Literature Review

Author: Ilya Grigorev, DS-01

Introduction

In this paper, I report major findings about standards and advances in designing time series forecasting models available in the literature. Additionally, two studies reporting gold as an asset are included to extend the knowledge about the field. The goal of the review is to propose a model for gold price short-term and long-term forecasting for the Russian market. The results demonstrate supremacy of the deep learning methods compared to statistical approaches and classical machine learning approaches. Ultimately, the CNN-LSTM model is suggested as the main predictor for short-term, and Transformer-based model is proposed to be a long-term forecasting tool.

Method

The research papers and review articles were searched using ScienceDirect to ensure credibility and relevance of material. The following search string was used for the machine learning models for stock prediction literature: “(Financial OR Stock) AND optimization AND (model OR method OR approach OR solution OR prediction) AND 'Risk Management'.” Additionally, the gold price analysis papers were searched using this query: “('precious metal') AND optimization AND (model OR method OR approach OR solution OR prediction) AND 'Risk Management').”

Results

“Two-stage stock portfolio optimization based on AI-powered price prediction and mean-CVaR models” – Wang et. al. [1]

1. Two types of research focus exist: predicting or optimizing portfolio.
2. Stock market data is often regarded as a time series, characterized by *volatility*, *nonlinearity*, *non-stationarity*, and *high noise.*
3. Filtering: traditional methods (moving average, exponential smoothing) and classical methods (Fourier transform, wavelet transform, empirical mode decomposition, kalman filter). The Kalman filter is not suitable for non-linear non-Gaussian noise. Savitzky–Golay (SG) filter.
4. Predicting stock prices methods: deep learning models + hyperparameter tuning using metaheuristic algorithms.
5. Portfolio optimization models:
    1. Mean-Variance (MV) model: emphasizes the maximum expected return while limiting the variance or reduces the variance given the minimum expected return. Limitation: the method does not directly minimize the volatility.

    2. Value-at-Risk (VaR) model: estimating the maximum loss within a specified range under normal market conditions. Limitations: no proper handling of tail losses and lack of convexity.
    3. Mean VaR models.
    4. Conditional VaR (CVaR) model: fixes the drawbacks of VaR.
    5. <u>Mean CVaR (mCVaR) models</u>.
    6. Lack of most approaches: no risk preferences (confidence intervals).
6. Stages:
    1. Smooth and filter stock price data.
    2. Tune and trane the deep learning model.
    3. Metrics of the predictions are obtained.
    4. mCVaR is used to determine the proportion of allocation for each stock in a portfolio.
7. Conclusions:
    1. Sparrow Search Algorithm (SSA) exhibited the highest accuracy and convergence speed among other NIC algorithms for hyperparameter tuning.
    2. The combination of filtering and hyperparameter tuning improved the results compared to the individual use of each.
    3. Predictive part is essential for decision making about the future, and allocation part is useful for specifying the amount of risk preferable to be during investment.
8. Limitations:
    1. Dataset size and diversity of stock markets.
    2. Stock prices as the single mean of input (other factors matter as well).
        1. e.g. transaction costs, liquidity constraints, and bid-ask spreads.
9. My observations:
    1. Apply ML approaches for portfolio allocation stage (recommender systems, train set of portfolios).
    2. NIC methods for portfolio allocation (max profit and min of risk).
10. Related paper:
    1. Applications of deep learning in stock market prediction by W. Jiang.

"Deep learning for algorithmic trading: A systematic review of predictive models and optimization strategies" – Bhuiyan et al. [2]

- DL models are suitable for the highly volatile and nonlinear nature of stock market data. Additionally, such models enable incorporating <u>a broader array of inputs</u>.
- Algorithmic trading involves computer methods to automatically trade on the financial market according to specific, pre-established criteria. Such strategies can execute trades at high rates or analyze complex data which is far beyond human capabilities.
- Key strategies: high-frequency trading (HFT), arbitrage, trend-following, mean reversion, market making, pairs trading.
- LSTMs: RNNs designed to overcome the vanishing gradient problem. Found effective for short periods and ineffective for long-term prediction.
- CNNs: architecture that treats time series as one-dimensional signal and filer it with layers. Useful to identify patterns (trend prediction), and for feature extraction.

- Autoencoders (feature transformation): effective feature extraction (denoising as well).
- Variational Autoencoders: probabilistic mapping from data to latent space. Allows to generate latent space data. Useful for denoising and generation of new instances in financial data analysis.
- GNNs are used in social graph representation of factors but has high computational complexity and design. Sensitive to sudden market fluctuations. Commonly accompanied with other models.
- Transformers: autoencoder models with "In the context of financial predictions, the encoder can analyze historical price movements and other relevant features, while the decoder can generate forecasts or trading signals." While Transformers excel at absolute price prediction, LSTMs are superior for predicting price differences.
- Reinforcement Learning: requires market interaction to learn optimal trading strategies.
- Challenges to be concerned about: noise filtration (VAEs, denoising), imputing missing data, overfit prevention (dropout, cross-validation, avoid over-reliance on historical data, regularization techniques: adversarial training and ensemble methods), interpretability inference (feature importance and attention mechanisms), and online learning (continuous update based on new market data).
- Current trends in research: transformers and hybrid models (e.g. CNN + LSTMs).
- Deploy: "Research into model compression techniques, such as pruning and quantization."

"Machine learning techniques via ensemble approaches in stock exchange index prediction: Systematic review and bibliometric analysis" – J. V. R. Ferro et al. [3]

- Ensemble models provide proper division of tasks for each type of approach, leveraging their capabilities in the final result.

"Applications of deep learning in stock market prediction: Recent progress" – W. Jiang [4]

- Efficient-Market Hypothesis: asset prices already reflect all available information.
- If the target is to predict the specific value of movement direction, then such a problem is classification.
- Prediction workflow:
  - Raw data: collection proper data as basis (intrinsic and extrinsic sources). Data types: market data (e.g. open/close price), text data (e.g. opinion → sentiment analysis), macroeconomics data (e.g. GDP), knowledge graph data (e.g. special kind of relationship between different companies or markets), image data (e.g. images for monitoring different situations), fundamental data (e.g. accounting data of companies), analytics data (e.g. handmade features from data exploration). Data length: short time period data (high overfit risk) and long time period data (harder to analyze).
  - Data processing: imputation (alignment of low and high frequency data), denoising (Wavelet transform and kNN-classifiers), feature extraction

(technical indicators for market data, NLP for text data, CNN for general feature extraction), dimensionality reduction (PCA and modifications, independent component analysis, autoencoder, restricted Boltzmann machine, empirical mode decomposition, sub-mode coordinate algorithm), feature selection (chi-squared method, maximum relevance and minimum redundancy, rough set attribute reduction, autocorrelation function, partial correlation function, analysis of variance, maximal information coefficient feature selection), feature scaling (normalization, standardization), data split (train-validation-test split, k-fold, modifications with rolling and successive training sets), data augmentation (usually, used for CNNs).
  - Prediction model: model types (DNN, CNN, RNN - standard models; GAN, transfer learning, RL - other models), baseline models (linear regression, autoregressive integrated moving average, generalized autoregressive conditional heteroskedasticity, logistics regression, SVM/SVR, kNN). About GAN: discriminative net D learns to distinguish whether a given data instance is real or not, and a generative net G learns to confuse D by generating high quality fake data).
  - Model evaluation: classification metrics (accuracy, precision, recall, sensitivity, specificity, F1 score, macro-average F-score, Matthews correlation coefficient, average AUC score), regression metrics (MAE, RMAE, normalized MSE, RMSE, relative RMSE, normalized RMSE, R^2), profit analysis (evaluation of profitability of the predicted-based trading, Sharpe Ratio), significance analysis (cross model comparison of difference in predictions - Kruskal-Wallis and Diebold-Mariano tests).
- Scraping sources: Yahoo! Finance, Tushare (Chinese stocks), IMF and World Bank (macroeconomics), Wikidata (relational data), Kaggle datasets (Two Sigma Financial Modelling Challenge).

"Adding precious metals to a risk avert Investor's portfolio – Is gold alone?" – D. Chattopadhyay [5]

- "Existing literature identifies gold as 'safe haven' since its returns remain **uncorrelated or negatively correlated** with that of market during crises."
- "the safe haven property of gold was found to be extremely short-lived and not to persist over longer market durations. Gold acts as a hedge against stocks in average market conditions, but the correlation is often weak or zero rather than strongly negative (Baur and Lucey, 2010; Coudert and Raymond, 2011). Bredin et al. (2015) found gold to act as a safe haven in short run but is less effective over long run. Beckmann et al. (2015) found gold's safe haven properties against stocks to be non-linear and time-varying that are strong during certain crises and weak during others. Safe haven effects are often not persistent. Similar results were obtained by Cheema et al. (2022) who found gold to be a safe haven during the GFC of 2007–08 crisis but lost this property during COVID-19 pandemic. Hood and Malik (2013) found the safe haven property to be valid only in high-volatility regimes and nowhere else. Pullen et al. (2014) emphasized that gold is not a consistently strong safe haven across all market downturns and may fail during certain crises."

“Does gold act as a hedge or a safe haven for stocks? A smooth transition approach,” – J. Beckmann [6]

- “In the era of globalization correlations among most types of assets increased dramatically, however gold is still known to be frequently uncorrelated with other assets.”
- Gold is known to be well-known to be durable, easily recognizable, storable, portable, divisible, and easily standardized (high liquidity).
- Gold has the potential to hedge against fluctuations of the exchange rate.
- “A hedge and a diversifier cannot shield a portfolio of exhibiting losses in times of extreme adverse market conditions, since both properties only work on average.”
- “but gold **does not act** as a hedge or a safe haven **for emerging economies** as well as for Australia, Canada, and Japan.”
- “We solely rely on monthly data in the following due to the fact that **both daily and weekly data appear to be too noisy** to capture the regimes we are interested in, and also due to the conjecture that there may be non-synchronicity issues, which are easier to neglect at a monthly frequency.”

Key Regional Outcomes

- An asset acts as a hedge if it is uncorrelated or negatively correlated with another asset or portfolio on average.
- An asset is regarded as a diversifier if it is positively (but not perfectly correlated) with another asset or portfolio on average.
- An asset is seen as a safe haven if it is uncorrelated or negatively correlated with another asset or portfolio in times of market stress or turmoil.
- **Strong Hedge Function**
  - Observed in: **EMU, Indonesia, Russia, Turkey**
  - Gold acts as a hedge against stock market downturns in these regions.
- **No Hedge Function**
  - Observed in: **China, Germany, World Index**
  - Gold does not provide a hedging benefit in these markets.
- **Strong Safe Haven Function (Negative ψ_2-coefficient)**
  - Observed in: **India, UK, World Index**
  - Gold serves as a safe haven during extreme market turmoil.
- **No Safe Haven Function**
  - Observed in: **EMU, Indonesia, Russia**
  - Gold fails to protect investors during severe market stress in these economies.

Discussion

The existing body of research underscores the complexity of financial time series modeling and portfolio optimization under uncertainty. Given the volatile, nonlinear, and noisy nature of stock market data, sophisticated filtering and predictive techniques are required to ensure reliable forecasts. A two-stage approach—first applying signal smoothing, followed by model-driven prediction and allocation—has demonstrated significant potential. In particular, the integration of signal processing methods, such as the Savitzky–Golay filter, with deep learning models optimized through metaheuristic algorithms like the Sparrow

Search Algorithm (SSA), leads to improved accuracy and convergence in forecasting. The second stage, involving portfolio allocation via mean-CVaR (mCVaR) optimization, effectively addresses risk preferences, particularly in capturing tail risks that conventional models (e.g., VaR, MV) fail to handle.

Deep learning methods—especially LSTM and CNN architectures—have been found effective in managing the inherent volatility and temporal dependencies of financial data. LSTMs are particularly suited for short-term price forecasting due to their ability to capture sequential dependencies over limited time spans, though they underperform in long-term scenarios. Conversely, CNNs facilitate pattern recognition and feature extraction from time series, and their use in hybrid models (e.g., CNN-LSTM) enables a layered understanding of temporal and spatial features. Transformer-based models, particularly in encoder-decoder configurations, have shown promising results in absolute price prediction, leveraging attention mechanisms for both interpretability and performance. Reinforcement learning and generative approaches (e.g., VAEs, GANs) offer additional benefits in strategy learning and data denoising, though their application remains more computationally demanding and sensitive to model specification.

Gold's role as a financial asset in forecasting and portfolio design exhibits temporal and regional heterogeneity. While it is often considered a hedge or safe haven, this function is neither persistent nor universal. The findings from Beckmann and Chattopadhyay highlight gold's safe haven property as short-lived and contingent on high-volatility regimes or specific regional market conditions. Notably, gold acts as a hedge in markets like Russia but lacks safe haven characteristics during stress periods. This non-stationarity and context dependence suggest that short-term forecasting of gold prices should consider high-frequency volatility regimes and regime-switching behaviors, while long-term modeling must accommodate the diminishing and inconsistent nature of its hedging capability.

In light of the discussed results, we propose two models. For short-term gold price forecasting, a CNN-LSTM hybrid model filtered through a denoising autoencoder is appropriate, as it leverages CNN's ability for feature extraction and LSTM's sequential modeling for immediate temporal dependencies. Hyperparameter tuning should be performed using SSA or other nature inspired algorithms to enhance predictive robustness. For long-term forecasting of the Russian stock market, where structural market shifts and long-term macroeconomic variables play a substantial role, a transformer-based model integrated with monthly data is recommended. This model should incorporate external data (e.g., macroeconomic indicators and relational knowledge).

Conclusion

To conclude, the proposal consists of a short-term forecasting model for gold using a CNN-LSTM-DAE architecture with SSA tuning, and a long-term transformer-based framework for the Russian stock market incorporating external data and mCVaR-based allocation. These models reflect the insights from existing literature: the importance of noise filtration, the specificity of model design to task horizon, and the non-uniform hedging properties of assets such as gold.

References

[1] Wang et al., “Two-stage stock portfolio optimization based on AI-powered price prediction and mean-CVaR models,” *Expert Syst. Appl.*, vol. 255, 2024. [Online] Available: https://www.sciencedirect.com/science/article/pii/S0957417424014222

[2] Bhuiyan et al., “Deep learning for algorithmic trading: A systematic review of predictive models and optimization strategies,” *Array*, vol. 26, 2025. [Online] Available: https://www.sciencedirect.com/science/article/pii/S2590005625000177

[3] J. V. R. Ferro et al., “Machine learning techniques via ensemble approaches in stock exchange index prediction: Systematic review and bibliometric analysis,” *Appl. Soft Comput.*, vol. 167, 2025. [Online] Available: https://www.sciencedirect.com/science/article/pii/S1568494624011335

[4] W. Jiang, “Applications of deep learning in stock market prediction: Recent progress,” *Expert Syst. Appl.*, vol. 184, 2021. [Online] Available: https://www.sciencedirect.com/science/article/pii/S0957417421009441

[5] D. Chattopadhyay, “Adding precious metals to a risk avert Investor's portfolio – Is gold alone?”, *Resources Police*, vol. 106, 2025. [Online] Available: https://www.sciencedirect.com/science/article/pii/S0301420725001692

[6] J. Beckmann, “Does gold act as a hedge or a safe haven for stocks? A smooth transition approach,” *Economic Modelling*, vol. 48, 2015. [Online] Available: https://www.sciencedirect.com/science/article/pii/S0264999314004015?ref=pdf_download&fr=RR-2&rr=951c08c87e5be2dd