The Economics and Finance Letters
2024 Vol. 11, No. 1, pp. 57-79
ISSNI: 2312-430X
ISSN(p): 2312-6310
DOI: 10.18488/29.v11i1.3667

Neural network and machine learning use cases: Indian bond market predictions

Jesmine Mary Antony[1]
Sundaram Natarajan[2+]

[1,2]Department of Commerce, School of Social Sciences and Languages, Vellore Institute of Technology, Tamil Nadu, 632014, India.
[1]Email: jesminemary.a2021@vitstudent.ac.in
[2]Email: nsundaram@vit.ac.in

(+ Corresponding author)

Article History
Received: 31 October 2023
Revised: 19 January 2024
Accepted: 2 February 2024
Published: 4 March 2024

Keywords
Artificial intelligence
Auto-regressive integrated moving average
Bond spread strategy
Bond yield forecasting
Deep long short-term memory recurrent neural network
Linear regression
Yield curve analysis.

JEL Classification:
C58; C45; C49; C53.

ABSTRACT

This study examines machine learning techniques to investigate how artificial intelligence (AI) affects predicting future trends in the bond market. The bond market offers a global perspective on capital costs for a business by establishing the fair value of the bond issue, which is based on multiple factors. The asset price market, which has employed machine learning (ML) and deep learning (DL) techniques to address the primary forecasting difficulty, surprisingly plays a significant role in predicting future bond market returns. As an outcome, if this gap can be forecast, it can act as the bond market's data-driven long-term direction and yield additional profits. Daily security-specific data for the 10-to-3-year Indian Treasury Bond (ITB) was gathered from 2013 to 2022 and is available in the global government bonds database. The researchers looked at how well the auto-regressive integrated moving average (ARIMA), linear regression, and deep recurrent neural network-long short-term memory (DLSTM) models could predict bond yields and returns in future bond markets. The empirical results demonstrate that the DLSTM models most fairly predict the price of government bonds over both the short and longer horizons when compared to ARIMA and linear regression.

Contribution/Originality: This paper focuses on the neural network effect of back-testing to confirm real bond interest rate volatility, different maturities, and credit rating risks to explain yield spread forecasting. The study helps business forecasters and institutional investors make well-informed decisions on the acquisition and disposal of financial assets.

1. INTRODUCTION

1.1. Financial Market

According to literature such as Kumar (2014) the financial market can be defined as any exchange where buyers and sellers exchange assets, including stocks, debt instruments, currencies, and derivatives. According toMadura (2015) the bond market, also known as the credit market or the debt market, is a financial market where participants can issue new debt in the primary market or buy and sell debt securities in the secondary market. Though notes, invoices, and other financial instruments for both public and private spending can also be used, bonding is the most popular type of this. Bond markets are the third basic component of financial systems, according to Thumrongvit, Kim, and Pyun (2013) and are becoming more and more important for the growth of the financial sector. Paul and Reddy (2022) elucidated how the financial sector grew enormously as it developed, satisfying the demands of the public and private domains. A portion of this rise could come from the bond market.

Thus, the researchers examined the Indian bond market in the direct financial markets. Given their generally lower volatility compared to stocks, bonds are frequently recommended as at least a portion of a diversified portfolio. While financial markets are usually connected to stock markets, Mišura (2016) noted that they are occasionally perceived as trade-only securities. Kim (2021) noted that financial time series display non-stationary characteristics, making financial market analysis a challenging task. A stronger force in the larger financial market results from the acknowledged successful compliance of machine learning (ML) and deep learning (DL) forecasting models in the yield price prediction of the bond market. Therefore, the literature's studies recommended employing analytical machine learning models to analyze the bond market in the financial markets. Because analytical modeling can immediately result in significant gains, a lot of research has been done in this area (Ederington, Guan, & Yang, 2015; Huynh, Hille, & Nasir, 2020; Kim, Kim, & Jung, 2021).

1.2. Indian Treasury/Government Bond

The size and liquidity of Indian government bonds make them an important part of the bond market. By contrasting them with other bonds, government bonds are often used to assess credit risk. Because bond valuation and interest rates (or yields) have an inverse connection, the bond market is often used to forecast changes in interest rates or the shape of the yield curve, the measure of "cost of funding." In light of this topic, government bonds are usually regarded as low-risk financial instruments. Jobst, Kunzel, Mills, and Sy (2008) investigated sovereign debt, another term for loans from the federal government. Governments issue government bonds to raise funds for particular projects or ongoing expenses. In India, a government bond is essentially an investor's guarantee that the issuer will repay principal and interest on bonds having a face value equivalent to the investor's investment. According to a 2008 study by Shankar and Bose (2008), government bonds are long-term investment instruments with maturities ranging from five to forty years. They are categorized under the general category of government securities (G-Sec) in India. Both the Indian federal government and the state governments can issue government bonds. Government bonds called State Development Loans (SDLs) are issues by state governments. Coupons, the term for interest rates on government bonds, are issued on a semi-annual basis and can be fixed or variable. The Indian government (GOI) regularly issues bonds on the market with fixed coupon rates. Debt securities are traded on the Bombay Stock Exchange (BSE) and the National Stock Exchange of India (NSE). Businesses as well as government agencies are present in the Indian bond market. Government bonds, Special GOI Securities, Treasury Bills (T-Bills), and SDLs are the main participants in the Indian government bond market. The Reserve Bank of India (RBI) is the government's fiscal agent and provides a range of government securities, according to Asher (2007). Longer-term government bonds, also known as dated securities, and shorter-term Treasury bills fall under this category. As stated by Parameswaran (2022) dated securities come in a variety of forms, such as nominal and inflation-indexed bonds, put/call options, coupons with fixed and variable interest rates, and specialty bonds (such as oil bonds). Fixed-rate bonds issued by the Indian government can be subject to interest rate risk, which occurs when interest rates rise and investors hold onto fixed-rate bonds that pay less than market value. Besides, inflation is a measure of price increases across the country, and not many bonds can keep up with it. Although the investor only receives a real return of 0.5% when market prices increase by 1.5%, a fixed-rate government bond in this case returns 2% every year. According to Zhou, Chen, Li, Zhang, and Zheng (2022) AI-driven networks are crucial to bond yield and are widely used in all security markets, including economic forecasts. In many different disciplines and applications today, machine learning (ML) and deep learning (DL)are becoming compatible. Using ML and DL, the data maximum might be lowered in a couple of seconds, helping the organization identify and address securities risks. According to Trippi and Turban (1992) the AI-powered system is excellent at learning from mistakes made in the past and avoiding making the same ones again. The security market's data collection, storage, and computation capabilities have continuously increased as a result of AI applications in securities trades. Therefore, this work's inclusive research question is:

Research Question: How well can AI inspired machine learning models predict the bond yield spread in the Indian market?

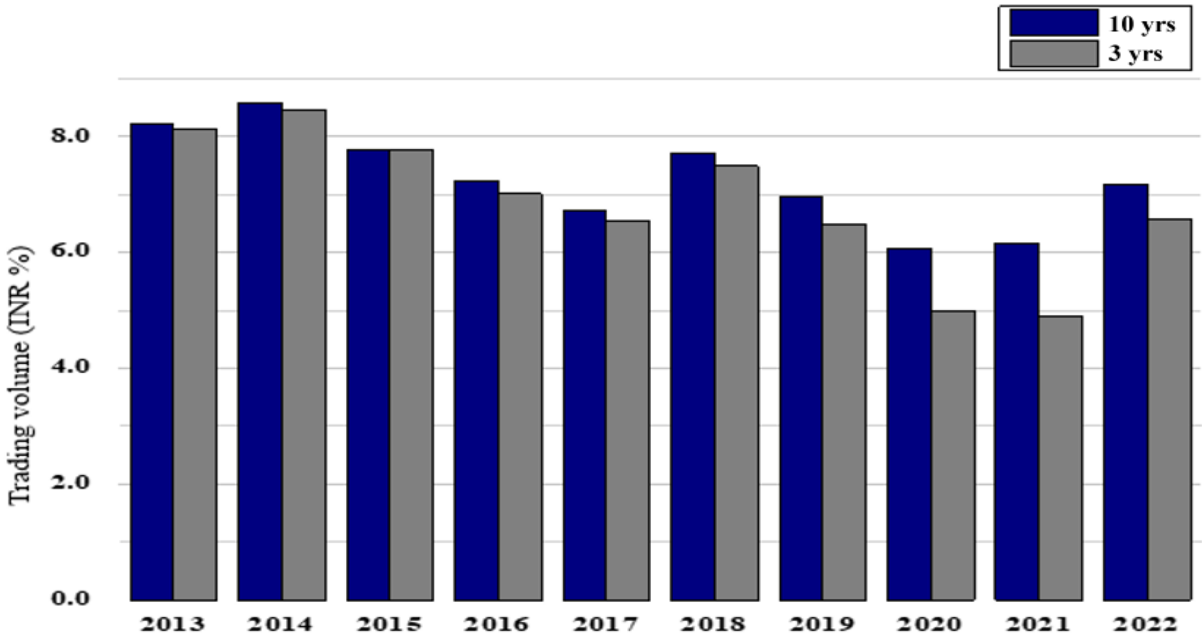

Figure 1. Trading volume of Indian treasury bonds.

The study uses one deep learning classifier and two machine learning models to forecast the relative daily performance of the 15000 greatest Indian government bond prices to address this research topic. In this regard, the research covered 10-year bond prices using techniques including ARIMA-2501, LSTM-2501, and linear regression-2501. In a similar vein, the pricing of 3-year bonds, such as ARIMA-2499, LSTM-2499, and linear regression-2499, respectively. The yield curve analysis and bond yield spread were also measured in the study. The study has five main contributions:

The researchers first displayed the slope of the interest rate differential between long- and short-term bonds using a stratified yield curve. Secondly, the bond spreads with positive values ranging from 8 to 150 bps are analyzed based on a range of economic factors and market conditions. Third, an analysis is done on how accurate each of the three machine learning models is at predicting bond prices. To assess the efficacy of the model, the study finally contrasts the findings of the MAE, MAPE, MSE, and RMSE error metrics. The deep long-short-term recurrent neural network (DLSTM-RNN) model is very good at predicting bond prices. It has an average mean absolute percentage error (MAPE) of 0.0099 in the loss function. This means that 98% of the model is highly recommended for predicting bond prices over 10 years and 99% is highly recommended for predicting bond prices over 3 years.

2. LITERATURE REVIEW

2.1. Traditional Statistical Techniques

So far, the majority of academics have examined the trading performance of bond interest rate prediction using statistical models, time series/cross-section models, other econometric models, or contemporary machine learning methods. Interest rates, which have a direct impact on their fluctuations, are what primarily determine bond prices. Using traditional financial models like the yield curve method, the conventional binary tree model, structure theory, Nelson-Siegel and Svensson's model, and time series models, many researchers have primarily focused on interest rates (Baghestani, 2009; Barr & Campbell, 1997; Black & Karasinski, 1991; Favero, 2013). Bernoth, Von Hagen, and Schuknecht (2012) used basic portfolio theory to study the relationship between government debt and deficits in European sovereign bond yields to evaluate the effect of fiscal policies on interest rates. The results show that measures of government debt, both before and after the establishment of the European Monetary Union (EMU), have a considerable impact on yield spreads. For the investigation of the bond market, some research employed conventional econometric and statistical techniques. Asgharian, Christiansen, and Hou (2015) discovered that the degree of macroeconomic uncertainty affects the long-term stock-bond correlation and bond volatility. Maghyereh and Awartani (2016) highlighted the significance of Islamic bonds (Sukuk) in comparison to international bond markets for foreign investors' strategic asset allocation and hedging. To assist investors in differentiating ratings from various credit rating agencies (CRAs), Livingston, Poon, and Zhou (2018) looked at the Chinese credit rating business and the bond market. The outcome does state, however, that Chinese CRAs' grading scales are not equivalent to those of foreign CRAs. Pendharkar and Cusatis (2018) looked into two different asset portfolios: the bond market index and the Standard & Poor's (S&P) 500 index. The study found that temporal-difference learning algorithms (TD (λ)) agents and high learning frequency, commonly referred to as adaptive learning, often beat single-asset stocks and bonds in terms of cumulative returns. Barr and Campbell (1997) used the observed prices of nominal and index-linked UK government bonds to forecast future real interest rates and inflation rates. The estimation technique corrects for errors in British index-linked bond indexation. All bonds are considered to have equal expected log returns, and inflation and real interest rates are assumed to follow simple time-series processes, the parameters of which can be deduced from the bond price cross-section. Hirshleifer and Luo (2001) looked at the profitable strategies employed in the competitive securities market as well as the overconfidence of traders. Banerjee and Pradhan (2022) studied the correlation between macroeconomic news and the intraday prices of the benchmark Indian government bond. To ascertain the significance, the study also applies robustness checks that consider the date, frequency, and dissemination of the news. Paul and Reddy (2022) use an autoregressive distributed lag (ARDL) bounds-testing co-integration method to look at short-and long-term effects of US quantitative easing (QE) on the benchmark 10-year Indian government bond rate. The analysis indicates that in the absence of QE, yields would have been less volatile. To preserve system liquidity and reduce interest rate volatility, the RBI was forced to alter its policy rate and carry out open market operations (OMOs). Using the capital asset pricing model (CAPM), Abad, Chuliá, and Gómez-Puig (2010) looked at how the returns on the government bonds in two groups of EU-15 members were affected by two sources of systematic risk: the world and the Eurozone. The results show that EMU risk variables have a greater potential to affect Euro markets than global risk factors. Gilmore, Lucey, and Boscia (2010) suggested analyzing the movement patterns of twenty developed North American, European, and Asian government bond market indexes using the minimum spanning tree (MST). The study explained how market links dynamically evolve, as well as how the MST and its corresponding hierarchical tree develop. Nonetheless, there is increasing movement in a portion of the European Union's (EU) bond market. Schuknecht, Von Hagen, and Wolswijk (2009) examined risk premiums paid by European central governments and local governments in Germany, Spain, and Canada using bond yield spread data. The investigation showed that sub-central governments in Germany and Spain pay premiums on interest rates connected to liquidity, while Canadian and German provinces and states that benefit from fiscal equalization have lower spreads. Using a traditional international capital asset pricing model, Pozzi and Wolswijk (2012) explained the time-varying integration of the government bond markets in the Euro area. The study found that by 2006, idiosyncratic features had all but vanished in all countries except Italy; nonetheless, Italy experienced its preemergence in response to the financial crisis that started in 2007. Second, national exposures to the shared international risk factor have converged across countries despite the crisis.

As Barr and Campbell (1997) showed, a straightforward, broadly applicable method can produce statistically sound and empirically appropriate estimates of the yield and forward rate curves. Applied to British index-linked and conventional bonds, this paradigm yields a plausible breakdown of forecast nominal rates into expected real rates and inflation. When utilizing the decomposition instead of only nominal interest rates, much more precise inflation forecasts are generated. Kolluri, Wahab, and Wahab (2015) examined India's two equity benchmarks, the

BSE-SENSEX and NSE (Indian stock market indices). Furthermore, a multivariate cointegration test was conducted between the stock and bond markets in India and those in the US, UK, China, Japan, and the Emerging Equity Markets index of Morgan Stanley. The findings support a positive role for India's equity and bond markets in the portfolios of foreign investors by demonstrating that there are only two cointegrating vectors at most between these five foreign equity indices and the country's stock and bond markets. This suggests that full cointegration is still far from achieved. According to Wahidin, Akimov, and Roca (2021) research, the modest increases in corporate borrowing, the real decline in financial corporate borrowing, and the consequent slowdown in private lending meant that bond markets achieved very little in the way of actual economic growth. The economy doesn't seem to be benefiting from the government's increased borrowing on the bond market. Banerjee and Pradhan (2022) looked at the effect of unexpected macroeconomic news on the return and volatility behaviour of intraday bond prices from new and different perspectives. According to Zaremba, Kizys, Aharon, and Umar (2022) report on the composite premium of multidimensional risk, the term premium is expected to rise in reaction to changes in the number of reported COVID-19 infections. This is likely due to increased investor perceptions of risk and uncertainty surrounding investments in the sovereign bond market.

Paul and Reddy (2022) used the autoregressive distributed lag (ARDL) bounds testing co-integration method to look at how US quantitative easing (QE) affected 10-year Indian government bond (IGB) rates in the short and long term. There is a long-term correlation between the 10-year IGB yield, the control variables, and QE. Both the rise in industrial production and the price of Brent crude oil have a major effect on yields. According to the counterfactual analysis, yield volatility would have decreased if QE hadn't been implemented. Asgharian et al. (2015) used the mixed data sampling (MIDAS) technique to explain how macroeconomic uncertainty affects the long-run correlation and volatility of stocks and bonds. The empirical findings demonstrate excellence-seeking behavior. The GDP growth rate is an alternative macroeconomic state variable with effects akin to those of macroeconomic uncertainty. Out-of-sample results are not particularly excellent when it comes to the effects of macroeconomic uncertainty. To give investors, governments, institutions, and research communities useful results for their well-informed decision-making, the research can therefore be broadened by utilizing ML models.

2.2. Artificial Intelligence Techniques

One of the most important technologies that will protect the Indian security market and enable it to compete and win in future wars is artificial intelligence (AI). AI is expected to revolutionize several industries and create new investment opportunities due to advancements in machine learning (ML), natural language processing (NLP), and computer vision, according to a poll conducted by the international market analysis research and consultancy company (IMARC). Infosys, Wipro, HCL, Tech Mahindra, Mindtree, Tata Elexi, Cyient, Kellton, and Persistent Systems are a few of the businesses mentioned. These AI firms have a track record of developing and deploying AI-related technologies, making them top achievers in India's AI market. According to Webster and Ivanov (2020) the global adoption of robotics and AI technologies is having an impact on the industrial robots in the manufacturing sector, as well as other financial activities like trading on financial markets, transportation using autonomous vehicles, customer relationship management using chatbots, legal services, and medical diagnosis and operation. Before investing in any firm, careful research and analysis are required to make an informed decision. Investors may benefit from the AI sector's growth and long-term sustainability. The researchers developed a forecasting model using returns from Indian government bonds and a range of machine-learning techniques. However, there are additional dangers related to investing in AI stocks in India, such as market volatility and company-specific hazards. It is imperative to conduct thorough research and consider factors such as a company's financial health, competitive landscape, and management team before investing. By doing this, investors can reduce risks and benefit from the growth of the artificial intelligence industry. The laws that regulate AI-related technology are always evolving. Before investing, think about India's regulatory landscape and how it can impact the company's potential for growth. Select companies that respect all relevant legal requirements and are informed of any updates that may impact their business operations. Shares of companies that develop and use artificial intelligence technologies are known as AI stocks. These businesses may specialize in machine learning (ML), robotics, natural language processing (NLP), or other AI-related disciplines. By purchasing Indian AI stocks, investors may participate in and perhaps profit from the expanding AI industry. The AI industry is expected to grow rapidly in the upcoming years, providing opportunities for profitable ventures. Numerous industries are adopting AI technologies, indicating a sizeable market for businesses operating in this field. Investing in AI stocks in India allows investors to diversify their holdings and gain exposure to a rapidly expanding industry. Demand for AI technology is increasing due to demands for accuracy, efficiency, and cost-effectiveness, which suggests long-term sustainability. AI will probably have a significant impact on future innovation, which will present interesting new opportunities for businesses operating in the sector.

According to recent studies,Gogas, Papadimitriou, Matthaiou, and Chrysanthidou (2015) and Jabeur, Sadaaoui, Sghaier, and Aloui (2020) have predicted different bond rates using machine learning approaches. To predict the asset pricing of the CAT bond market, Götze, Gürtler, and Witowski (2023) used neural networks, random forests, and linear regression models. Chen, Leung, and Daouk (2003) researched the use of neural networks to predict trading returns in the financial market of Taiwan. Huang, Chen, Hsu, Chen, and Wu (2004) assessed the models of support vector machines (SVM) and backpropagation neural networks (BNN) that were employed to assess the prediction accuracy of credit ratings in the US and Taiwan markets. Kim (2021) investigated back-testing techniques used in the government bond market to forecast trading performance in the Korean bond markets. The spread between long- and short-term bond yields is a trading tool that traders use to lower market risks. As a result, the spread may be predicted, serve as the data-driven long-term direction of the bond market, and generate additional profits.

Using machine learning and back-testing, this research develops a novel adaptive trading system for the bond market. Huynh et al. (2020) studied the fourth industrial revolution of artificial intelligence (AI) and green bonds. The study finds that, in terms of asset volatility, short-term volatility transmission is stronger than long-term transmission. The study concluded that there is an inherent self-transmitting risk in the portfolio that necessitates careful diversification.

When Kim et al. (2021) looked at how well forecasting techniques worked in the corporate bond markets, they found that neural network projections were the most accurate over both short- and long-term periods. The authors view the limitations of the Livingston et al. (2018) study findings as a motivator for their investigation. The study concludes that bond ratings are less useful since a crude and coarse rating scale combines bonds with significantly varied default risks into the same rating category without establishing a differentiation based on default risk. This does not assist the growth of a strong and healthy bond market.

Maghyereh and Awartani (2016) conducted a comparative analysis of research on the transmission of returns and volatility in the sukuk and bond markets. Pendharkar and Cusatis (2018) examine a personal retirement portfolio consisting of two assets and recommend a few reinforcement learning agents for asset trading. The results of the continuous action adaptive knowledge agent are quite encouraging. This agent beats both the S&P 500 and 10-year T-note portfolio and the S&P 500 and Aggregate Bond Index (Agg)portfolio by a significant margin. Even though it will be very tough to outperform the agent that now leads the field in terms of performance, future researchers may take into consideration model changes to increase the performance of this agent. The yield spread on the Indian bond market and the accuracy of several machine learning models' predictions are taken into account in this study. While retail investors are mostly prominent in the stock markets in India, institutional investors are active in both the debt and stock markets. Interest rate conditions in the economy, inflation, and stock market risk all have a significant impact on bond market yields.

3. METHODOLOGY

The data used to develop the forecasting model for the secondary market for Indian government bonds is explained in this section. The researchers first describe the procedure for selecting the sample. The variables utilized in the empirical analysis are introduced in the second phase. Third, the data's descriptive statistics are shown. Participants in the Indian bond market frequently use the nseindia.com database, which is where th e dataset come from. The factors were selected based on suggestions from Google, and the researchers collected real data over ten years, from January 1, 2013, to December 30, 2022. The appropriate sample sizes for the 10-3-year bond are 2501 (10 years) and 2499 (3 years), respectively. The current model, distinct from prior research, serves as an integration of advanced deep neural networks and linear, sub-linear machine learning to verify the appropriate bond yield spread forecasting strategy. The existing study focused on either bond yield or bond spread prediction. The numerous economic input indicators are considered to strengthen the current study. Input data was collected for the following reasons: In the beginning, interest rates and currency rates were linked. In general, increased interest rates boost currency demand, which raises the currency's value and lowers the exchange rate. The bond and stock markets are related because of the limited availability of investment funds. Bond prices typically decline as bond rates rise because investors shift their attention from the stock market to bonds in search of more secure investment returns.

3.1. Data Collection and Expansion

The collected dataset was implemented to calculate the total market value of the Indian Treasury bond in the financial market. Figure 1 shows the total market value of the Indian government bond market from 2013 to 2022. However, investors can be affected by economic factors, in turn which affect the bond value in the financial market.

Table 1. Details of Indian treasury bond datasets are available for training and testing purposes.

Name of the bond	Dataset collection period	Model	Total nos. of available data	Total nos. of data produced	No. of training data	No. of testing data
3-years bond	01-01-2013 to 30-12-2022	Linear regression	2499	2499	1874	625
		LSTM	2499	2499	1940	559
		ARIMA	2499	2499	1999	500
10-years bond	01-01-2013 to 30-12-2022	Linear regression	2501	2501	1875	626
		LSTM	2501	2501	1941	560
		ARIMA	2501	2501	2000	501
			15000	15000	1874	3371

Theworldgovernmentbonds.com website, a popular resource for traders in the Indian bond market, provided the statistics for Table 1. The researchers gathered actual data over a 10-year (January 1, 2013, to December 30, 2022), following the recommendations of the domain expert. The first dataset consists of two bonds, one of which was traded on the secondary market for three years, 2499, and the other for ten years, 2501. The dependent variable in our analysis is the daily yield on government bonds that have been available since January 1, 2013. The following motives led to the collection of input data: First of all, there is a relationship between interest rates and currency rates. Increased demand for a particular currency leads to an increase in its value and a decrease in the exchange rate when interest rates are higher. All observations with missing or implausible data are excluded to prepare the data set for additional analysis. The bond and stock markets are related since there is a limited quantity of money available for investments. Furthermore, when the gap between Indian and US government bonds narrows, the Reserve Bank of India (RBI) will undoubtedly raise interest rates in India. Bondholders may suffer losses if interest rates rise in India and bond values decline, making it less appealing for investors to buy existing bonds. The researchers also gathered additional economic data that is relevant to real dealers, as well as indications of each nation's overall economic status and the bond market, to better assess the possible effects of numerous variables. As shown in Tables 2 and 3, the researchers ultimately collected 62 input variables, which included the RBI's monetary policy, the state of the government's budget, the global market, the economy, and inflation. Table 4 lists the many economic indicators that the Indian bond market has an impact on. Input factors that are designed to replicate some of the effects of historical data from 2013 to 2022 have an impact on the bond price. The frequency of the variable fluctuated daily, monthly, quarterly, and annually.

3.2. Designing Data with Statistics

We employed the 10- to 3-year government bond price trading day travel market value data from January 1, 2013, to December 30, 2022. There are two sets of opening and closing prices for each trading day. The model yield is 15000, and the total dataset available for the investigation is 15000. This comprises 3371 (20%) of validation and testing data and 11629 (80%) of training data. Table 1 lists the length of time that the data was collected overall, the bond's label, the number of training and testing patterns, and the observed dataset that the models employed.

3.3. Software Database

Using Anaconda3 (64-bit), the researchers gathered, processed, and evaluated data for the whole study. In this case, Anaconda is an excellent option for those who are not familiar with Python or data science. NumPy, Pandas, Cython, Urllib3, and Joblib are some of the Python software packages used for data processing and feature development. Using the matplotlib, pmdarima, and scikit libraries, all other machine learning models are built and trained. Keras is used with the TensorFlow backend to create deep learning models.

Table 2. List of variables affecting the Indian government bond market.

Category	Frequency	Features	Variables acronym (Refer to Table 3)
Exchange rate	5	Daily	INR/USD, GBP/INR, EUR/INR, AUD/INR, CAD/INR
Equities market volatility	5	Daily	MS, GD, MC, CP, MPC
Return trends	8	Daily	IR, inflation, EG, Sensex, Nifty, FII, DII, CPI
Monetary policy of RBI	12	Bi-monthly	GDP, RR, RRR, SLR, CRR, TI, MS, PS, SEG, FE, AoC, IR
Fiscal position of government	5	Yearly	GR, GE, PD, COVID-19, GDP
Global market	6	Daily	IFEI, BASM, FPIs, FME, Gold, COVID-19
Factors that affect bond prices	11	Daily	IR, BM, BCR, BS, MC, inflation, par value, yield, IR, BR, COVID-19
Interest rate	2	Monthly	CR, MR.
Maturity of the bond	42	Monthly	3mths, 6mths, 1yr, 2yrs, 3yrs, ...40yrs.
Daily financial indicator	6	Daily	S&P BSE, S&D, Yield, IR, BR, PV
Economic indicator	5	Daily	FIIs, DIIs, Politics, GDP, COVID-19
Inflation	3	Daily	GDP, COVID-19, MEF
Stock market index	7	Daily	Sensex, S&P BSE, NIFTY 50, NIFTY 100, NIFTY 150, NIFTY 200, NIFTY 250

Table 3. Descriptions of variables acronyms.

Acronym	Description	Acronym	Description	Acronym	Description
GDP	Gross domestic product	BM	Bond maturity	GR	Government receipts
RR	Repo rate	BCR	Bond credit rating	GE	Government expenditure
RRR	Reverse repo rate	BS	Bond structure	PD	Public debt
SLR	Statutory liquidity ratio	MC	Market conditions	Govt	Government
CRR	Cash reserve ratio	IFEI	Increased foreign exchange inflows	IR	Interest rate
TI	Tame inflation	BASM	Bonds affect the stock market	PV	Price volatility
MS	Money supply	FPIs	Foreign portfolio bonds	AoC	Availability of credit
PS	Price stability	FME	Free market economy	FIIs	Foreign institutional

Acronym	Description	Acronym	Description	Acronym	Description
					investors
SEG	Stable economic growth	S&D	Supply and demand	DIIs	Domestic institutional investors
FE	Full employment	BR	Bond rating	MS	Market sentiments
GD	Geopolitical developments	MC	Market cycles	CP	Company performance
MPC	Monetary policy changes	CPI	Consumer price index	CR	Coupon rate
MR	Market rate	MEF	Macroeconomic factors	GBP	Great Britain Pounds
CAD	Canadian dollars				

Table 4. Descriptions of lead indicators in the Indian government bond market.

Indicator	Type	Description
Interest rate	Prime mover	Interest rates were controlled and maintained below what the market would bear. The central bank was able to borrow money at a favorable interest rate, particularly by investing in Treasury bills, and debt monetization was both automated and limitless (Reddy, 2002).
Spread	Prime mover	Long-term yields differ from short-term yields (Krishna & Nag, 2022).
US bond yield affects Indian bond market	Prime mover	The Indian equity market is impacted by changes in US treasury bond yields due to Foreign Portfolios Investments (FPIs). When yields rise, FPIs tend to pull their money out of bonds, which lowers the P/E multiple of the US market (Prasanna & Sowmya, 2017).
Repo rate	Prime mover	Raising the repo rate harms debt mutual funds because their primary investment is in bonds, and rising repo rates often result in falling bond yields (Arrata, Nguyen, Rahmouni-Rousseau, & Vari, 2020).
Money supply	Prime mover	Another monetary indicator that encourages growth is narrow money (Paramanik & Kamaiah, 2017).
Politics	Other	Stability, government intervention, elections, budgets, and other factors all have a big influence on the economy and financial markets. Political changes and budget announcements have a significant impact on the stock market and increase market volatility to an exceptionally high degree (Wisniewski, 2016).
High inflation and external imbalances	Other	The delay in reducing the monetary and fiscal stimulus contributed to the high inflation and external imbalances (Ederington et al., 2015).
Exchange raterupee/U.S. dollar (RS/USD)	Other	Interest rates and currency rates are connected. The macro vulnerability of the economy is indicated by a significant decline in the exchange rate (Paramanik & Kamaiah, 2017; Sharm, 2016).
Liquidity risk	Other	Credit risk is the likelihood of a market failure in the absence of buyers or sellers. This could occur due to a lack of depth in the market, a small number of buyers and sellers, or a low trading volume for the bond (Gubareva, 2021).

3.4. Models

The researchers used the most recent model innovation in machine learning while also introducing the deep LSTM recurrent neural network, ARIMA model, and linear regression as a hybrid model to capture enhanced forecasting skills.

3.4.1. Linear Regression

In machine learning, linear regression has been used. This statistical technique was developed to examine the explanatory power of input factors by weighing their weight and focusing on numerical variables. The multiple linear regression models are in their general form, as shown in Equation 1 (Yan & Su, 2009).

$$\mathcal{Y} = f(x) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_n x_n + \varepsilon, \quad (1)$$

Where $\mathcal{Y}$ is the target variable, $\beta_0, \beta_1, \beta_2 \ldots, \beta_p$ are weighted linear regression coefficients that yield precise predictions, $x_1, x_2, .., x_n$ are training set input variables, and ε represents the model error. In the meantime, the cost function that equals is minimized to find a best-fit line, as shown in Equation 2 (Raschka & Mirjalili, 2019).

$$\varepsilon(\beta_1, \ldots, \beta_D, b) = \frac{1}{N}\sum_{i=1}^{N}(\mathcal{Y}^i - t^i)^2 \quad (2)$$

The linear regression's goal is to minimize ε by selecting $\beta_1, \dots, \beta_D, and\ b$. As a result, the cost function is the mean squared error (MSE), which is the sum of the squared differences between the true value and the prediction.

3.4.2. Deep LSTM Neural Network

An RNN type called Long Short-Term Memory (LSTM) allows the network to remember both short- and long-term standards. The most popular method for predicting financial time series data is to employ LSTM models. An LSTM layer is created by the network made up of LSTM units. Each cell in the LSTM units has an input gate, an output gate, and a forget gate. These gates are in charge of controlling the data movements. Every cell remembers the expected standards during random time intermissions via these topographies. The LSTM unit's onward pass technique is illustrated in Equations 3–7 (Sezer, Gudelek, & Ozbayoglu, 2020).

$$f_t = \sigma_g \left(W_f x_t + U_f h_{t-1} + b_f\right) (3)$$
$$i_t = \sigma_g \left(W_i x_t + U_i h_{t-1} + b_i\right) (4)$$
$$o_t = \sigma_g \left(W_o x_t + U_o h_{t-1} + b_o\right) (5)$$
$$c_t = f_t * c_{t-1} + i_t * \sigma_c \left(W_c x_t + U_c h_{t-1} + b_c\right) (6)$$
$$h_t = o_t * \sigma_h \left(c_t\right) (7)$$

Where x_t ;f_t ;i_t; o_t ; h_t ; c_t ; σ_g ; σ_c ,σ_h ;* ; W, U and, b denoted the LSTM unit's input vector, forget gate, inputs gate, output gates, output vector, cell conditions vector, sigmoid function, hyperbolic tangent functions, component-wise (Hadamard) creation, learning weight matrices, and learning parameters for the bias vector. Equation 8 displays the total error, which is the sum of the errors (Sezer et al., 2020).

$$\frac{\partial E}{\partial W} = \sum_{t=1}^{T} \frac{\partial E_t}{\partial W} (8)$$

LSTM is the name of a particular kind of RNN. Consequently, the favored optimization schemes and weight updates are similar. The number of hidden layers, the number of units in each layer, the activation functions, the learning rate, the momentum values, the number of epochs, the batch size, the decay rate, the optimization algorithms, the sequence length for the LSTM, the gradient clipping, the gradient normalization, and the dropout are all the same hyperparameters of an LSTM as they are of an RNN. The same hyperparameter optimization techniques that are used for RNNs may also be utilized to determine the ideal hyperparameters for an LSTM. The researchers utilized root mean square error (RMSE) and mean absolute percentage error (MAPE) as performance benchmarks (Sezer et al., 2020).

3.4.3. ARIMA

An efficient solution to the time series problem is the complex Auto-Regressive Integrated Moving Average (ARIMA) model. In this article, the ARIMA model also determines the yield on the bond market. To calculate how unexpected occurrences affect market yield, exogenic imitation variables have to be included in the ARIMA model's provisional adjustment equation. According to Liu et al. (2022) the variance of the random error period in the ARIMA model can be represented by the q command distribution lag model of the square of the error period, Equation 10, and the stationary random variable xt can be specified as the p command autoregressive procedure,Equation 9.

$$xt = \beta 0 + \beta 1xt - 1 + \beta 2xt - 2 + \cdots + \beta pxt - p + ut \qquad (9)$$
$$\sigma 2t = E(ut2) = \alpha 0 + \alpha 1u2\ t - 1 + \alpha 2u2\ t - 1 + \alpha qu2\ t - q \qquad (10)$$

Whereas $xt = \beta 0 + \beta 1xt - 1 + ut$ is an illustration of an autoregressive model. Because they show the dependent variable's time path concerning its previous values, the latter are frequently referred to as dynamic models.

3.5. Training of the Proposed Prediction Metrics

The training of the models for the simulation uses 80% of the feature datasets at random. To assess the machine learning techniques, the planned model for this investigation will precisely match the data. The deep LSTM-recurrent neural network, ARIMA, and linear regression are excellent for data-fitting analysis because they use the training data as the test data set. The mean absolute error (MAE), mean absolute percentage error (MAPE), mean square error (MSE), and root-mean-square error (RMSE) are used to measure the inaccuracies in bond price forecasts. Error measures were employed in the discussion section to assess the study's prediction accuracy. The actual numbers were well coordinated, and these indexes yielded outcomes exactly as expected. The number of days that must lapse before a forecast may be made determines the training parameters. The learning principles of the linear regression model are given by Equations 1 and 2, those of the deep LSTM model by Equations 3-8 and those of the ARIMA model by Equations 9-10. Every training pattern is meticulously applied to the relevant models, and the related error levels are widely acknowledged as performance metrics. The mean squared error (MSE) is also generated for each set of patterns.

3.5.1. Scale-Dependent Errors

Hyndman and Koehler (2006) state that the scale errors are identical to those of the raw data. This constraint prohibits the comparison of series with various scales using accuracy metrics that are particularly dependent on these errors. Therefore, the root mean square error (RMSE), mean squared error (MSE), and mean absolute error (MAE) can all be defined using scale-dependent measures, as seen in Equations 11-14 (Vijh, Chandola, Tikkiwal, & Kumar, 2020). The following equation is verified by measuring the mean absolute error (MAE) score as the average of the absolute error values.

$$\frac{1}{N}\sum_{i=1}^{N}|y_i - \hat{y_i}| \quad (11)$$

The mean squared error (MSE) is equal to the prediction error per square, as obtained by using the following formula.

$$\frac{1}{N}\sum_{i=1}^{N}(y_i - \hat{y_i})^2 \quad (12)$$

The root-mean-squared error (RMSE) is a well-liked general-purpose error metric for numerical forecasts, and it may be calculated using the following formula.

$$\sqrt{\frac{1}{N}\sum_{i=1}^{N}(y_i - \hat{y_i})^2} \quad (13)$$

Equations 11 through 13 show, where y_i stands for the actual closing selling price, $\hat{y_i}$ for the prediction price, and $\frac{1}{N}$ denoting the total number of datasets utilized in the research.

3.5.2. Percentage Errors

The scale independence of the percentage mistakes used to forecast performance across different scaled datasets. Most statisticians calculate the Root Mean Square Percentage Error (RMSPE) in triplicate using the following mathematical equation, which was published by Vijh et al. (2020) to attain precision.

$$\sqrt{\frac{1}{N}\sum_{i=1}^{N}\left[\frac{y_i - \hat{y_i}}{\hat{y_i}}\right]^2} \times 100 \quad (14)$$

Where $\frac{1}{N}$ denotes the total number of datasets used by the models, and y_i represents the actual value, $\hat{y_i}$ represents the closing value.

Using a comparison between the time series' target values and associated expectations, the researchers were able to identify several cost issues with regression machine learning techniques. Each metric has the same importance in assessing how good the prediction models are, even when the calculated values of the outcomes of using the two metrics vary. Take note that the production statistics frequently display different scales for the ideal price projection. It is advised to utilize RMSPE or any other % error metrics to examine the relative inaccuracy of several models for bond yield prediction.

3.6. Evaluation of the Prediction Metrics

The remaining 20% of the feature data is used to assess the model's prediction performance after the training phase. Finding the testing values of bond yield projections is the main objective of the performance evaluation, which compares the training data with the actual (observed) data. Big data companies frequently use regression machine-learning algorithms to produce projections for a range of industries, according to Xu and Chan (2019). The root of the mean squared error (RMSE) and the mean absolute percentage error (MAPE) are used to assess a forecast's accuracy.

4. RESULTS AND DISCUSSION

The researchers compare the prediction algorithms' results in terms of predictive accuracy and investment performance using the machine-based trading strategy in this part. The current study's findings are contrasted with those from related research using both linear and non-linear models. The simulation findings of all three models—linear regression, long- and short-term memory neural networks, and autoregressive moving averages—predicted the accuracy level of the bond price. The researchers first show that the three machine learning algorithms' anticipated predictions are accurate. The results of the MAE, MAPE, MSE, and RMSE error metrics are then compared to evaluate the model's performance. Prediction accuracy results are further analyzed for each model's performance over time using statistical metrics. Lastly, the researchers provided an explanation of how well one model outperformed all other machine learning models in terms of total prediction performance. The 10-3-year sovereign spread is the outcome of the prediction model for the Indian bond market used in this study, which takes arithmetical financial data as input.

4.1. Yield Curve Analysis

Accurate predictions of the returns on financial assets and their movements are crucial for safeguarding portfolios and helping investors lower their exposure to potential losses. Both governments and investors find it difficult to evaluate risk and predict future earnings with greater accuracy. Due to their potential to lower prediction errors, machine learning (ML) approaches have drawn a lot of attention lately for modeling financial time series to achieve this (see, for example, (Campbell & Taksler, 2003; Culkin & Das, 2017; De Spiegeleer, Madan, Reyners, & Schoutens, 2018; Ghoddusi, Creamer, & Rafizadeh, 2019; Khandani, Kim, & Lo, 2010)). According to Choudhry (2003) the yield curve represents the term structure of interest rates with steady variation. Longer-term interest rates are often higher than short-term interest rates. Generally speaking, the yield curve slopes upward as the duration increases. As a result, the yield difference (or spread) between longer and shorter bonds should be positive. The yield curve may be flat or reversed otherwise. The 3-year and 10-year bond durations that are used in determining the convexity of the curve were incorporated into the current analysis.

The yield curve flattens out by the end of 2021 as a result of a 102-bps decrease in the 10-3-year ITB spread, as seen in Figure 2. The yield curve has the highest slope near the end of 2015. Falling rates of inflation or concerns about a rise in the federal funds rate are examples of normal conditions. A flattening yield curve is a sign that investors are undervaluing the market in the long run due to macroeconomic concerns. Therefore, the difference in interest rates between long- and short-term bonds serves as a crucial indicator of the yield curve's slope and implicitly indicates that the yield curve's slope is an important indicator.

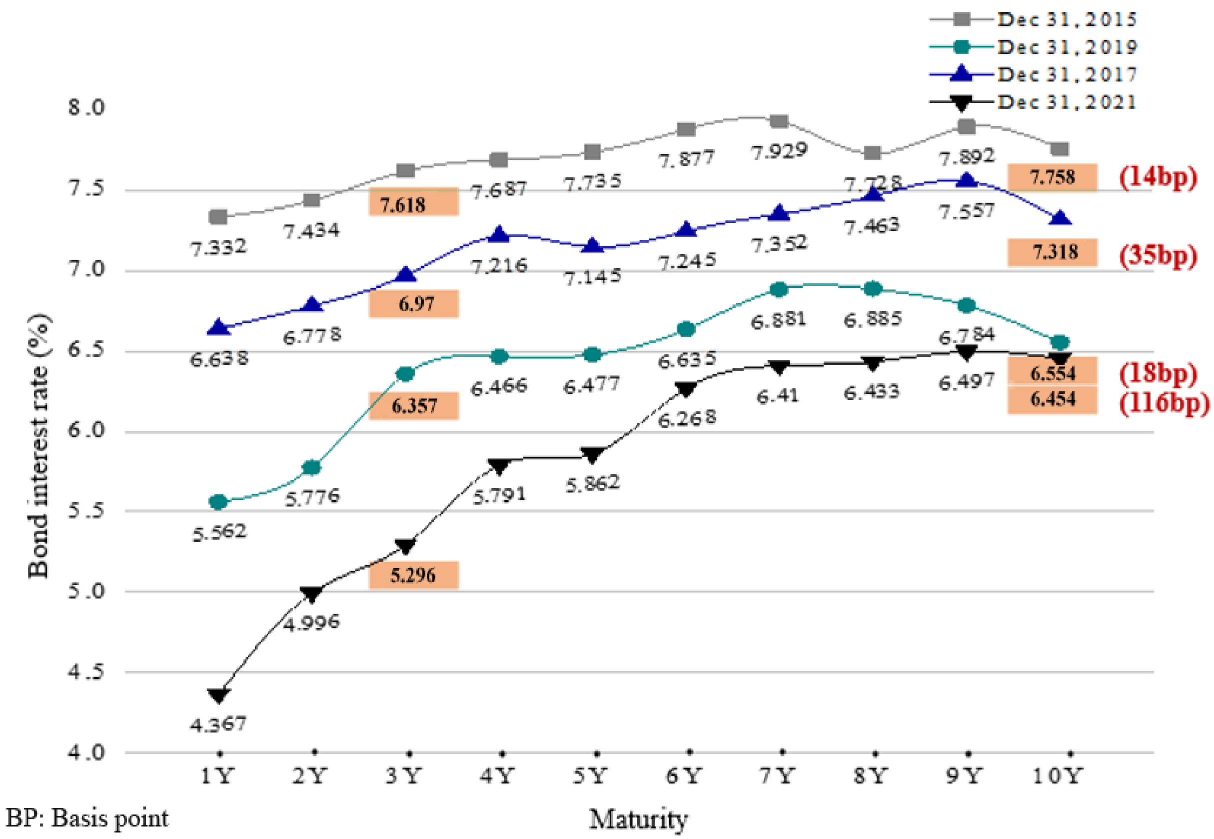

Figure 2. Yield curves for Indian government bonds.

4.2. Yield Spread Strategy

Interest rates and stock volatility were identified by Kim and Stock (2014) as significant determinants of yield spreads. The yield curve spread is composed of two components, which are both helpful forecasting inputs, according to Krishna and Nag (2022): the predicted real interest rate and the expected inflation. The estimated real rate is based on forecasts of the path of monetary policy and, in turn, real growth in the future. Given the high correlation between inflation and activity, inflation may also hold signals about possible future growth. As a result, the spread-the difference between short- and long-term yields-determines expectations for future interest rate changes.

This is a straightforward signal that policymakers may use to develop plans for influencing the economy through the implementation of corrective measures. The present analysis used the word bond yield in place of the commonly used concept of spread, which is the difference between the ask and bid prices. The 3- and 10-year bond rates in the Indian bond market are typically influenced by the 10-to 3-year bond spread, as seen in Figure 3.

Figure 3 shows that, for ten years before November 19, 2013, the values of the 10-year Indian government bond spreads were between 8 and 10 basis points. The majority of these spreads' values are positive, yet after the same intervals, they show 100 bps negative values.

The market circumstances and investors' pessimistic view of economic progress were the reasons for the negative spreads. As Keim and Stambaugh (1986) showed, there appears to be an inverse link between asset price levels and the conditional expected return differential between short-term and long-term bonds.

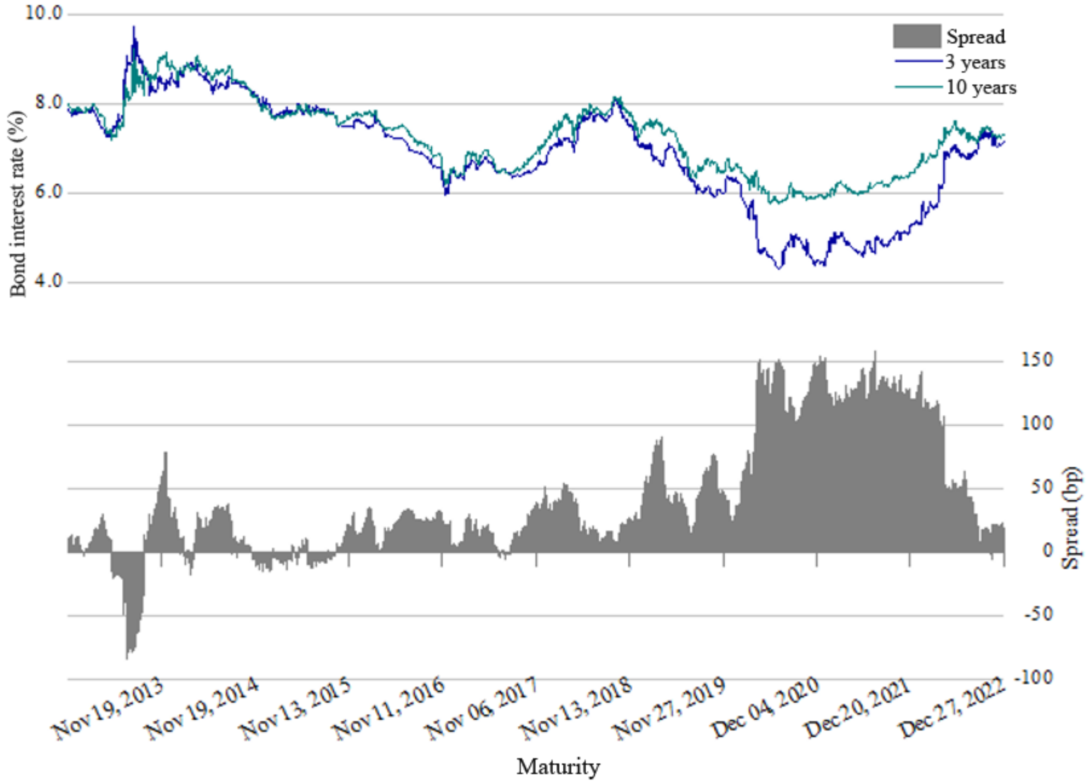

Figure 3. 10-3-year Indian government bond spread.

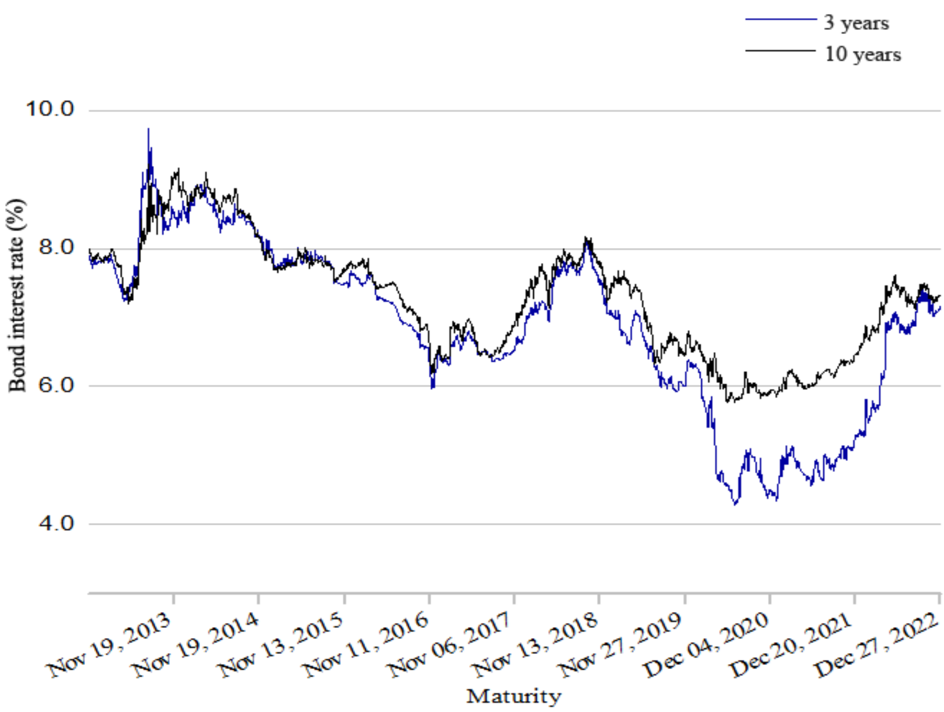

Figure 4. 10-3-year government bond yield history.

Since 2020–21, some economic data and market conditions indicating that investors are receiving positive perspectives on economic growth have caused the 10–3-year bond spreads, which are relatively given the positive values of 8–150bps, to increase. Once there are no more upper bounds for 2021–2022, it indicates that the market conditions make value forecasting extremely difficult. Figure 4 shows the 10- to 3-year government bond yield movements from 2013 to 2022.

4.3. Model Accuracy

An analysis of the various simulation results produced in the previous section is provided in this section. Although all bond yields, models, and comprehensive estimations have been simulated and results have been generated, space limitations only permit the presentation of typical scenarios in this paper.

Table 5. Results summary of linear regression, deep LSTM neural network, and ARIMA models for 10-3-years government bond yield prediction.

Methods	Metrics	MAE	MAPE	MSE	MSLE	RMSE	R2
Linear	10-yrs	0.0057	0.0289	0.0001	--	0.0110	--
	3-yrs	0.0062	0.0319	0.0003	--	0.0163	--
DLSTM	10-yrs	0.0651	0.0099	--	0.0001	0.0481	0.9795
	3-yrs	0.1241	0.0033	--	0.0007	0.1084	0.9743
ARIMA	10-yrs	0.8800	1.7600	1.1700	--	1.0800	--
	3-yrs	1.4300	2.3100	3.2100	--	1.7900	--

The results of the forecasts are compiled in Table 5.The bold values represent the 10- to 3-year MAPE projections of all three models. The mean absolute percent error (MAPE) deep LSTM neural network outperformed the linear regression model by 65.74% and the sub-linear ARIMA model by 99.44% in 10-year bond case forecasting. The mean absolute percent error (MAPE) deep LSTM neural network outperformed the linear regression model by 89.66% and the sub-linear ARIMA model by 99.86% in 3-year bond case forecasting. For a 10-year bond, the linear regression model's MAPE is 0.0289, the sub-linear ARIMA model's MAPE is 1.7600, and the LSTM neural network model's MAPE is 0.0099. For a 3-year bond, the LSTM neural network model's MAPE between the actual and predicted price is 0.0033, the linear regression model's is 0.0319, and the sub-linear ARIMA model's is 2.3100.

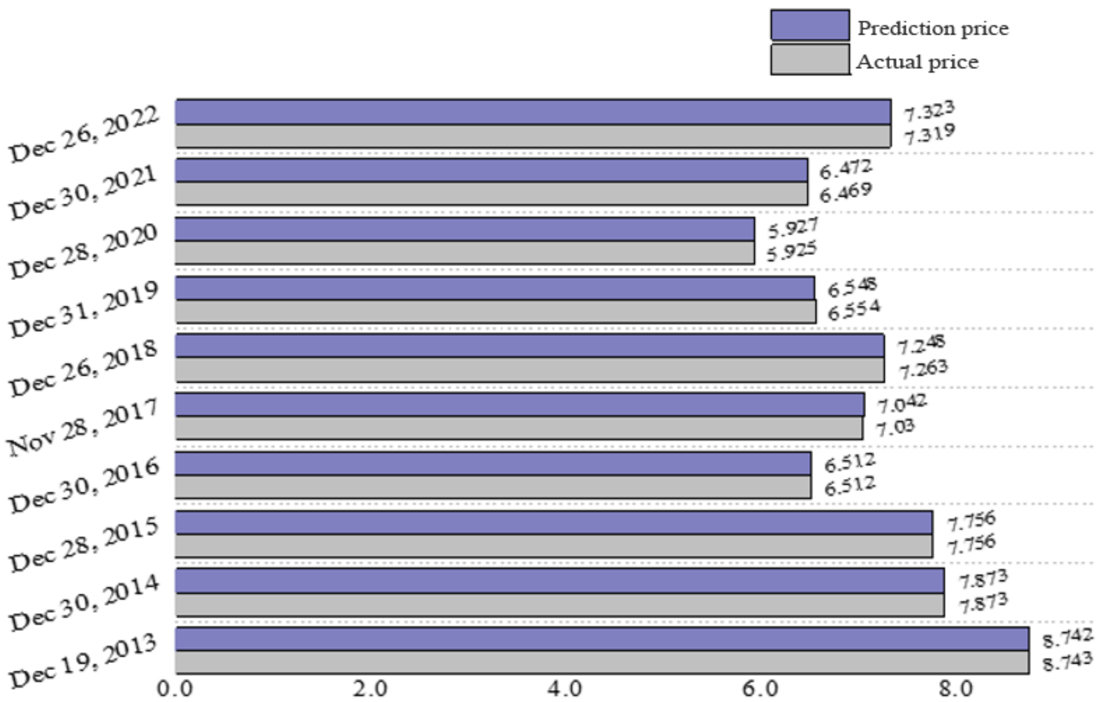

Figure 5. 10-year government bond yield prediction using the linear regression model.

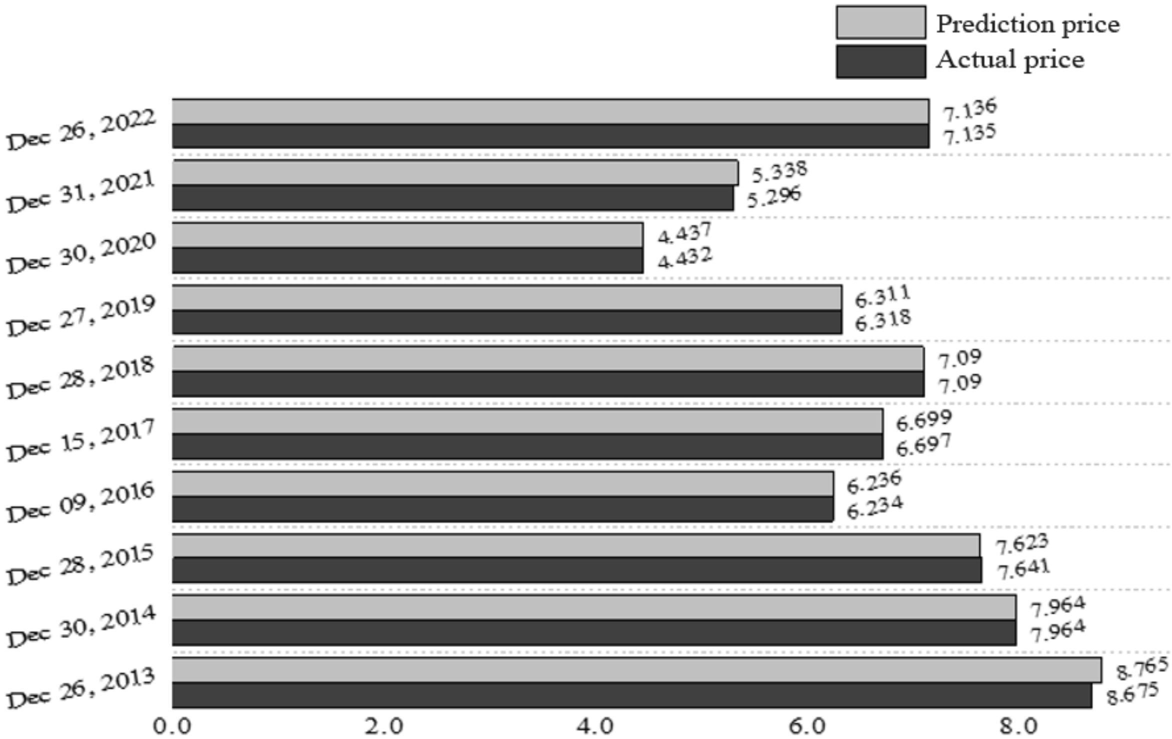

Figure 6. 3-year government bond yield prediction using the linear regression model.

Using a linear regression model for the 10- and 3-year bonds, the researchers discovered that there are significant variations in the statistical approach. The above Figures 5 and 6 display the comprehensive simulation results of the suggested linear regression prediction model for 10-3-year bond yield during the testing period. It is clear from the comparison results shown in the figures that there is a relationship between the actual and expected time series. The linear regression model's descriptive statistics are shown in Tables 6 and 7.

Table 6. Linear 10-year descriptive statistics.

Benchmarks	Mean	Std.	Min.	25%	50%	75%	Max.	Count
Price	7.269	0.839	5.760	6.528	7.379	7.809	9.228	2501
Open	7.273	0.836	5.764	6.530	7.387	7.809	9.228	2501
High	7.282	0.833	5.775	6.538	7.399	7.816	9.228	2501
Low	7.258	0.843	5.747	6.518	7.360	7.804	9.228	2501

Table 7. Linear 3-year descriptive statistics.

Benchmarks	Mean	Std.	Min.	25%	50%	75%	Max.	Count
Price	6.861	1.122	4.277	6.145	7.021	7.779	9.728	2499
Open	6.881	1.203	4.335	6.212	7.049	7.784	9.728	2499
High	6.892	1.197	4.335	6.240	7.053	7.789	9.728	2499
Low	6.855	1.220	4.277	6.138	7.015	7.778	9.728	2499

Below, Figures 7 and 8 display the comprehensive simulation results of the suggested deep LSTM neural network model for 10-3-year bond yield during the testing era. The comparison results shown in the figures, which clearly show which time series is more relevant-the real or anticipated one- demonstrate the deep neural network model's capability. Tables 8 and 9 depict the descriptive statistics of the deep LSTM model.

Figure 7. 10-year government bond yield prediction using DLSTM model.

Figure 8. 3-year government bond yield prediction using DLSTM model.

Table 8. DLSTM 10-year bond descriptive statistics.

Benchmarks	Mean	Std.	Min.	25%	50%	75%	Max.	Count
Price	7.269	0.839	5.760	6.528	7.379	7.809	9.228	2501
Open	7.273	0.836	5.764	6.53	7.387	7.809	9.228	2501
High	7.282	0.833	5.775	6.538	7.399	7.816	9.228	2501
Low	7.258	0.843	5.747	6.518	7.360	7.804	9.228	2501

Table 9. DLSTM 3-year bond descriptive statistics.

Benchmarks	Mean	Std.	Min.	25%	50%	75%	Max.	Count
Price	6.861	1.217	4.277	6.145	7.021	7.779	9.728	2499
Open	6.881	1.203	4.335	6.212	7.049	7.784	9.728	2499
High	6.891	1.197	4.335	6.24	7.053	7.789	9.728	2499
Low	6.855	1.221	4.277	6.138	7.015	7.777	9.728	2466

Figures 9 and 10 display the comprehensive simulation results of the suggested sub-linear ARIMA model for 10- to 3-year bond yield during the testing period. Tables 10 and 11 depict the descriptive statistics of the ARIMA model.

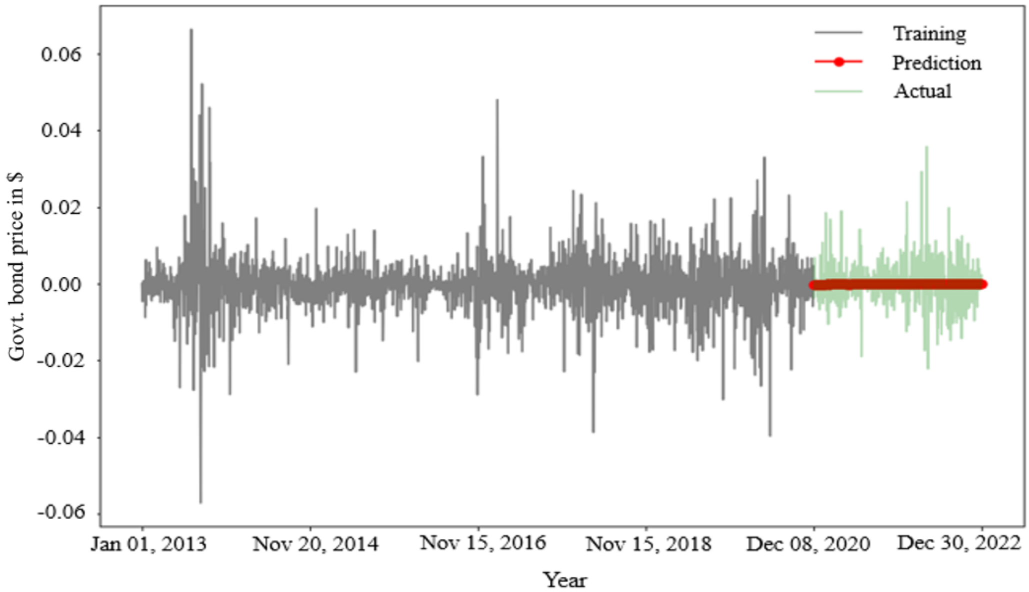

Figure 9. 10-year government bond yield prediction using the ARIMA model.

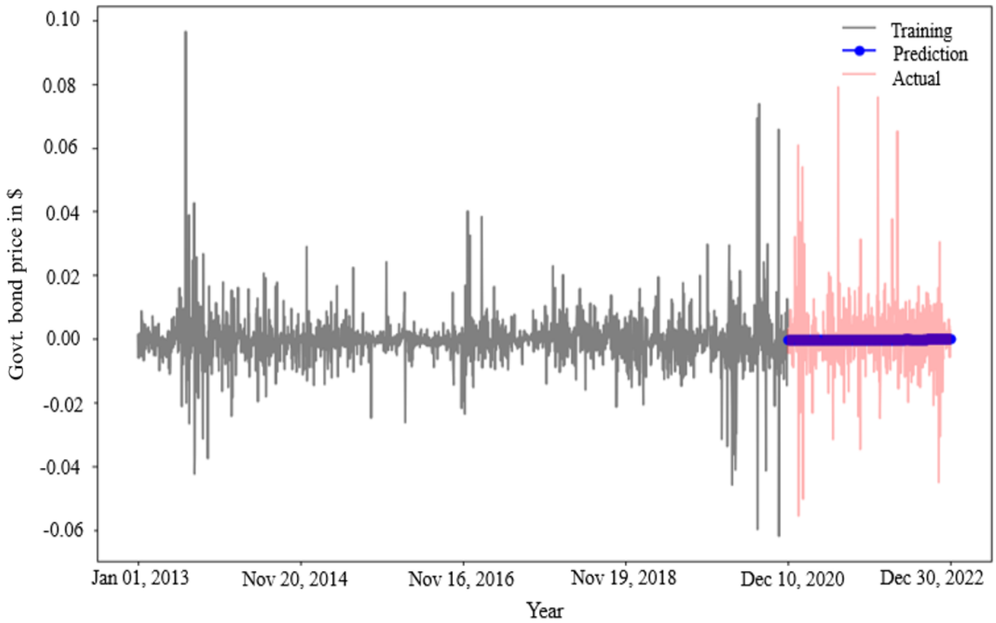

Figure 10. 3-year government bond yield predictions using the ARIMA model.

Table 10. Descriptive statistics of 10-year bond yield using ARIMA model.

Benchmarks	Mean	Std.	Min.	25%	50%	75%	Max.	Count
Price	7.269	0.839	5.760	6.528	7.379	7.809	9.228	2501
Open	7.273	0.836	5.764	6.530	7.387	7.809	9.228	2501
High	7.282	0.833	5.775	6.538	7.399	7.816	7.804	2501
Low	7.258	0.843	5.747	6.518	7.360	7.804	9.228	2501

Table 11. Descriptive statistics of 3-year bond yield using ARIMA model.

Benchmarks	Mean	Std.	Min.	25%	50%	75%	Max.	Count
Price	6.861	1.217	4.277	6.145	7.021	7.779	9.728	2499
Open	6.881	1.203	4.335	6.212	7.049	7.784	9.728	2499
High	6.891	1.197	4.335	6.240	7.053	7.789	9.728	2499
Low	6.855	1.221	4.277	6.138	7.015	7.778	9.728	2499

Table 12. Actual and prediction results of bond yield assessments.

ML-model	Year	Difference	53	326	330	703	728	820	850	982	1009	1097	1417	2083	2114	2255	2392	2399
Sample index																		
Linear regression	10-yrs	Act. price	7.422	6.225	6.250	6.505	6.501	6.518	6.882	7.408	7.853	7.772	6.860	6.126	6.005	6.469	7.440	7.240
		Pred. price	7.416	6.231	6.253	6.506	6.511	6.529	6.908	7.426	7.863	7.787	6.860	6.112	6.011	6.472	7.446	7.264
	3-yrs	Act. price	7.330	4.858	4.930	6.110	6.313	6.110	6.590	7.163	7.651	7.672	6.691	4.859	4.676	5.296	6.785	6.865
		Pred. price	7.333	4.877	4.946	6.113	6.313	6.148	6.593	7.166	7.669	7.676	6.692	4.864	4.681	5.338	6.788	6.887
DLSTM	10-yrs	Act. price	7.422	6.225	6.250	6.505	6.501	6.518	6.882	7.408	7.853	7.772	6.860	6.126	6.005	6.469	7.440	7.240
		Pred. price			Trained datasets					Trained datasets				6.017	5.933	6.414	7.410	7.365
	3-yrs	Act. price	7.330	4.858	4.930	6.110	6.313	6.110	6.590	7.163	7.651	7.672	6.691	4.859	4.676	5.296	6.785	6.865
		Pred. price								Trained datasets				4.826	4.642	5.213	6.739	6.683
ARIMA	10-yrs	Act. price	7.422	6.225	6.250	6.505	6.501	6.518	6.882	7.408	7.853	7.772	6.860	6.126	6.005	6.469	7.440	7.240
		Pred. price								Trained datasets				5.869	5.843	5.744	5.696	5.695
	3-yrs	Act. price	7.330	4.858	4.930	6.110	6.313	6.110	6.590	7.408	7.651	7.672	6.691	4.859	4.676	5.296	6.785	6.865
		Pred. price			Trained datasets					Trained datasets				4.402	4.369	4.236	4.171	4.129

Table 12, which employs all three models, presents the comparisons of various actual and predictive daily bond performance metrics. The primary difference between the actual and predictable price results for 10–3 years of the various bond yield index samples is discussed.

5. CONCLUSION

Implications: The researchers presented a cutting-edge study that was specifically devoted to the bond market. They used it to forecast interest rate spreads between two bonds with various maturities, define a trading strategy based on the predicted values, and assess the prediction models using machine learning and deep learning techniques. To clarify, the researchers proposed use cases for deep learning and machine learning in the Indian bond market that are based on algorithms. The 10- to 3-year Indian Treasury bond spread was reconstructed and collated to take into consideration the time difference between 2013 and 2022, when it was impacted by various economic events. Additionally, the researchers created additional input variables using an exponential function or the value change over the previous few business days to reflect the time series characteristics of the financial market. The parameters of three predictive models were pre-optimized. In the case of a 10-year bond, the researchers discovered that the deep LSTM neural network outperformed linear regression by 65.74% and ARIMA by 99.44%. The prediction models, according to the researchers, were all rather robust and did not depend on the dataset. The deep LSTM neural network outperformed Linear Regression by 89.66% and ARIMA by 99.86% in the case of a 3-year bond. The deep LSTM neural network performed the best in terms of bond price prediction.

Limitations: It is rare to study machine learning applications in the bond market since, in contrast to the stock market, it has high entry barriers and is not friendly to ordinary investors.

Future research: This work is expected to function as an initial and useful research foundation in the future by utilizing machine learning to predict the interest rates of real bonds and a deep neural network model to validate significant earnings. Moreover, a range of data types, including news, bond issuance documents, economic indicators, and documents about monetary policy, contain essential information for assessing each nation's viewpoint on the state of the world economy and the characteristics of the financial markets. As a result, development via different fusion research and interest rate forecasting is possible.

Suggestion: The Bond market traders can utilize this research as a data-driven platform for decision-making to refine their trading strategies or gain a broader understanding of the financial domain. Effective oversight, including budgetary planning, is another application that can benefit all countries and organizations.

Funding: This study received no specific financial support.
Institutional Review Board Statement: Not applicable.
Transparency: The authors state that the manuscript is honest, truthful, and transparent, that no key aspects of the investigation have been omitted, and that any differences from the study as planned have been clarified. This study followed all writing ethics.
Data Availability Statement: The corresponding author can provide the supporting data of this study upon a reasonable request.
Competing Interests: The authors declare that they have no competing interests.
Authors' Contributions: Both authors contributed equally to the conception and design of the study. Both authors have read and agreed to the published version of the manuscript.

REFERENCES

Abad, P., Chuliá, H., & Gómez-Puig, M. (2010). EMU and European government bond market integration. *Journal of Banking & Finance*, *34*(12), 2851-2860. https://doi.org/10.1016/j.jbankfin.2009.10.009

Arrata, W., Nguyen, B., Rahmouni-Rousseau, I., & Vari, M. (2020). The scarcity effect of QE on repo rates: Evidence from the Euro area. *Journal of Financial Economics*, *137*(3), 837-856. https://doi.org/10.1016/j.jfineco.2020.04.009

Asgharian, H., Christiansen, C., & Hou, A. J. (2015). Effects of macroeconomic uncertainty on the stock and bond markets. *Finance Research Letters*, *13*, 10-16. https://doi.org/10.1016/j.frl.2015.03.008

Asher, M. G. (2007). Reforming governance and regulation of urban cooperative banks in India. *Journal of Financial Regulation and Compliance*, *15*(1), 20-29. https://doi.org/10.1108/13581980710726769

Baghestani, H. (2009). Forecasting in efficient bond markets: Do experts know better? *International Review of Economics & Finance*, *18*(4), 624-630. https://doi.org/10.1016/j.iref.2008.10.007

Banerjee, A. K., & Pradhan, H. (2022). Intraday analysis of macroeconomic news surprises, and asymmetries in Indian benchmark bond. *Finance Research Letters*, *45*, 102135. https://doi.org/10.1016/j.frl.2021.102135

Barr, D. G., & Campbell, J. Y. (1997). Inflation, real interest rates, and the bond market: A study of UK nominal and index-linked government bond prices. *Journal of Monetary Economics*, *39*(3), 361-383. https://doi.org/10.1016/s0304-3932(97)00027-5

Bernoth, K., Von Hagen, J., & Schuknecht, L. (2012). Sovereign risk premiums in the European government bond market. *Journal of International Money and Finance*, *31*(5), 975-995. https://doi.org/10.1016/j.jimonfin.2011.12.006

Black, F., & Karasinski, P. (1991). Bond and option pricing when short rates are lognormal. *Financial Analysts Journal*, *47*(4), 52-59. https://doi.org/10.2469/faj.v47.n4.52

Campbell, J. Y., & Taksler, G. B. (2003). Equity volatility and corporate bond yields. *The Journal of Finance*, *58*(6), 2321-2350. https://doi.org/10.1046/j.1540-6261.2003.00607.x

Chen, A.-S., Leung, M. T., & Daouk, H. (2003). Application of neural networks to an emerging financial market: Forecasting and trading the Taiwan stock index. *Computers & Operations Research*, *30*(6), 901-923. https://doi.org/10.1016/s0305-0548(02)00037-0

Choudhry, M. (2003). *The bond and money markets: Strategy, trading, analysis*. Great Britain: Butterworth-Heinemann.

Culkin, R., & Das, S. R. (2017). Machine learning in finance: The case of deep learning for option pricing. *Journal of Investment Management*, *15*(4), 92-100.

De Spiegeleer, J., Madan, D. B., Reyners, S., & Schoutens, W. (2018). Machine learning for quantitative finance: Fast derivative pricing, hedging and fitting. *Quantitative Finance*, *18*(10), 1635-1643. https://doi.org/10.1080/14697688.2018.1495335

Ederington, L., Guan, W., & Yang, L. Z. (2015). Bond market event study methods. *Journal of Banking & Finance*, *58*, 281-293. https://doi.org/10.1016/j.jbankfin.2015.03.013

Favero, C. A. (2013). Modelling and forecasting government bond spreads in the Euro area: A GVAR model. *Journal of Econometrics*, *177*(2), 343-356. https://doi.org/10.1016/j.jeconom.2013.04.004

Ghoddusi, H., Creamer, G. G., & Rafizadeh, N. (2019). Machine learning in energy economics and finance: A review. *Energy Economics*, *81*, 709-727. https://doi.org/10.1016/j.eneco.2019.05.006

Gilmore, C. G., Lucey, B. M., & Boscia, M. W. (2010). Comovements in government bond markets: A minimum spanning tree analysis. *Physica A: Statistical Mechanics and its Applications*, *389*(21), 4875-4886. https://doi.org/10.1016/j.physa.2010.06.057

Gogas, P., Papadimitriou, T., Matthaiou, M., & Chrysanthidou, E. (2015). Yield curve and recession forecasting in a machine learning framework. *Computational Economics*, *45*, 635-645. https://doi.org/10.1007/s10614-014-9432-0

Götze, T., Gürtler, M., & Witowski, E. (2023). Forecasting accuracy of machine learning and linear regression: Evidence from the secondary CAT bond market. *Journal of Business Economics*, 1-32. https://doi.org/10.1007/s11573-023-01138-8

Gubareva, M. (2021). The impact of COVID-19 on liquidity of emerging market bonds. *Finance Research Letters*, *41*, 101826. https://doi.org/10.1016/j.frl.2020.101826

Hirshleifer, D., & Luo, G. Y. (2001). On the survival of overconfident traders in a competitive securities market. *Journal of Financial Markets*, *4*(1), 73-84. https://doi.org/10.1016/s1386-4181(00)00014-8

Huang, Z., Chen, H., Hsu, C.-J., Chen, W.-H., & Wu, S. (2004). Credit rating analysis with support vector machines and neural networks: A market comparative study. *Decision Support Systems*, *37*(4), 543-558. https://doi.org/10.1016/s0167-9236(03)00086-1

Huynh, T. L. D., Hille, E., & Nasir, M. A. (2020). Diversification in the age of the 4th industrial revolution: The role of artificial intelligence, green bonds and cryptocurrencies. *Technological Forecasting and Social Change*, *159*, 120188. https://doi.org/10.1016/j.techfore.2020.120188

Hyndman, R. J., & Koehler, A. B. (2006). Another look at measures of forecast accuracy. *International Journal of Forecasting*, *22*(4), 679-688. https://doi.org/10.1016/j.ijforecast.2006.03.001

Jabeur, S. B., Sadaaoui, A., Sghaier, A., & Aloui, R. (2020). Machine learning models and cost-sensitive decision trees for bond rating prediction. *Journal of the Operational Research Society*, *71*(8), 1161-1179. https://doi.org/10.1080/01605682.2019.1581405

Jobst, A., Kunzel, P., Mills, P., & Sy, A. (2008). Islamic bond issuance: What sovereign debt managers need to know. *International Journal of Islamic and Middle Eastern Finance and Management*, *1*(4), 330-344. https://doi.org/10.1108/17538390810919637

Keim, D. B., & Stambaugh, R. F. (1986). Predicting returns in the stock and bond markets. *Journal of Financial Economics*, *17*(2), 357-390. https://doi.org/10.1016/0304-405x(86)90070-x

Khandani, A. E., Kim, A. J., & Lo, A. W. (2010). Consumer credit-risk models via machine-learning algorithms. *Journal of Banking & Finance*, *34*(11), 2767-2787. https://doi.org/10.1016/j.jbankfin.2010.06.001

Kim, D. H., & Stock, D. (2014). The effect of interest rate volatility and equity volatility on corporate bond yield spreads: A comparison of noncallables and callables. *Journal of Corporate Finance*, *26*, 20–35. https://doi.org/10.1016/j.jcorpfin.2014.02.005

Kim, J., Kim, D. H., & Jung, H. (2021). Applications of machine learning for corporate bond yield spread forecasting. *The North American Journal of Economics and Finance*, *58*, 101540. https://doi.org/10.1016/j.najef.2021.101540

Kim, M. (2021). Adaptive trading system integrating machine learning and back-testing: Korean bond market case. *Expert Systems with Applications*, *176*, 114767. https://doi.org/10.1016/j.eswa.2021.114767

Kolluri, B., Wahab, S., & Wahab, M. (2015). An examination of co-movements of India's stock and government bond markets. *Journal of Asian Economics*, *41*, 39-56. https://doi.org/10.1016/j.asieco.2015.10.001

Krishna, G. D., & Nag, B. (2022). Predictive power of yield curve–Evidence from India. *IIMB Management Review*, *34*(1), 54-67. https://doi.org/10.1016/j.iimb.2022.03.001

Kumar, R. (2014). *Strategies of banks and other financial institutions: Theories and cases*. Oxford: Elsevier.

Liu, F., Kong, D., Xiao, Z., Zhang, X., Zhou, A., & Qi, J. (2022). Effect of economic policies on the stock and bond market under the impact of COVID-19. *Journal of Safety Science and Resilience*, *3*(1), 24-38. https://doi.org/10.1016/j.jnlssr.2021.10.006

Livingston, M., Poon, W. P., & Zhou, L. (2018). Are Chinese credit ratings relevant? A study of the Chinese bond market and credit rating industry. *Journal of Banking & Finance*, *87*, 216-232. https://doi.org/10.2139/ssrn.2893776

Madura, J. (2015). Financial markets and institutions. In (pp. 11 – 32). Stamford: Cengage Learning.

Maghyereh, A. I., & Awartani, B. (2016). Dynamic transmissions between Sukuk and bond markets. *Research in International Business and Finance*, *38*, 246-261. https://doi.org/10.1016/j.ribaf.2016.04.016

Mišura, U. S. (2016). Financial mathematics. In (pp. 6 – 7): ISTE Press.

Paramanik, R. N., & Kamaiah, B. (2017). An empirical analysis of Indian business cycle dynamics. *Economic Annals*, *62*(213), 7-25. https://doi.org/10.2298/eka1713007p

Parameswaran, S. K. (2022). *Fundamentals of financial instruments: An introduction to stocks, bonds, foreign exchange, and derivatives*: John Wiley & Sons.

Paul, M., & Reddy, K. S. (2022). US QE and the Indian bond market. *Journal of Quantitative Economics*, *20*(1), 137-157. https://doi.org/10.1007/s40953-021-00257-9

Pendharkar, P. C., & Cusatis, P. (2018). Trading financial indices with reinforcement learning agents. *Expert Systems with Applications*, *103*, 1-13. https://doi.org/10.1016/j.eswa.2018.02.032

Pozzi, L., & Wolswijk, G. (2012). The time-varying integration of euro area government bond markets. *European Economic Review, 56*(1), 36-53. https://doi.org/10.1016/j.euroecorev.2011.05.006

Prasanna, K., & Sowmya, S. (2017). Yield curve in India and its interactions with the US bond market. *International Economics and Economic Policy, 14*, 353-375.

Raschka, S., & Mirjalili, V. (2019). Python machine learning: Machine learning and deep learning with Python, scikit-learn, and TensorFlow 2. In (pp. 71 – 83): Packt Publishing Ltd.

Reddy, Y. V. (2002). *Issues and challenges in the development of the debt market in India. BIS Papers, 11*. Retrieved from https://www.bis.org/publ/bppdf/bispap11j.pdf

Schuknecht, L., Von Hagen, J., & Wolswijk, G. (2009). Government risk premiums in the bond market: EMU and Canada. *European Journal of Political Economy, 25*(3), 371-384. https://doi.org/10.1016/j.ejpoleco.2009.02.004

Sezer, O. B., Gudelek, M. U., & Ozbayoglu, A. M. (2020). Financial time series forecasting with deep learning: A systematic literature review: 2005–2019. *Applied Soft Computing, 90*, 106181. https://doi.org/10.1016/j.asoc.2020.106181

Shankar, R., & Bose, S. (2008). *Auctions of government securities in India: An analysis*: Reserve Bank of India.

Sharm, A. (2016). Exchange rate and business cycles in India: An analysis of lead and lag behavior. *ZENITH International Journal of Multidisciplinary Research, 6*(9), 70-88.

Thumrongvit, P., Kim, Y., & Pyun, C. S. (2013). Linking the missing market: The effect of bond markets on economic growth. *International Review of Economics & Finance, 27*, 529-541. https://doi.org/10.1016/j.iref.2013.01.008

Trippi, R. R., & Turban, E. (1992). *Neural networks in finance and investing: Using artificial intelligence to improve real-world performance.* New York, United States: McGraw-Hill, Inc.

Vijh, M., Chandola, D., Tikkiwal, V. A., & Kumar, A. (2020). Stock closing price prediction using machine learning techniques. *Procedia Computer Science, 167*, 599-606. https://doi.org/10.1016/j.procs.2020.03.326

Wahidin, D., Akimov, A., & Roca, E. (2021). The impact of bond market development on economic growth before and after the global financial crisis: Evidence from developed and developing countries. *International Review of Financial Analysis, 77*, 101865. https://doi.org/10.1016/j.irfa.2021.101865

Webster, C., & Ivanov, S. (2020). Robotics, artificial intelligence, and the evolving nature of work In (pp. 127-143). Cham: Springer International Publishing.

Wisniewski, T. P. (2016). Is there a link between politics and stock returns? A literature survey. *International Review of Financial Analysis, 47*, 15-23. https://doi.org/10.2139/ssrn.2690852

Xu, S., & Chan, H. K. (2019). Forecasting medical device demand with online search queries: A big data and machine learning approach. *Procedia Manufacturing, 39*, 32-39. https://doi.org/10.1016/j.promfg.2020.01.225

Yan, X., & Su, X. (2009). *Linear regression analysis: Theory and computing*. Singapore: World Scientific.

Zaremba, A., Kizys, R., Aharon, D. Y., & Umar, Z. (2022). Term spreads and the COVID-19 pandemic: Evidence from international sovereign bond markets. *Finance Research Letters, 44*, 102042. https://doi.org/10.2139/ssrn.3727882

Zhou, J., Chen, C., Li, L., Zhang, Z., & Zheng, X. (2022). FinBrain 2.0: When finance meets trustworthy AI. *Frontiers of Information Technology & Electronic Engineering, 23*(12), 1747-1764. https://doi.org/10.1631/fitee.2200039