Katarzyna Topolska
Politechnika Wrocławska
katarzyna.topolska@pwr.edu.pl

Intelligent Management of Library Collections: The Application of Neural Networks for Predicting Book Demand

(Inteligentne zarządzanie zbiorami bibliotecznymi: zastosowanie sieci neuronowych do przewidywania popytu na książki)

Słowa kluczowe: biblioteka, prognozowanie, sieci neuronowe, analiza szeregów czasowych

Abstrakt: Celem przeprowadzonych badań była analiza i prognozowanie zapotrzebowania na zasoby biblioteczne z wykorzystaniem metod klasycznych oraz modeli opartych na sztucznej inteligencji. Przeanalizowano dane historyczne dotyczące wypożyczeń książek, identyfikując wzorce sezonowości i zmieniające się preferencje tematyczne użytkowników. Porównano skuteczność takich metod, jak regresja liniowa, LSTM, GRU, CNN oraz MLP w prognozowaniu zapotrzebowania. Wyniki wskazują na wyraźną przewagę modeli neuronowych, zwłaszcza LSTM i CNN, w dokładności predykcji oraz ich potencjalne zastosowanie w optymalizacji zarządzania zbiorami bibliotecznymi.

Keywords: library, forecasting, neural networks, time series analysis

Abstract: The aim of the study was to analyze and forecast library resource demand using both classical methods and artificial intelligence-based models. Historical book borrowing data was examined to identify seasonal patterns and changing user preferences. The effectiveness of methods such as linear regression, LSTM, GRU, CNN, and MLP was compared in forecasting demand. The results demonstrate a clear advantage of neural network models – particularly LSTM and CNN – in terms of prediction accuracy and their potential to optimize library collection management.

Introduction

Modern libraries face challenges resulting from a growing number of users, evolving reading trends, and limited financial resources [2]. Traditional collection management methods are increasingly inadequate in the face of dynamic changes in the information environment [17]. In response to these challenges, there is a growing interest in the application of artificial intelligence in librarianship [8]. One of the most promising approaches is the use of artificial neural networks to forecast book demand [9]. LSTM (Long Short-Term Memory) models are particularly effective in analyzing time-series data, such as borrowing history [6]. Research has demonstrated that LSTM models outperform traditional methods, including linear regression and ARIMA, in terms of forecasting accuracy [13]. Neural networks allow for more precise alignment of library acquisitions with users' actual needs [12], helping to reduce both overstocking and shortages of high-demand titles [20]. As a result, libraries can allocate budgets and spatial resources more efficiently [4]. The combination of machine learning and historical data analysis also enables the prediction of seasonal borrowing patterns. For instance, during the summer holiday period, demand for travel literature tends to rise – an insight that can inform acquisition planning [29].

Academic libraries are increasingly implementing predictive systems to support collection management aligned with the educational and research needs of universities [15]. Neural networks can also be applied to analyze user data, such as thematic preferences and borrowing history [1; 27, s. 157-194; 28, s. 273-308]. This approach enables the development of more personalized services [10]. Studies have shown that combining CNN and LSTM models allows for the simultaneous analysis of books' textual features and temporal borrowing patterns. Such solutions have already been implemented in libraries in South Korea and the Netherlands [7].

AI-based systems enable data visualization and the creation of intuitive dashboards to support librarians' decision-making processes [16]. The implementation of such tools fosters more informed acquisition decisions [19]. Forecasting automation also facilitates the management of large collections in central and specialized libraries [14]. Additionally, AI supports the planning of collection transfers between municipal library branches based on local user preferences [25]. Predictive models enhance the efficiency of managing both physical and digital resources [5]. Thanks to deep learning technologies, it is also possible to forecast the popularity of new publications based on reviews, trends, and demographic data [18]. This approach combines sentiment analysis with predictive analytics [3]. However, successful implementation requires appropriate organizational and technical preparation [26]. Data quality, structure, and compliance with privacy regulations are of key importance [11]. Furthermore, libraries must ensure that staff possess the necessary competencies to interpret and utilize AI-generated insights [25]. Training and collaboration with academic institutions can play a crucial role in this process [6].

The literature emphasizes that the implementation of AI in libraries is not an end in itself, but a means to improve service quality [2]. Users expect timely access to relevant resources, and predictive technologies can help achieve this goal [13]. The introduction of artificial intelligence supports the concept of a proactive library that is responsive to changing user needs [4]. AI does not replace librarians but assists them in the decision-making process [15]. Managing collections with the support of neural networks is also part of the broader trend of digital transformation in public institutions [9]. As centers of education and access to knowledge, libraries should take a leading role in this transformation [29].

Examples of implementation demonstrate that even small libraries can successfully utilize open-source AI tools such as TensorFlow, PyTorch, or Keras [14]. However, it is essential to properly adapt the models to local conditions. AI can also be used to predict which books will be needed in the context of specific social or academic events [1], allowing libraries to respond proactively – before demand actually arises [20]. The development of deep learning technologies opens new perspectives for librarianship, making it increasingly user- and data-oriented [17].

This article presents specific examples of neural network models, their implementation, and the outcomes of their application in various types of libraries [8]. The aim of the research is to develop, implement, and evaluate the effectiveness of artificial neural network models – particularly recurrent neural networks (RNNs) and Long Short-Term Memory (LSTM) networks – in forecasting book demand in both public and academic libraries.

The research objectives are as follows:

1. To identify borrowing patterns over time and across subject areas based on historical data.
2. To compare the effectiveness of various predictive models, including both traditional statistical methods and machine learning approaches.
3. To examine the impact of prediction on collection management, particularly in the areas of acquisition optimization, collection rotation, and budget planning.
4. To assess the feasibility of implementing neural network models within library systems, considering institutional constraints such as data availability, IT infrastructure, and staff competencies.
5. To identify potential barriers and limitations associated with the implementation of AI in the library environment.

Methods

The study employed four popular neural network architectures to forecast the demand for library resources: MLP, CNN, LSTM, and GRU. Each of these methods has a distinct structure and data processing mechanism, which influences their effectiveness in time series analysis.

MLP (Multilayer Perceptron) is a classical, fully connected neural network composed of an input layer, one or more hidden layers, and an output layer. Since MLP lacks sequential memory, analyzing time series data requires pre-processing inputs using a sliding window approach. Due to its simplicity, MLP performs well on data with relatively low temporal complexity but may struggle to capture long-term dependencies [24].

CNN (Convolutional Neural Network), although commonly used in image analysis, has also been applied to sequential data processing. CNNs utilize convolutional filters to detect local patterns in data, making them effective in recognizing seasonality and recurring structures within time series. Thanks to their ability to process data in parallel, CNNs are typically faster to train than recurrent networks, although their capacity to model very long-term temporal dependencies may be limited [23].

LSTM (Long Short-Term Memory) is a type of recurrent neural network designed specifically to capture long-term dependencies in sequential data. Through memory cells and gating mechanisms – input, forget, and output gates – LSTM can retain relevant information across many time steps. It is particularly effective in tasks requiring the tracking of long sequences and the identification of subtle trends or delayed effects within the data [21].

GRU (Gated Recurrent Unit) is a simplified variant of LSTM, characterized by fewer parameters, which results in faster training times and reduced data requirements. GRUs employ gating mechanisms similar to those of LSTM, but in a more streamlined form by combining certain functions. In many cases, GRUs achieve performance comparable to LSTMs, especially when dealing with shorter time series or smaller datasets [22].

Experimental evaluation

The conducted experiments followed an experimental protocol based on historical annual data, covering actual demand indicator values in the library system over the past ten years. The objective was to compare the effectiveness of various forecasting models in predicting average annual demand. The models also incorporated independent contextual features.

Features found to significantly influence the dependent variable in library systems ($p < 0.05$) include: month or season, day of the week, number of enrolled students per semester, type of study program, academic year or semester, number of new books added to the system, title popularity measured by the number of searches, number of events organized by the library, average weekly opening hours, availability of e-books and digital resources, number of pending reservations, average book holding time by users, user academic level (e.g., undergraduate, master's, PhD), availability of computers and study spaces in the reading room, presence of promotional or informational campaigns, average satisfaction rating of library services, occurrence of holiday breaks or academic recesses, user activity level in the previous month, number of available copies of a given title, and presence of sanitary or organizational restrictions (e.g., during a pandemic).

The experimental protocol consisted of the following stages:

1. Data Collection: Data were obtained from the library system, including the number of borrowings, reservations, catalog queries, as well as seasonality and periods of increased activity (e.g., exam sessions, semester start). The data were anonymized and transformed into a time series format.
2. Data Splitting: The dataset was chronologically divided into a training set (80%) and a test set (20%). For model validation, a walk-forward validation approach was employed, in which models were trained on historical data and tested on subsequent years sequentially.
3. Model Training: Five models were evaluated: CNN, LSTM, GRU, MLP, and linear regression. Each model was trained using identical input parameters, standardized values, and a fixed time window length of 12 months (up to 2024). The models were trained using the Mean Squared Error (MSE) loss function and the Adam optimizer.
4. Model Evaluation: For each model, the predicted average annual demand, accuracy (defined as the inverse of the relative error compared to the actual value), and a 95% confidence interval (estimated via bootstrap simulation) were calculated.
5. Experiment Replications: To increase the reliability of the results, each experiment was repeated 30 times with random weight initialization. The reported values represent the averages across all repetitions.

This protocol ensured an objective and reproducible comparison of the models while preserving the real-world temporal structure and seasonality inherent in library data.

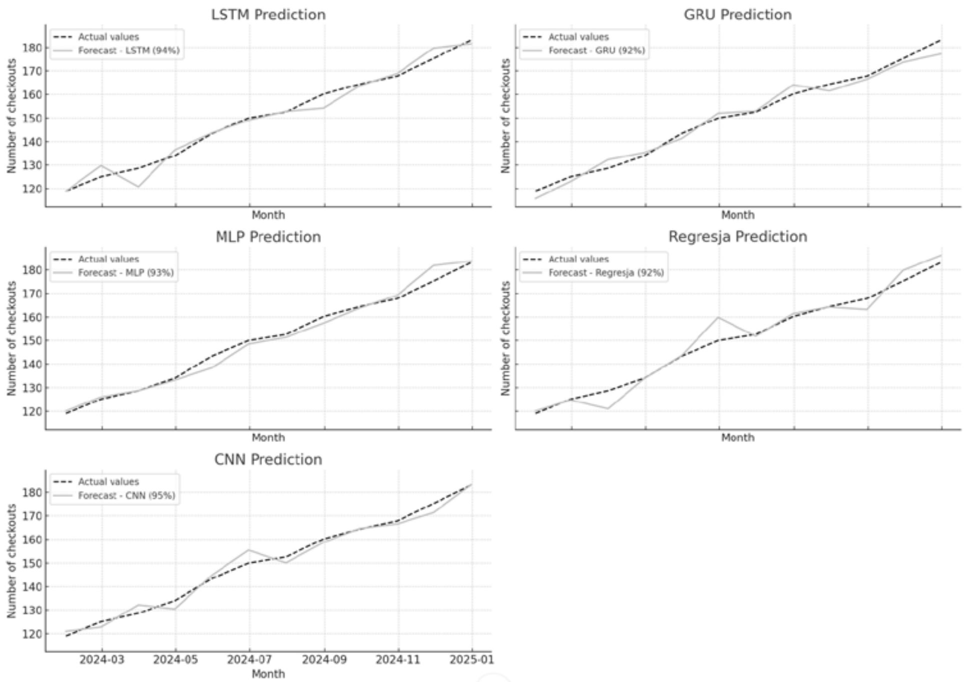
The results presented in Table 1 and Figure 1 compare the predictive performance of the five forecasting models in terms of average annual forecasts, accuracy, and stability of results.

Table 1. Statistics of Average Annual Forecasts

Model	Average Forecast	Average Actual Value	Accuracy (%)	Confidence Interval (±)
LSTM	151.6	152.9	94.00	±5.88
GRU	147.8	152.9	92.00	±7.84
MLP	150.8	152.9	93.00	±6.86
Regression	146.7	152.9	92.00	±7.84
CNN	153.9	152.9	95.00	±4.90

Source: own work.

Figure 1. Results of the Experimental Evaluation of Book Borrowing Predictions



Source: own work.

The best-performing model was CNN, with a forecast of 153.9, closest to the actual value of 152.9. It achieved the highest accuracy (95%) and the narrowest confidence interval (±4.90). The second-best model was LSTM, which produced a forecast of 151.6, an accuracy of 94%, and moderate stability (±5.88). The MLP model reached 93% accuracy with a confidence interval of ±6.86, indicating good but slightly lower effectiveness. The GRU and linear regression models showed the lowest accuracy (92%) and the widest confidence intervals (±7.84), with forecasts deviating most from the actual value.

Based on these results, it can be concluded that deep learning models – particularly CNN – offer the highest precision and stability in forecasting, making them especially suitable for applications requiring high reliability. In the context of library systems, these models can be effectively employed for tasks such as forecasting demand for specific titles, predicting system load during different periods, managing staffing resources, and optimizing acquisitions of new materials.

It is recommended to adopt the CNN model as the primary forecasting tool in library management support systems, while considering LSTM as an alternative in scenarios where analyzing sequential usage trends is particularly important.

Research conclusions

The aim of the study was to analyze historical data on book borrowings across various time periods, including months, semesters, and years. Significant temporal patterns were identified, such as seasonality (e.g., increased demand for textbooks before exam sessions) and shifts in the popularity of specific topics (e.g., a rise in interest in non-fiction literature during particular periods). Time series analysis algorithms were employed to generate forecasts of book demand based on past data. The analysis also captured changes in users' thematic preferences, as reflected in borrowing trends.

The objective was to compare the effectiveness of classical forecasting methods, such as linear regression, ARIMA, and exponential smoothing, with modern AI-based approaches, including neural networks (LSTM, GRU). The study revealed that traditional methods, despite their simplicity, struggled to capture complex patterns in the data – such as seasonal variability and intricate interactions between book topics. In contrast, machine learning models – especially LSTM networks – proved significantly more effective, delivering more accurate forecasts and better adaptability to changing trends. The performance of other neural architectures, such as CNNs, was also evaluated and compared to that of LSTMs.

The research further aimed to assess the impact of predictive modeling on library collection management. Book demand forecasts enabled the optimization of acquisitions, preventing situations where books remained unused on shelves or, conversely, popular titles were unavailable. Predictive models supported improved planning of new acquisitions and facilitated collection rotation by indicating which books should be relocated between branches or removed from the system. Forecasting also contributed to more efficient budget planning by aligning expenditures more closely with actual user needs.

The study additionally examined the feasibility of implementing neural network models in libraries, considering institutional factors such as data availability, IT infrastructure, and staff competencies. It was found that while AI technologies offer significant benefits, in many cases infrastructure modernization is necessary to implement these solutions effectively. Moreover, appropriate training must be provided to librarians, enabling them to understand and utilize the insights generated by predictive models. Implementing AI also requires investment in robust data management systems to ensure data quality and compliance with data protection regulations.

Several key barriers and limitations related to AI implementation in libraries were identified. These included limited access to high-quality data – essential for training neural models – and the high costs associated with acquiring and deploying AI technologies.

Additionally, many libraries face a shortage of qualified personnel capable of handling and interpreting algorithmic outputs. Another significant challenge is users' concern regarding privacy and the security of their personal data, which poses an additional obstacle to the widespread adoption of AI technologies.

Summary

The study aimed to analyze historical book borrowing data and evaluate the effectiveness of various demand forecasting methods. Distinct seasonal and thematic patterns were identified, such as increased demand for textbooks before exam sessions. Classical statistical approaches were compared with modern AI-based models, including LSTM, GRU, and CNN neural networks. Neural models, particularly LSTM, demonstrated superior effectiveness in predicting fluctuating borrowing trends. The application of these forecasts facilitated improved collection management, reducing both surpluses and shortages of books. Additionally, it had a positive impact on budget planning and resource redistribution across library branches.

However, several challenges were noted, including the need for IT infrastructure modernization and staff training. The study also emphasized the critical importance of data quality and compliance with GDPR regulations. Despite these barriers, the implementation of AI in libraries was deemed feasible and potentially highly beneficial. Further investment in the development of predictive systems and enhanced access to library data is strongly recommended.

References

1. ARUNKUMAR O. N., DIVYA O. Deep Learning Techniques for Demand Forecasting: Review and Future Research Opportunities. *Information Resources Management Journal.* 2022, nr 2, s. 1–24. ISSN 1533-7979.
2. BANDARA Kasun, SHI Peibei, BERGMEIR Chritoph, HEWAMALAGE Hansika, TRAN Quoc, SEAMAN Brian. *Sales Demand Forecast in E-commerce using a Long Short-Term Memory Neural Network Methodology* [Dokument elektroniczny]. 2019. Tryb dostępu: https://arxiv.org/abs/1901.04028. Stan z dnia 04.07.2025.
3. DAIMARI Dersin, NARZARY Mwnthai, MAZUMDAR Nabajyoti, NAG Amitava. A Machine Learning Based Book Availability Prediction Model for Library Management System. *Library Philosophy and Practice*. 2021, winter. ISSN 1522-0222.
4. DAVE Emmanuel, LEONARDO Albert, JEANICE Marethia, HANAFIAH Novita. Forecasting Indonesia Exports using a Hybrid Model ARIMA-LSTM. *Procedia Computer Science.* 2021, vol. 179, s. 480–487. ISSN 1877-0509.
5. DUDLA Anil Kumar, BATHULA Thirupathi Rao, BATHINI Rangaswamy, KAGITHA Meghana. An Efficient Approach for Food Demand Forecasting Using an Ensemble Technique and Statistical Analysis. In PAREEK Prakash, MISHRA Sumita, REIS Manuel J. C. S., GUPTA Nishu (ed.). *Cognitive Computing and Cyber Physical Systems. 5th EAI International Conference, IC4S 2024, Bhimavaram, India, April 5–7, 2024, Proceedings, Part-I.* Cham: Springer Professional, 2025, s. 61–77. ISBN 978-3-031-77075-3.
6. GOŁĄBEK Marta, SENGE Robin, NEUMANN Rainer. *Demand Forecasting using Long Short-Term Memory Neural Networks* [Dokument elektroniczny]. 2020. Tryb dostępu: https://arxiv.org/abs/2008.08522. Stan z dnia 22.07.2025.
7. HE Chaoyang i in. FedML: *A Research Library and Benchmark for Federated Machine Learning* [Dokument elektroniczny]. Tryb dostępu: https://arxiv.org/abs/2007.13518. Stan z dnia 10.07.2025.
8. HOPP Daniel. Economic Nowcasting with Long Short-Term Memory Artificial Neural Networks (LSTM). *Journal of Official Statistics*. 2022, nr 3, s. 847–873. ISSN 0282-423X.
9. IMANI Maryam. Fuzzy-based weighting long short-term memory network for demand forecasting. *The Journal of Supercomputing*. 2023, vol. 79, s. 435–460. ISSN 0920-8542.
10. IQBAL Naeem, JAMIL Faisal, AHMAD Shabir, KIM Dohyeun. Toward effective planning and management using predictive analytics based on rental book data of academic libraries. *IEEE ACCESS.* 2020, vol. 8, s. 81978–81996. ISSN 2169-3536.
11. JYOTI, PANKAJ Kumar. Reshaping the library landscape: Exploring the integration of artificial intelligence in libraries. *IP Indian Journal of Library Science and Information Technology*. 2024, nr 1, s. 29–36. ISSN 2582-1555.
12. KHADEMIZADEH Shahnaz, NEMATOLLAHI Zahra, DANESH Farshid. Analysis of book circulation data and a book recommendation system in academic libraries using data mining techniques. *Library and Information Science Research.* 2022, nr 4. ISSN 0740-8188.
13. KIM Sooi, ABEDINIANGERABI Bahram. Forecasting Architecture Billings Index Using Time Series Models. In *Construction Research Congress* [Dokument elektroniczny]. 2022, s. 391–402. Tryb dostępu: https://ouci.dntb.gov.ua/en/works/4NGLk18l/. Stan z dnia 22.07.2025.
14. KOUSHIK Roy, ABTAHI Ishmam, KAZI Abu Taher. *Demand Forecasting in Smart Grid Using Long Short-Term Memory* [Dokument elektroniczny]. 2021. Tryb dostępu: https://arxiv.org/abs/2107.13653. Stan z dnia 22.07.2025.
15. LEI Chunrui, ZHANG Heng, WANG Zhigang, MIAO Qiang. Deep Learning for Demand Forecasting: A Framework Incorporating Variational Mode Decomposition and Attention Mechanism. *Processes*. 2025, nr 2, s. 1–21. ISSN 2227-9717.
16. RENUKA B. C., PRAVEEN G. M., SARASWATHI H. S. Predicting book availability in library management systems using machine learning. *International Journal of Creative Research Thoughts*. 2025, nr 1, s. 583–589. ISSN 2320-2882.

17. RUSSELL Stuart J., NORVIG Peter. *Artificial Intelligence: A Modern Approach*. Upper Saddle River: Prentice Hall, 2020. ISBN 978-0134610993.
18. SAFONOV Kirill. *Neural Network Approach to Demand Estimation and Dynamic Pricing in Retail* [Dokument elektroniczny]. 2024. Tryb dostępu: https://arxiv.org/abs/2412.00920. Stan z dnia 15.07.2025.
19. SALAMANIS Athanasios, XANTHOPOULOU Georgia, KEHAGIAS Dionysios, TZOVARAS Dimitrios. LSTM-Based Deep Learning Models for Long-Term Tourism Demand Forecasting. *Electronics*. 2022, nr 22. ISSN 2079-9292.
20. TAVAKOLI Neda. Prediction and Estimation of Book Borrowing in the Library: Machine Learning. *Informatica*. 2021, nr 1, s. 163–168. ISSN 1854-3871.
21. TOPOLSKI Mariusz, TOPOLSKA Katarzyna. Algorithm for Constructing a Classifier Team Using a Modified PCA (Principal Component Analysis) in the Task of Diagnosis of Acute Lymphocytic Leukaemia Type B-CLL. In PÉREZ García H., SÁNCHEZ González L., CASTEJÓN Limas M., QUINTIÁN Pardo H., CORCHADO Rodríguez E. (eds.). *Hybrid Artificial Intelligent Systems. HAIS 2019*. Cham: Springer, 2019, s. 614–624. ISBN 978-3-030-29858-6.
22. TOPOLSKI Mariusz. Algorithm of multidimensional analysis of main features of PCA with blurry observation of facility features detection of carcinoma cells multiple myeloma. In BURDUK Robert, KURZYŃSKI Marek, WOŹNIAK Michał (eds.). *Progress in Computer Recognition Systems. CORES 2019*. Cham: Springer, 2020, s. 286–294. ISBN 978-3-030-19737-7.
23. TOPOLSKI Mariusz. Application of the Stochastic Gradient Method in the Construction of the Main Components of PCA in the Task Diagnosis of Multiple Sclerosis in Children. In KRZHIZHANOVSKAYA Valeria i in. (ed.). *Computational Science – ICCS 2020*. Cham: Springer, 2020, s. 35–44. ISBN 978-3-030-50422-9.
24. TOPOLSKI Mariusz. The modified principal component analysis feature extraction method for the task of diagnosing chronic lymphocytic leukemia type b-CLL. *Journal of Universal Computer Science*. 2020, nr 6, s. 734–746. ISSN 0948-6968.
25. VAN LINT Joost W. C., HOOGENDOORN Serge Paul, VAN ZUYLEN Henk J. Freeway Travel Time Prediction with State-Space Neural Networks: Modeling State-Space Dynamics with Recurrent Neural Networks. *Transportation Research Record.* 2002, nr 1, s. 30–39. ISSN 0361-1981.
26. VERMA Manish. Novel Study on AI-Based Chatbot (ChatGPT) Impacts on the Traditional Library Management. *International Journal of Trend in Scientific Research and Development.* 2023, nr 1, s. 961–964. ISSN 2456-6470.
27. WOJCIECHOWSKA Maja (red.). *Marketing w działalności bibliotecznej.* Warszawa: Wydawnictwo Naukowe i Edukacyjne SBP, 2023. ISBN 978-83-89316-82-0.
28. WOJCIECHOWSKA Maja (red.). *Zarządzanie biblioteką.* Warszawa: Wydawnictwo Naukowe i Edukacyjne SBP, 2019. ISBN 978-83-65741-26-4.
29. XIAOHUA Shi, CHEN Hao, DING Yue, HONGTAO Lu. Library book recommendation with CNN-FM deep learning approach. *Library Hi Tech.* 2023, nr 5, s. 1559–1578. ISSN 0737-8831.