &= \sigma\left(W_f \cdot [h_{t-1}, X_t] + b_f\right) \\ i_t &= \sigma(W_i \cdot [h_{t-1}, X_t] + b_i) \\ \tilde{C}_t &= \tanh(W_C \cdot [h_{t-1}, X_t] + b_C) \\ C_t &= f_t \times C_{t-1} + i_t \times \tilde{C}_t \\ o_t &= \sigma(W_o \cdot [h_{t-1}, X_t] + b_o) \\ h_t &= o_t \times \tanh(C_t) \end{aligned} \tag{1}$$

where:

- $f_t, i_t, \tilde{C}_t, C_t, o_t, h_t$ are the forget gate, input gate, cell candidate, cell state, output gate, and hidden state at time t respectively.
- W_f, W_i, W_C, W_o and b_f, b_i, b_C, b_o are the weights and biases for the respective gates.
- σ is the sigmoid function.

B. Sentiment Index

In this paper, the Baidu index, a popular sentiment index in China, gauges the public sentiment by analyzing search query frequencies and user behaviors on the Baidu search engine. To harmonize the scale between stock prices and the Baidu Index during pre-processing, Min-Max scaling is applied. Recognizing the potential latency in sentiment influence on stock prices, a time-lagged version of the Baidu Index is integrated. This latency captures the inherent delay in investor reactions, attributable to the time taken to interpret and act upon sentiment-driven information, thereby encapsulating market inertia.

For model input augmentation, at each time step t, our LSTM model receives a concatenated input vector, represented as:

$$X_t = [P_t, B_t] \tag{2}$$

Where:

- P_t is the stock price at time t.
- B_t represents the Baidu Sentiment Index at time t.

The LSTM, characterized by its inherent memory, updates its cell state c, factoring in both the stock prices and sentiment index, thereby accounting for sentiment influence across sequences. During the LSTM's forward propagation, this merged input not only impacts immediate outputs but also modifies the hidden states. Consequently, the sentiment information guides predictions and tweaks the LSTM's internal memory. The chosen loss function, Mean Squared Error (MSE), benchmarks model forecasts against real stock prices. During backpropagation, the gradients, inherently influenced by the Baidu Index, fine-tune the model's weightings throughout its training phase.

C. Grid Search Optimization

Grid search is an exhaustive search approach employed for hyperparameter tuning, ensuring optimal model performance. By systematically working through multiple combinations of hyperparameter sets, it evaluates which combination gives the best performance based on a scoring technique [15, 16].

In the optimization process of the LSTM model, several pivotal hyperparameters are meticulously evaluated. The batch size, dictating the number of samples processed before a model update, emerges as a key factor under consideration. The efficacy of various optimizers, notably RMSprop and Adam, is also examined [17]. These algorithms adjust the model's weights in alignment with the data and its corresponding loss function. To fortify the model against overfitting, dropout layers are integrated within the LSTM's architecture [18]. The dropout rate specifies the percentage of neurons randomly nullified during each training phase, enhancing model generalizability. A detailed visualization of the grid search, illustrating the synergy of these hyperparameters and their bearing on model performance, is provided in Fig. 2.

D. Forecasting Evaluation Criterias

To gauge the accuracy and reliability of the model's forecasts, followed by Paudel et al, [19], several statistical metrics are employed:

R^2 (Coefficient of Determination):

$$R^2 = 1 - \frac{\sum_{i=1}^{n}\left(Y_i - \hat{Y}_i\right)^2}{\sum_{i=1}^{n}(Y_i - \bar{Y})^2} \tag{3}$$

RMSE (Root Mean Squared Error):

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(Y_i - \hat{Y}_i\right)^2} \tag{4}$$

MAE (Mean Absolute Error):

$$MAE = \frac{1}{n}\sum_{i=1}^{n}\left|Y_i - \hat{Y}_i\right| \tag{5}$$

MAPE (Mean Absolute Percentage Error):

$$MAPE = \frac{100}{n}\sum_{i=1}^{n}\left|\frac{Y_i-\hat{Y}_i}{Y_i}\right| \quad (6)$$

where Y_t is the actual value at time $t, \hat{Y}_t$ is the predicted value at time , and $\bar{Y}$ represents the mean of the actual values.

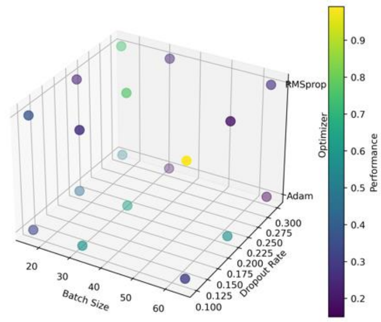

Fig. 2. 3D visualization of grid search.

E. Summary

The proposed technique for predicting stock prices by incorporating the Sentiment Index employs a systematic process, seamlessly integrating the power of Long Short-Term Memory (LSTM) networks and the insights from the Sentiment Index.

Initially, input data comprising of stock prices and a one-day lagged Sentiment Index is collected. This data undergoes normalization to ensure that the range of values is consistent, thereby facilitating effective training. A time-series dataset is then constructed to ensure sequences of data are fed to the LSTM network, aiding in the capture of temporal dependencies.

The dataset gets bifurcated into training and test sets. The LSTM model, designed with two LSTM layers and a dense output layer, is structured to efficiently incorporate and process this information. The first LSTM layer, activated by the tanh function and equipped with dropout, returns sequences which are fed into the second LSTM layer. This layer not only processes the sequences but also integrates the Sentiment Index data. The dense layer at the end aids in producing the final prediction.

Upon constructing this model, it undergoes compilation wherein various aspects like optimizer, loss function, and evaluation metrics are defined. To determine the most optimal parameters for training, a grid search technique is employed. This process considers parameters such as optimizer choice, dropout rate, and batch size, ensuring the model's robustness and accuracy. The model, trained with the best parameters identified by the grid search, is then used to make predictions on the test set.

Once predictions are made, they are de-normalized to revert them to their original scale. Finally, these predictions are evaluated using several metrics to determine the model's efficacy. The metrics include RMSE, MAE, MAPE, and R^2.

The entire technique, encapsulated in the flowchart (see Fig. 3), presents a comprehensive approach to effectively predict stock prices by harnessing the dual power of LSTM networks and sentiment indices.

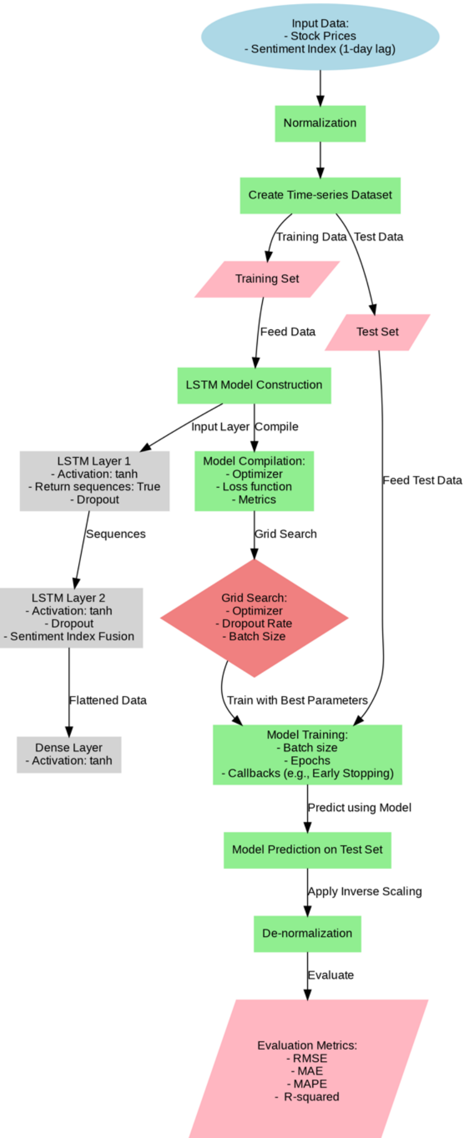

Fig. 3. Flowchart of the proposed technique.

IV. Numerical Results

A. Data Description

The study employs daily stock prices from August 1, 2022, to August 1, 2023, of seven eminent banks listed on the Shanghai Stock Exchange, including ICBC, ABC, BoCom, CCB, CMB, SPDB, and IB (see Fig. 4). This data, sourced from Yahoo Finance, assures precision. Considering Baidu's significant presence in China, its index is a robust reflection of the public's sentiment towards these banks, sentiment indicators is derived from the Baidu index specific to each bank depicted in Fig. 5. For modeling purposes, the data is split: 80% for training and 20% for testing. Preprocessing steps involve data normalization to ensure consistency.

Fig. 4. Stock price.

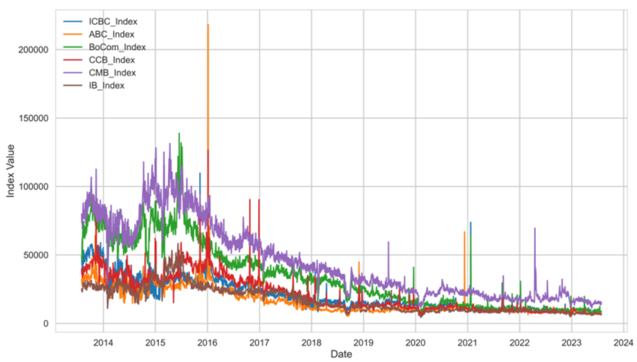

Fig. 5. Stock sentiment index.

B. Hyperparameter Determination

Using grid search, various combinations of batch size, dropout rate, and optimizer type are examined. Each bank's data undergoes this analysis individually, catering to the distinct behavior of each bank's stock prices. The optimal hyperparameters obtained are delineated below in Table I.

C. Forecasting Performance

Table II showcases the forecasting performance of both the LSTM and Sentiment-integrated LSTM (SB-LSTM) models for each bank. For ICBC, the SB-LSTM model exhibits a remarkable increase in R^2 by 6.67% compared to the standard LSTM, and the RMSE also drastically reduces from 0.0510 to 0.0107. In the case of ABC, the R^2 in the SB-LSTM model increases by 2.01% and sees a significant reduction in RMSE from 0.0411 to 0.0185. BoCom's results indicate that the SB-LSTM has an RMSE of 0.0338, an improvement from 0.0553 as observed in the standard LSTM.

Interestingly, for CCB, the standard LSTM outperforms the SB-LSTM in terms of R^2 and RMSE metrics. This suggests that the integration of sentiment does not consistently lead to improved forecasts for every stock. For CMB, the SB-LSTM model dramatically reduces the RMSE from 0.7648 to 0.2316. Similarly, IB's forecasting with the SB-LSTM exhibits a substantial enhancement in R^2, marking an increase of 7.29% compared to the LSTM.

TABLE I. OPTIMAL HYPERPARAMETERS FOR EACH BANK

Stock	Hyperparameter	LSTM	SB-LSTM
ICBC	Batch Size	16	16
	Dropout rate	0.3	0.1
	Optimizer	Adam	Adam
ABC	Batch Size	16	16
	Dropout rate	0.1	0.2
	Optimizer	Adam	Adam
BoCom	Batch Size	16	16
	Dropout rate	0.3	0.3
	Optimizer	Adam	Adam
CCB	Batch Size	16	16
	Dropout rate	0.1	0.3
	Optimizer	Adam	Adam
CMB	Batch Size	16	16
	Dropout rate	0.1	0.3
	Optimizer	Adam	Adam
IB	Batch Size	16	16
	Dropout rate	0.1	0.2
	Optimizer	Adam	Adam

TABLE II. MODEL PERFORMANCE

Stock	Model	R^2	RMSE	MAE	MAPE
ICBC	LSTM	93.03%	0.0510	0.0410	0.9805%
	SB-LSTM	99.70%	0.0107	0.0091	0.2197%
ABC	LSTM	97.48%	0.0411	0.0295	1.0268%
	SB-LSTM	99.49%	0.0185	0.0171	0.6091%
BoCom	LSTM	98.69%	0.0553	0.0379	0.8234%
	SB-LSTM	99.51%	0.0338	0.0320	0.7278%
CCB	LSTM	95.99%	0.0598	0.0419	0.7625%
	SB-LSTM	93.35%	0.0771	0.0767	1.4282%
CMB	LSTM	98.55%	0.7648	0.5662	1.5184%
	SB-LSTM	99.87%	0.2316	0.1973	0.5328%
IB	LSTM	92.66%	0.3340	0.2675	1.6014%
	SB-LSTM	99.95%	0.0274	0.0229	0.1372%

In essence, while the SB-LSTM model generally showcases superior accuracy and precision in forecasting for most stocks, there are notable exceptions, such as CCB. This underscores the idea that sentiment integration typically enhances prediction capabilities, but the individual characteristics of stocks must be taken into account.

D. Classification Analysis

Fig. 6 delineates the confusion matrices for the LSTM and SB-LSTM models across six prominent banks: ICBC, BoCom, ABC, CCB, CMB, and IB. For the ICBC stock, the LSTM model yielded 138 true negatives and 84 true positives but was marred by 131 false positives and 132 false negatives.

Conversely, the SB-LSTM showed a marked improvement, boasting 247 true negatives, 212 true positives, and significantly fewer false positives and negatives. A similar trend was observed for BoCom, where the SB-LSTM displayed enhanced accuracy with 232 true negatives and 227 true positives, overshadowing the LSTM's 134 true negatives and 102 true positives. ABC's LSTM model presented 166 true negatives and 58 true positives, but the SB-LSTM again outperformed by producing 242 true negatives and 189 true positives. This pattern of SB-LSTM superiority persisted with CCB, where the LSTM's 144 true negatives and 86 true positives were eclipsed by the SB-LSTM's 251 true negatives and 210 true positives. For CMB, while the LSTM model recorded 144 true negatives and 111 true positives, the SB-LSTM surged ahead with 257 true negatives and 220 true positives. Lastly, IB's results from the LSTM, comprising 132 true negatives and 96 true positives, were also surpassed by the SB-LSTM's impressive 257 true negatives and 225 true positives.

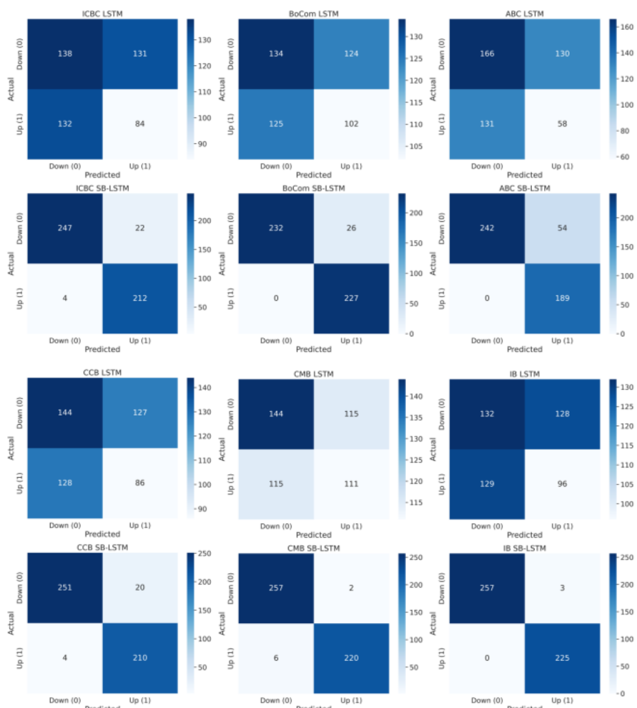

Fig. 6. Confusion matrix.

In summary, the SB-LSTM models consistently demonstrated higher classification accuracy across all stocks when juxtaposed with the LSTM models, emphasizing their superior capability in predicting stock price movements.

V. Conclusion

This study ventured into the realm of sentiment-driven stock prediction, specifically within China's banking sector, utilizing a Sentiment-integrated LSTM (SB-LSTM) model. Our findings predominantly showcased the SB-LSTM model's superior accuracy over traditional LSTM in forecasting stock prices, echoing the potential of sentiment analysis in financial forecasting as suggested by previous studies [20, 21, 22]. Particularly, the use of the Baidu Index highlighted how sentiment data can enhance predictive accuracy. The consistent enhancement in classification accuracy across most stocks by the SB-LSTM models further corroborates the promise of integrating sentiment analysis with LSTM networks for more precise stock price forecasting. While the use of the Baidu Index effectively demonstrated how sentiment data can enhance predictive accuracy, a key limitation lies in the reliance on a single sentiment data source, which may not capture the full spectrum of market sentiments. Additionally, the model's efficacy outside the Chinese banking sector remains untested, suggesting the need for further research to evaluate its applicability across different sectors and geographical regions. This endeavor not only enriches the literature but also paves the way for future explorations in leveraging unstructured sentiment data for financial forecasting amidst the digital transformation sweeping across the financial sector. This transformation is increasingly recognized as crucial in the evolving landscape of financial analytics.

References

[1] Shiller, R. J. (2014). Speculative Asset Prices. *American Economic Review*, *104*(6), 1486–1517. https://doi.org/10.1257/aer.104.6.1486

[2] Li, X., Xie, H., Lau, R. Y. K., Wong, T.-L., & Wang, F.-L. (2018). Stock Prediction via Sentimental Transfer Learning. *IEEE Access*, *6*, 73110–73118. https://doi.org/10.1109/ACCESS.2018.2881689

[3] Lang, Q., Wang, J., Ma, F., Huang, D., & Mohamed Ismail, M. W. (2023). Is Baidu index really powerful to predict the Chinese stock market volatility? New evidence from the internet information. *China Finance Review International*, *13*(2), 263–284.

[4] Fang, J., Lau, C.-K. M., Lu, Z., Tan, Y., & Zhang, H. (2019). Bank performance in China: A perspective from bank efficiency, risk-taking and market competition. *Pacific-Basin Finance Journal*, *56*, 290–309. https://doi.org/10.1016/j.pacfin.2019.06.011

[5] Peng, Z.-Y., & Guo, P.-C. (2022). A Data Organization Method for LSTM and transformer when predicting Chinese banking stock prices. *Discrete Dynamics in Nature and Society*, *2022*, 1–8. https://doi.org/10.1155/2022/7119678

[6] Khadjeh Nassirtoussi, A., Aghabozorgi, S., Ying Wah, T., & Ngo, D. C. L. (2015). Text mining of news-headlines for FOREX market prediction: A Multi-layer Dimension Reduction Algorithm with semantics and sentiment. *Expert Systems with Applications*, *42*(1), 306–324. https://doi.org/10.1016/j.eswa.2014.08.004

[7] Kumar, B. S., & Ravi, V. (2016). A survey of the applications of text mining in financial domain. *Knowledge-Based Systems*, *114*, 128–147. https://doi.org/10.1016/j.knosys.2016.10.003

[8] Li, Q., Wang, J., Wang, F., Li, P., Liu, L., & Chen, Y. (2017). The role of social sentiment in stock markets: A view from joint effects of multiple information sources. *Multimedia Tools and Applications*, *76*(10), 12315–12345. https://doi.org/10.1007/s11042-016-3643-4

[9] Li, X., Xie, H., Chen, L., Wang, J., & Deng, X. (2014). News impact on stock price return via sentiment analysis. *Knowledge-Based Systems*, *69*, 14–23. https://doi.org/10.1016/j.knosys.2014.04.022

[10] Guo, K., Sun, Y., & Qian, X. (2017). Can investor sentiment be used to predict the stock price? Dynamic analysis based on China stock market. *Physica A: Statistical Mechanics and Its Applications*, *469*, 390–396. https://doi.org/10.1016/j.physa.2016.11.114

[11] Smales, L. A. (2017). The importance of fear: Investor sentiment and stock market returns. *Applied Economics*, *49*(34), 3395–3421. https://doi.org/10.1080/00036846.2016.1259754

[12] Xing, F. Z., Cambria, E., & Zhang, Y. (2019). Sentiment-aware volatility forecasting. *Knowledge-Based Systems*, *176*, 68–76. https://doi.org/10.1016/j.knosys.2019.03.029

[13] Swathi, T., Kasiviswanath, N., & Rao, A. A. (2022). An optimal deep learning-based LSTM for stock price prediction using twitter sentiment analysis. *Applied Intelligence*, *52*(12), 13675–13688. https://doi.org/10.1007/s10489-022-03175-2

[14] Hochreiter, S., & Schmidhuber, J. (1997). Long Short-Term Memory. *Neural Computation*, *9*(8), 1735–1780. https://doi.org/10.1162/neco.1997.9.8.1735

[15] Kao, L.-J., Chiu, C.-C., Lu, C.-J., & Yang, J.-L. (2013). Integration of nonlinear independent component analysis and support vector regression for stock price forecasting. *Neurocomputing*, 99, 534–542. https://doi.org/10.1016/j.neucom.2012.06.037

[16] Luo, S., & Tian, C. (2020). Financial high-frequency time series forecasting based on sub-step grid search long short-term memory network. *IEEE Access*, *8*, 203183–203189. https://doi.org/10.1109/access.2020.3037102

[17] Mahardika T, N. Q., Fuadah, Y. N., Jeong, D. U., & Lim, K. M. (2023). PPG signals-based blood-pressure estimation using grid search in hyperparameter optimization of CNN–LSTM. *Diagnostics*, 13(15), 2566. https://doi.org/10.3390/diagnostics13152566

[18] Priyadarshini, I., & Cotton, C. (2021). A novel LSTM–CNN–grid search-based deep neural network for sentiment analysis. *The Journal of Supercomputing*, *77*(12), 13911–13932. https://doi.org/10.1007/s11227-021-03838-w

[19] Paudel, S., Pudasaini, A., Shrestha, R. K., & Kharel, E. (2023). Compressive strength of concrete material using machine learning techniques. *Cleaner Engineering and Technology*, *15*, 100661. https://doi.org/10.1016/j.clet.2023.100661

[20] Huang, M. Y., Rojas, R. R., & Convery, P. D. (2019). Forecasting stock market movements using Google Trend searches. *Empirical Economics*, *59*(6), 2821–2839. https://doi.org/10.1007/s00181-019-01725-1

[21] Liang, C., Tang, L., Li, Y., & Wei, Y. (2020). Which sentiment index is more informative to forecast stock market volatility? evidence from China. *International Review of Financial Analysis*, *71*, 101552. https://doi.org/10.1016/j.irfa.2020.101552

[22] Ito, T., Masuda, M., Naito, A., & Takeda, F. (2021). Application of google trends-based sentiment index in exchange rate prediction. *Journal of Forecasting*, *40*(7), 1154–1178. https://doi.org/10.1002/for.2762