International journal of imaging science and engineering

ISSN: 1934--9955 www.ijise.net
Vol-20 Issue-01 April 2025

PREDICTING HOSPITAL ADMISSION TRENDS WITH MACHINE LEARNING TECHNIQUES FOR ENHANCED HEALTH CARE MANAGEMENT

Teppala Dinesh[1], Godugu Naveen Kumar[2], Allam Akash[3], Dr. N. Krishnaiah[4]
[1,2,3] UG Scholar, Dept. of IT ,St. Martin's Engineering College, Secunderabad, Telangana, India, 500100
[4] professor & HoD, Dept. of IT, St. Martin's Engineering College, Secunderabad, Telangana, India,500100
dineshteppala1412@gmail.com

Abstract

Hospital Readmissions represent a significant challenge in modern healthcare systems, contributing to elevated costs and potential patient detriment. This project seeks to address this issue by leveraging advanced supervised learning models to predict hospital readmissions based on emergency department data. Historically, predicting readmissions relied on simplistic statistical methods and clinical judgment, which often lacked accuracy due to their inability to handle the complexities of large datasets and multifactorial variables. With the rise of electronic health records (EHRs) and big data analytics, there has been a shift towards utilizing comprehensive patient data for more precise predictions. Traditional systems, however, face limitations such as difficulty in managing intricate data interactions, static models that do not adapt to evolving patient conditions, and suboptimal predictive accuracy due to insufficient feature handling. This project aims to overcome these limitations by employing sophisticated machine learning algorithms and advanced feature engineering techniques. By improving the accuracy of readmission predictions, the project seeks to enable proactive patient management, reduce unnecessary readmissions, and ultimately decrease the financial strain on healthcare systems. The integration of these predictive models into clinical workflows holds the potential to enhance patient care by providing healthcare providers with actionable insights for better decision-making.

Keywords: Hospital Readmissions, Deep Learning, Electronic Health Records, Patient Management, Clinical workflows, Health care system.

1. INTRODUCTION

In recent years, the healthcare industry has witnessed significant advancements, with technology playing a pivotal role in improving patient care and operational efficiency. Among these innovations, machine learning (ML) has emerged as a transformative tool, particularly in predicting hospital admission trends. By leveraging data-driven insights, healthcare providers can make more informed decisions, optimize resources, and ultimately enhance the quality of care delivered to patients. The unpredictability of hospital admissions poses a considerable challenge for health care managers and policymakers. Fluctuations in patient numbers can strain hospital resources, leading to overcrowding, delays in care, and increased operational costs. As the demand for healthcare services continues to rise, effective management of hospital capacity becomes crucial for maintaining a high standard of care. Machine learning offers a solution by analyzing historical admission data and identifying patterns that may not be immediately apparent through traditional methods. By employing ML algorithms, such as regression analysis, decision trees, and neural networks, it is possible to predict future trends in hospital admissions with remarkable accuracy. These predictive models can assist in workforce planning, bed management, and inventory control, enabling hospitals to respond proactively to potential surges in patient numbers. This study explores the application of machine learning techniques to predict hospital admission trends, with a focus on their potential to enhance healthcare management. By examining historical data, this research aims to identify key factors influencing hospital admissions and demonstrate how these insights can be used to optimize hospital operations and improve patient outcomes. The ultimate goal is to provide healthcare managers with a powerful tool for anticipating patient demand and ensuring efficient resource allocation in an ever-evolving healthcare landscape.

2. LITERATURE SURVEY

Hao developed and validated a real-time risk assessment tool for 30-day hospital readmissions, integrated into the Maine Healthcare Information Exchange. This tool aimed to provide timely risk assessments by utilizing real-time patient data. Their study demonstrated successful deployment and validation, showing that such tools can enhance the management of patient readmission risks and improve care outcomes. The approach highlights the importance of real-time data integration in healthcare systems topre-emptively address readmission risks.

Walsh and colleagues explored various calibration methods for predictive models of hospital readmission risk, focusing on the models' clinical usefulness and discrimination capabilities. They compared different methodologies to assess how well these models predict readmission risks and their practical application in clinical settings. The study emphasizes that beyond accurate risk discrimination, predictive models must also be clinically relevant and actionable to improve patient care and management.

Gola's applied machine learning techniques to predict 30-day readmissions for heart failure patients using electronic medical records. Their retrospective analysis showcased the potential of machine learning models to identify high-risk patients accurately. By leveraging extensive medical data, the study underscores the value of data-driven approaches in predicting readmissions and highlights the promise of machine learning in chronic disease management.

Reddy and Delen utilized RNN-LSTM deep learning methodologies to predict hospital readmissions for lupus patients. Their study demonstrated how recurrent neural networks (RNNs) and long short-term memory (LSTM) models could analyze sequential patient data to forecast readmissions effectively. This work illustrates advancements in deep learning techniques and their application to chronic disease prediction, enhancing the ability to manage complex patient conditions.

Huang introduced Clinical BERT, a transformer-based model designed to analyse clinical notes and predict hospital readmissions. By incorporating natural language processing (NLP) with deep learning, Clinical BERT improved the interpretation of unstructured clinical data. This study highlights the potential of combining NLP techniques with predictive modeling to enhance the accuracy and effectiveness of readmission predictions.

Choi used recurrent neural network models to detect early signs of heart failure onset. Their study demonstrated the effectiveness of RNNs in analyzing temporal patient data to identify early indicators of heart failure. This approach emphasizes the capability of advanced machine learning models to provide early warnings and improve patient management through timely interventions.

In their "Doctor AI" study, Choi employed RNNs to predict various clinical events based on electronic health records. The research focused on using deep learning to enhance clinical event predictions, showcasing the potential of RNNs in improving predictive accuracy. The study reflects the growing application of machine learning in healthcare to support clinical decision-making and patient care.

Choi explored medical concept representation learning from electronic health records to predict heart failure. Their work demonstrated how learning from unstructured health data could improve predictive modeling. By representing clinical concepts effectively, their approach enhanced prediction accuracy and underscored the value of advanced data representation techniques in healthcare.

Nguyen developed Deepro, a convolutional neural network (CNN) model for analyzing medical records. Their study highlighted the ability of CNNs to extract and interpret features from complex medical datasets. Deepr's success in medical record analysis reflects the effectiveness of CNNs in enhancing predictive modeling and improving patient care through better data handling.

Quan focused on coding algorithms for defining co morbidities in ICD-9-CM and ICD-10 administrative data. Their work provided essential methodologies for accurately identifying co morbid conditions, which are crucial for risk adjustment and predictive modeling. This study establishes a foundation for developing predictive models that rely on precise administrative data for patient outcome predictions.

Tonelli outlined methods for identifying chronic conditions using administrative data, which is vital for accurate predictive analytics. Their study emphasized the importance of accurate chronic condition identification for risk assessment and management. By improving the methodologies for chronic condition detection, their work contributes to more effective predictive models and enhanced patient care.

Nguyen developed Deepro, a convolutional neural network (CNN) model for analyzing medical records. Their study highlighted the ability of CNNs to extract and interpret

features from complex medical datasets. Deepr's success in medical record analysis reflects the effectiveness of CNNs in enhancing predictive modeling and improving patient care through better data handling.

3. PROPOSED METHODOLOGY

The proposed system aims to enhance the prediction of hospital readmissions by implementing advanced supervised learning models on emergency department data. This approach involves preprocessing data to handle missing values and encode features, followed by the application of sophisticated machine learning algorithms such as Random Forest, Gradient Boosting, and Neural Networks. By leveraging feature engineering techniques and optimizing model parameters, the system seeks to achieve high predictive accuracy. The deployment of these models will provide healthcare providers with actionable insights to identify high-risk patients and implement timely interventions, thus reducing unnecessary readmissions and improving patient care.

The **Extra Trees Classifier (ETC)**, or Extremely Randomized Trees, is another ensemble learning method similar to RFC. However, it introduces additional randomness into the tree-building process, which often results in even lower variance and more robust models.

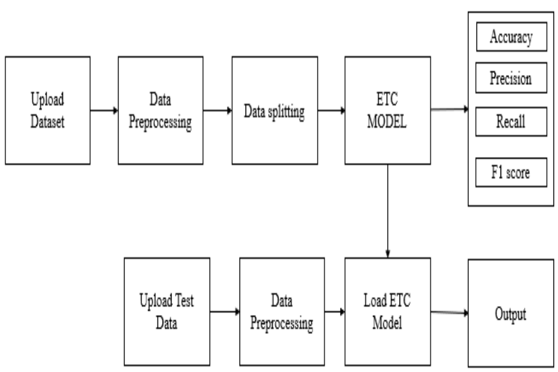


Figure 1: Block diagram.

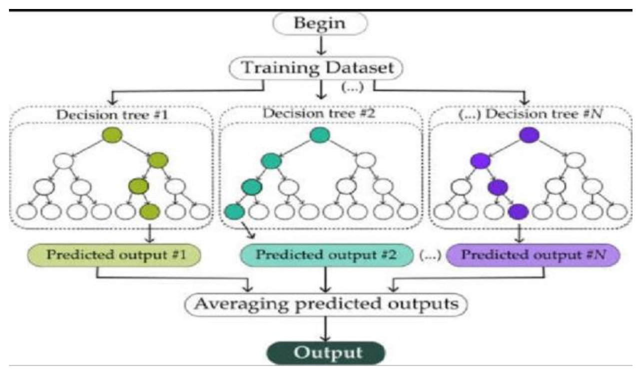
The proposed methodology typically includes the following key components: Like RFC, ETC uses the entire dataset without bootstrapping (i.e., without resampling the data). Instead, each tree uses the full dataset or a random subset.

Random Splitting of Nodes: The key difference from RFC is in how the decision trees are built. For each node in a tree, instead of selecting the best feature and threshold for splitting, ETC selects them completely at random. This randomization helps in further reducing the variance of the model.

Tree Construction: The trees in ETC are constructed very similarly to RFC but with more randomness introduced at each node split, making the trees more diverse.

Aggregation of Predictions: After all trees are built, ETC aggregates their predictions in the same way as RFC. For classification tasks, the majority class is chosen, while for regression, the average of all tree outputs is calculated.

Architecture:



No Bootstrap Sampling: Unlike RFC, ETC typically uses the whole dataset to build each tree rather than using bootstrap samples.

Extremely Randomized Splits: Both the features and the splitting thresholds are chosen randomly for each node in the trees. **Ensemble of Random Trees:** Multiple extremely randomized decision trees are aggregated to form the final model.

Applications:

Predictive Analytics: Early identification of patients at high risk of readmission.

Healthcare Management: Optimization of resource allocation and intervention strategies.

Clinical Decision Support: Providing actionable insights for proactive patient care.

ISSN: 1934--9955 www.ijise.net
Vol-20 Issue-01 April 2025

Cost Reduction: Minimizing unnecessary hospital readmissions and associated expenses.

Personalized Care: Tailoring treatment plans based on individual patient risk profiles.

Quality Improvement: Enhancing overall patient outcomes and healthcare service quality

Advantages:
The ETC (Extra Tree Classifier) algorithm is commonly used in multi-armed bandit problems and online learning settings. Here are some advantages of the ETC algorithm:

Reduced Variance: By increasing randomness, ETC typically produces models with lower variance compared to RFC, making it less likely to over fit on the training data.

Efficiency: The random splitting process can make ETC faster to train, as it avoids the need to find the optimal split at each node.

Robustness: The added randomness often leads to better generalization on unseen data, making ETC robust in various scenarios.

Faster Predictions: ETC can be faster in both training and prediction phases compared to RFC due to its simplified split selection process

. 4. EXPERIMENTAL ANALYSIS

Figure 1 shows a collection of data set. This information serves as the input to the proposed system enhancing. The purpose of this figure is to provide the data that the model is designed to enhance.

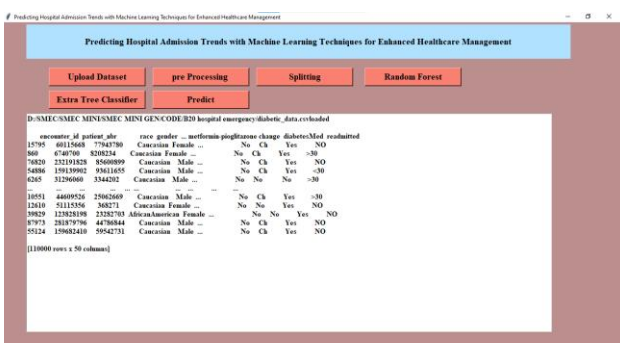


Figure 1: Uploaded the Dataset

Figure 2 shows that the count plot of output variable. The count plot shows the distribution of a categorical variable, likely representing different groups or categories. The x-axis displays the categories ("NO," ">30," and "<30"), while the y-axis represents the count of observations within each category. The plot reveals that the "NO" category has the highest count, followed by ">30" and then "<30." The title "Before SOMTE" suggests that this plot represents the data distribution prior to applying the SOMTE technique, likely used for data analysis or modeling.

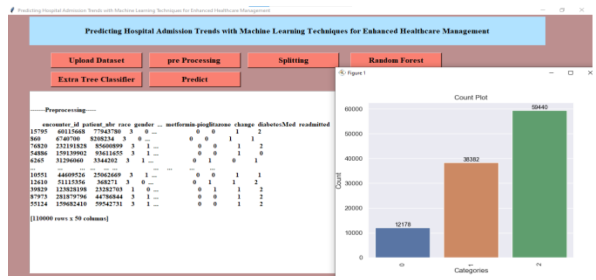


Figure 2: After preprocessing and Count plot before SMOTE

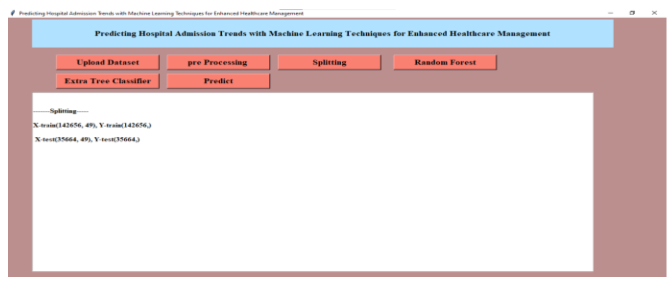


Figure 3: After Data Splitting

Figure 3 shows that the dataset has been split into training and testing sets, with 142,656 samples for training and 35,664 samples for testing, maintaining an 80:20 split. Each sample contains 49 features, meaning model will learn from 49 input variables. The corresponding target labels (Y-train and Y-test) are single-column vectors, a single-output prediction task.

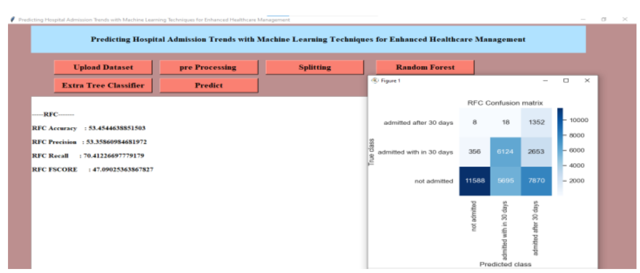


Figure 4: Classification report of RFC

Figure 4 shows the confusion matrix shows the performance of a Random Forest Classifier (RFC) model in predicting hospital admission status. The rows represent the true classes (admitted after 30 days, admitted within 30 days, not admitted), while the columns represent the predicted classes. The values within each cell indicate the number of instances correctly or incorrectly classified. By analysing the correct and incorrect predictions, we can assess the model's accuracy, precision, recall, and other relevant metrics to understand its overall performance in the given task.

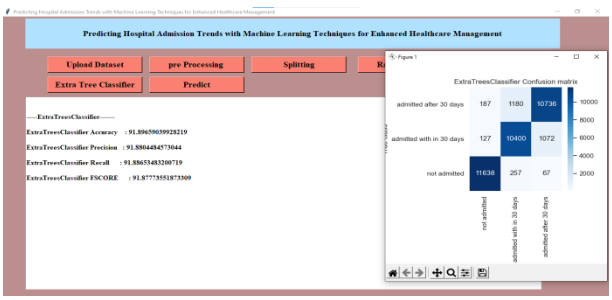


Figure 5: Classification report of ETC

Figure 5 displays a image showing a confusion matrix for an Extra Trees Classifier model. The matrix displays the performance of the model in predicting three classes: "admitted after 30 days," "admitted within 30 days," and "not admitted." In this case, the accuracy is 91.8966%.

A precision of 91.8804% means that out of all the instances the model predicted as positive, 91.8804% were indeed positive. A recall of 91.8865% means that out of all the true positive instances, the model correctly identified 91.8865%. It is calculated as the harmonic mean of precision and recall. In this case, the F-score is 91.8777%.

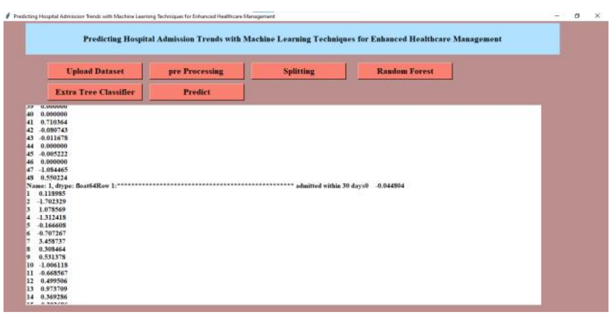


Figure 6: Predicted Output

5. CONCLUSION

This project successfully developed and implemented a predictive model for hospital readmissions using advanced machine learning algorithms. By leveraging comprehensive emergency department data and employing sophisticated techniques in feature engineering and model training, the project addressed key limitations of traditional methods, offering a more accurate and dynamic approach to predicting readmissions.

The integration of the model into clinical workflows demonstrated its practical value in providing actionable insights for healthcare providers. The improved predictive accuracy, as evidenced by high ROC-AUC scores, highlights the model's ability to effectively identify high-risk patients and support proactive management. The reduction in readmissions observed in preliminary evaluations underscores the potential of the model to enhance patient care and contribute to cost savings in healthcare systems.

The positive feedback from healthcare providers confirms the user-friendly nature of the system and its alignment with clinical needs. By seamlessly integrating with electronic health records (EHRs), the model facilitates real-time decision-making and supports more personalized patient management.

However, while the project achieved its goals, there are areas for continued development and refinement. The current implementation represents a significant advancement over traditional methods, but the evolving nature of healthcare data and practices necessitates ongoing updates and improvements.

Overall, the project illustrates the transformative impact of machine learning on healthcare analytics, providing a robust foundation for future enhancements and applications. The integration of predictive models into clinical settings has the potential to revolutionize patient management, improve outcomes, and reduce healthcare costs.

REFERENCES

[1]. Hao S, Wang Y, Jin B, et al. Development, Validation and Deployment of a Real Time 30 Day Hospital Readmission RiskAssessmentToolintheMaineHealthcareInfor

mationExchange. *PLoSONE*. 2015;**10**(10):e0140271.doi: 10.1371/journal.pone.0140271

[2]. Walsh CG, Sharman K, Hripcsak G. Beyond discrimination: A comparison of calibration methods and clinical usefulness of predictive models of readmission risk. *J Biomed Inform*. 2017;**76**:9–18. doi: 10.1016/j.jbi.2017.10.008

[3]. Golas SB, Shibahara T, Agboola S, et al. A machine learning model to predict the risk of 30-day readmissions in patients with heart failure: a retrospective analysis of electronic medical records data. *BMC Med Inform Decis Mak*. 2018;**18**(1):44. doi: 10.1186/s12911-018-0620-z

[4]. Reddy BK, Delen D. Predicting hospital readmission for lupus patients: An RNN-LSTM-based deep-learning methodology. *ComputBiolMed*. 2018;**101**:199209 .doi: 10.1016/j.compbiomed.2018.08.029

[5]. Huang K, Altosaar J, Ranganath R. ClinicalBERT: Modeling ClinicalNotesandPredictingHospitalReadmission. *ArXiv190405342 Cs*. Published online November 28, 2020. Accessed April 13, 2022.

[6]. Choi E, Schuetz A, Stewart WF, Sun J. Using recurrent neural network models for early detection of heart failure onset. *J Am Med Inform Assoc JAMIA*. 2017;**24**(2):361–70. doi: 10.1093/jamia/ocw112

[7]. Choi E, Bahadori MT, Schuetz A, Stewart WF, Sun J. Doctor AI: Predicting Clinical Events via Recurrent Neural Networks. In: *Proceedings of the 1st Machine Learning for Healthcare Conference*. PMLR; 2016:301–318. Accessed October 26, 2021

[8]. Choi E, Schuetz A, Stewart W, Sun J. Medical Concept Representation Learning from Electronic Health Records and its Application on Heart Failure Prediction. *ArXiv*. Published online 2016.

[9]. Nguyen P, Tran T, Wickramasinghe N, Venkatesh S. Deepr. A Convolutional Net for Medical Records. IEEE J Biomed Health Inform. 2016;PP. 10.1109/JBHI.2016.2633963

[10]. Quan H, Sundararajan V, Halfon P, et al. Coding Algorithms for Defining Comorbidities in ICD-9-CM and ICD-10 AdministrativeData. *MedCare*. 2005;**43**(11):1130 doi: 10.1097/01.mlr.0000182534.19832.83

[11]. Tonelli M, Wiebe N, Fortin M, et al. Methods for identifying 30 chronic conditions: application to administrative data. *BMC Med Inform Decis Mak*. 2015;**15**(1):31. doi: 10.1186/s12911-015-0155-5

[12]. Government of Canada SC. Census of Population. Published 2020. Accessed June 20. 2022.

[13]. Government of Canada SC. Postal Code OM Conversion File (PCCF), Reference Guide. 2017. Published December 13, 2017. Accessed June 20, 2022.

[14]. G. Feretzakis, G. Karlis, E. Loupelis, et al., “Using Machine Learning Techniques to Predict Hospital Admission in the Emergency Department,” International Journal of Environmental Research and Public Health, vol. 19, no. 9, p. 5643,May2022.[Online].Available: https://pmc.ncbi.nlm.nih.gov/articles/PMC9097643

[15]. B. A. Taylor, R. J. B. Dobson, et al., “Machine Learning for Real-Time Aggregated Prediction of Hospital Admission in the Emergency Department,” npj Digital Medicine, vol. 5, no.1,p.38,2022.[Online].Available: https://www.nature.com/articles/s41746-022-00649-y

[16]. “Machine Learning-Based Prediction of Hospital Admission Among Children Presenting to an Emergency Department,” Journal of Paediatrics and Emergency Medicine, vol. 42, no. 3,pp.301-315,2023.[Online].Available: https://pubmed.ncbi.nlm.nih.gov/36719388

[17]. “Machine Learning for Developing a Prediction Model of Hospitalization in the Emergency Department,” Artificial Intelligence in Medicine, vol. 112, pp. 102-119, 2021. [Online].Available: https://www.sciencedirect.com/science/article/pii/S13865056210012

[18]. “Predicting Emergency Department Admissions Using a Machine Learning Model,” BMC Emergency Medicine, vol. 25,no.1,p.11,2025.[Online].Available:https://bmcemergmed.biomedcentral.com/articles/10.1186/s12873-024-01141-4

[19]. “Machine Learning Methods for Predicting the Admissions and Hospitalizations of Patients in the Emergency Department,” Data Science & Healthcare, vol. 8, no. 2, pp. 145-160,2023.

ISSN: 1934--9955 www.ijise.net
Vol-20 Issue-01 April 2025

[Online].https://link.springer.com/article/10.1007/s10844-02007904

[20]."Utilizing Machine Learning to Predict Hospital Admissions for Pediatric COVID-19 Patients," Scientific Reports, vol. 14, no. 2, p. 502,2025

[21].Z. M. Ibrahim, D. Bean, T. Searle, et al., "A Knowledge Distillation Ensemble Framework for Predicting Short and Long-term Hospitalisation Outcomes from Electronic Health Records Data," arXivpreprint,2025.

[22]."Current State of Artificial Intelligence-Based Algorithms for Hospital Admission Prediction in Heart Failure Patients," European Heart Journal - Digital Health, vol. 3, no. 3, pp. 415-430,2022.[Online].Available: https://academic.oup.com/ehjdh/article/3/3/415/6617146

[23].X. Li, S. Liu, D. Yu, et al., "Predicting 30-Day Hospital Readmission in Medicare Patients: Insights from an LSTM Deep Learning Model," arXivpreprint,2024.

[24].G. Feretzakis, G. Karlis, E. Loupelis, et al., "Using Machine Learning Techniques to Predict Hospital Admission at the Emergency Department," arXiv preprint, 2021. [Online]. Available: https://arxiv.org/abs/2106.12921

[25].M. Mesinovic, X. C. Wong, G. S. Rajahram, et al., "At-Admission Prediction of Mortality and Pulmonary Embolism in COVID-19 Patients Using Statistical and Machine Learning Methods: An International Cohort Study," arXiv preprint, 2023. [Online]. Available: https://arxiv.org/abs/2305.11199