International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 9
***Article Received:** 25 July 2023 **Revised:** 12 September 2023 **Accepted:** 30 September 2023*

A Model for Predicting E-Commerce Product Returns Using Hybrid CNN-GRU

A.Arulmurugan[1], Arul N[2], Meera Alphy[3], Vidya Rajasekaran[*4], Kuncham Sreenivasa Rao[5], Karumuri Sri Rama Murthy[6]

[1] Department of Computing Technologies, School of Computing
[1] SRM Institute of science and Technology, kattankulathur , chennai,Tamilnadu, India-603203
[1] arulmura1@srmist.edu.in

[2] Department of Computer Science and Engineering
[2] St.Peter's Engineering College,Hyderabad, India
[2] arulthala82@gmail.com

[3] Department of Computer Science and Engineering,
[3] Mahatma Gandhi Institute of Technology, Gandipet, Hyderabad,India
[3] drmeeraalphy@gmail.com

[4] Department of Information Technology
[4] B.S. Abdur Rahman Crescent Institute of Science and Technology, Chennai, India
[*4] Corresponding author e-mail: vidyarajesh23@gmail.com

[5] Department of Computer Science and Engineering,
[5] Faculty of Science and Technology, (IcfaiTech), ICFAI Foundation for Higher Education, Hyderabad, India
[5] ksrao517@gmail.com

[6] Department of Computer Science and Engineering,
[6] VNR Vignana Jyothi Institute of Engineering and Technology, Hyderabad,India
[6] sriram.jntuphd@gmail.com

Abstract—The aim of this work is to predict the product return rate of e-commerce using deep learning algorithms. A novel hybrid model combining Convolutional Neural Network(CNNs) and Gated Recurrent Units (GRUs) is proposed to predict e-commerce product return rates. The developed model is trained and validated on a large e-commerce dataset with features like consumer demographic informations, product details, transaction and product return details, and consumer feedbacks. The return history of every product is learned and products least and most sold is analyzed. Based on all this analysis the return of products in the future is predicted. The proposed work demonstrates that the hybrid CNN-GRU model outperforms conventional models and standalone CNN and GRU architectures with an accuracy of 83%. Also focus is made on understanding the features influencing the product returnsthat help firms to make data driven decisions and minimize product returns . This type of predictive models can be applied to enhance the business strategies in making informed decisions enhancing the overall satisfaction of the consumer and improving the revenue of the business.

Keywords-E-Commerce; Return To Origin; Prediction Model; Deep Learning; Convolutional Neural Network; Gated Recurrent Units.

I. INTRODUCTION

The e-commerce industry has undergone an unprecedented growth and tremendous revolution and is developed into a vivacious and sternly competitive business by changing the way consumers obtain goods and services. The online shopping continues to expand in popularity throughout the world and change the shopping experience of the consumers, the e-commerce businesses must make measures to balance the possibilities of the problems in order to assure the customer prospect and uphold the business development. The e-commerce businesses face several challenges to make their business successful and several issues affect the revenue of the business. The returns disturb the supply chain operations increasing the financial and logistical burden on the retailers. The e-commerce sector makes use of predictive and machine learning techniques to overcome the challenges and understanding the complex data patterns.

One of the main issues focused in this paper is the return of products. The product returns affect the revenue of the business, the operational effectiveness, and customer satisfaction need to be concentrated and handling them is one of the major issues that e-commerce enterprises face in recent days. The vast growth of the e-commerce industry has been fueled by the ease, accessibility, and wide range of goods provided by online retailers. This online market faces several challenges and several set of difficulties along with its numerous benefits. One of the major concerns of all the issues is product return. Product returns stand out among these issues as a crucial problem that may have an extensive pressure on both the monetary constancy and consumer relations of e-commerce businesses.

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 9
***Article Received:** 25 July 2023 **Revised:** 12 September 2023 **Accepted:** 30 September 2023*

The product return issues faced by e-commerce industries are driving the online merchants to develop strategies and effectively handle product return issues [1]. Small return rates are not a big deal but high return rates greatly affect the businesses productivity as well as its buyer's contentment. The development of prognostic models that can forecast return rates are very helpful to firms in this situation.

Product returns are typical in the retail industry, but they may be a complicated problem in the world of online shopping. The easiness of online shopping, which enables customers to buy things without feeling the product, provides a setting where refunds are a necessary component of the business model. Excessive return rates can provide serious financial and logistical issues for e-commerce companies, even though they are a necessary part of consumer rights and customer-centric practices. Because there is no actual touch with the items before purchase, product returns pose a complicated problem in the e-commerce environment. Customers frequently rely on product descriptions, photos, and reviews to make wise purchases. The lack of in-person inspection, however, can lead to a greater percentage of product returns compared to traditional brick-and-mortar retail, even with these facilities. E-commerce companies must strike a balance between assuring customer pleasure, preserving operational efficiency, and lowering the financial burden of excessive returns. There are several research papers proposed for predicting the product returns [2] [3] [4].

The complex world of e-commerce product returns has given rise to predictive modeling as a useful tool. These models enable e-commerce companies to forecast return rates and obtain a better knowledge of the variables affecting customer decisions to return goods. These data enable businesses to fine-tune their product descriptions, better inventory management, and improve overall consumer experiences, eventually resulting to lower return rates and higher profitability. Predictive models that can foresee and comprehend product return rates have developed into indispensable tools for e-commerce businesses in order to properly manage this dilemma [5]. By decreasing returns and boosting customer satisfaction, these models give organizations the ability to manage inventory proactively, improve product descriptions, and enhance customer service.

There are several models proposed using CNN algorithm for predictions [16] [17] as shown in Fig. 1. But the development of hybrid models is more accurate compared to the conventional ones. Using a hybrid model that unites the strength of CNN and GRU, this research proposes a unique method for forecasting e-commerce product return rates.

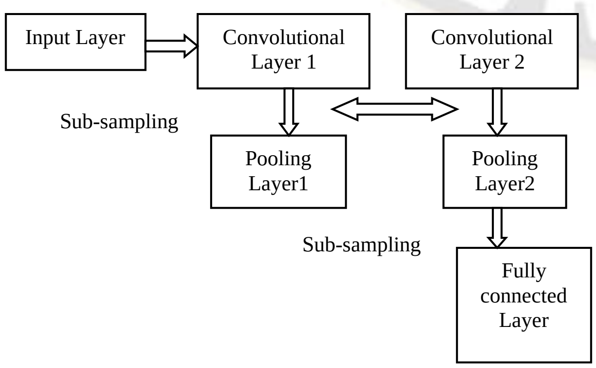

Figure 1. Classical CNN model.

This hybrid model is proposed to address the urgent demand for superior prediction models in e-commerce. This hybrid model adopts a holistic viewpoint, acknowledging that a variety of elements, including visual features, textual content, and temporal dynamics, affect return rates. Our approach strives to give a full understanding of the complicated nature of e-commerce product returns by merging CNNs to extract spatial characteristics from product photos and GRUs to analyze sequential data, which includes customer reviews and product descriptions. Fig. 2 depicts the GRU architecture used in the hybrid model.

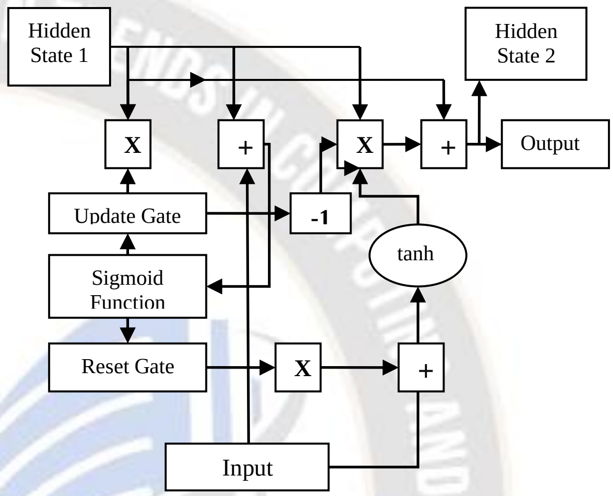

Figure 2. GRU Architecture

The hybrid model presented in this study acknowledges the need for a varied strategy to comprehend product returns in e-commerce. It requires the integration of several data sources to develop a comprehensive knowledge of the elements influencing returns rather than being restricted to just one form of data, such as customer evaluations or product photos. Therefore, our model aims to take advantage of the advantages of GRUs to analyze sequential data, such as customer reviews and product descriptions, and CNNs to extract spatial information from product photos.

The next sections of this article will present an in-depth examination of the hybrid CNN-GRU model, focusing on its architecture, the training methods used, and the assessment outcomes. Our goal is to demonstrate the unique model's higher prediction skills in comparison to conventional approaches and standalone CNN or GRU architectures. As a result, we give e-commerce businesses with a strong tool for making data-driven choices, improving operational efficiency, and ultimately lowering return rates while increasing consumer happiness.

The work proposes a significant step in the ongoing attempts of finding the difficulties in anticipating the return of products of e-commerce businesses. A promising approach is developed which enables the online firms to continuously mitigate the risk by applying the proposed hybrid CNN-GRU methodology.

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 9
***Article Received:** 25 July 2023 **Revised:** 12 September 2023 **Accepted:** 30 September 2023*

II. LITERATURE REVIEW

The authors in [6] explore the challenges occurred due to the increased rate of product returns from the consumers back to the origin in e-commerce firms. The primary focus of the work is to explore strategies to limit the product returns.

The work in [7] investigates the vendor's strategy to enhance the sales by integrating the product return costs. It applies both centralized and de-centralized channels with an extensive cost agreement.

The work in paper [8] highlights the return logistics and their issues with consumer satisfaction. The paper states the importance of businesses to adapt new techniques in handling challenges in returns of products. A systematic approach is developed for handling product returns, with minimized costs. The authors also surveyed several companies facing product returns and their challenges.

The work concentrates on e-commerce products quality management through sentiment analysis [9]. The insights on products reviews are extracted and deep learning techniques are applied to it. The developed deep learning model helps to explore the opinion of the consumers using deep learning algorithms resulting in good accuracy level.

The authors in [10] address the issues of the existing forecasting model to handle dynamic changes and non-linearity of the time-series data. The author has proposed a hybrid model using deep neural networks to overcome these challenges. The developed hybrid model uses a feature extraction module to extract the trends and patterns using time-series data of varying times. The author developed a three-dimensional sales forecasting model and it outperforms the existing model and adapts the new techniques and results in improved sales forecasting capabilities.

There are several research papers proposed by combining GRU with other algorithms. The inventory management for forecasting future sales is focused in this work [11]. Their proposed hybrid approach using GRU outstand the other prediction models. The further usage of posterior test confirms the effectiveness and accuracy of the proposed model. The research work in [12] proposes building load forecasting model using Convolutional Neural Network and GRU. The proposed model outperforms the existing ones with lower prediction losses. The hybrid approach results superior to conventional models. The paper in [13] introduces a novel classifier for Human Activity Recognition. A hybrid of CNN along with GRU is proposed to provide a solution. The developed model integrates with adaptive user interfaces and pattern based surveillance that outperforms similar structured architectures. The work in [14] proposes a model to predict traffic patterns in Intelligent Transport Systems (ITS). A comprehensive algorithm is proposed for integrating the heterogeneous data using exploratory data analysis tools. Then several deep learning algorithms are applied and hybrid combinations are developed and LSTM-GRU results in high accuracy compared to the proposed models. This work [15] discusses the applications of deep learning algorithms and the prediction of stock market indices in finance sector. Several models are proposed and the hybrid model outperforms the other models in terms of accuracy and error rates.

III. PROPOSED METHODOLOGY

The proposed model uses the strengths of both the CNNs and GRUs. CNN extracts information influencing consumer's decision. The GRUs processes the data on consumer reviews, product descriptions, and product return details. The below Fig. 3 denotes the architectural representation of the CNN-GRU proposed model.

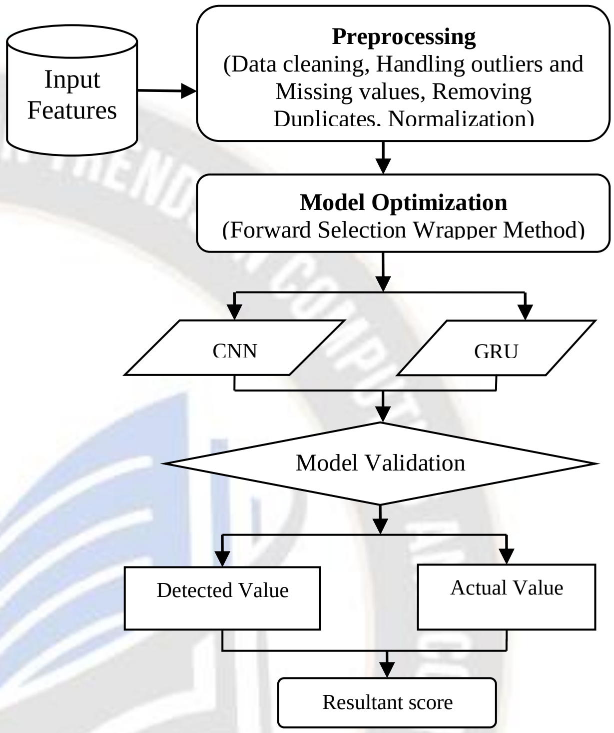

Figure 3. Architecture of the proposed model based on CNN-GRU for product return prediction

A. Data Preprocessing

The data preprocessing involves several stages starting from data cleaning to remove the outliers and inconsistencies in the dataset. It is a crucial step that aims to find and rectify the errors and enhance the accuracy and reliability of the dataset. The dataset is checked for missing values and the missing values were removed. The duplicates were removed. Then data type conversion as numeric format is carried out to make the dataset suitable for the model. The final dataset is normalized to make sure that the values are in a similar scale to fit the developed model.

B. Hybrid Model Architectural Design

The CNN and GRU model is combined to develop a hybrid model for comprehensive understanding of the factors leading to product returns. The below Fig. 3 represent the architectural block of the hybrid model.

C. Feature Extraction

The process of creating a subset of pertinent and instructive features from a larger set of data is termed as feature extraction. Feature extraction is performed to reduce the dimensionality of the proposed hybrid model and improve the models efficiency. It also leads to better predictive

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 9
***Article Received:** 25 July 2023 **Revised:** 12 September 2023 **Accepted:** 30 September 2023*

performance. In wrapper method the feature selection is carried out to evaluate the different subsets of features by training and evaluating it. The idea behind applying forward technique is that they are capable of finding feature subsets for specific models and datasets. Also the interactions between the features are considered and also provide a method to optimize the performance of the feature selection process.

D. Model Training and Validation

The process of model training, testing and validating are vital components of the predictive model process.
The dataset is split as 80% for training the predictive model, 10% for validating the model and remaining 10% is for testing the model. The model is trained on the training dataset and learns from features and target variables to make accurate predictions. Validation is done to optimize the model. The performance of the proposed hybrid model is evaluated using evaluation metrics scores. Finally the hybrid model is evaluated on the test dataset to evaluate the unbiased estimation of the performance on unseen data and generalize it. The overall process flow diagram of the proposed hybrid model is represented in Fig.4.

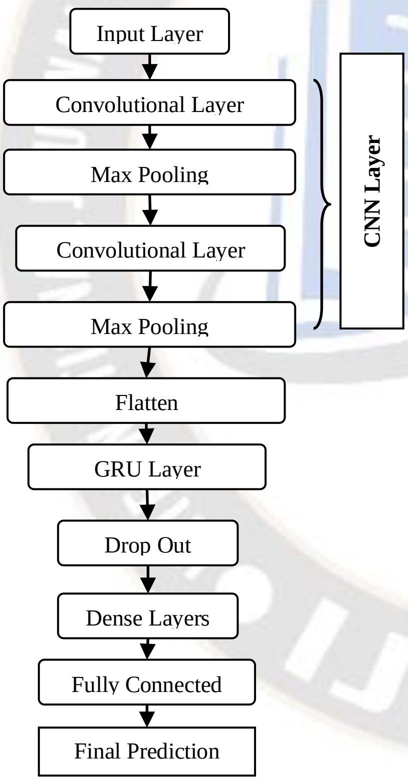

Figure 4. Flow diagram of the proposed model

E. Feature Attribution

The process of determining the impact of the features in predicting the outcome is termed as feature attribution. It is used in extracting the most relevant features contributing to the model prediction. We apply permutation feature importance technique for evaluating the importance of the unique features of the proposed hybrid model. The calculated performance metrics of the hybrid model is used as a baseline for this model. At a particular time individual feature randomly permutes its values in the dataset. So there is an interrupt between the feature variables and the target one. The selected features are shuffled in a random manner keeping the other features remain unchanged. A new score for the performance metric of the model is evaluated using the dataset using the perturbed feature. The impact of the individual features can be learnt using permutation feature importance model. The permutation explains the features influencing the models behavior and decision.

IV. Results And Discussion

The model is proposed by using hybrid CNN-GRU for finding product return rates. The performance of the developed model is evaluated using the training, validation and test dataset. The optimization is carried out through Adam algorithm. The experiment results are shown in the Table 1. It is observed that the hybrid CNN- GRU model outperforms the other models.

TABLE 1: COMPARISON TABLE OF PERFORMANCE EVALUATION OF THE PROPOSED MODELS

Evaluate	Models	Accuracy	Sensitivity	Specificity	Precision
Individual Training and Test Set	**GRU**	76.32	72.64	73.53	75.67
	CNN	72.16	72.12	71.35	72.15
	CNN–GRU	**86.77**	**87.14**	**87.58**	**85.21**
Cross Validation	**GRU**	80.11	72.13	75.46	77.56
	CNN	76.16	73.24	77.17	76.02
	CNN – GRU	**87.24**	**87.14**	**87.01**	**87.28**

The dataset is split as 80:10:10 for the training, Validating and testing data. CNN-GRU model has resulted with an accuracy of 87.24%. The resulting accuracy is comparatively better than the other proposed models. The results of evaluation metrics are illustrated in Table 1 and from that it can be finalized that hybrid CNN-GRU model outstands compared to the other ones as shown in Table 1.

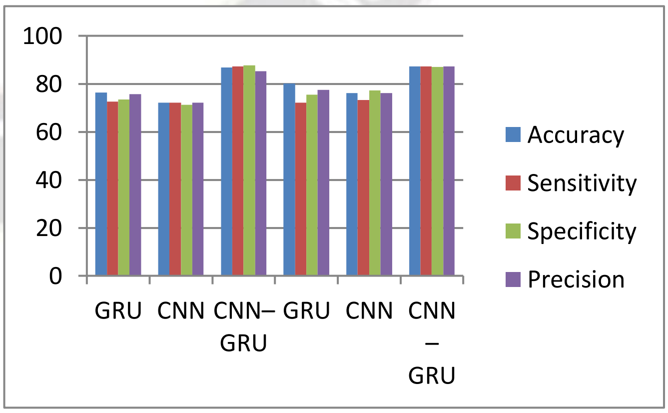

Figure 5: Performance results of proposed models

The graphical representation based on the evaluation metrics like precision, specificity, sensitivity, and accuracy is represented in Fig.5. It clearly states that CNN-GRU hybrid model competently overcomes the alternative proposed models with enhanced accuracy score of 87.24% and the other models CNN and GRU are lesser than 80%.

V. Conclusion

The necessity of product return rate prediction cannot be overstated as e-commerce businesses continue to thrive and expand. The proposed approach has the potential to empower online e-commerce retailers by allowing them to make hands-on steps to lesser the return rates, to improve customer happiness, and restructure their business processes. The model serves as an intended benefit in the competitive e-commerce business zone in addition to being a forecasting system. The CNN-GRU model acts as a base for prospect research and development. Integrating of further data sources, research of alternate deep learning architectures, and more advanced approach for model analysis and optimization may be incorporated in future revisions. In an active e-commerce world where consumer experience and operational effectiveness are vital, the hybrid CNN-GRU approach is a solid choice. It is ready to support online businesses in their efforts to prosper, provide a faultless shopping experience, and progress their return handling measures.

References

[1] Priya Ambilkar, Vishwas Dohale, Angappa Gunasekaran & Vijay Bilolikar (2022) Product returns management: a comprehensive review and future research agenda, International Journal of Production Research, 60:12, 3920-3944, DOI: 10.1080/00207543.2021.1933645

[2] Li Y, Martínez-López FJ, Feng C, Chen Y. Green Communication for More Package-Free Ecommerce Returns. *Journal of Theoretical and Applied Electronic Commerce Research*. 2022; 17(4):1450-1472. https://doi.org/10.3390/jtaer17040073

[3] M. Radhi and G. Zhang, "Optimal cross-channel return policy in dual-channel retailing systems," International Journal of Production Economics, vol. 210, pp. 184–198, 2019.

[4] Qingyun Xu, Zhen Shao, Lin Zhang, Yi He, Optimal livestream selling strategy with buy-online-and-return-in-store, Electronic Commerce Research and Applications,Volume 61,2023,101307,ISSN 1567-4223 https://doi.org/10.1016/j.elerap.2023.101307

[5] Rajasekaran, V., & Priyadarshini, R. (2021). An e-Commerce prototype for predicting the product return phenomenon using optimization and regression techniques. *Communications in Computer and Information Science*, 230-240. https://doi.org/10.1007/978-3-030-88244-0_22

[6] Han, H. (2019) Review and Prospect on Return Problems of E-Commerce Platform. Open Journal of Business and Management, 7, 837-847. doi: 10.4236/ojbm.2019.72057.

[7] Bieniek, M. Returns handling in e-commerce: How to avoid demand negativity in supply chain contracts with returns?. Electron Commer Res (2023). https://doi.org/10.1007/s10660-023-09689-2

[8] Dobroselskyi, M., Madleňák, R., & Laitkep, D. (2021). Analysis of return logistics in e-Commerce companies on the example of the Slovak Republic. *Transportation Research Procedia*, *55*, 318-325. https://doi.org/10.1016/j.trpro.2021.06.037

[9] Yi Liu, Jiahuan Lu, Jie Yang, Feng Mao. Sentiment analysis for e-commerce product reviews by deep learning model of Bert-BiGRU-Softmax[J]. Mathematical Biosciences and Engineering, 2020, 17(6): 7819-7837. doi: 10.3934/mbe.2020398

[10] Efat, Md & Hájek, Petr & Abedin, M. & Azad, Rahat & Jaber, Md & Aditya, Shuvra & Hassan, M. Kabir. (2022). Deep-learning model using hybrid adaptive trend estimated series for modelling and forecasting sales. Annals of Operations Research. 10.1007/s10479-022-04838-6.

[11] Chunqiang Lu, Gang Shang, Liyun Xu, Huan Shao and Beikun Zhang , Sales Volume Forecast of Typical Auto Parts Based on BiGRU: A Case Study, E3S Web Conf., 409 (2023) 04008. DOI: https://doi.org/10.1051/e3sconf/202340904008

[12] Ming-Chuan Chiu, Hsin-Wei Hsu, Ke-Sin Chen, Chih-Yuan Wen,A hybrid CNN-GRU based probabilistic model for load forecasting from individual household to commercial building, Energy Reports,Volume 9, Supplement 10,2023,Pages 94-105,ISSN 2352-4847,https://doi.org/10.1016/j.egyr.2023.05.090.

[13] Dua N, Singh S, Semwal V and Challa S. (2022). Inception inspired CNN-GRU hybrid network for human activity recognition. Multimedia Tools and Applications. 10.1007/s11042-021-11885-x. 82:4. (5369-5403). Online publication date: 1-Feb-2023.

[14] Zafar N, Haq IU, Chughtai J-u-R, Shafiq O. Applying Hybrid Lstm-Gru Model Based on Heterogeneous Data Sources for Traffic Speed Prediction in Urban Areas. *Sensors*. 2022; 22(9):3348. https://doi.org/10.3390/s22093348

[15] Song H, Choi H. Forecasting Stock Market Indices Using the Recurrent Neural Network Based Hybrid Models: CNN-LSTM, GRU-CNN, and Ensemble Models. *Applied Sciences*. 2023; 13(7):4644. https://doi.org/10.3390/app13074644

[16] Liu, Xia. (2022). E-Commerce Precision Marketing Model Based on Convolutional Neural Network. Scientific Programming. 2022. 1-11. 10.1155/2022/4000171.

[17] Z. Zhuo, S. Chen, and K. Y. Chau, "A new model of manufacturer's optimal product supply strategy in the context of precision marketing: based on real demand pattern," Mathematical Problems in Engineering, vol. 59, no. 02, pp. 62–98, 2020.

[18] Adrian MICU & Marius GERU & Alexandru CAPATINA & Constantin AVRAM & Robert RUSU & Andrei Alexandru PANAIT, 2019. "Leveraging e-Commerce Performance through Machine Learning Algorithms," Economics and Applied Informatics, "Dunarea de Jos" University of Galati, Faculty of Economics and Business Administration, issue 2, pages 162-171. DOI: https://doi.org/10.35219/eai1584040947

[19] Petroșanu D-M, Pîrjan A, Căruțașu G, Tăbușcă A, Zirra D-L, Perju-Mitran A. E-Commerce Sales Revenues Forecasting by Means of Dynamically Designing, Developing and Validating a Directed Acyclic Graph (DAG) Network for Deep Learning. *Electronics*. 2022; 11(18):2940. https://doi.org/10.3390/electronics11182940

[20] Chen X, Long Z. E-Commerce Enterprises Financial Risk Prediction Based on FA-PSO-LSTM Neural Network Deep Learning Model. *Sustainability*. 2023; 15(7):5882. https://doi.org/10.3390/su15075882