2025 3rd International Conference on Intelligent Systems, Advanced Computing and Communication (ISACC) | 979-8-3315-2389-3/25/$31.00 ©2025 IEEE | DOI: 10.1109/ISACC65211.2025.10969378

Enhancing Investment Decisions: A Machine Learning Approach to Recommending Stocks and Mutual Funds

1st Mrinmoy Borah
Royal School of Information Technology
The Assam Royal Global University
Guwahati, India
mrinmoyborahdev@gmail.com

2nd Afsana Laskar
Royal School of Information Technology
The Assam Royal Global University
Guwahati, India
laskar.afs@gmail.com

***Abstract*—Financial literacy is an essential competency in contemporary society, enabling individuals to make informed financial decisions and attain economic well-being. Nevertheless, despite its importance, attaining financial literacy continues to pose a significant problem for numerous individuals. Multiple factors contribute to this challenge, including the intricacy of financial ideas and the extensive amount of information accessible. The financial landscape is ever-changing, with new products, laws, and investing methods frequently arising. Adapting to these changes can be daunting for the typical person. To tackle this issue, this paper investigates various machine learning and deep learning algorithms and techniques with the goal of developing a recommendation system that can suggest stocks and mutual funds in the Indian financial market to users. Recommendation systems represent a successful commercial application of machine learning. It selectively filters information based on the user's preferences from the vast array of possibilities and suggests options that are optimal or highly probable for the user. We employed Cosine similarity and Pearson correlation to ascertain similarity metrics between two stocks and mutual funds, in conjunction with a Euclidean distance-based method, KNN. We have incorporated three clustering algorithms: K-Means, Agglomerative, and DBSCAN to categorize related mutual funds. We used LSTM for predicting stock prices and Autoencoders for making recommendations. Among all clustering techniques, DBSCAN had the highest silhouette score of 0.85. The LSTM model exhibited a mean training loss of 0.000226 and a mean validation loss of 0.000658 for Nifty-50 businesses, derived from training and testing performed on data spanning the last 20 years.**

***Index Terms*—recommendation, cosine similarity, pearson correlation, KNN, K-Means, agglomerative, DBSCAN, LSTM, autoencoders**

I. INTRODUCTION

The expertise and competencies required for effective financial management have grown progressively essential in today's intricate economic landscape. Nonetheless, it continues to be an unexpectedly intricate and frequently misconstrued subject. Despite the efforts of financial institutions and governments to advocate for financial literacy, numerous individuals continue to grapple with fundamental financial concepts, resulting in suboptimal financial decisions. A contributing factor to this lack of comprehension is the complex structure of finance itself. Terminology and complexities frequently obscure financial instruments such as equities, fixed-income securities, and insurance contracts, rendering them challenging for the typical individual to understand. Moreover, conventional education institutions frequently overlook financial literacy, concentrating predominantly on academic disciplines. This renders folks inadequately prepared to manage their finances proficiently. Consequently, numerous individuals depend on disinformation, fallacies, or instinctual sensations while making financial choices.

Our primary goal is to create a stock and mutual fund recommendation system that provides suggestions to users based on their interests, resources, and risk tolerance. Recommendation systems are models that deliver individualized information and suggestions to users, considering their prior selections and preferences. The recommendation system executes information filtering to enhance user satisfaction and conserves time by alleviating the arduous process of research and item selection [1]–[3]. Commonly used recommendation methods encompass content-based filtering, collaborative filtering, and hybrid filtering. Content-based filtering suggests items by assessing their resemblance to items with which a user has already engaged or rated favorably. It examines the attributes of things, including genres, keywords, or characteristics, and formulates a user profile derived from their historical preferences. Collaborative filtering suggests things based on the interests of analogous users. It examines the historical engagements of users with things, such ratings or purchases, and identifies users with analogous preferences. The system subsequently recommends products favored by analogous users [4]. Hybrid filtering integrates the advantages of both content-based and collaborative filtering to deliver more precise and tailored recommendations [5]–[7]. This research explores various algorithms to build a recommendation system that suggests stocks or mutual funds to users based on their preferences [8]. To develop an optimal recommendation system, we employed various machine learning and deep learning algorithms, including Cosine similarity, Pearson correlation, K-nearest neighbors, K-means clustering, Agglomerative hierarchical clustering, Density-Based Spatial Clustering of Applications with Noise (DBSCAN), Autoencoders, and Long Short-Term Memory (LSTM).

The following sections of the paper are organized as follows: Section 2 examines the existing literature. Section 3 outlines the research methodology. Section 4 presents the findings. Section 5 presents the limitations and finally, Section 6 presents a synopsis of the paper.

II. Literature Review

Recommendation systems have discretely influenced our decisions for ages, ranging from a bookstore proposing a classic based on our previous purchase to a buddy endorsing a film we may appreciate. Currently, driven by machine learning, these systems have developed into advanced instruments that influence our digital interactions. In the domain of e-commerce, recommendation systems are the overlooked catalysts of sales. Through the analysis of user behavior and preferences, they generate tailored product recommendations, enhancing client satisfaction and augmenting revenue. In the entertainment sector, they compile playlists, recommend films, and customize content to suit personal preferences. Recommendation systems, encompassing personalized investments and adaptive asset allocation, can enhance economic growth by promoting innovations and advancing financial inclusion within the financial sector.

Studies on recommendation systems for financial services has highlighted their uses and essential role in banking, the stock market, insurance, and real estate. It also identified the two predominant issues with the recommendation system: the user cold start problem and the item cold start problem [9]. Researchers employed several machine learning algorithms to develop stock recommendation and risk management systems. These encompassed hierarchical clustering, K-nearest neighbors (KNN), singular value decomposition (SVD), and association rule mining (ARM) [10]. Numerous studies employed deep learning algorithms, including the bilateral variational autoencoder (BiVAE) model, to produce personalized stock recommendations and utilized the variational autoencoder (VAE) model within the content-based recommendation framework to analyze the representation of latent features in financial instruments [11]. Researchers employed a deep learning framework, the EPRSA model, which utilizes query sequence data from a user's click-through personalized recommender system alongside topic-based information regarding the user's financial history and product choices. This method employed a framework utilizing LSTM models, a structure grounded in LDA topic models, and a self-attention mechanism to autonomously discern the distribution of both short-term and long-term user preferences for financial items [12], [13]. A restricted number of studies focused on stock market prediction and financial management decision-making, employing RCA, Bidirectional Long Short-Term Memory Network (BiLSTM), and Deep Q Network (DQN) in conjunction [14]. A study employed Autoregressive Integrated Moving Average (ARIMA) for a stock recommendation system. Time series analysis and forecasting employ ARIMA as a statistical technique. It is a robust instrument for forecasting future values of a time series based on historical behavior, in conjunction with the subsequent machine learning algorithms for backend processes: Logistic Regression, Multinomial Naive Bayes, and Passive Aggressive Classifier algorithms for stock recommendation systems [15]. The utilization of LSTM and CNN-BiLSTM hybrid models in financial market prediction has been addressed by only a few researchers, who have underscored their potential to discern nonlinear dynamics in fluctuating markets [16]. A study examining stock price predictions employed MLP, RNN, and LSTM approaches on the National Stock Exchange (NSE) of India and the New York Stock Exchange (NYSE). Remarkably, the CNN model excelled, effectively forecasting NYSE stock prices even though it was trained exclusively on data from the NSE [17].

Few studies focused on personalized, sentiment and trust based recommendation systems. Personalized Dynamic Recommender System for Investors (PDRSI). The suggested PDRSI takes into account two personal characteristics of investors: dynamic preferences and historical interests, as well as two temporal environmental factors: recent debates on social media and the current market data. The Sentiment and Price Combined Model (SPCM), employing BERT transformer models and sophisticated machine learning methodologies, integrates sentiment attributes and price elements to predict stock price variations [18]–[20]. A Large Language Model (LLM) is a form of artificial intelligence capable of comprehending and producing human language text. It operates by examining extensive datasets of language to discern patterns and links among words and sentences. A study employed A-LLMRec, which demonstrates proficiency in both cold and warm scenarios. The main objective of the study was to enable a large language model (LLM) to directly harness the collaborative knowledge embedded in a pre-trained, state-of-the-art collaborative filtering recommendation system (CF-RecSys), thereby synergistically utilizing the LLM's emergent capabilities alongside the high-quality user and item embeddings established by the CF-RecSys [21].

This study focuses on various machine-learning and deep-learning algorithms to build a recommendation system that can recommend stocks and mutual funds in the Indian financial market to users based on their past preferences.

III. Methodology

A. Data Collection

The procedure of gathering information is referred to as data collection. It entails collecting pertinent data from diverse sources to train and evaluate algorithms. The caliber and volume of data profoundly influence the model's efficacy.

- **Mutual Funds Dataset I**: This dataset was obtained from Kaggle and comprises information about mutual funds documented in April 2023. This dataset includes 814 mutual funds and 20 attributes, such as mutual fund names, minimum SIP, minimum investments, expense ratio, sortino ratio, sharpe ratio, alpha, beta, risk level, fund size, fund manager, fund age, standard deviation, AMC names, ratings, category, subcategory, and returns over 1 year, 3 years, and 5 years.

- **Mutual Funds Dataset II**: We gathered this dataset from Kaggle, containing information about mutual funds recorded during April 2023. This dataset comprises 2556 mutual funds and 11 features, which include the names of the funds, their categories, types, benchmarks, net asset values, expense ratios, compound interest over 6 months, 1 year, and 3 years, as well as minimum SIP and investment amounts.
- **Stocks Dataset I**: We carefully curated this dataset for the stock recommendation system. We collected our data from the NSE (National Stock Exchange) and money-control website and built a custom dataset on the data collected. The dataset includes the following features: stock name, sectors, market capitalization, risk, volatility, and one-year return (2023-2024) for over 200 companies across more than 80 sectors.
- **Stocks Dataset II**: This dataset was obtained from Kaggle and contains information on Nifty-50 companies spanning the last 20 years, specifically from 1st January, 2000 to 30th April, 2021. This dataset comprises 50 subdatasets, each featuring the following attributes: open, close, high, low, adjusted close, volume, and date for each company.

B. Preprocessing

Preprocessing is an essential phase in the machine learning workflow. It entails the cleansing, transformation, and preparation of raw data to render it appropriate for machine learning algorithms. The objective is to enhance the quality and consistency of the data, resulting in more precise and dependable models.

- **Mutual Funds Dataset I**: Following preprocessing, we eliminated missing data, duplicates, and unnecessary features. The preprocessed dataset comprises 612 mutual funds and 16 attributes: mutual fund names, minimum SIP, minimum investment, expense ratio, sortino ratio, sharpe ratio, alpha, beta, risk level, AMC names, ratings, category, subcategory, and returns for 1 year, 3 years, and 5 years. We employed one-hot encoding to convert the categorical variables of AMC names, category, and subcategory into numerical values. After converting all necessary features into numerical values, we employed the scikit-learn library for normalization.
- **Mutual Funds Dataset II**: In accordance with the prior dataset, Mutual Funds Dataset I, we have preprocessed our current dataset, which comprises 1,766 mutual funds and 9 features. The characteristics encompass the names of mutual funds, their classifications, benchmarks, net asset values, fee ratios, compound interest rates for 6 months and 1 year, minimum systematic investment plans, and minimum investment amounts.
- **Stocks Dataset I**: This custom dataset has been verified for high data quality, containing no missing values or outliers. We calculated the volatility of each stock based on daily returns throughout the three-year period from 2021 to 2024 and categorized the stocks as high risk if the volatility exceeds 0.40, moderate risk if the volatility exceeds 0.25 and less than 0.40, and low risk otherwise. Ultimately, we executed one-hot encoding on categorical features and normalized our dataset utilizing the Scikit-learn library.
- **Stocks Dataset II**: The dataset included no missing entries or duplicates; we converted the Date column's datatype from object to datetime64 and sorted it in ascending order. We employed the MinMaxScaler from the Scikit-learn library to normalize the "close" feature of the dataset to a designated range, typically between 0 and 1.

C. Training and Testing the Model

Conventional train-test splitting is inapplicable for constructing an item-based content recommendation system, as cosine similarity and pearson correlation compute similarity scores between items based on their properties, independent of prior user interaction data. For clustering methods such as K-Means, we employed the elbow curve to determine the ideal value of K; for agglomerative clustering algorithm and DBSCAN, we utilized the silhouette score to create high-quality clusters. For KNN and LSTM, we employed a train-test split, allocating 80% of the data for training and 20% for testing. We utilized the Python packages Pandas, Scikit-learn, and TensorFlow for the training and evaluation of our models.

D. Algorithms

- **Cosine Similarity**: Cosine similarity is a mathematical metric that quantifies the similarity between two vectors in a multi-dimensional space. We compute it by ascertaining the cosine of the angle between the two vectors. It elucidates the link between two items by examining the "direction" they indicate, rather than merely contrasting their individual values. The formula of cosine similarity is as follows:

$$\cos(\theta) = \frac{\mathbf{A} \cdot \mathbf{B}}{||\mathbf{A}|| \cdot ||\mathbf{B}||}$$

- **Pearson Correlation**: Pearson correlation is a statistical metric that quantifies the linear association between two variables. It is a value ranging from -1 to 1 that signifies the magnitude and orientation of the relationship. -1 indicates a negative correlation, meaning that when one variable increases, the other falls linearly. 0 indicates the absence of linear correlation, signifying no association between the variables, while +1 denotes perfect positive correlation, implying that an increase in one variable results in a proportional increase in the other. The formula of pearson correlation is as follows:

$$r = \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^{n}(x_i - \bar{x})^2}\sqrt{\sum_{i=1}^{n}(y_i - \bar{y})^2}}$$

- **K-nearest Neighbors**: KNN is a simple, non-parametric, supervised learning machine learning algorithm, applicable to both classification and regression problems. It works on the premise that comparable data points typically yield comparable results. After determining the distance between the new point and every data point that already exists, it chooses the K points that are the closest. The categorization is based on the majority class of these K neighbors, or the regression prediction is based on the average value.
- **K-Means Clustering**: K-Means Clustering is an unsupervised machine learning approach designed to aggregate analogous data points into clusters. It is a prevalent method for categorizing data into separate groups according to its characteristics. It is a centroid-centric algorithm, wherein each cluster is linked to a centroid. The algorithm accepts an unlabeled dataset as input, partitions the dataset into K clusters, then iterates the process until optimal clusters are identified. This algorithm must ascertain the value of K in advance.
- **Agglomerative Clustering**: Agglomerative clustering, or hierarchical clustering, is an unsupervised machine learning technique that organizes comparable data points into clusters. This is a bottom-up strategy, initiating with each data point as an individual cluster and thereafter merging the nearest cluster pairs iteratively until all data points are consolidated into a single cluster or a predetermined number of clusters is achieved.
- **DBSCAN**: DBSCAN, an acronym for Density-Based Spatial Clustering of Applications with Noise, identifies clusters of densely packed points, or high-density zones. In contrast to distance-based clustering techniques such as K-means, DBSCAN does not necessitate the prior specification of the number of clusters. DBSCAN can identify clusters of various shapes, in contrast to distance-based algorithms that often presume spherical clusters.
- **Autoencoders**: Autoencoders are neural networks designed to replicate input data. They comprise two primary components: an encoder and a decoder. The encoder transforms the input data into a lower-dimensional latent space representation, while the decoder reconstructs the original input data from this latent representation. The principal elements of an autoencoder are the encoder, decoder, and loss function. An encoder translates input data into a latent space. The decoder reconstructs the original input data from the latent space, while the loss function quantifies the discrepancy between the original input and the reconstructed output.
- **LSTM**: Long Short-Term Memory, is a recurrent neural network (RNN) architecture designed to process sequential data, such as time series or text, where temporal links between elements are essential. When traversing time in reverse, conventional RNNs encounter the ”vanishing gradient” issue, resulting in gradients diminishing to a level that hinders effective learning. LSTM networks, conversely, excel at acquiring long-term dependencies. The gated method enables LSTMs to preserve information for extended durations, rendering them appropriate for tasks like as language modeling, machine translation, and financial forecasting.

IV. Results and discussions

- **Mutual Funds Dataset I**: We utilized cosine similarity, pearson correlation, KNN, K-Means clustering, agglomerative clustering, and DBSCAN on the Mutual Funds Dataset I. Table 1 delineates the suggestions of the cosine similarity method, whereas Table 2 delineates the recommendations of the pearson correlation algorithm. We supplied the ”Franklin Build India Fund” mutual fund as input for these two algorithms, and both algorithms suggested mutual funds that nearly mirrored the original fund. We delineate the features of Table 1 and Table 2 as follows: ”SN” denotes Scheme Name, ”MSIP” refers to Minimum SIP, ”MI” signifies Minimum Investment, ”RL” indicates Risk Levels, ”R” represents Ratings, and ”R5” pertains to Returns over 5 years.
  A common challenge in recommendation systems is the cold start problem. A cold start in recommendation systems occurs when there is insufficient data on new users or objects to produce accurate recommendations. The lack of historical interaction data impedes the capacity to provide individualized recommendations, as the algorithm cannot effectively identify user preferences or item popularity. We employed three clustering algorithms: K-Means, agglomerative, and DBSCAN to tackle the cold start issue. We utilized the silhouette score to assess the efficacy of a clustering technique. The silhouette score is maximized in DBSCAN, achieving a value of 0.85 for two clusters, whereas it is minimized in K-Means and the agglomerative method. Fig. 1. depicts the dendrogram of agglomerative clustering. Classifying things based on their attributes enables the system to aggregate similar items together. The system can assign a new item to an appropriate cluster, thereby enhancing suggestions for users who have expressed interest in similar products. Clustering approaches enhance the management of sparse interaction data by identifying similar users or products, hence increasing recommendation accuracy in cold start scenarios.
- **Mutual Funds Dataset II**: We utilized autoencoders and cosine similarity in this dataset. Autoencoders facilitate clustering by converting high-dimensional, intricate data into a compact, lower-dimensional representation that is more amenable to clustering methods. We trained our autoencoder utilizing mean squared error (MSE) as the loss function for 50 epochs, yielding a final training loss of 0.0030 and a validation loss of 0.0034. The low mean square error (MSE) values indicate that the autoencoder effectively reconstructed the input features. The compressed representations retain a significant amount of the original data's information, rendering them effective for clustering. Ultimately, we utilized cosine similarity to

encode the characteristics. We provided the ”Aditya Birla SL Money Manager Fund (G)” as an input, and our model recommended three mutual funds that closely resembled the input, based on a comparison of their features. Table 3 delineates the recommendations of autoencoders and cosine similarity algorithms. In this table, ”S” refers to Scheme Names, ”C” refers to Categories, ”NAV” stands for Net Asset Value, ”C1” refers to Compound Interest over 1 year in per cent, ”ER” stands for Expense Ratio, and ”MI” stands for Minimum Investments.

- **Stocks Dataset I**: We applied cosine similarity and pearson correlation to this dataset. Table 4 presents stock recommendations based on cosine similarity, while Table 5 provides recommendations derived from pearson correlation. Below, we outline the features of Tables 4 and 5. ”Stock,” ”Sector,” ”MktC” denote Market Capitalization, categorized into three types: large cap (LC), medium cap (MC), and small cap (SC). ”R” signifies Risk, classified as high (H), moderate (M), and low (L). ”V” indicates Volatility. ”R(23-24)” refers to the Returns for the year 2023-2024 of a specific stock. We utilized ”TATPOWER” as input for the cosine similarity algorithm and ”SOBHA” as input for the pearson correlation algorithm.
- **Stocks Dataset II**: We utilized the Long-Short-Term Memory (LSTM) algorithm to forecast stock market prices and implemented the mean squared error (MSE) as the loss function during training and validation. The Mean Squared Error (MSE) is the average of the squared discrepancies between the actual and anticipated values. The LSTM model attained a mean training loss of 0.000226 and a mean validation loss of 0.000658. Although both losses are minimal, the model appears to efficiently capture trends; nonetheless, the disparity suggests it may encounter difficulties with very volatile patterns or unforeseen market shocks. This demonstrates the model’s efficacy in capturing historical trends while also identifying areas for enhancement to improve generality. Fig. 2. describes the LSTM training and prediction of company ”TCS” and Fig. 3. describes the LSTM training and prediction of company ”RELIANCE”. Fig. 4. depicts mean training loss and validation loss over Nifty-50 companies.

TABLE I
RECOMMENDATIONS OF MUTUAL FUNDS USING COSINE SIMILARITY ALGORITHM

SN	MSIP	MI	RL	R	R5
Franklin Build India Fund	500	5000	3	3	13.3
Franklin India Opportunities Fund	500	5000	3	0	11.5
Franklin India Technology Fund	500	5000	3	0	15.9
Franklin India Tax-shield Direct Growth	500	500	3	3	11.4

TABLE II
RECOMMENDATIONS OF MUTUAL FUNDS USING PEARSON CORRELATION ALGORITHM

SN	MSIP	MI	RL	R	R5
Franklin Build India Fund	500	5000	3	3	13.3
SBI Infrastructure Fund	500	5000	3	4	13.0
Templeton India Value Fund	500	5000	3	3	12.3
Franklin India Smaller Companies Fund	500	5000	3	2	11.2

TABLE III
RECOMMENDATIONS OF AUTOENCODERS AND COSINE SIMIALRITY ALGORITHM

SN	C	NAV	C1	ER	MSIP
Aditya Birla SL Money Manager Fund(G)	Debt - Money Market Fund	309.99	5.18	0.33	1000
ICICI Pru Money Market Fund(G)	Debt - Money Market Fund	318.04	5.10	0.31	1000
ICICI Pru Money Market Fund(W-IDCW)	Debt - Money Market Fund	100.74	5.10	0.31	1000
ICICI Pru Money Market Fund(DD-IDCW)	Debt - Money Market Fund	100.13	5.10	0.31	1000

TABLE IV
RECOMMENDATIONS OF STOCKS USING COSINE SIMILARITY ALGORITHM

Stock	Sector	MktC	R	V	R(23-24)
TATAPOWER	Integrated Power Utilities	LC	M	0.38	56.28
LTIM	Computers - Software & Consulting	LC	M	0.33	43.98
TORNTPHARM	Pharmaceuticals	LC	M	0.38	48.65
TECHM	Computers - Software & Consulting	LC	M	0.28	27.70

TABLE V
RECOMMENDATIONS OF STOCKS USING PEARSON CORRELATION ALGORITHM

Stock	Sector	MktC	R	V	R(23-24)
SOBHA	Residential Commercial Projects	SC	H	0.48	72.42
NBCC	Civil Construction	SC	H	0.45	99.15
J&KBANK	Private Sector Bank	SC	H	0.49	119.63
ARROWGREEN	Packaging	SC	H	0.97	144.20

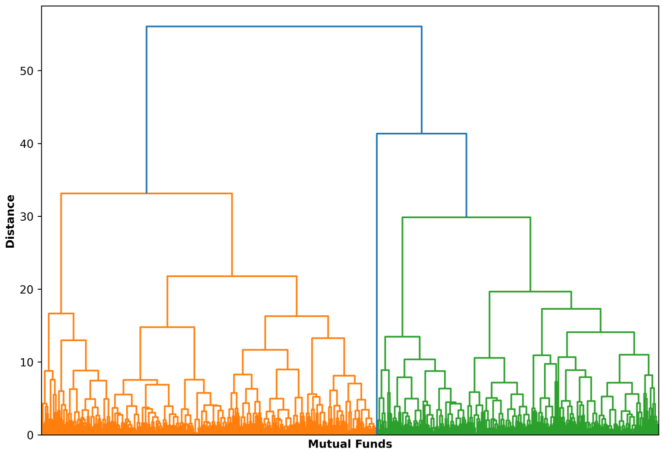

Fig. 1. Dendrogram of agglomerative clustering

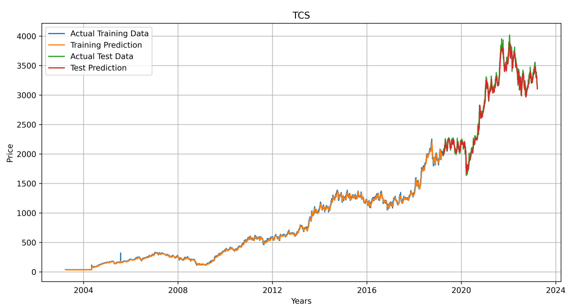

Fig. 2. LSTM model's performance on TCS stock data

Fig. 3. LSTM model's performance on RELIANCE stock data

V. Limitations

The recommendation system is subject to limitations and potential biases due to the fluid and unpredictable nature of financial markets. While analyzing a 20-year historical stock market dataset provides a broad perspective, it may not fully capture the impact of extraordinary market events, economic shifts, or geopolitical dynamics that can significantly

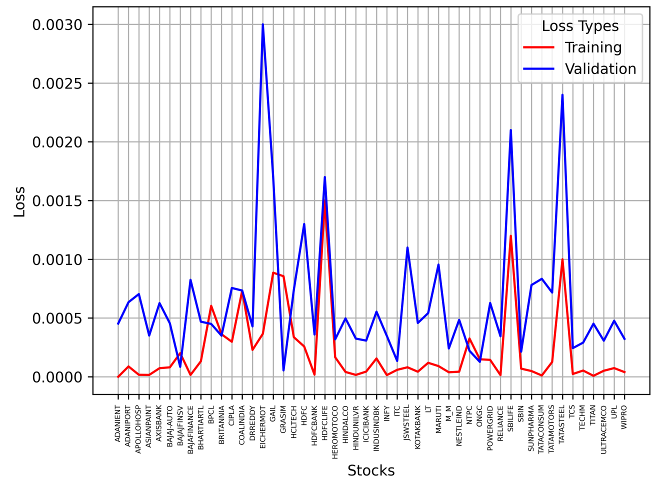

Fig. 4. LSTM model's training loss vs. validation loss

influence the performance of stocks and mutual funds. Biases may also arise from the selection of features and metrics, such as risk levels, returns, and correlation measures, which might overlook qualitative factors like investor sentiment or regulatory changes. These limitations highlight the need for continuous algorithm refinement, real-time data integration, and consideration of external variables to enhance the system's effectiveness and reliability in dynamic financial contexts.

VI. Conclusion And Future Scope

Stock and mutual fund recommendation systems can improve financial literacy by offering tailored investment guidance based on an individual's risk tolerance, financial objectives, and investment timeline. This can clarify the investment process and render it more accessible to those with diverse levels of financial acumen. It is essential to remember that these systems function as tools, and their application should enhance prudent financial counsel and thorough study.

We sourced our data from Kaggle, NSE and moneycontrol, conducted preprocessing, and implemented the following algorithms: Cosine similarity, Pearson correlation, KNN, K-means, Agglomerative, DBSCAN, LSTM, and Autoencoders. The stock and mutual fund recommendation systems accurately suggested comparable stocks based on the provided stock and mutual fund inputs. We employed clustering methods to address the cold start problem and utilized LSTM for stock price forecasting. Of all clustering techniques, DBSCAN had the highest silhouette score of 0.85. The LSTM model demonstrated a mean training loss of 0.000226 and a mean validation loss of 0.000658 for Nifty-50 companies.

Future research may include recommendations for stocks and mutual funds derived from real-time data and an analysis of evolving market trends. We are considering evaluating the user's financial's and suggesting stocks based on their income and expenditure trends.

REFERENCES

[1] Hyeyoung Ko, Suyeon Lee, Yoonseo Park, and Anna Choi. A survey of recommendation systems: recommendation models, techniques, and application fields. *Electronics*, 11(1):141, 2022.

[2] Zeshan Fayyaz, Mahsa Ebrahimian, Dina Nawara, Ahmed Ibrahim, and Rasha Kashef. Recommendation systems: Algorithms, challenges, metrics, and business opportunities. *applied sciences*, 10(21):7748, 2020.

[3] Lalita Sharma and Anju Gera. A survey of recommendation system: Research challenges. *International Journal of Engineering Trends and Technology (IJETT)*, 4(5):1989–1992, 2013.

[4] Jesús Bobadilla, Jorge Dueñas-Lerín, Fernando Ortega, and Abraham Gutierrez. Comprehensive evaluation of matrix factorization models for collaborative filtering recommender systems. *arXiv preprint arXiv:2410.17644*, 2024.

[5] Lipi Shah, Hetal Gaudani, and Prem Balani. Survey on recommendation system. *International Journal of Computer Applications*, 137(7):43–49, 2016.

[6] Luong Vuong Nguyen, Quoc-Trinh Vo, and Tri-Hai Nguyen. Adaptive knn-based extended collaborative filtering recommendation services. *Big Data and Cognitive Computing*, 7(2):106, 2023.

[7] Maxim Naumov, Dheevatsa Mudigere, Hao-Jun Michael Shi, Jianyu Huang, Narayanan Sundaraman, Jongsoo Park, Xiaodong Wang, Udit Gupta, Carole-Jean Wu, Alisson G Azzolini, et al. Deep learning recommendation model for personalization and recommendation systems. *arXiv preprint arXiv:1906.00091*, 2019.

[8] Nandini Seal and Soumya Mukherjee. A study on the performance of mutual funds of indian amcs. *Management Journal for Advanced Research*, 2:16–23, 2022.

[9] Marwa Sharaf, Ezz El-Din Hemdan, Ayman El-Sayed, and Nirmeen A El-Bahnasawy. A survey on recommendation systems for financial services. *Multimedia Tools and Applications*, 81(12):16761–16781, 2022.

[10] Reyes Michaela Denise Gonzales and Carol Anne Hargreaves. How can we use artificial intelligence for stock recommendation and risk management? a proposed decision support system. *International Journal of Information Management Data Insights*, 2(2):100130, 2022.

[11] Narada Wijerathne, Jamini Samarathunge, Krishalika Rathnayake, Supuni Jayasinghe, Sapumal Ahangama, Indika Perera, Vinura Dhananjaya, and Lushanthan Sivaneasharajah. Deep learning based personalized stock recommender system. In *International Conference on Neural Information Processing*, pages 362–374. Springer, 2023.

[12] Xiaowen Yue. Application of ai technology in personalized recommendation system for financial services. *Applied Mathematics and Nonlinear Sciences*, 9(1), 2023.

[13] Vikram Sharma, Manik Rakhra, and Gauri Mathur. Hybrid approaches for stocks prediction and recommendation system. In *E3S Web of Conferences*, volume 453, page 01047. EDP Sciences, 2023.

[14] Yijing Huang and Vinay Vakharia. Deep learning-based stock market prediction and investment model for financial management. *Journal of Organizational and End User Computing (JOEUC)*, 36(1):1–22, 2024.

[15] Jashanjit Singh and John Jenq. A stock recommendation system. *EPiC Series in Computing*, 91:164–171, 2023.

[16] Haotian Zheng, Jiang Wu, Runze Song, Lingfeng Guo, and Zeqiu Xu. Predicting financial enterprise stocks and economic data trends using machine learning time series analysis. *Applied and Computational Engineering*, 87:26–32, 2024.

[17] Bishnu Padh Ghosh, Mohammad Shafiquzzaman Bhuiyan, Debashish Das, Tuan Ngoc Nguyen, Rasel Mahmud Jewel, Md Tuhin Mia, Duc M Cao, and Rumana Shahid. Deep learning in stock market forecasting: Comparative analysis of neural network architectures across nse and nyse. *Journal of Computer Science and Technology Studies*, 6(1):68–75, 2024.

[18] Takehiro Takayanagi, Chung-Chi Chen, and Kiyoshi Izumi. Personalized dynamic recommender system for investors. In *Proceedings of the 46th International ACM SIGIR Conference on Research and Development in Information Retrieval*, pages 2246–2250, 2023.

[19] Jiawei Wang and Zhen Chen. Spcm: A machine learning approach for sentiment-based stock recommendation system. *IEEE Access*, 2024.

[20] C Prem Sankar, R Vidyaraj, and K Satheesh Kumar. Trust based stock recommendation system–a social network analysis approach. *Procedia Computer Science*, 46:299–305, 2015.

[21] Sein Kim, Hongseok Kang, Seungyoon Choi, Donghyun Kim, Minchul Yang, and Chanyoung Park. Large language models meet collaborative filtering: an efficient all-round llm-based recommender system. In *Proceedings of the 30th ACM SIGKDD Conference on Knowledge Discovery and Data Mining*, pages 1395–1406, 2024.

[22] Chaher Alzaman. Deep learning in stock portfolio selection and predictions. *Expert Systems with Applications*, 237:121404, 2024.

[23] Swarna Prabha Jena, Ankit Kumar Yadav, Dhiraj Gupta, and Bijay Kumar Paikaray. Prediction of stock price using machine learning techniques. In *2023 IEEE 2nd International Conference on Industrial Electronics: Developments & Applications (ICIDeA)*, pages 169–174. IEEE, 2023.

[24] Artur Bugaj and Weronika T Adrian. Real-time recommendation system for stock investment decisions. In *WEBIST*, pages 490–493, 2021.

[25] Aakash Agarwal. Study of machine learning algorithms for potential stock trading strategy frameworks. *International Journal of Financial, Accounting, and Management*, 3(3):275–287, 2021.

[26] Chae-eun Park, Dong-seok Lee, Sung-hyun Nam, Soon-kak Kwon, et al. Implementation of fund recommendation system using machine learning. *Journal of Multimedia Information System*, 8(3):183–190, 2021.

[27] Nithya Sampath, Jayakumar Sadhasivam, R Raj Kumar, M Sathish Kumar, Balajee Jeyakumar, and PV PraveenSundar. Analyzing financial data and mutual funds recommendation by using big data analytics. *Journal of Computational and Theoretical Nanoscience*, 16(5-6):2414–2418, 2019.

[28] Weina Wang and Krishn Kumar Mishra. A novel stock trading prediction and recommendation system. *Multimedia Tools and Applications*, 77(4):4203–4215, 2018.