Jack Bosco
TFJ-DRL With Confusion

1 Introduction

Algorithmic trading is a dynamic perception and decision-making problem which has become increasingly popular in recent years. The task is to buy, sell or hold stocks on an exchange given a stream of quantitative statistics, called indicators, from noisy financial time series data (ticks) and the goal is to maximize profit while mitigating risk [1].

Autoregressive deep learning (DL) approaches for algorithmic trading involve predicting a stock's price at a future tick and trading based on that prediction. These approaches utilize deep learning models to perceive and represent noisy hyperdimensional data as latent vectors allowing for intricate pattern recognition [2]. While DL shows promise, the predictive task is problematic since the model stops short of making the trades itself, meaning an automated trading system requires additional prior knowledge and extra parameter tuning to make trading decisions [3]. Moreover, the predictability of the stock market has always been controversial [4] and trading based on predictions is just as dubious.

Another approach is reinforcement learning (RL) which maximizes an objective function by training an agent to choose from a set of actions at a given state [5]. In algorithmic trading, RL agents choose to either buy, sell or hold a stock based on the of historical stock data in order to maximize profit, minimize risk or reduce transaction costs. The state space is problematic, however, since direct RL approaches such as Q-Learning struggle generalizing high complexity state spaces such as historical stock data [6], [7].

Deep reinforcement learning (DRL) combines DL and RL approaches to perceive the environment and make continuous decisions, respectively. The DL model learns to represent the current window as latent vectors, thus reducing the state space and allowing the RL agent to converge [8]. For example, Time-feature Aware Jointly Deep Reinforcement Learning (TFJ-DRL) is a trading-specific DRL approach which adaptively weights relevant features with a gating component and adaptively weights ticks with a temporal attention mechanism [9] . While its robustness is promising, TFJ-DRL poses new challenges since the DL module must now balance between accurately capturing the context and presenting a familiar hidden state to the RL agent.

2 Motivation and Problem Statement

Algorithmic trading relies on the interplay between complex predictive and decision-making models, operating on highly noisy and volatile financial time-series data. While deep learning (DL) and reinforcement learning (RL) methods have made significant strides in modeling these environments, the inherent opacity of these models, especially in high-stakes financial contexts, presents a substantial risk. A lack of transparency in predictions or actions can lead to catastrophic financial losses if the model operates in unrecognized anomalous conditions.

Explainable AI (XAI) addresses these challenges by providing insights into model behaviors and predictions, fostering trust, and enabling human oversight. Explainable AI techniques seek to unveil the reasoning behind a model's predictions and provide a means for tracing those predictions back to the input [10]. In algorithmic trading, explainability signals enhance the human-interpretability of the model's familiarity with the current market state, thus promoting trust in the model's decisions. This dual purpose of providing real-time indicators of reliability and enabling model diagnostics underscores the role of XAI in advancing safer and more accountable algorithmic trading systems.

Anomaly detection serves as a practical XAI tool in algorithmic trading by flagging conditions where model performance may degrade. In TFJ-DRL, the RL agent relies on the latent environment vector produced by the DL model to make optimal trading decisions. Anomalies in this latent space indicate conditions the model has not seen during training, potentially leading to unreliable actions [11]. By integrating anomaly detection into the system, we gain a real-time explainability signal indicating the model's familiarity with the current market state. This not only improves transparency but also acts as a safeguard, enabling traders or

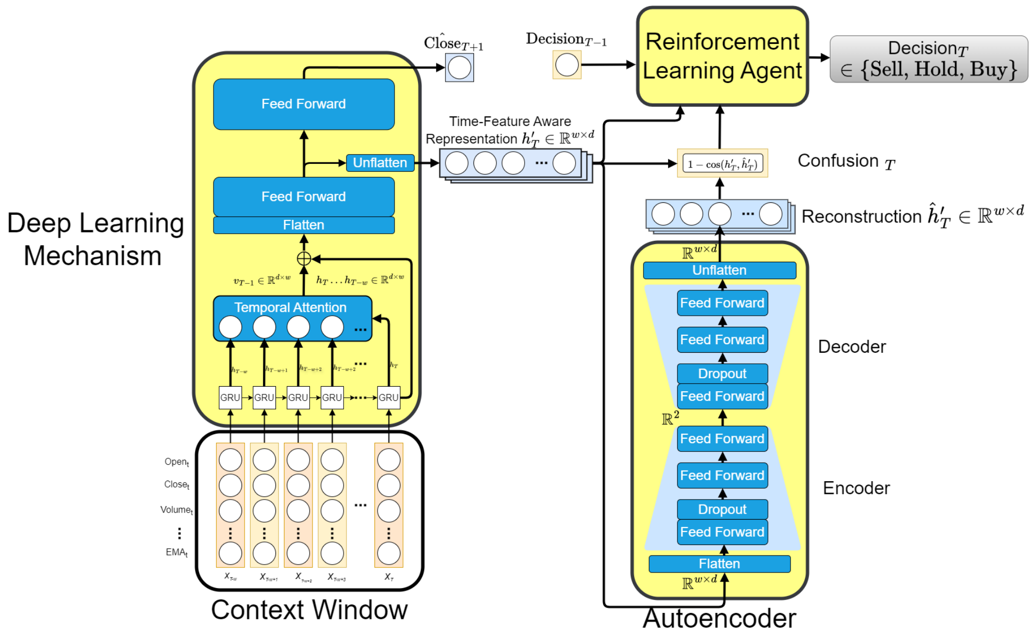

Figure 1: TFJ-DRL Architecture with autoencoder mechanism for confusion

automated systems to intervene when anomalies are detected, thus reducing the risk of significant financial losses.

Autoencoders are unsupervised learning units which compress information by passing it through a series of nonlinear transformations. They learn to express this compressed representation by reconstructing the original data from the compressed state. Autoencoders have seen widespread usage in anomaly detection due to their tendency to emphasize common patterns in their reconstruction while suppressing rare or anomalous patterns [12]–[14]. The compressed representation in the bottleneck layer captures the most essential features of the input data, allowing the decoder to reconstruct observations that align with the learned distribution. Anomalies, being outside this distribution, incur higher reconstruction loss because they deviate significantly from the patterns learned by the autoencoder. This property makes autoencoders particularly effective for unsupervised anomaly detection in multivariate time-series data, where patterns may be complex and non-linear.

While previous studies on XAI for deep reinforcement learning models (i.e. TFJ-DRL) are promising, the application of latent reconstruction loss as an XAI metric remains unstudied [10]. In this study we introduce confusion as an explainable signal gauging TFJ-DRL's familiarity with the current context. We modify the original architecture to include an autoencoder mechanism which takes the embedded environment state h_t and reconstructs it into h'_t. Confusion is then the reconstruction loss:

$$\text{Confusion}_t = 1 - \cos(h'_t, \hat{h}'_t) \tag{1}$$

Confusion quantifies the familiarity of the DL model's latent representation, signaling the RL agent's potential to operate suboptimally in anomalous states. This enhances the model's decision-making reliability by incorporating a self-diagnostic mechanism. Confusion thus bridges the gap between latent representation accuracy and actionable model transparency, aligning well with the dual goals of XAI and performance in algorithmic trading.

3 Methods

The original TFJ-DRL implementation lacked a public codebase, necessitating a manual reconstruction of the architecture and its experiments. Our implementation mirrors the original structure as described in Lei et al.

(2020) and integrates a Gated Recurrent Unit (GRU) network and Temporal Attention Mechanism (TAM) within the deep learning module. This module outputs a latent vector representation of historical data, which serves as the input environment state for the reinforcement learning (RL) agent. For the RL agent, we utilize a single Recurrent Neural Network (RNN) layer with $\mathbb{R}^{1\times d}$, activated using a tanh nonlinearity, allowing it to process the expanded environment state when confusion is incorporated.

For the confusion module, we explore two main implementations:

1. **Single Feed-Forward Autoencoder**: This forms the baseline for evaluating reconstruction loss.

2. **Advanced Architectures**: Future iterations include robust collaborative autoencoders (RCA) which minimize overfitting through mutual learning among multiple autoencoders [15], Recurrent Autoencoder Ensembles (RAE) effective for anomaly detection in multivariate time-series data [16], and Convolutional Recurrent Autoencoders (CRA) which have been shown to improve anomaly detection in the univariate time-series setting [17].

The integration of confusion into TFJ-DRL decision-making is tested using both ad-hoc and post-hoc approaches. The ad-hoc approach involves appending the confusion metric to the environment vector, thereby expanding the RL agent's input dimensionality to $R^{(w+1)\times d}$. In contrast, the post-hoc approach uses confusion as a threshold-based heuristic for modifying the RL agent's decisions in real-time, pulling positions when confusion exceeds predefined upper or lower quantiles. This threshold is established by analyzing training-time confusion distributions, as illustrated in Figure 2.

4 Preliminary Results

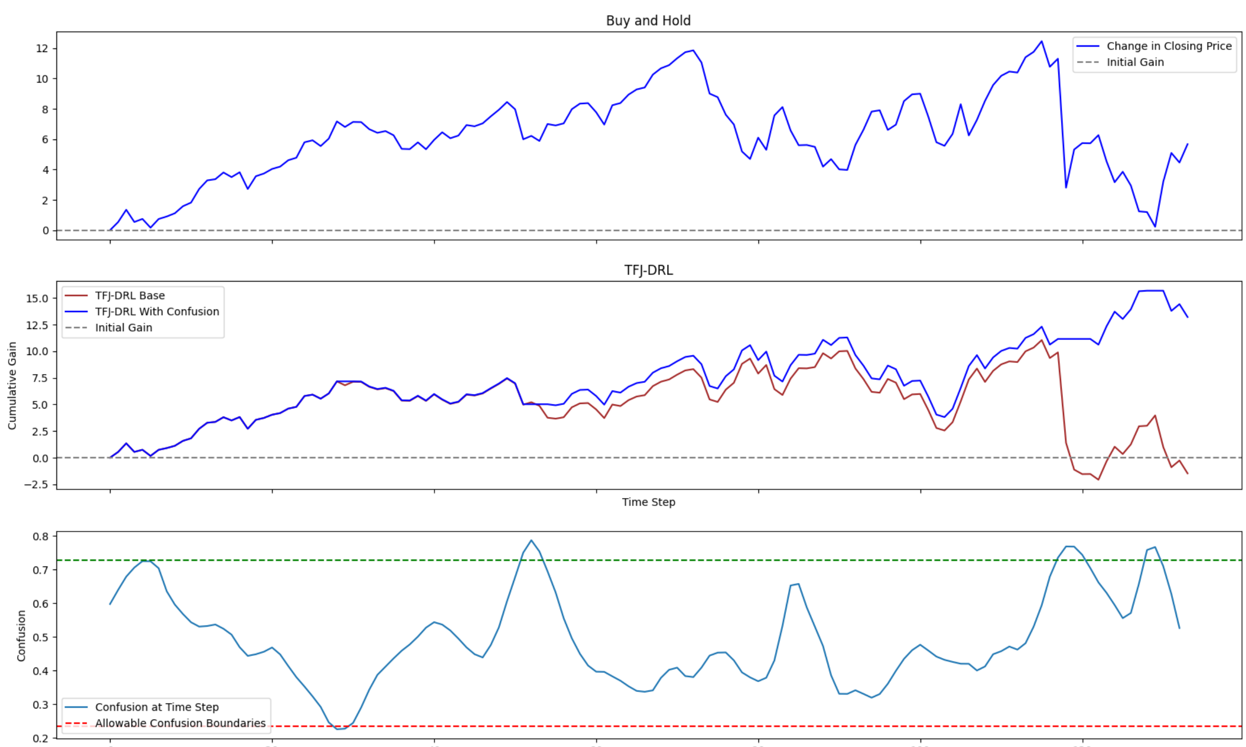

Figure 2: PnL across Buy and Hold, TFJ-DRL and post-hoc TFJ-DRL+Confusion strategies for trading stock on Cooper Companies Inc. between Jan. 1st 2018 and Jan. 1st 2019

Preliminary experiments indicate that confusion enhances both explainability and decision-making. Results for The Cooper Companies, Inc. (COO) between January 2018 and January 2019 are summarized in Figure 2. The Buy and Hold strategy shows moderate returns aligned with price fluctuations, while TFJ-DRL (Base) outperforms the market but suffers during volatile conditions. Notably, TFJ-DRL with Confusion (Ad-Hoc) maintains higher stability by avoiding catastrophic losses observed in volatile periods, as it exits positions when confusion surpasses acceptable bounds.

While not shown, the integration of confusion also improves the Sharpe Ratio by reducing risk exposure during uncertain market states. These results demonstrate the potential of confusion as a dual-purpose metric for improving transparency and robustness in algorithmic trading strategies.

References

[1] A. Sokolovsky and L. Arnaboldi, "A generic methodology for the statistically uniform & comparable evaluation of Automated Trading Platform components," *Expert Systems with Applications*, vol. 223, p. 119 836, Aug. 2023, ISSN: 0957-4174. DOI: `10.1016/j.eswa.2023.119836`. [Online]. Available: `https://www.sciencedirect.com/science/article/pii/S0957417423003378` (visited on 11/24/2024).

[2] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," en, *Nature*, vol. 521, no. 7553, pp. 436–444, May 2015, Publisher: Nature Publishing Group, ISSN: 1476-4687. DOI: `10.1038/nature14539`. [Online]. Available: `https://www.nature.com/articles/nature14539` (visited on 11/14/2024).

[3] E. Chong, C. Han, and F. C. Park, "Deep learning networks for stock market analysis and prediction: Methodology, data representations, and case studies," *Expert Systems with Applications*, vol. 83, pp. 187–205, Oct. 2017, ISSN: 0957-4174. DOI: `10.1016/j.eswa.2017.04.030`. [Online]. Available: `https://www.sciencedirect.com/science/article/pii/S0957417417302750` (visited on 11/14/2024).

[4] E. F. Fama, "Random Walks in Stock Market Prices," *Financial Analysts Journal*, vol. 21, no. 5, pp. 55–59, 1965, Publisher: CFA Institute, ISSN: 0015-198X. DOI: `10.2469/faj.v51.n1.1861`. [Online]. Available: `https://www.jstor.org/stable/4469865` (visited on 11/24/2024).

[5] S. Thrun and M. L. Littman, "Reinforcement Learning: An Introduction," English, *AI Magazine*, vol. 21, no. 1, pp. 103–103, Mar. 2000, Publisher: John Wiley & Sons, Inc., ISSN: 07384602. [Online]. Available: `https://go.gale.com/ps/i.do?p=AONE&sw=w&issn=07384602&v=2.1&it=r&id=GALE%7CA61573878&sid=googleScholar&linkaccess=abs` (visited on 11/15/2024).

[6] J. Moody and M. Saffell, "Learning to trade via direct reinforcement," *IEEE Transactions on Neural Networks*, vol. 12, no. 4, pp. 875–889, 2001. DOI: `10.1109/72.935097`.

[7] Y. Deng, F. Bao, Y. Kong, Z. Ren, and Q. Dai, "Deep Direct Reinforcement Learning for Financial Signal Representation and Trading," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 28, no. 3, pp. 653–664, 2017. DOI: `10.1109/TNNLS.2016.2522401`.

[8] V. Mnih, K. Kavukcuoglu, D. Silver, *et al.*, *Playing Atari with Deep Reinforcement Learning*, arXiv:1312.5602, Dec. 2013. DOI: `10.48550/arXiv.1312.5602`. [Online]. Available: `http://arxiv.org/abs/1312.5602` (visited on 11/15/2024).

[9] K. Lei, B. Zhang, Y. Li, M. Yang, and Y. Shen, "Time-driven feature-aware jointly deep reinforcement learning for financial signal representation and algorithmic trading," en, *Expert Systems with Applications*, vol. 140, p. 112 872, Feb. 2020, ISSN: 09574174. DOI: `10.1016/j.eswa.2019.112872`. [Online]. Available: `https://linkinghub.elsevier.com/retrieve/pii/S0957417419305822` (visited on 10/11/2024).

[10] F. Di Martino and F. Delmastro, "Explainable AI for clinical and remote health applications: A survey on tabular and time series data," en, *Artificial Intelligence Review*, vol. 56, no. 6, pp. 5261–5315, Jun. 2023, ISSN: 1573-7462. DOI: `10.1007/s10462-022-10304-3`. [Online]. Available: `https://doi.org/10.1007/s10462-022-10304-3` (visited on 10/11/2024).

[11] L. Wong, D. Liu, L. Berti-Equille, S. Alnegheimish, and K. Veeramachaneni, "AER: Auto-Encoder with Regression for Time Series Anomaly Detection," in *2022 IEEE International Conference on Big Data (Big Data)*, Dec. 2022, pp. 1152–1161. DOI: `10.1109/BigData55660.2022.10020857`. [Online]. Available: `https://ieeexplore.ieee.org/document/10020857` (visited on 11/18/2024).

[12] T. Kieu, B. Yang, C. Guo, *et al.*, "Robust and Explainable Autoencoders for Unsupervised Time Series Outlier Detection—Extended Version," 2022, Publisher: arXiv Version Number: 1. DOI: 10.48550/ARXIV.2204.03341. [Online]. Available: https://arxiv.org/abs/2204.03341 (visited on 11/18/2024).

[13] G. Li and J. J. Jung, "Deep learning for anomaly detection in multivariate time series: Approaches, applications, and challenges," en, *Information Fusion*, vol. 91, pp. 93–102, Mar. 2023, ISSN: 15662535. DOI: 10.1016/j.inffus.2022.10.008. [Online]. Available: https://linkinghub.elsevier.com/retrieve/pii/S1566253522001774 (visited on 10/11/2024).

[14] M. Sakurada and T. Yairi, "Anomaly Detection Using Autoencoders with Nonlinear Dimensionality Reduction," in *Proceedings of the MLSDA 2014 2nd Workshop on Machine Learning for Sensory Data Analysis*, ser. MLSDA'14, New York, NY, USA: Association for Computing Machinery, Dec. 2014, pp. 4–11, ISBN: 978-1-4503-3159-3. DOI: 10.1145/2689746.2689747. [Online]. Available: https://dl.acm.org/doi/10.1145/2689746.2689747 (visited on 11/18/2024).

[15] B. Liu, D. Wang, K. Lin, P.-N. Tan, and J. Zhou, "RCA: A Deep Collaborative Autoencoder Approach for Anomaly Detection," en, ISSN: 1045-0823, vol. 2, Aug. 2021, pp. 1505–1511. DOI: 10.24963/ijcai.2021/208. [Online]. Available: https://www.ijcai.org/proceedings/2021/208 (visited on 11/14/2024).

[16] T. Kieu, B. Yang, C. Guo, and C. S. Jensen, "Outlier Detection for Time Series with Recurrent Autoencoder Ensembles," en, in *Proceedings of the Twenty-Eighth International Joint Conference on Artificial Intelligence*, Macao, China: International Joint Conferences on Artificial Intelligence Organization, Aug. 2019, pp. 2725–2732, ISBN: 978-0-9992411-4-1. DOI: 10.24963/ijcai.2019/378. [Online]. Available: https://www.ijcai.org/proceedings/2019/378 (visited on 11/14/2024).

[17] C. Yin, S. Zhang, J. Wang, and N. N. Xiong, "Anomaly Detection Based on Convolutional Recurrent Autoencoder for IoT Time Series," *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, vol. 52, no. 1, pp. 112–122, Jan. 2022, Conference Name: IEEE Transactions on Systems, Man, and Cybernetics: Systems, ISSN: 2168-2232. DOI: 10.1109/TSMC.2020.2968516. [Online]. Available: https://ieeexplore.ieee.org/document/8986829 (visited on 11/14/2024).