Deep learning for decision making and the optimization of socially responsible investments and portfolio

Nhi N.Y. Vo\textsuperscript{a}, Xuezhong He\textsuperscript{b}, Shaowu Liu\textsuperscript{a}, Guandong Xu\textsuperscript{a,\ast}

\textsuperscript{a} Advanced Analytics Institute, University of Technology Sydney, 2-12 Blackfriars Street, Chippendale NSW 2008, Australia
\textsuperscript{b} Business School, University of Technology Sydney, 14-28 Ultimo Rd, Ultimo NSW 2007, Australia

\textbf{ARTICLE INFO}

\textbf{Keywords:}
Socially responsible investment
Portfolio optimization
Multivariate analytics
Deep reinforcement learning
Decision support systems

\textbf{ABSTRACT}

A socially responsible investment portfolio takes into consideration the environmental, social and governance aspects of companies. It has become an emerging topic for both financial investors and researchers recently. Traditional investment and portfolio theories, which are used for the optimization of financial investment portfolios, are inadequate for decision-making and the construction of an optimized socially responsible investment portfolio. In response to this problem, we introduced a Deep Responsible Investment Portfolio (DRIP) model that contains a Multivariate Bidirectional Long Short-Term Memory neural network, to predict stock returns for the construction of a socially responsible investment portfolio. The deep reinforcement learning technique was adapted to retrain neural networks and rebalance the portfolio periodically. Our empirical data revealed that the DRIP framework could achieve competitive financial performance and better social impact compared to traditional portfolio models, sustainable indexes and funds.

\section{1. Introduction}

Traditionally, investors have focused on the investment returns by actively looking at the financial reports to find the best performing stocks. With the recent mindset change towards sensitive topics like global warming or refugees, investors are becoming concerned with other aspects of companies rather than just earnings. They are shifting their investment towards companies which are actively doing good things for the environment, contributing to the society and operating with transparency. According to the 2018 Biennial Report On US Sustainable, Responsible and Impact Investing Trends\textsuperscript{[51]}, Socially Responsible Investment (SRI) assets accounted for $12 trillion out of $47 trillion in total assets under professional management in the United States in 2018, representing a sharp increase of 38\% since 2016.

Conventional investment and portfolio theory focuses on financial performance, i.e., the returns and risks of the portfolio\textsuperscript{[55]}. Direct application of the theory might not be suitable for SRI because it focuses more on non-monetary objectives\textsuperscript{[6]}. Therefore, socially responsible investors need a modified version of the modern portfolio theory that can serve their purpose better\textsuperscript{[38]}. Besides, SRI investors currently have to read Corporate Social Responsibility (CSR) reports to find good companies to invest in, which is time-consuming and difficult. The lack of effective quantitative approaches for SRI makes it more difficult for not only professional investors, but also the vast majority of lay investors. Therefore, this research will provide an easy and automated way of doing such investments in an ethical manner, which greatly benefits their decision-making and secures the optimal investment returns. This is one of the main motivational purposes for this research.

Recently, the Global Reporting Initiatives (GRI) and the United Nation Sustainable Development Goals (SDGs) have provided standardized metrics and frameworks for companies to disclose more information regarding their sustainability practices\textsuperscript{[14]}. For example, Environmental, Social and Governance (ESG) metrics of companies have been derived from reports and news articles (e.g. CSR reports, news articles, carbon disclosure project ratings), evaluating the company in different prospects (e.g. air emissions and waste management, employee health and safety control, board transparency and diversity) including their controversies (e.g. involvement in adult entertainment or gambling). These metrics have been consolidated into the combined ESG ratings (see Fig. 1). The availability of ESG ratings has led to an emerging research topic in SRI portfolio.

With the availability of ESG metrics, quantitative methods can now be applied effectively to address the SRI portfolio construction problem. Current data mining approaches in this research field face a number of challenges. The first challenge is the accuracy of multivariate time series predictions. Stock return forecasts have been extensively studied with various quantitative finance and machine learning models\textsuperscript{[24,}
Most of these works have been focused on univariate time series predictions because it is expected that multivariate data would contain too much noise for the neural network to perform well [20, 33]. However, stock movements in the financial market are highly correlated; thus a multivariate model can learn these deep insights better than a combination of univariate networks. Following recent advances in neural networks research, especially Long Short-Term Memory networks (LSTM) [26], the application of deep learning in the predictive investment field has become an alternative approach to the traditional financial model. In this paper, we propose a novel Multivariate BiLSTM neural network to predict multiple time series for stock returns.

The second challenge faced by current approaches in SRI is the application of multi-objective portfolio construction. Existing portfolio optimization methods are evolving around the standard Mean-Variance (MV) Portfolio [32], which focuses on maximizing returns and minimizing risks. To incorporate corporate responsibility performance into our optimization problem, we introduced a modified MV model for SRI portfolio construction by integrating ESG ratings.

The third challenge for SRI portfolio is building a model that can adapt to market movements. As SRI in particular, and financial investment in general, are sensitive to market volatility, model parameters should be tuned up periodically to achieve both financial performance and ESG rating objectives. By adopting reinforcement learning techniques, we introduced a Deep Responsible Investment Portfolio model to retrain the prediction model and rebalance the portfolios effectively and autonomously.

An advantage of our proposed approach, which incorporates a multivariate BiLSTM neural network and MV-ESG, is that the framework can be generalized and extended to other scenarios with a similar multivariate prediction and multi-objective optimization problem. The developed deep reinforcement learning framework could also accommodate different neural networks and AI algorithms to tackle other types of complex and highly intercorrelated problems.

The main contributions of our research are:

- A novel deep responsible investment portfolio framework to integrate deep neural networks, multi-objective optimization, and reinforcement learning. The framework could be applied to other similar contexts of multivariate predictive analytics.
- A novel DRIP model that can forecast the returns quarterly and yearly on investment instead of just daily, which is a more realistic scenario for investors. The model has been fully tested and deployed on real-life datasets containing 100 stocks over a period of 30 years.
- The first report (to the best of our knowledge) leverages deep learning and incorporates ESG ratings into a portfolio optimization model.

This paper is organized as follows. In Section 1, we introduce the current background of socially responsible investment portfolios and their limitations, as well as give our motivation for our proposed reinforcement learning solution. In Section 2, we review the literature on socially responsible investment research, and the application of deep learning by focusing on recent methodologies that are closely related to this paper. In Section 3, we introduce the technical details of our DRIP model and in Section 4, we present our empirical studies that applied our novel algorithms to real-life financial datasets to construct socially responsible investment portfolios, evaluate the performance against some baseline models and explore the potential for further and related research. We provide our conclusions in Section 5.

2. Literature Review

2.1. Socially Responsible Investment

The optimization of financial portfolios has been researched extensively. Many approaches have been developed to build decision support systems for stock trading. This includes standard mathematical finance modeling, e.g. Mean-Variance (MV) [32], AutoRegressive Moving Average (ARMA) and Generalized AutoRegressive Conditional Heteroskedasticity (GARCH) models [18], text mining of financial news [35] and social media [25].

However, limited research has been carried out on socially responsible investment. Although the socially responsible investment was proposed in the 1980s [22], it only became a topic of interest for academia and industry in the past decade [15]. During this time, research has correlated ESG ratings with the financial performance of companies [16, 23] or socially responsible funds [2, 28, 34]. The availability of environmental, social and governance (ESG) ratings has enabled more research and application in this area in academia [54] and industry [51].

Many sustainability funds have offered portfolios with certain values to attract investors to SRI. In management funds [46], there has been an increasing demand from sustainably conscious investors to have more SRI options [37]. Multiple sustainable indexes and funds have been constructed based on areas of investor interests (e.g. water treatment, clean tech, renewable energy, gender equality and diversity). The literature has shown that companies or sustainable funds with higher ESG ratings can outperform the lower ones financially in long-term investments [19].

The literature of qualitative research in SRI has focused on reviewing the performance of companies [4] and socially responsible
indexes or funds [49], and not on a data-driven approach to incorporate sustainability into an investment system. Some of the research has criticized the current stock screening process of SRI funds [53] and has proposed that the full integration of ESG ratings would be more beneficial [1]. These findings underpin the main motivation for our research to develop a framework with full integration of ESG ratings. Our paper contributes to the current knowledge of the application of deep learning for the prediction of stock returns and ESG-based SRI portfolio optimization.

2.2. Deep Learning for Stock Returns Forecasting

Researchers have undertaken extensive studies to solve the time series forecasting problem of stock returns using deep learning [9, 13, 33]. Many have suggested that different types of Recurrent Neural Networks (RNN) outperform traditional financial time series models in different markets [3, 7, 44]. RNN contains feedback loops in its recurrent layer, which enables the storage of information in the “memory cell” over time. However, it does not perform well when the learning requires long-term temporal dependencies.

Long Short-Term Memory (LSTM) is a special type of RNN that has been proven to be effective in text mining to predict stock returns [30]. LSTM contains “memory cells” that are able to retain information for longer periods of time [26]. Consequently, LSTM often performs better in sequential data and financial time series predictions compared with RNN [27, 36], particularly in the SRI context where investors are concerned more about long-term returns rather than the volatility of the short-term market.

Researchers have also compared the performance of different RNN architectures like LSTM and Gated Recurrent Unit (GRU) networks [40]. Others have suggested that Bi-directional LSTM (BiLSTM) might be a better option in a similar sequence prediction problem [8]. While the LSTM and GRU, with the unidirectional flow of information, might be adequate in most sequence prediction problems, the BiLSTM model reads the data one more time backward [42] which helps improve prediction accuracy, particularly in forecasting sequential data like financial time series.

Recently it was suggested that back-testing results could have given rise to false positives due to the normalization of testing data and prediction of the next time step only [43]. The next-time step prediction is only suitable for high-frequency trading strategies using intra-day data, such as foreign exchange markets. In SRI, investors are more interested in long-term returns on investment. Conversely, research has been conducted on the long-term prediction for financial indexes with 1-year and 2-year time gaps, suggesting that long-term forecasting is possible for stock returns [17].

Our paper contributes to the current deep learning methodologies through the design of a novel BiLSTM neural network that predicts a long-term multivariate time series. To avoid false positive results, the financial returns data is not normalized and the model predicts multiple steps ahead. By constructing the baseline models using different types of LSTM networks as undertaken previously, we evaluate the prediction accuracy of the LSTM networks in the forecasting of SRI stock returns.

2.3. Portfolio Optimization

Few socially responsible investment models have been developed and proposed that utilize ESG ratings [52], [21], for example, suggested a modification to the standard portfolio selection model with ESG scores. They utilized the Mean-Variance Stochastic Goal Programming (MV-SGP) model with a statistical approach for ESG screening on stocks based on scores and controversy risk. However, they did not consider predictive analytics; they only used past returns and volatility to test their hypotheses. Furthermore, they did not validate their models with real financial data.

Multiple optimization functions are available, including the Expectation Maximization (EM) algorithm [12], quasi-Newton [5] or Powell methods [39]. However, most of them are not multi-objective or allow the special limit conditions that are required in a complex context like in socially responsible investment. The Sequential Least SQares Programming (SLSQP) method proposed by [29], for example, can be used to minimize a function of various variables with a different combination of bounds, equality and inequality constraints. However, its greedy behavior leads to a skewed distribution for the weights of stocks in the portfolio. This is not an optimum choice for investors who are worried about non-diversified portfolios with extreme exposure risk.

We have developed a financial model to construct a socially responsible investment portfolio that incorporates the Mean-Variance portfolio theory and ESG ratings (MV-ESG). Our model is not based on the ESG screening approach. Instead, it filters and leverages the ESG ratings in a multi-objective optimization function based on the SLSQP method. It also considers both past and predicted the future performance of stocks in a portfolio selection. This is one of the first mathematical models for constructing a socially responsible investment portfolio that achieves both better ESG ratings and competitive financial performance.

3. Methodology

Our DRIP framework consists of three main components: a multivariate BiLSTM neural network to predict stock returns quarterly and yearly; these predicted values are then combined with ESG ratings in our MVP-ESG model for portfolio construction; reinforcement learning techniques are then leveraged to automatically retrain the prediction models and re-balance our MVP-ESG portfolios after each period. The full reinforcement learning DRIP framework is as shown in Fig. 4.

3.1. Multivariate BiLSTM Neural Networks

Standard feature engineering often includes a normalization step, which transforms the data range to [0,1]. This common approach can help to improve the prediction accuracy of the neural networks. However, in the time series model, this approach implicitly tells the trained model the movement range of future stock prices, which makes out-of-bag testing results unrealistic accurate. We processed the input data for our neural networks in a different approach. In our DRIP model, we did not normalize data but instead fed the stock returns directly into the neural networks. We also trained the model to predict values with a longer time gap instead of a next period prediction, which is a more suitable scenario for stock investors in real-life trading.

Let $p_i(t)$ be the price at time $t = (1,...,T)$ for stock $i = (1,...,N)$. $\Delta t$ was the time gap $(1 < \Delta t < T)$. The return $r_i(t)$ for stock $i$ at time $t$ was $r_i(t) = p_i(t) - p_i(t - \Delta t)$. In the DRIP model, we used the sliding window technique to perform a rolling forecast. Let $\delta t$ be the sliding window size. The train features matrix $X_i(t)$ and return vector $Y_i(t)$ for stock $i$ at time $t$ were:

$$X_i(t) = \begin{bmatrix}
\eta(t - \Delta t) & \eta(t - \Delta t + 1) & \cdots & \eta(t - T) \\
\eta(t - \Delta t + 1) & \eta(t - \Delta t + 1 + 1) & \cdots & \eta(t - T + 1) \\
\vdots & \vdots & \ddots & \vdots \\
\eta(t) & \eta(t - \Delta t) & \cdots & \eta(t)
\end{bmatrix}$$

(1)

$$Y = \begin{bmatrix}
\eta(t - \Delta t) \\
\eta(t - \Delta t + 1) \\
\vdots \\
\eta(t + \Delta t)
\end{bmatrix}$$

(2)

As suggested by [36], LSTM networks would outperform other neural networks in solving similar problems due to its information persistence characteristic. We considered three types of LSTM neural networks:
• **LSTM**, initially proposed by [26], is a special kind of RNN, which is capable of learning long-term dependencies. For each input vector $x_t$ at time step $t$, LSTM network uses multiple gating functions: the input gate $i_t$, forget gate $f_t$, and output gate $o_t$, together with a memory cell $C_t$ to preserve long-term information and keeps track of its flow. The forget gate $f_t$ and input gate $i_t$ generated at each time step $t$ are defined as follows:

$$f_t = \sigma(W_f[h_{t-1}, x_t] + b_f)$$  \hspace{1cm} (3)

$$i_t = \sigma(W_i[h_{t-1}, x_t] + b_i)$$  \hspace{1cm} (4)

In the next step, a tanh layer generates a new memory cell $C_t$. LSTM then updates the old memory cell $C_{t-1}$ and generates the output gate $o_t$ and hidden state $h_t$:

$$i_t = \tanh(W_i[h_{t-1}, x_t] + b_i)$$  \hspace{1cm} (5)

$$C_t = f_t \odot C_{t-1} + i_t \odot \tilde{C}_t$$  \hspace{1cm} (6)

$$o_t = \sigma(W_o[h_{t-1}, x_t] + b_o)$$  \hspace{1cm} (7)

$$h_t = o_t \odot \tanh(C_t)$$  \hspace{1cm} (8)

where $\sigma$ is the sigmoid function and $\odot$ is the element-wise multiplication. $W$ is the weight matrix and $b$ is the bias vector to be learned by the LSTM at each specific gate.

**BiLSTM** is a variation of the bidirectional RNN, firstly introduced by [42]. It concatenates a forward and backward unidirectional LSTM on the stock return time series $\text{Combined}(h_t) = [h^f_t, h^b_t]$. Unidirectional LSTM only preserves long-term information of the past, while BiLSTM can preserve information from both past and future by using the combined two hidden states $\text{Combined}(h_t)$.

**GRU** is a more recent alteration of LSTM, suggested by [10]. It concatenates both the forget gate $f_t$ and input gate $i_t$ into a single update gate $z_t$, and merges the cell state $C_t$ and hidden state $h_t$. The architecture of GRU is simpler than the standard LSTM one. The hidden state $h_t$ generated at each time step $t$ is defined as follows:

$$z_t = \sigma(W_z[h_{t-1}, x_t])$$  \hspace{1cm} (9)

$$r_t = \sigma(W_r[h_{t-1}, x_t])$$  \hspace{1cm} (10)

$$\tilde{h}_t = \tanh(W_c[h_{t-1}, x_t])$$  \hspace{1cm} (11)

$$h_t = (1 - z_t) \odot h_{t-1} + z_t \odot \tilde{h}_t$$  \hspace{1cm} (12)

By simplifying the architecture of the LSTM, GRU may learn the data at the combined gate. However, this single update gate might not learn some hidden information effectively. Hence, the performance of GRU networks may be less effective in forecasting long-term time series (Fig. 2).

For our DRIP model, we designed a special type of BiLSTM to perform multivariate time series prediction. The data input shape for the multivariate BiLSTM neural network was in the form of a three-di-dimensional matrix with sizes $(T - \Delta t - \delta t, \delta t, N)$, where $N$ was the number of stocks in total, $\delta t$ was the sliding window size, and $\Delta t$ was the prediction time gap (see Fig. 4).

We also replicated neural network models with LSTM and GRU networks as in [40] to predict returns for every single stock in the portfolio. We constructed the neural networks with recurrent layers using the Adam optimizer from the “Keras” package [11]. This network also contained a dense layer and a final output layer with the “linear” activation function to predict the stock returns $r_t(t + \Delta t)$ in $\Delta t$ periods of time.

### 3.2. MV-ESG Model

The Mean Variance portfolio (MV) of [32] has always been the standard portfolio selection model. Its mathematical principle is constructed by two main components: maximizing the return $r_p$ and minimizing the risk $\sigma_p$. The output of this optimization process is the efficient frontier, which is a set of investment portfolios with a greater return than any other with the same or less risk, and a lower risk than any other with the same or greater return. For illustration, the efficient frontier is plotted in Fig. 3 with the risk on the horizontal axis and the return on the vertical axis.

The optimal portfolio based on the efficient frontier is commonly known as the maximum Sharpe portfolio (MAX-S), where the portfolio has a maximum Sharpe ratio calculated as $S_p = (r_p - r_f)/\sigma_p$. For the MAX-S portfolio, considering the risk free rate $r_f$ (normally the return on bond investment or the bank interest rate), it minimizes the negative Sharpe Ratio [45]:

$$\min(-S_p) = \min\left(-\frac{r_p - r_f}{\sigma_p}\right)$$  \hspace{1cm} (13)

$$r_p = \sum_{i=1}^{N} w_i r_i$$  \hspace{1cm} (14)

$$\sigma_p = \sqrt{\sum_{i=1}^{N} \sum_{j=1}^{N} w_i w_j \sigma_{ij}}$$  \hspace{1cm} (15)

where $w_i$ and $w_j$ are the weights of stock $i$ and $j$, with the boundary limit $w_i w_j \in [0,1]$, and $\sigma_{ij}$ is the covariance matrix of the two stock $i$ and $j$ in the portfolio. The initial weight of each stock in the computation will be equally allocated according to the total number of stocks $N$ in the portfolio, $w_i(0) = w(0) = 1/N$.

In our MV-ESG model, we built a multi-objective algorithm based on the SLSQP method [29] with three objectives: maximizing returns, minimizing volatility and maximizing ESG ratings. This algorithm minimized: $\min_{\text{weqg}} || - G ||$ with $G$ being a three-dimensional matrix of constraints of the three objectives and weqg being the ESG weights subject to boundary limits inferred from the companies’ ESG ratings.

For comparison, we constructed a maximum ESG portfolio (MAX-ESG) for investors with low risk averse to compare with the standard MAX-S portfolio. In MAX-ESG, we minimized the negative Sharpe Ratio with the portfolio ESG ratings ($ESG_p$) as a new variable of the objective function.

$$\min(-\tilde{S}_p) = \min\left(-\frac{ESG_p r_p - r_f}{\sigma_p}\right)$$  \hspace{1cm} (16)

$$ESG_p = \sum_{i=1}^{N} \frac{\text{weqg}_i ESG_i + ESG_j}{2}$$  \hspace{1cm} (17)

where $ESG_i$ was the combined ESG ratings of company $i$ in the past year, $ESG_j$ was the combined ESG ratings at the current prediction year, and weqg was the ESG weight of stock $i$ in the portfolio.

In the traditional MV model, $r_p$ and $\sigma_p$ are the past returns $r_t$ and volatility $\sigma_t$, which is often called ex-post MV. In recent years, researchers and investors have been using the expected returns $\bar{r}$ and volatility $\bar{\sigma}$. This approach called ex-ante MV is more suitable for predictive analytics in real-world financial trading. In our MV-ESG model, we combined both ex-post MV and ex-ante MV for portfolio selection and replaced the standard weight boundary with our ESG ones calculated based on the combined ESG ratings for each stock. Our MV-ESG model was computed using:
where $r_i$ and $\hat{r}_i$ were the ex-post and ex-ante returns, $\sigma_{ij}$ and $\hat{\sigma}_{ij}$ were the ex-post and ex-ante covariance matrix of the two stock $i$ and $j$ in the portfolio. $w_{esg_i}$ and $w_{esg_j}$ were the ESG weight of stock $i$ and $j$ in the portfolio, with the boundary limit $w_{esg_i} \in [0,1]$ for the company with the highest combined ESG score, then gradually decreasing to $w_{esg N} \in [0,0]$ for the company with the lowest combined ESG score. This means the allocation of the company “$N$” in the portfolio was zero, indicating no investment in that company. The initial weight of each stock in the computation was not to be equally allocated but assigned according to the ESG ratings.

### 3.3. DRIP Model with Reinforcement Learning

We combined the multivariate BiLSTM neural networks and the MV-ESG models into a single integrated reinforcement learning model named Deep Responsible Investment Portfolio (DRIP). Starting with a set of agent states $S$ and a set of possible portfolio allocation sets $A$, we had the probability of the DRIP model selecting the specific portfolio allocation (the “action”) $a$ when in state $s$ at time step $t$ as:

$$p(a|s) = \frac{\text{Pr}(a = s_t = a)}{\text{Pr}(s_t)}$$

We defined a simple state-value function $V_s^* \hat{\gamma}$ as the expected reward starting with the state $s_0 = s$ and $R_t$ denoting the reward function calculated as the sum of future discounted rewards:

$$V_s^* = E[Re] = E \left[ \sum_{t=0}^{\infty} \gamma^t R_t | s_0 = s \right]$$

where $\gamma \in [0,1]$ was the discount rate. $\hat{\gamma}$ was the ESG-adjusted Sharpe Ratio [45], and $\text{MSE}$ was the mean squared error of the prediction model. The DRIP model found a set of portfolio allocation to maximize...
the expected return. After each time gap $\Delta t$, DRIP retrained the prediction model with new stock prices data and then, together with the portfolio performance and stock weights from the previous period of time, constructed a new portfolio with updated allocation weights. The reinforcement learning was repeated on a predefined period basis, to improve the prediction model accuracy and the performance of the portfolio over time. The design of DRIP enabled its self-learning with the least human involvement as possible. The reinforcement learning model is shown in Fig. 4.

4. Empirical Experiment

We designed experiments to test our proposed model in two parts: 1) DRIP model forecasts for quarterly and yearly returns of multivariate stock timeseries during the three year period from 2016 to 2018; and 2) Socially Responsible Portfolios optimization using the predicted returns and reinforcement learning DRIP model framework.

4.1. Datasets

Currently, there are various ESG rating services available [41], many of which offer a subscription fee for data access which limits its availability to the public. In 2018 however, Yahoo Finance made some of the ESG ratings obtained from Sustainalytics [48] available publicly. In this research, we utilized Yahoo Finance to obtain both financial stock prices and public ESG rating datasets in our reinforcement learning DRIP framework.

We downloaded the daily closing prices of all stocks in the Standard and Poor 500 list (S&P500) from the past 30 years from 31 December 1988 to 31 December 2018. In order to ensure a sufficient number of data points, we removed all the stocks which did not have a market price on 31 December 1988, which left us with 262 companies. SRI investors do not invest in companies with low ESG ratings; therefore we used a simple stock screening process to remove these unwanted stocks. From the shortlisted 262 stocks, we selected the top 100 companies with the highest combined ESG ratings according to Sustainalytics [48] available publicly. In 2018 however, Yahoo Finance made some of the ESG ratings obtained from Sustainalytics [48] available publicly. In this research, we utilized Yahoo Finance to obtain both financial stock prices and public ESG rating datasets in our reinforcement learning DRIP framework.

We downloaded the historical window of size $T - \Delta t - 30$ from 31 December 2018 onwards from Yahoo Finance. We then used the rolling forecast approach to further split the data in the testing period into validation and test sets. For quarterly return prediction, we used “Q4/2015” and “Q1/2016” as the validation and test set for the first period. We then moved to the next quarter period until “Q3/2018” and “Q4/2018” as validation and test sets in the last period. We applied the same data splitting process to the yearly return prediction dataset.

4.2. Evaluation Metrics

To test our DRIP model, we used the Mean Absolute Error (MAE) and the Root Mean Squared Error (RMSE) as the evaluation metrics for the absolute value prediction:

\[
\text{MAE} = \frac{1}{N} \sum_{i=1}^{N} |\bar{r}_i - r_i|
\]  
\[
\text{RMSE} = \sqrt{\frac{1}{N} \sum_{i=1}^{N} (\bar{r}_i - r_i)^2}
\]  

where $N = 100$ was the total number of stocks in the portfolio and $\bar{r}_i$ and $r_i$ were the predicted and actual return of stock $i$ for that period.

We also converted the predicted value to a binary label to evaluate the performance of uprend or downturn forecast using the prediction accuracy metric and the Area Under the Curve (AUC) scores with the Receiver Operating Characteristic (ROC) curve. The lower MAE and RMSE together with the higher prediction accuracy and AUC scores indicate the better performance of the prediction model. Our baseline models for comparison are the LSTM and GRU neural networks as in [40] and a univariate standard BiLSTM model (Uni).

To evaluate the performance of our socially responsible portfolios using MV-ESG model, we compared its Sharpe Ratio against those of the standard MV portfolios and the reported financial performance from similar sustainable indexes and funds. The Sharpe Ratio was defined as $S = (r_p - r_f)/\sigma_p$, where $r_p$ was the portfolio annualized return, $\sigma_p$ was the portfolio annualized volatility, and $r_f = 2\%$ was the nominal risk-free rate. A better performing portfolio had a higher Sharpe Ratio, which yielded higher returns if the risks were similar or a lower risk if the returns were the same.

The sustainable indexes for comparison were: Dow Jones Sustainability World Index (DJSI World), Dow Jones Sustainability World Diversified Select Index (DJSI WD), and S&P500 ESG Factor Weighted Index (S&P500 ESG). All indexes data were obtained on the 31 December 2018 from S&P Dow Jones Indices, a division of S&P Global. The sustainable Exchange Traded Funds (ETF) with their symbol codes in the brackets were: iShares Global Clean Energy ETF (ICLN), Invesco Solar ETF (TAN), iShares MSCI USA ESG Select ETF (SUSA) and Workplace Equality Portfolio (EQLT). All funds data were obtained on 31 December 2018 from Morningstar.
4.3. Prediction Model Results

First of all, we tested the performance of DRIP model on the prediction of quarterly returns. The hyperparameters in our neural networks were set as: the number of units in the deep learning layers equaled to 100, batch size equaled to 1, the loss was the mean squared error and random seed equaled 0. We used Adam optimizer with learning rate $\eta = 0.001$, $\beta_1 = 0.9$, $\beta_2 = 0.999$, fuzz factor $\epsilon = 1e^{-7}$, and decay equaled to 0. We also set the number of epochs equaled to 10 with a checkpoint after each epoch and only saved the best model for prediction.

Our experiment setup was as follows: $\Delta t = 63$, $\delta t = 63$ and the time gap was set to 63 representing the total number of trading days in a quarter. This meant that the model predicted the prices and returns three months ahead in time. After each period, the model was retrained and validated with the out-of-bag three-month data and predicted the next return in 63 days. The testing data for each quarter of each year from 2016 to 2018 were referred to as “Q1”, “Q2”, “Q3” and “Q4” respectively.

We then tested the performance on the prediction of yearly returns with $\Delta t = 252$, $\delta t = 252$ representing the 252 trading days in a typical year. The other setup was the same as in the quarterly returns prediction model. The empirical results in Table 1 showed the performance evaluation for the quarterly and yearly return prediction models using the multivariate financial time series as input. The reported RMSE and AUC Scores were averages for 100 stocks in each time period.

Our DRIP models, which used multivariate financial returns as the input significantly outperformed the other baseline models for most prediction periods in term of MAE and RMSE. We can conclude that the prediction model using multivariate financial returns and BiLSTM neural networks in our design was a better solution for this predictive analytic problem. Focusing on the trend prediction accuracy, except for the slightly worse results in “Q2/16” and “Q3/18”, our DRIP models that used BiLSTM achieved higher prediction accuracy and AUC scores regardless of the time periods or of the quarterly or yearly returns. These results demonstrated the effectiveness of our approach, that the reinforcement learning had successfully captured the underlying hidden information in the inter-correlated multivariate series and improved itself over time.

The value predictions of quarterly returns generally had lower MAE and RMSE than they yearly forecast. This result was expected as the time gap was smaller; hence, it was easier to forecast the absolute stock return values. The ROC curves in Fig. 5 showed a performance lift in the quarterly return prediction model compared to other baselines for the entire 3-year testing period. Conversely, the trend prediction was more accurate in yearly return models, which proved that the reinforcement

Table 1

<table>
<thead>
<tr>
<th>Mean Absolute Error (MAE)</th>
<th>Q1/16</th>
<th>Q2/16</th>
<th>Q3/16</th>
<th>Q4/16</th>
<th>Q1/17</th>
<th>Q2/17</th>
<th>Q3/17</th>
<th>Q4/17</th>
<th>Q1/18</th>
<th>Q2/18</th>
<th>Q3/18</th>
<th>Q4/18</th>
</tr>
</thead>
<tbody>
<tr>
<td>DRIP</td>
<td>0.0547</td>
<td>0.0880</td>
<td>0.0578</td>
<td>0.0814</td>
<td>0.0600</td>
<td>0.0551</td>
<td>0.0555</td>
<td>0.0601</td>
<td>0.0692</td>
<td>0.0667</td>
<td>0.0771</td>
<td>0.0948</td>
</tr>
<tr>
<td>Uni</td>
<td>0.0656</td>
<td>0.0930</td>
<td>0.0678</td>
<td>0.0999</td>
<td>0.0633</td>
<td>0.0704</td>
<td>0.0679</td>
<td>0.0683</td>
<td>0.0714</td>
<td>0.0882</td>
<td>0.0854</td>
<td>0.1096</td>
</tr>
<tr>
<td>LSTM</td>
<td>0.0771</td>
<td>0.0960</td>
<td>0.0734</td>
<td>0.1043</td>
<td>0.0677</td>
<td>0.0767</td>
<td>0.0707</td>
<td>0.0718</td>
<td>0.0650</td>
<td>0.0986</td>
<td>0.0788</td>
<td>0.1404</td>
</tr>
<tr>
<td>GRU</td>
<td>0.0652</td>
<td>0.0950</td>
<td>0.0723</td>
<td>0.1139</td>
<td>0.0623</td>
<td>0.0793</td>
<td>0.0776</td>
<td>0.0729</td>
<td>0.0801</td>
<td>0.0993</td>
<td>0.1003</td>
<td>0.0936</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Root Mean Squared Error (RSME)</th>
<th>Q1/16</th>
<th>Q2/16</th>
<th>Q3/16</th>
<th>Q4/16</th>
<th>Q1/17</th>
<th>Q2/17</th>
<th>Q3/17</th>
<th>Q4/17</th>
<th>Q1/18</th>
<th>Q2/18</th>
<th>Q3/18</th>
<th>Q4/18</th>
</tr>
</thead>
<tbody>
<tr>
<td>DRIP</td>
<td>0.0672</td>
<td>0.1080</td>
<td>0.0754</td>
<td>0.1209</td>
<td>0.0768</td>
<td>0.0741</td>
<td>0.0750</td>
<td>0.0790</td>
<td>0.0858</td>
<td>0.0843</td>
<td>0.1202</td>
<td>0.1165</td>
</tr>
<tr>
<td>Uni</td>
<td>0.0822</td>
<td>0.1189</td>
<td>0.0863</td>
<td>0.1534</td>
<td>0.0806</td>
<td>0.0933</td>
<td>0.0887</td>
<td>0.0939</td>
<td>0.0926</td>
<td>0.1120</td>
<td>0.1333</td>
<td>0.1402</td>
</tr>
<tr>
<td>LSTM</td>
<td>0.0957</td>
<td>0.1225</td>
<td>0.0911</td>
<td>0.1637</td>
<td>0.0847</td>
<td>0.0991</td>
<td>0.0944</td>
<td>0.1029</td>
<td>0.0849</td>
<td>0.1233</td>
<td>0.1238</td>
<td>0.1739</td>
</tr>
<tr>
<td>GRU</td>
<td>0.0812</td>
<td>0.1254</td>
<td>0.0914</td>
<td>0.1708</td>
<td>0.0801</td>
<td>0.1041</td>
<td>0.0952</td>
<td>0.0981</td>
<td>0.1057</td>
<td>0.1238</td>
<td>0.1536</td>
<td>0.1191</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Area Under the Curve (AUC) Scores</th>
<th>Q1/16</th>
<th>Q2/16</th>
<th>Q3/16</th>
<th>Q4/16</th>
<th>Q1/17</th>
<th>Q2/17</th>
<th>Q3/17</th>
<th>Q4/17</th>
<th>Q1/18</th>
<th>Q2/18</th>
<th>Q3/18</th>
<th>Q4/18</th>
</tr>
</thead>
<tbody>
<tr>
<td>DRIP</td>
<td>0.8392</td>
<td>0.5809</td>
<td>0.8307</td>
<td>0.8286</td>
<td>0.8495</td>
<td>0.7659</td>
<td>0.7521</td>
<td>0.8170</td>
<td>0.8165</td>
<td>0.8045</td>
<td>0.7024</td>
<td>0.9047</td>
</tr>
<tr>
<td>Uni</td>
<td>0.7443</td>
<td>0.5803</td>
<td>0.7876</td>
<td>0.6603</td>
<td>0.8522</td>
<td>0.6052</td>
<td>0.6297</td>
<td>0.7486</td>
<td>0.7441</td>
<td>0.6477</td>
<td>0.6952</td>
<td>0.8502</td>
</tr>
<tr>
<td>LSTM</td>
<td>0.7115</td>
<td>0.5946</td>
<td>0.7719</td>
<td>0.6320</td>
<td>0.8790</td>
<td>0.5069</td>
<td>0.5739</td>
<td>0.7001</td>
<td>0.7516</td>
<td>0.5692</td>
<td>0.7549</td>
<td>0.7278</td>
</tr>
<tr>
<td>GRU</td>
<td>0.6821</td>
<td>0.5954</td>
<td>0.7522</td>
<td>0.5202</td>
<td>0.8281</td>
<td>0.5427</td>
<td>0.5632</td>
<td>0.7287</td>
<td>0.6642</td>
<td>0.5692</td>
<td>0.6884</td>
<td>0.8820</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Trend Prediction Accuracy (%)</th>
<th>Q1/16</th>
<th>Q2/16</th>
<th>Q3/16</th>
<th>Q4/16</th>
<th>Q1/17</th>
<th>Q2/17</th>
<th>Q3/17</th>
<th>Q4/17</th>
<th>Q1/18</th>
<th>Q2/18</th>
<th>Q3/18</th>
<th>Q4/18</th>
</tr>
</thead>
<tbody>
<tr>
<td>DRIP</td>
<td>77%</td>
<td>50%</td>
<td>74%</td>
<td>76%</td>
<td>80%</td>
<td>66%</td>
<td>70%</td>
<td>77%</td>
<td>72%</td>
<td>76%</td>
<td>73%</td>
<td>85%</td>
</tr>
<tr>
<td>Uni</td>
<td>69%</td>
<td>50%</td>
<td>73%</td>
<td>62%</td>
<td>77%</td>
<td>58%</td>
<td>65%</td>
<td>75%</td>
<td>65%</td>
<td>64%</td>
<td>70%</td>
<td>76%</td>
</tr>
<tr>
<td>LSTM</td>
<td>64%</td>
<td>53%</td>
<td>76%</td>
<td>61%</td>
<td>75%</td>
<td>52%</td>
<td>66%</td>
<td>75%</td>
<td>66%</td>
<td>57%</td>
<td>74%</td>
<td>60%</td>
</tr>
<tr>
<td>GRU</td>
<td>67%</td>
<td>57%</td>
<td>70%</td>
<td>48%</td>
<td>75%</td>
<td>56%</td>
<td>58%</td>
<td>73%</td>
<td>58%</td>
<td>58%</td>
<td>62%</td>
<td>84%</td>
</tr>
</tbody>
</table>

Fig. 5. ROC curves.
learning model could filter out the market noise in short-term price changes. Overall, our DRIP model effectively and accurately predicted the annual returns in all three years and the quarterly returns in 10 out of 12 testing periods. It showed that our prediction model was not overfitted to a certain dataset period, and it could be generalized for similar applications.

4.4. Robustness Test

To test the robustness of our model, we first benchmarked the prediction model using different combinations of the neural network hyperparameters. We split the dataset into train, validation and test sets with the ratio 8:1:1. The hyperparameter sets were: number of units in the deep learning layers was in [100,200,300], batch size (BS) was in [1,10,20] and learning rate (LR) was in [0.0001,0.001,0.01] accordingly. The MAE and prediction accuracy results of both validation and test set are presented in Table 2.

Table 2 showed that different hyperparameter sets could result in varied MAE and prediction accuracy. The gap between validation and test results are not significantly large, which indicates that our model was not overfitted. Our setting to generate the best results in the test set was: number of units equaled 100, batch size equaled 1 and learning rate equaled 0.01. In our rolling forecast and reinforcement learning model, the hyperparameters could be automatically tuned using grid search after each period.

We then used this set of hyperparameters to test the prediction model on three different datasets with 50, 100 and 200 randomly selected stocks (denoted as "Random50", "Random100", "Random200"). We also split these datasets into train, validation and test sets with the ratio 8:1:1. The results of this experiment are presented in Table 3.

Table 3 showed that our model still achieved a good prediction accuracy in randomly selected stock datasets. It is worth noticing that the MAE and the prediction accuracy are not worsened for the larger dataset but varied due to the randomness of stock selection. These results indicated that our prediction model is robust and generalizable with different data sizes.

4.5. Portfolio Optimization Model Results

We used the predicted returns from the DRIP model as input for our MV-ESG model to construct socially responsible investment portfolios. We constructed the MAX-ESG portfolios using predicted returns. The nominal risk free rate was set to 2%, $r_f = 0.02$. After obtaining the stock allocation in each portfolio, we calculated the actual annualized returns and volatility using real stock prices for that period. The annualized returns, volatility, Sharpe Ratio and ESG Score given in Table 4 were averaged for the entire year, for each year in the testing period.

The results showed that our MAX-ESG portfolios had consistently higher ESG ratings (3 to 5 points above). Even though the MAX-S portfolios had better financial returns in 2016 and 2018, they also showed a relatively higher volatility level. Conversely, our MAX-ESG portfolios still achieved great financial returns with lower risk. The Sharpe Ratios of the MAX-ESG portfolios were higher than those of the MAX-S ones for 2017 and 2018. In 2017, the MAX-ESG portfolio achieved a better financial return 50.78% at a lower risk level 19.19%, compared with 47.76% return at 19.37% volatility in the MAX-S portfolio. These findings showed that achieving a socially responsible investment portfolio, with higher ESG ratings, and without the sacrifice of a large financial return, was achievable with our MV-ESG model.

We compared the performance of our final MAX-ESG portfolio with reported financial returns in 2018 obtained from similar sustainable indexes and funds. The results in Table 5 show that our portfolio outperformed other indexes and funds in terms of financial performance and achieved the Sharpe Ratio of 2.0634. Our portfolio had the best 3-year annualized return of 35.28%. Particularly in 2018, all indexes and funds had negative returns because many large stocks were in the downturn. Our MAX-ESG portfolio was still able to achieve a positive return. This was mainly because the portfolio constructed was based on the maximization of Sharpe Ratio in the MV-ESG model, which optimally selects stocks with higher returns.

Our model's 3-year annualized volatility was in third place with 16.13%. This higher level of risk aligned with common investment knowledge on diversification [47]. Because the indexes often consist of a larger number of stocks, they generally had lower risks. However, the level diversification of our SRI portfolio was sufficient for individual investors.

Our best MAX-ESG portfolio, for example, consisted of 7 stocks in 2016, 18 stocks in 2017 and 12 stocks in 2018 with the allocation as shown in Fig. 6. The model could be enhanced to construct a more diversified portfolio for sustainable investment funds with further constraints on weights.

Since all these indexes and funds published different types of sustainability metrics, we could not directly compare our portfolio ESG.
ratings to their benchmarks. We also could not report the net returns on investment due to the lacking of fund fees and tax calculation. In general, these results showed the effectiveness of our DRIP framework, not only for the optimization of socially responsible investment portfolios but also for financial stock investments in general.

4.6. Reinforcement Learning Test

Our reinforcement learning DRIP framework could be used to construct socially responsible portfolios with higher ESG ratings that still achieved competitive financial returns. Our DRIP model could predict multiple time steps ahead, which is an important feature for stock investors. Furthermore, the model significantly outperformed the univariate networks in both prediction accuracy and training speed with the same epoch size in terms of both prediction accuracy and training speed. In our experiments, it took one hour to perform reinforcement learning with the multivariate BiLSTM: a combination of 100 univariate neural networks with 10 epochs typically takes 100-times more training duration compared to our approach. This finding could lead to a better computationally efficient approach because the multivariate BiLSTM

| Table 5 |
| Benchmarking MAX-ESG portfolio with sustainable indexes and funds in 2018. |
| | Period returns | 3-year Annualized |
| | 2016 | 2017 | 2018 | Return | Volatility | Sharpe ratio |
| MAX-ESG | 28.47% | 50.78% | 26.60% | 35.28% | 16.13% | 2.0634 |
| S&P500 | 11.29% | 23.28% | −3.25% | 10.41% | 10.88% | 0.7727 |
| S&P500 ESG | 14.52% | 21.24% | −8.44% | 9.11% | 11.76% | 0.6043 |
| DJSI World | 8.23% | 27.98% | −8.03% | 9.39% | 11.52% | 0.6418 |
| DJSI WD | 10.71% | 24.00% | 9.54% | 14.75% | 10.48% | 1.2166 |
| ICLN | −16.91% | 21.48% | −9.02% | −1.48% | 17.11% | −0.2036 |
| TAN | −43.23% | 54.39% | −25.66% | −4.83% | 23.00% | −0.2971 |
| SUSA | 12.15% | 22.53% | −5.65% | 9.68% | 11.42% | 0.6722 |
| EQLT | 13.93% | 21.33% | 9.22% | 14.83% | 11.80% | 1.0870 |

Fig. 6. MAX-ESG portfolio allocation (labels are trade symbols of companies).
takes N times less in total training duration.

We also tested the performance of reinforcement learning by comparing the results to those without prediction model retraining and portfolio rebalancing (Non-RL). In the “Non-RL” framework, we still used the Multivariate BiLSTM networks for prediction model and the portfolio rebalancing (Non-RL). In the “Non-RL” framework, we were working with multivariate time series, retraining models and rebalancing portfolios were proven to be essential. Therefore, our reinforcement learning approach within the DRIP system was suitable for this time-sensitive data analytics problem.

### 4.7. Discussion of Research

Overall, our research demonstrated a promising trend in applying deep learning techniques for the selection of socially responsible investment portfolios. With the current progress in artificial intelligence, we believe it will bring further breakthroughs in socially responsible investment research. Our research will not only contribute directly to current literature in various disciplines but also translate into benefits for responsible investors, funds or indexes in markets. In the AI research field, our prediction model with a BiLSTM network could serve as a baseline for further research of long-term stock return forecasting using neural networks. In this research, we only used a single type of neural networks and structured data (stock prices and ESG ratings) as input. Studying the different variations and combination of the deep neural networks, as well as incorporating unstructured data (news, company reports or social media content) with text mining approaches in SRI, was beyond the scope of this paper. However, our framework was designed with the flexibility to adopt different data mining approaches, neural networks or optimization algorithms in our future research.

Please be aware that by focusing on policies and rewards, our system might fail under extreme situations, e.g. a financial crisis. Our model’s performance and applicability are subject to the hypothesis of stable company performances and normal financial market scenario. To safeguard investments in such cases, we would need additional failsafe measures when applying our model in practice.

From the financial research aspect, the MV-ESG model was one of the first to combine ESG ratings with a math finance model. Further research on how to incorporate this with other quantitative finance models such as GARCH [18] would be relevant to both SRI scholars and investors. As many researchers are working on similar approaches for oil price forecasting [31], we believe a further investigation into this direction would be beneficial for SRI researchers. For simplicity purpose, we did not take into account income tax rates, inflation rates, trading fees, and other financial fund management costs. Further calculation of these fees would help the model implementation in the real-world investment scene. Moreover, ESG ratings are not the only metrics to measure corporate social responsibility. The integration of hundreds of ESG sub-categorial ratings (e.g. greenhouse gas emissions or community support) could improve the model significantly. In our future research, we could study the potential of using deep learning approaches for a personalized stock recommendation system in the SRI context.

### 5. Conclusions

Socially responsible investment is an emerging research topic with potential for long-term social impact. In this research, we proposed the DRIP model, which leveraged deep learning techniques to predict financial returns and construct a socially responsible investment portfolio. Validated with real-world data, our DRIP model, with a multivariate time series model, was able to accurately predict the stock returns three months ahead of time. It is possible that our framework could be generalized to build decision-support systems for similar multivariate prediction problems.

The socially responsible portfolios that we constructed using our novel MV-ESG model and reinforcement learning achieved much higher ESG ratings and a competitive financial performance overall compared with standard MV portfolio models and similar sustainable indexes and funds. With this rising trend in socially responsible investment, financial capital will diverge into good companies that contribute to a cleaner environment and a better society. This research also highlights a new direction for the use of more advanced deep learning approaches for quantitative finance research.

### Acknowledgment

This research is partly supported by Australian Research Council Linkage Project Scheme under Grant LP170100891. We also thank University of Technology Sydney for supporting us with the infrastructures and computing power for empirical studies. We would also like to thank the editor and anonymous reviewers for their constructive and insightful feedbacks.

### References


### Table 6

Reinforcement learning test results.

<table>
<thead>
<tr>
<th>Prediction model</th>
<th>MAE</th>
<th>RSME</th>
<th>AUC</th>
<th>Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>DRIP</td>
<td>0.0807</td>
<td>0.1117</td>
<td>0.9354</td>
<td>88%</td>
</tr>
<tr>
<td>Non-RL</td>
<td>0.1998</td>
<td>0.1499</td>
<td>0.5515</td>
<td>55%</td>
</tr>
<tr>
<td>MAX-ESG portfolio</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DRIP</td>
<td>Return</td>
<td>Volatility</td>
<td>Sharpe ratio</td>
<td>ESG score</td>
</tr>
<tr>
<td></td>
<td>35.28%</td>
<td>16.13%</td>
<td>2.0634</td>
<td>73</td>
</tr>
<tr>
<td>Non-RL</td>
<td>5.30%</td>
<td>14.00%</td>
<td>0.2357</td>
<td>68</td>
</tr>
</tbody>
</table>