**Student ID: 202395913**

**Student name: Dang Bao Nhi Tran**

**Assignment 1.**

1. **Introduction**

Stock price forecasting is a rewarding task in finance, where accurate predictions can support decision-making to possibly gain profit and reduce uncertainty. **Long Short-Term Memory (LSTM)** networks have been effective in time-series analysis, as shown by studies like Fischer and Krauss [1], A. Moghar et al [2] and Yasmine Ansari [3]. On the other hand, with recent development of quantum computation and algorithms, subfields like **Quantum Machine Learning (QML)**, which leverages quantum properties like **superposition** and **entanglement** can create new opportunities to leverage LSTM's learning capabilities with quantum circuits. Particularly, studies by Chen et al [4] and Cao et al [5] shows that QLSTM offers faster training and competitive performance while require less parameters. Inspired by recent works (Chen et al., 2022; Cao et al., 2023), which demonstrated QLSTM's faster training and competitive performance, this study explores further improvements by integrating advanced quantum encoding and variational gates. Additionally, to gain a deeper understanding of how LSTM and QLSTM make predictions, **interpretability methods** such as **LIME [6]** and **Saliency Maps** are applied. These methods help to interpret the decision-making process of classical LSTM model. Furthermore, **IMV-LSTM (Interpretable Multi-Variate LSTM)** [7] is introduced to address the interpretability challenges in complex models like LSTM. The dataset used in this study is preprocessed version of the stock prices of Merck & Co. Inc. from 2009 to 2020, comprising 34 columns and 2,717 records [8].

**Note: The report only provides brief analysis and evaluation, for the detail code implementation and detailed interpretation please refer to** [**this GitHub repository**](https://github.com/baonhi3008/AdvanceML_ASM1.git)

1. **Classical Machine Learning Models**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | MSE | R-squared | MAPE | Parameters used |
| LSTM | 12.73 | 0.94 | 2.96% | 3217 |
| Enhanced QLSTM (CRX and Amplitude embedding) | 6.2441 | 0.97 | 2.00% | 2047 |
| IMV\_LSTM | 8,3 | 0.96 | 2.75% | 37474 |

1. **LSTM (Long Short-Term Memory):** According to the result, LSTM delivered the weakest performance. It struggled to capture complex relationships, as indicated by higher MSE and MAPE. Interpretability methods reveal its heavily rely on technical indicator like ema, momentum, log momentum and time lagged features for stock price prediction. Specifically, LIME showed that these features significantly drive upward or downward trends, while MACD and volumes also plays a role in influencing predictions. Saliency maps confirmed the importance of momentum and moving averages (ma7, ma21) with market indices like NASDAQ and NYSE providing broader context. Stock-specific features had minimal influence.

2. **IMV-LSTM (Interpretable Multi-Variate LSTM):** The IMV LSTM is an enhanced version of classical LSTM by making it interpretable, especially in multi-variable series. The model introduces concept of variable-wise hidden states, which capture the dynamics of individual variables and differentiate their contribution to the prediction. Additionally, the mixture attention mechanism is used to create the generative process of target variable, making it easier to understand about each feature ‘contributions to the final prediction at different time steps. Particularly, with the high alpha (temporal attention) and beta (feature attention) values, features like 12ema, momentum, and ema are crucial both across time and play a critical role in stock price predictions, meaning they are consistently important throughout the model’s process.

1. **Quantum Long-Short Term Memory with linear embedding layer.**

Quantum LSTM (QLSTM) represents an extension of LSTM by incorporating Variational Quantum Circuits (VQC), which replace classical layers with quantum components. However, one challenge with QLSTM is the inefficient use of qubits due to the encoding and decoding processes in the VQC. When the number of input features exceeds the hidden state dimensions, some quantum information is wasted.

To address this, this work introduces a linear embedding layer which acts as feature compressor to transform input features into desired dimension which will ensure more efficient qubit usage. Specifically, a shared linear embedding layer is inserted before the VQC, and four separate linear layers are used after each VQC to remap the output, reducing information loss and improving learning capacity.

Another important aspect of QLSTM is about the VQC, and it generally consists of three components:   
- ***Encoding layer***: transform the classical input data into corresponding quantum states which will be further proceeded by the quantum circuits. This work explores the use of two different encoding methods such as Amplitude encoding and rotation-X then rotation-Y encoding.

-***Variational layer***: is key to accurate learning, as it handles qubit entanglement and rotation for complex information mapping. This work explores both CNOT and control-rotation-X gates to see how different quantum entanglement and rotation impact the model learning ability. Additionally, circuit block connectivity with cyclic interactions strengthens qubit connectivity, boosting entangling capacity while keeping training complexity low.

- ***Measurement layer***: convert quantum states back into classical data for post processing to perform prediction.

VQC is important to quantum-hybrid machine learning model, especially in this case it enables QLSTM to leverage quantum mechanics, such as superposition and entanglement, to represent and process complex patterns in time-series data more efficiently than classical methods.

**Performance:** According to table above, proposed model - enhanced QLSTM with the use of *amplitude embedding* significantly outperformed both classical LSTM and IMV LSTM. However, QLSTM’s performance depends on the quantum encoding technique used and the variational gates used. For instance, with amplitude encoding QLSTM tend to demonstrate closer prediction than the rotation-X then rotation-Y, as amplitude embedding is designate to work effectively in context of small number of quantum bit used and CRX also improves entanglement and expressibility and offering more parameters for broader quantum state exploration.

**Quantum advantageous:**

1. Faster Convergence: Because of characteristic of quantum parallelism that allows QLSTM to explore a broader solution space faster than classical models, leading to more efficient training, it displays in the loss plot with lower loss since the first epochs.

2. Higher Accuracy: Quantum superposition enables QLSTM to model complex dependencies between variables, making accurate predictions closer to true value for volatile stock data.

3. Less parameter: The classical LSTM requires a 1.6 times higher parameter compared to the enhanced QLST, this demonstrates that eQLSTM is an effective model with higher accuracy while maintaining significantly lesser number of parameter.

1. **Conclusion:**

This study demonstrates that enhanced QLSTM outperforms both classical and IMV LSTM, showcasing the potential of quantum computing in financial forecasting. QLSTM’s performance relies on the selection of quantum encoding and variational gates to determine the expressibility and qubit entanglement. Moreover, the application of LIME and Saliency maps to the classical LSTM provided insights into its decision-making process, revealing its dependence on general market trends and technical indicators. An extra implementation of IMV-LSTM makes it easier to interpret the LSTM and future understand about the problem context, and identifying key features at different time step, support to confirm about some features, and offering a more transparent look into time-series forecasting. In conclusion, the result display that quantum LSTM is a promising technology that worth for future research, particularly in stock price forecasting, further exploration of **quantum gates** and **encoding techniques** is warranted to fully unlock QLSTM’s potential. On another hand, interpretability remains a key factor in understanding and improving machine learning models

1. **References:**

[1] 1. Fischer, T., & Krauss, C. (2018). Deep learning with long short-term memory networks for financial market predictions, European Journal of Operational Research, 270 (2), 654-669.

[2] Y. Ansari, (2024) "Multi-Cluster Graph (MCG): A Novel Clustering-based Multi-Relation Graph Neural Networks for Stock Price Forecasting" in IEEE Access.

[3] A. Moghar and M. Hamiche, "Stock Market Prediction Using LSTM Recurrent Neural Network," Procedia Computer Science, vol. 170, pp. 1168-1173, 2020, doi: 10.1016/j.procs.2020.03.049.

[4] Chen, S. Y.-C., Yoo, S., & Fang, Y.-L. (2022). Quantum Long Short-Term Memory, International Conference on Acoustics, Speech, and Signal Processing, 2022.

[5] Y. Cao, X. Zhou, X. Fei, H. Zhao, W. Liu, and J. Zhao, “Linear- layer-enhanced quantum long short-term memory for carbon price forecasting,” Quantum Machine Intelligence, vol. 5, jul 2023

[6] M. T. Ribeiro, S. Singh, and C. Guestrin, “Why should I trust you?: Explaining the predictions of any classifier,” in Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, San Francisco, CA, USA, Aug. 2016, pp. 1135–1144.

[7] Guo, J., Chen, Z., & Cao, J. (2020). IMV-LSTM: Interpretable multi-variable LSTM for robust time series prediction, Proceedings of the 36th International Conference on Machine Learning, Long Beach, California, PMLR 97, 2019

[8] D. Dulal, "SoftServe\_QLSTM," GitHub repository, 2023. [Online]. Available: https://github.com/DikshantDulal/SoftServe\_QLSTM.