2025 3rd International Conference on Artificial Intelligence, Blockchain, and Internet of Things (AIBThings) | 979-8-3315-8856-4/25/$31.00 ©2025 IEEE | DOI: 10.1109/AIBThings66987.2025.11296145

Addressing Gender Imbalance in Cirrhosis Prediction with CTGAN and Transformer-Based Generative Models

Divya Saxena
Knight Foundation School of Computing and Information Sciences
Florida International University
Miami, USA
dsaxe001@fiu.edu

Naphtali Rishe
Knight Foundation School of Computing and Information Sciences
Florida International University
Miami, USA
rishe@fiu.edu

Abstract— **Cirrhosis is a severe liver condition characterized by irreversible scarring due to chronic liver damage. Early detection is critical for effective treatment and improved patient outcomes. This study investigates the application of various machine learning models for cirrhosis prediction, including Random Forest (RF), Logistic Regression (LR), Convolutional Neural Networks (CNN), Multilayer Perceptron (MLP), and a hybrid CNN-Long Short-Term Memory (CNN-LSTM) model. To address the gender imbalance in the dataset—where male samples outnumber female—synthetic data generation techniques are applied, with a focus on Conditional Generative Adversarial Networks (CTGAN). Two hybrid balancing approaches are also explored: one combining CTGAN with Edited Nearest Neighbors (ENN) and the other with the Synthetic Minority Over-sampling Technique (SMOTE). This paper introduces a novel generative framework, the Transformer-Conditional Generative Adversarial Network (TransCTGAN), which incorporates Transformer-based self-attention layers into both the generator and discriminator of the CTGAN architecture. The effectiveness of synthetic data is assessed through model performance improvements after dataset balancing. Our results demonstrate that CTGAN-based augmentation significantly boosts predictive accuracy, showing the potential of synthetic data in enhancing machine learning models. While not all models show marked improvement, the approach effectively mitigates class imbalance without introducing bias toward the male-dominated class.**

Keywords—Cirrhosis, Machine Learning, Random Forest, Logistic Regression, CNN, MLP, CNN-LSTM, CTGAN, Synthetic Data, Gender Imbalance, GAN, SMOTE, ENN, Transformer

I. INTRODUCTION

Cirrhosis, a chronic illness marked by progressive liver scarring, is usually the consequence of long-term harm brought on by conditions like nonalcoholic fatty liver disease, hepatitis B or C, or heavy alcohol use. Cirrhosis can cause liver failure, requiring a transplant, or even be fatal if it is not detected in time. Early and precise diagnosis is essential because the disease is incurable in its advanced stages, underscoring the significance of creating trustworthy prediction models.

As long as enough high-quality data is available, machine learning (ML) provides strong capabilities for finding patterns in complicated medical datasets, allowing for the fully accurate early diagnosis of diseases like cirrhosis. However, issues like imbalances in datasets, especially regarding demographic representation, and limited data availability continue to be major obstacles. For instance, women are often underrepresented in clinical datasets, leading to biased models that may underperform for this group.

To address the issues of data scarcity and imbalance, synthetic data generation has emerged as a viable solution. Generative Adversarial Networks (GANs), especially Conditional GANs (CTGANs), can generate realistic new samples conditioned on variables such as gender, thereby helping to balance datasets and improve model robustness and fairness. CTGAN was used to create synthetic female data because the original cirrhosis dataset included an excessive number of male individuals. By producing synthetic records that closely resemble the statistical characteristics of the original female data, this method reduces gender-based biases and raises the general caliber and equity of model predictions.

Through a generator-discriminator framework, CTGAN learns to generate synthetic data that is more realistic by using the discriminator's feedback. This method was used to create artificial female samples that more accurately reflected the features of the original dataset. Two hybrid models—ECTGAN, which combines CTGAN with Edited Nearest Neighbors, and SCTGAN, which combines CTGAN with the SMOTE technique—were created in order to further improve data quality and class balance. These methods enhance synthetic data by reducing noise and reinforcing class balance, with their effectiveness assessed by comparing model performance before and after data balancing.

Additionally, a novel approach called TransCTGAN—a Transformer-based conditional GAN—was introduced. This model combines the generative power of GANs with the Transformer architecture's strength in recognizing patterns, producing more representative datasets that support improved model performance and fairness across gender groups.

Following data enhancement and gender balance correction, several machine learning models were evaluated for cirrhosis prediction. These included both traditional algorithms and advanced deep learning techniques. Specifically, the models assessed were Random Forest (RF), Logistic Regression (LR), Convolutional Neural Networks (CNN), Multilayer Perceptrons (MLP), and a hybrid CNN-LSTM model. Each of these algorithms offers unique advantages, from LF's effectiveness in high-dimensional spaces to CNN-LSTM's ability to capture spatial and sequential patterns in the data.

This study's main goal is to evaluate how well these various machine learning models predict cirrhosis while addressing gender inequality by creating fake data. This entails assessing how gender-balanced data affects model accuracy, contrasting deep learning architectures with conventional models, and determining whether synthetic data enhances predictability and justice, especially for marginalized groups.

II. RELATED WORK

A. Machine Learning Models for Cirrhosis Prediction

With the growing use of machine learning in healthcare, predictive modeling for cirrhosis has gained more interest. Conventional models like logistic regression, decision trees, and support vector machines (SVM) have been often used. For example, Utku et al. [1] discovered that while SVM and Decision Trees performed well on datasets related to liver disorders, they were inflexible when dealing with unbalanced data. Similarly, on liver disease prediction tasks, Guo et al. [2] demonstrated that ensemble models, in particular Random Forest and AdaBoost, had superior generalization than single classifiers.

More recent work has explored deep learning methods. Lanjewar et al. [4] employed CNNs and MLPs to capture nonlinear interactions and improve prediction accuracy over traditional models. A CNN-LSTM hybrid model that incorporates temporal and spatial patterns in medical data was proposed by Kumar et al. [14] and has shown promising results for progressive disorders such as cirrhosis. Notwithstanding these advancements, these models sometimes fail to account for the difficulties presented by unbalanced datasets, which results in inaccurate forecasts for women and other minority classes.

B. Data Imbalance in Medical Prediction Tasks

In clinical machine learning, class imbalance is still a major problem. Certain patient groups are frequently overrepresented in medical databases, which skew model projections. Although synthetic oversampling occasionally created noise, Zhang et al. [5] discovered that SMOTE enhanced recall for minority groups when applied to correct class imbalance in disease datasets. To reduce redundancy, more recent approaches have merged oversampling with data cleaning methods such as Tomek Links or Edited Nearest Neighbors proposed by Agyemang et al. [15]. However, most conventional sampling methods frequently fail to maintain feature interdependencies, particularly in complicated medical data, and presume linear class borders. A move toward generative models for more realistic dataset balance has been spurred by these constraints.

C. GANs and CTGAN for Synthetic Medical Data Generation

GANs have emerged as a robust solution for generating synthetic samples in healthcare. GANs, particularly Conditional GANs (CGANs), can generate class-specific data while maintaining the statistical relationships of the original dataset as discussed by Ahmed et al.[16].

The problems that traditional GANs encounter when working with mixed categorical and continuous variables are addressed by CTGAN, which was developed especially for tabular data and first presented by Xu et al. [8]. The usefulness of CTGAN in healthcare has been confirmed by recent research. To maintain data fidelity and privacy, Alqulaity et al. [9] created artificial patient records for uncommon illnesses using CTGAN. Majeed et al. [10] used CTGAN to create artificial female samples in a dataset of liver illness that was dominated by men to address the gender imbalance. Their results demonstrated the feasibility of CTGAN for demographic balance, demonstrating a notable improvement in fairness without compromising model accuracy.

D. Gender Imbalance in Healthcare Datasets

There is ample evidence of gender bias in medical datasets. Women's underrepresentation in clinical datasets contributes to the underdiagnosis of cirrhosis and other liver illnesses discussed by Straw et al., [12]. As a result, women's diagnosis accuracy is decreased by predictive models that are tailored for male patients. Hupont et al. [11] addressed this issue by proposing gender-aware neural networks that mitigate bias by incorporating demographic variables during training. Musa et al. (2023) investigated domain adaptation strategies to lessen demographic bias in the prediction of chest X-rays.

Still, a lot of these methods rely on labeled data, which is hard to come by for minority populations. By producing realistic synthetic data that is suited to underrepresented communities, generative models such as CTGAN present a possible, data-driven substitute.

E. Hybrid and Transformer-Based Approaches for Data Balancing

To improve data quality, recent research has investigated merging CTGAN with oversampling or cleaning methods. To improve model generalization, Zhou et al. [18] suggested CTGAN-SMOTE, which produces high-fidelity samples without overfitting. Similarly, ECTGAN, which incorporates Edited Nearest Neighbors to lower noise in the synthetic data pipeline, was introduced by Shi et al. [7].

By addressing gender imbalance in cirrhosis datasets using ECTGAN, SCTGAN, and a novel TransCTGAN framework, our study builds on previous findings and ensures high-fidelity data augmentation and enhanced fairness across ML models.

F. Gaps in Current Research and Our Contributions

While prior studies confirm the effectiveness of ML models in disease prediction and the usefulness of GAN-based methods for synthetic data generation, several gaps remain. First, hardly much research use sophisticated generative models to directly address gender imbalance in cirrhosis datasets. Second, hybrid models like ECTGAN and SCTGAN are still understudied in this field, despite the use of CTGAN. Finally, transformer-based GANs have not yet been evaluated for liver disease prediction, despite their promising performance in general datasets.

By implementing and contrasting ECTGAN, SCTGAN, and TransCTGAN for gender-balanced data augmentation, assessing the impact of balancing on the performance of deep learning models and standard machine learning, and proving increases in fairness and prediction accuracy across gender groups, our study closes these disparities.

III. METHODOLOGY

A. Dataset

Clinical cirrhosis data, including age, gender, medical history, and liver function tests, are included in the dataset used in this investigation [6]. There is a gender imbalance in this publicly accessible Kaggle dataset, which includes both male and female patients. 424 PBC patients who were referred to the Mayo Clinic over a ten-year period were eligible for a D-penicillamine randomized, placebo-controlled study. Of these, 312 instances with largely complete data took part in the experiment, and 112 patients who chose not to participate agreed to baseline assessments and survival monitoring. Six of the 312 study participants were lost to follow-up soon after diagnosis, therefore the dataset includes an additional 106 cases.

B. Data Preprocessing

Prior to training the models, the dataset undergoes several preprocessing steps:

1) Handling missing values: Missing entries are imputed using mean imputation, the median, or the mode, depending on the feature type

2) Feature scaling: Continuous features are normalized using Min-Max scaling to ensure they lie within the same range.

3) Encoding categorical variables: Categorical features are encoded using one-hot encoding.

4) Gender ratio: There were 374 males vs 44 females before the synthetic data was added.

C. Synthetic Data Generation using various models

A generative model based on deep learning, CTGAN, generates synthetic data that closely resembles the original dataset. CTGAN creates more female samples while maintaining the natural distribution of other variables to balance gender in our data. The generator creates gender-conditioned synthetic data using a conditional configuration. We have used three approaches using hybrid models:

1) ECTGAN

a) Enhanced Synthetic Data Quality: The combination of CTGAN for generating synthetic data and ENN for cleaning ensures that the synthetic data is not only realistic but also clean and free from outliers.

b) Class Imbalance Handling: CTGAN can help augment the minority class in imbalanced datasets, and ENN ensures that the generated samples are reliable and robust.

c) Improved Model Training: By using cleaner synthetic data, models trained on this data will likely perform better and be more generalized to unseen data.

Using ECTGAN, we then generated synthetic female data to balance the gender ratio to approximately 333 females and 374 males. This balanced distribution helps reduce gender bias in the dataset, allowing the model to learn more fairly and perform more robustly across both genders in real-world applications.

2) SCTGAN

a) Generate Synthetic Data with CTGAN: Use CTGAN to generate realistic synthetic data based on the original dataset, including both the majority and minority classes.

b) Apply SMOTE on the Generated Data: After generating synthetic data using CTGAN, apply SMOTE to further augment the minority class. This allows us to create even more synthetic examples for the minority class, which can improve the performance of models trained on this augmented data.

We were able to correctly balance the gender distribution to 704 females and 704 males after creating synthetic data. By addressing the initial gender bias, this equal representation enhances the model's capacity to learn equally for both groups. Consequently, this methodology facilitates the creation of a more impartial, egalitarian, and trustworthy model for practical healthcare applications.

3) TransCTGAN

CTGAN can be enhanced by integrating Transformer capabilities to explore whether it yields better results. To incorporate Transformer layers into the CTGAN implementation, we can modify both the generator and discriminator components by adding self-attention layers, which are central to Transformer models. Since the Transformer architecture is particularly effective for sequence-based data, embedding the attention mechanism into the existing CTGAN framework could potentially improve the model's ability to capture complex dependencies between features in tabular data. Here is how we have added Transformer capabilities to CTGAN:

a) Adding Transformer Encoder Layer: We have modified the generator and discriminator by introducing a Transformer encoder layer.

b) Adjusting the Generator and Discriminator: Add new layers in the existing network architecture of CTGAN to integrate Transformer mechanisms.

c) Model Architecture:

- Transformer Layers: We've introduced the TransformerEncoderLayer and TransformerEncoder from torch.nn. These layers use self-attention to process the input data.
- Generator (TransformerGenerator): A fully connected layer (fc1) transforms the input into a dimension suitable for Transformer processing. The TransformerEncoder is applied to the data. A final fully connected layer (fc2) produces the output that matches the data dimensions.
- Discriminator (TransformerDiscriminator): The input is reshaped to match the expected input dimensions of the Transformer. Like the generator, the data passes through the Transformer encoder. The output layer (fc2) produces a scalar value indicating whether the input is real or fake.

Using TransCTGAN, we generated synthetic data to balance the gender distribution to approximately 374 females and 393 males. This near-equal ratio helps correct the original imbalance, ensuring fairer representation of both genders. By addressing gender bias, the model can learn more equitably, resulting in a more robust and unbiased tool for real-world healthcare applications.

A. Models for Prediction

We will be utilizing the machine learning models listed below to assess and compare the real dataset with the synthetic data generated by ECTGAN, SCTGAN, and TransCTGAN. The goal is to evaluate how well synthetic data mirrors the characteristics and patterns of the real data. By employing a variety of models, we have aimed to gain a comprehensive understanding of how synthetic datasets perform in terms of accuracy, generalization, and their ability to replicate the structure and distributions present in real-world data. This evaluation will help us determine the effectiveness of ECTGAN, SCTGAN, and TransCTGAN in producing realistic synthetic data that can be used for various machine learning tasks. The models are Random Forest, Logistic Regression, Convolutional Neural Networks, Multilayer Perceptron, and CNN-LSTM Hybrid.

IV. RESULTS

A. *Training Stability Analysis via Generator and Discriminator Loss Curves*

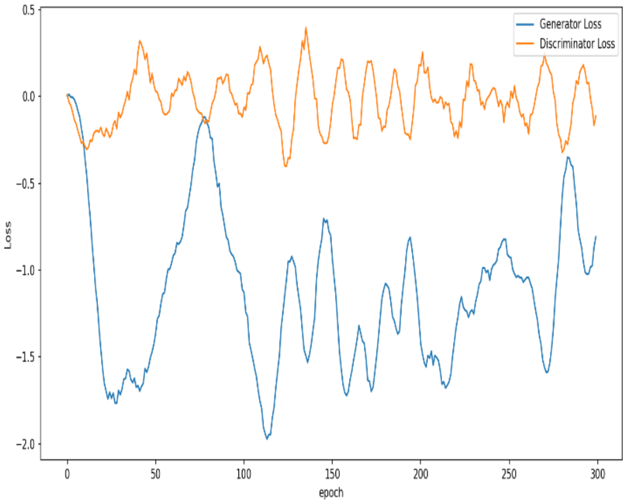

Fig. 1 Generator and Discriminator Loss in ECTGAN

Fig. 1, Fig. 2 and Fig. 3 present the generator and discriminator loss values over 300 training epochs for ECTGAN, SCTGAN, and TransCTGAN, respectively. Across all three models, the training behavior shows signs of stability and convergence, indicating effective GAN optimization.

In Fig. 1 (ECTGAN), the generator loss stabilized at a negative value, while the discriminator loss fluctuated around zero. This behavior suggests that the generator learned to produce realistic outputs over time, while the discriminator maintained a balanced ability to distinguish between real and generated samples—an indicator of training equilibrium.

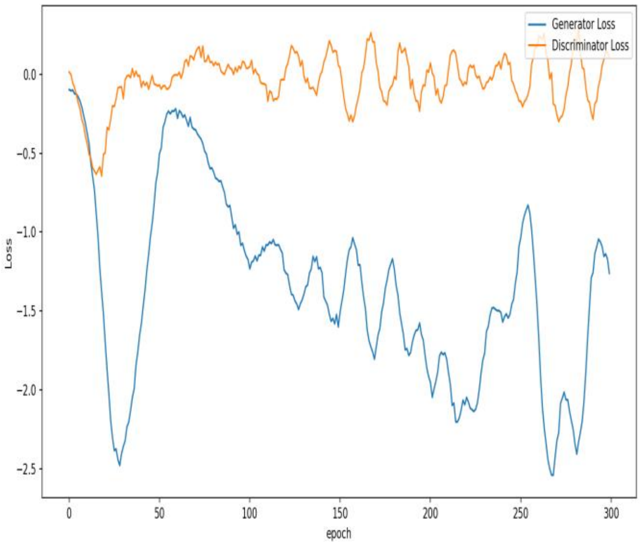

Fig. 2 Generator and Discriminator Loss in SCTGAN

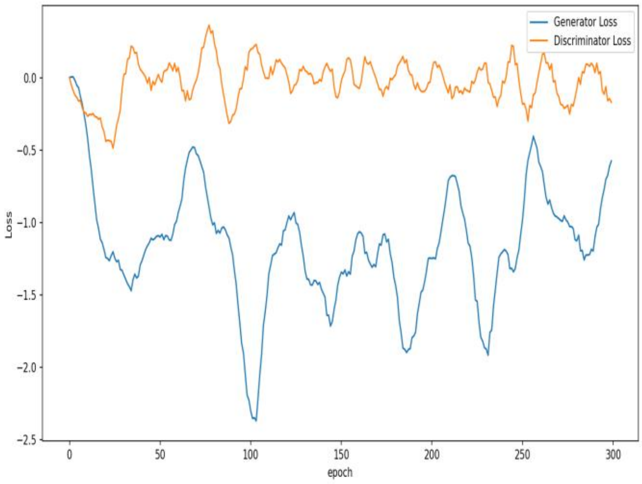

Fig. 3 Generator and Discriminator Loss in TransCTGAN

Similarly, in Fig. 2 SCTGAN shows the generator loss also stabilizing at a negative value, with the discriminator loss consistently hovering near zero. This pattern reflects stable training dynamics and supports the effectiveness of the SCTGAN architecture in maintaining adversarial balance during the training process.

In Fig. 3 TransCTGAN, the generator loss initially started at zero and gradually stabilized at a negative value, while the discriminator loss fluctuated slightly below zero. This consistent trend further confirms effective and stable model training, as both networks reach a point where neither dominates the other, allowing the generator to improve without overwhelming the discriminator.

Overall, the generator and discriminator loss curves across all three models indicate convergent behavior and training stability, which are crucial for generating high-quality synthetic data using GANs.

B. *Model Performance With and Without Synthetic Data*

We first evaluated the models on the original imbalanced dataset prior to generating any synthetic data and evaluated their performance using various metrics. These metrics include the Receiver Operating Characteristic (ROC) curve, accuracy, confusion matrix, and others. By assessing these evaluation metrics, we were able to establish a baseline for the model's performance on the imbalanced data, which would serve as a reference point for comparing the results after introducing the synthetic data generated by ECTGAN, SCTGAN, and TransCTGAN. This step allowed us to understand how well the models perform in the presence of class imbalance and provided insight into the areas where improvement might be needed.

After generating synthetic data using ECTGAN, SCTGAN and TransCTGAN to address the gender imbalance, we have retrained the machine learning models on the newly balanced dataset. This approach led to significant improvements in several performance metrics. In cases where the performance did not show substantial improvement, it remained the same, suggesting that the synthetic data generated closely mirrors the real dataset. Additionally, this result indicates that the synthetic data effectively mitigates the gender imbalance, maintaining the integrity of the data distribution while enhancing model performance. This suggests that the generated synthetic data is both realistic and useful for training, without introducing significant bias.

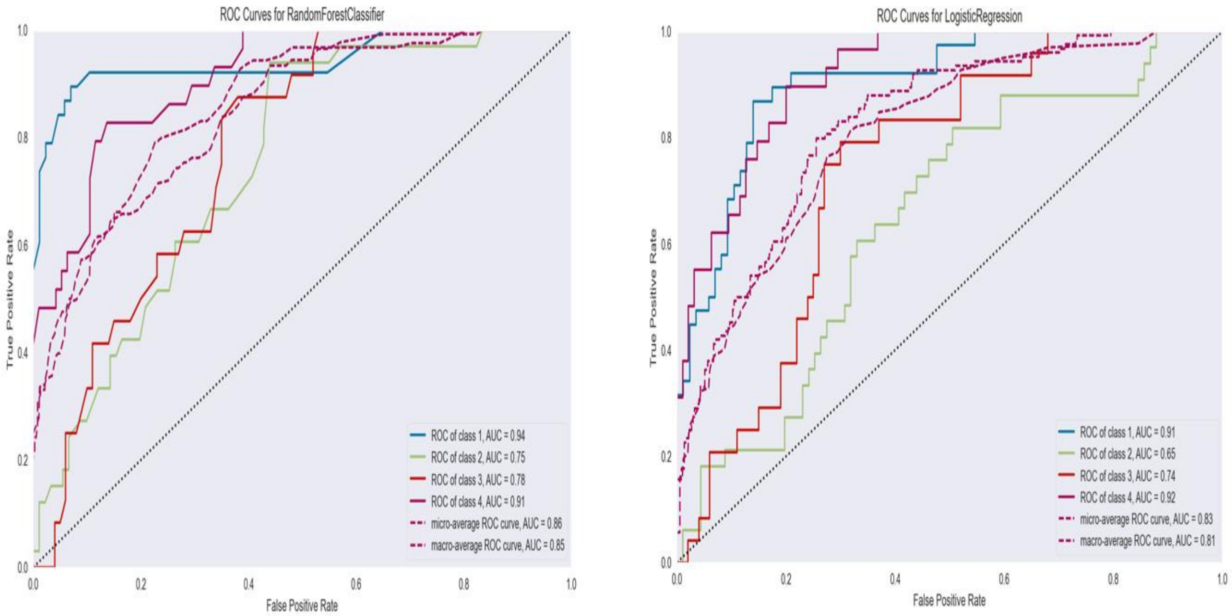

Fig. 4 ROC curves for Real Dataset using Random Forest and Logistic Regression—class-wise, micro-average, and macro-average

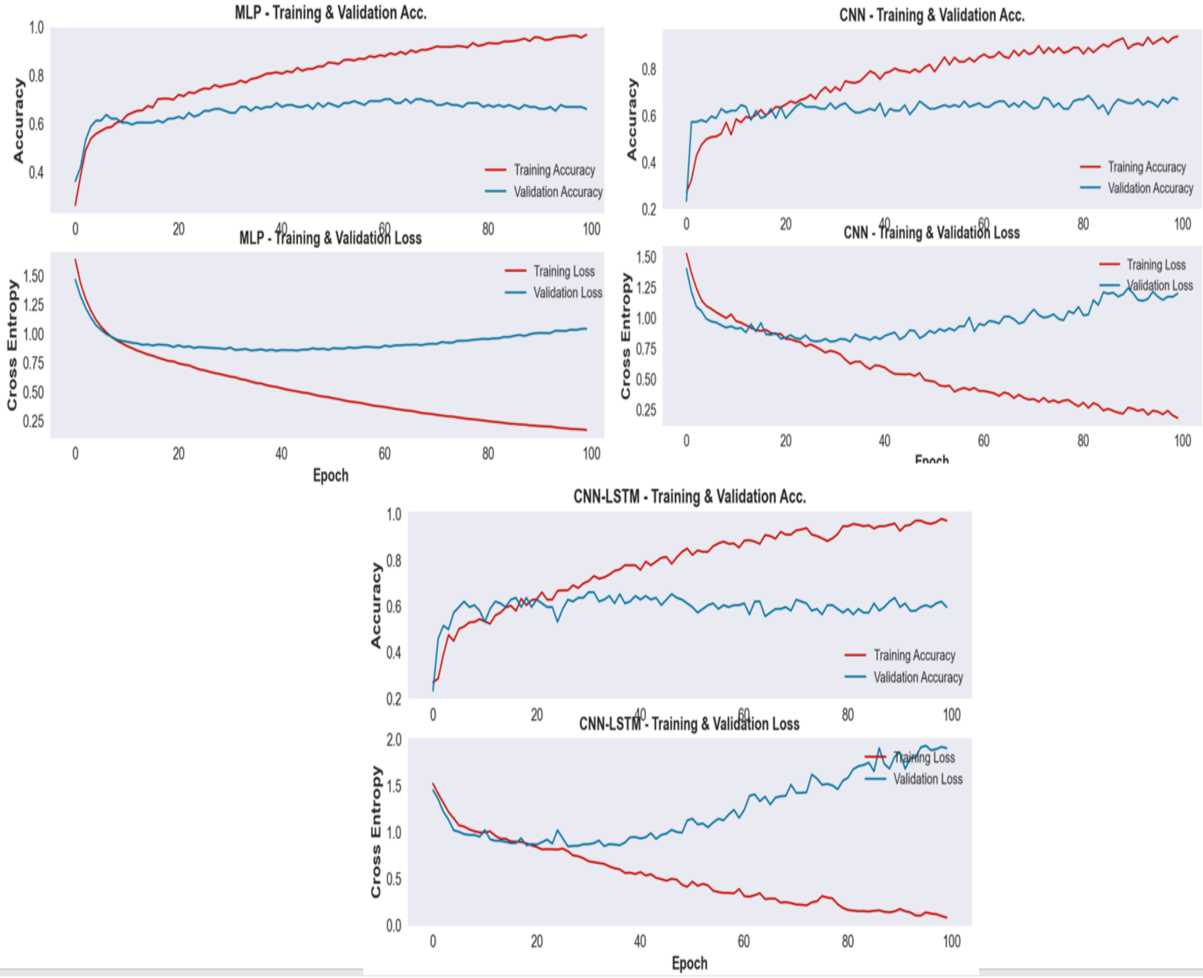

Fig. 5 Accuracy and Loss for Real Dataset showing training and validation for MLP, CNN, and CNN-LSTM

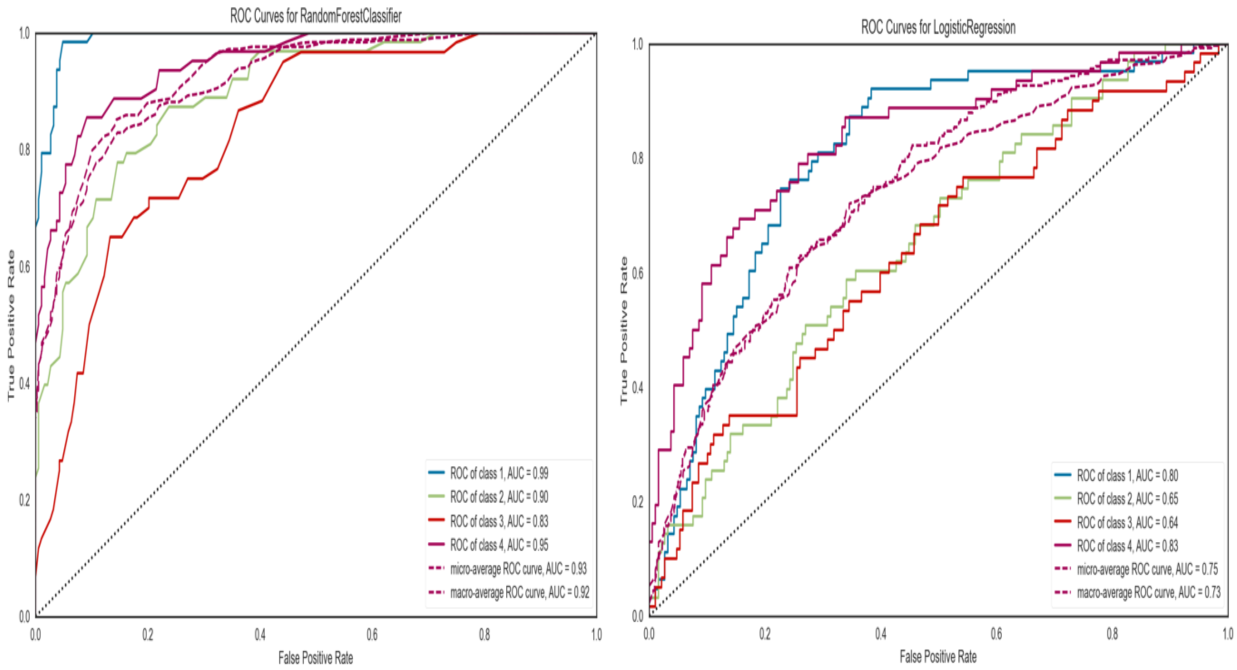

Fig. 6 ROC curves for Synthetic Dataset (ECTGAN) using RF and LR- class wise, micro-average and macro-average

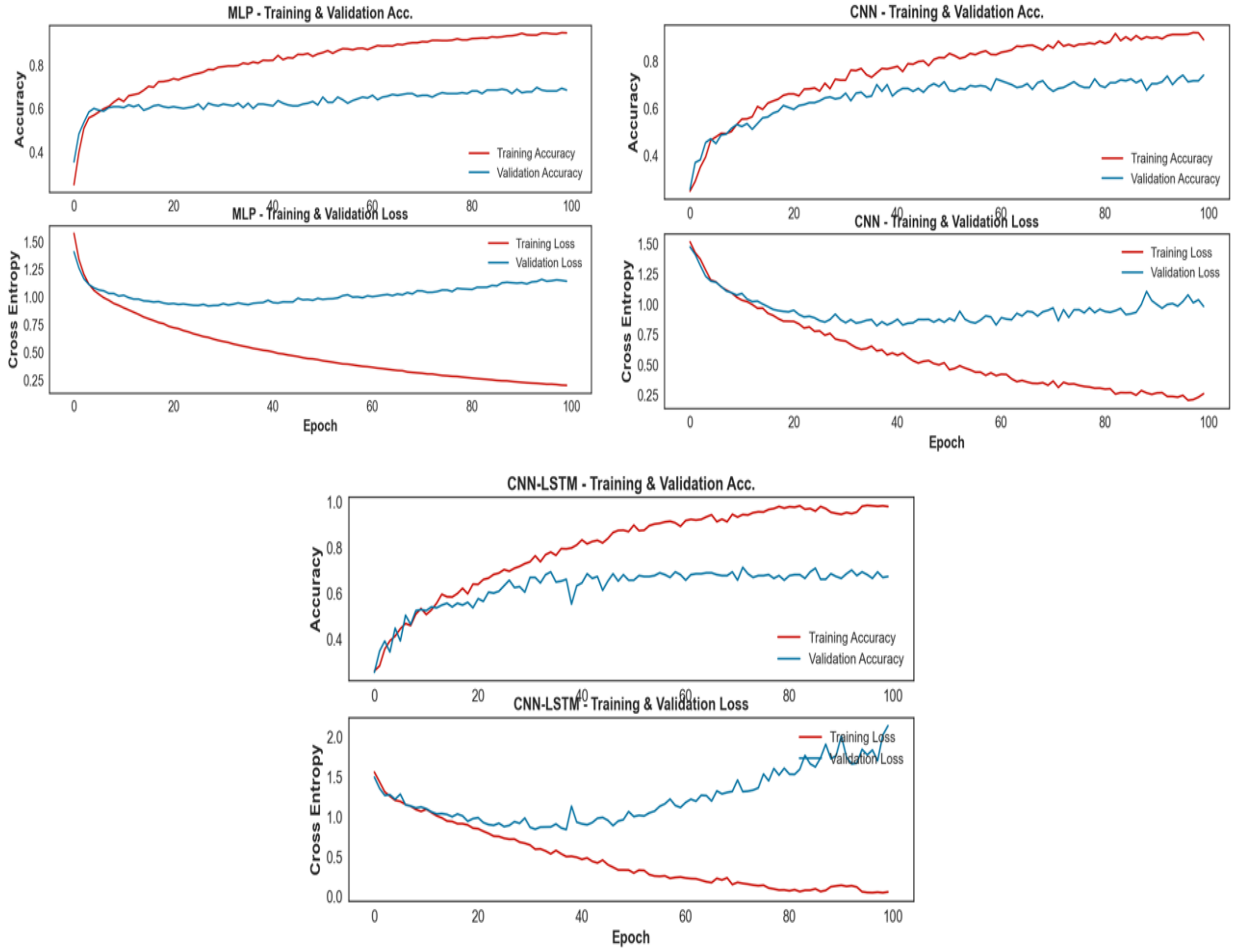

Fig. 7 Accuracy and Loss for Synthetic Dataset (ECTGAN) showing training and validation for MLP, CNN, and CNN-LSTM

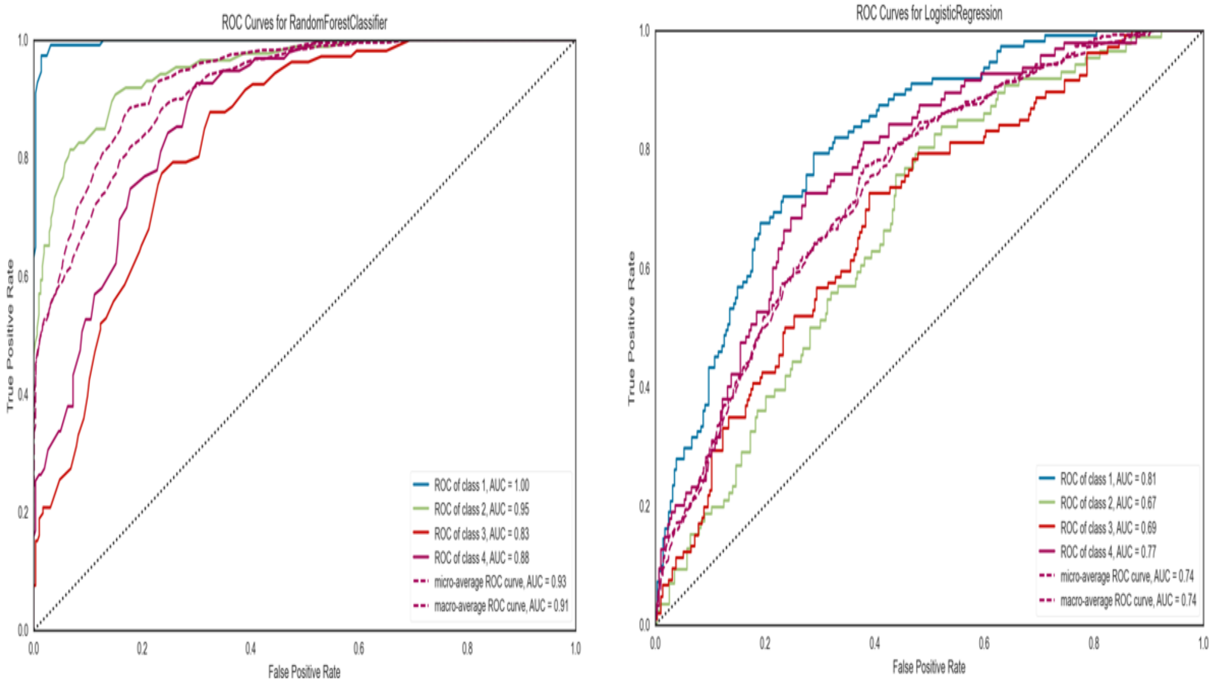

Fig. 8 ROC curves for Synthetic Dataset (SCTGAN) using RF and LR -- class-wise, micro-average, and macro-average

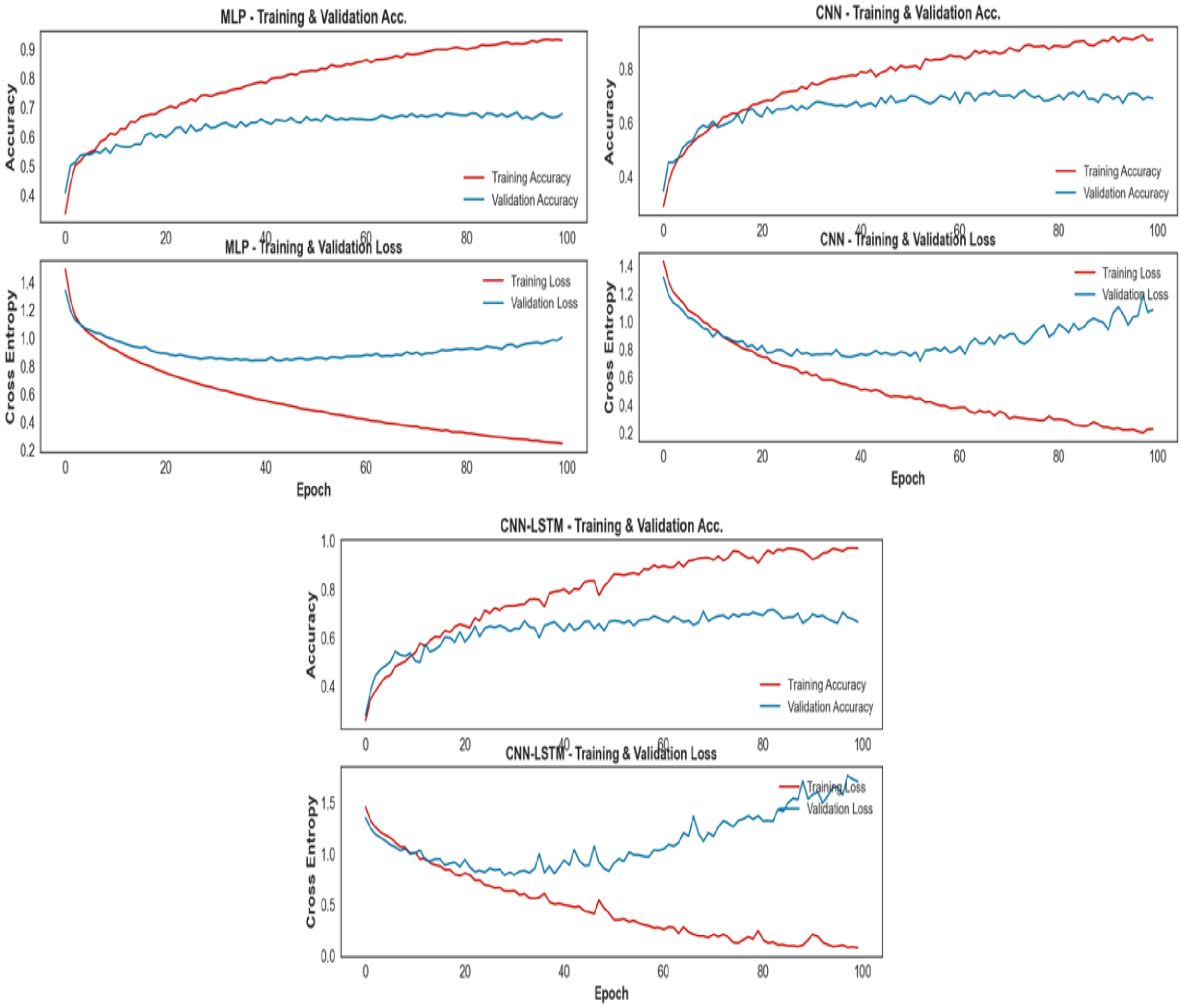

Fig. 9 Accuracy and Loss for Synthetic Dataset (SCTGAN) showing training and validation for MLP, CNN, and CNN-LSTM

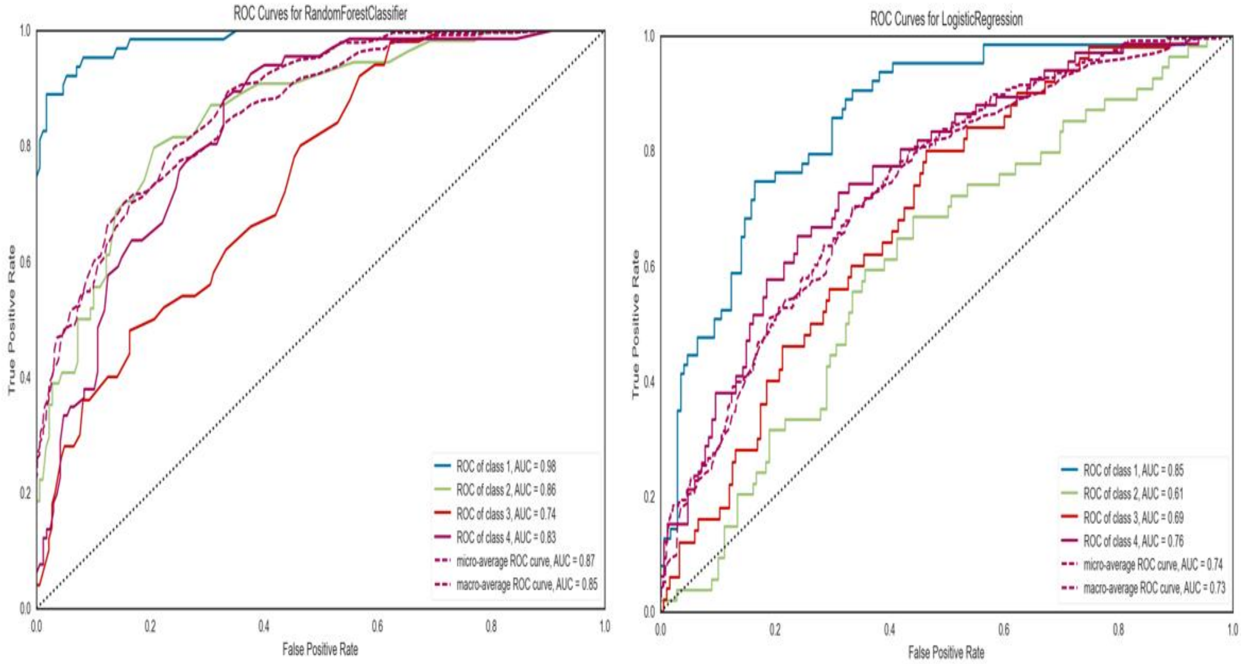

Fig. 10 ROC curves for Synthetic Dataset (TransCTGAN) using RF and LR -- class-wise, micro-average, and macro-average

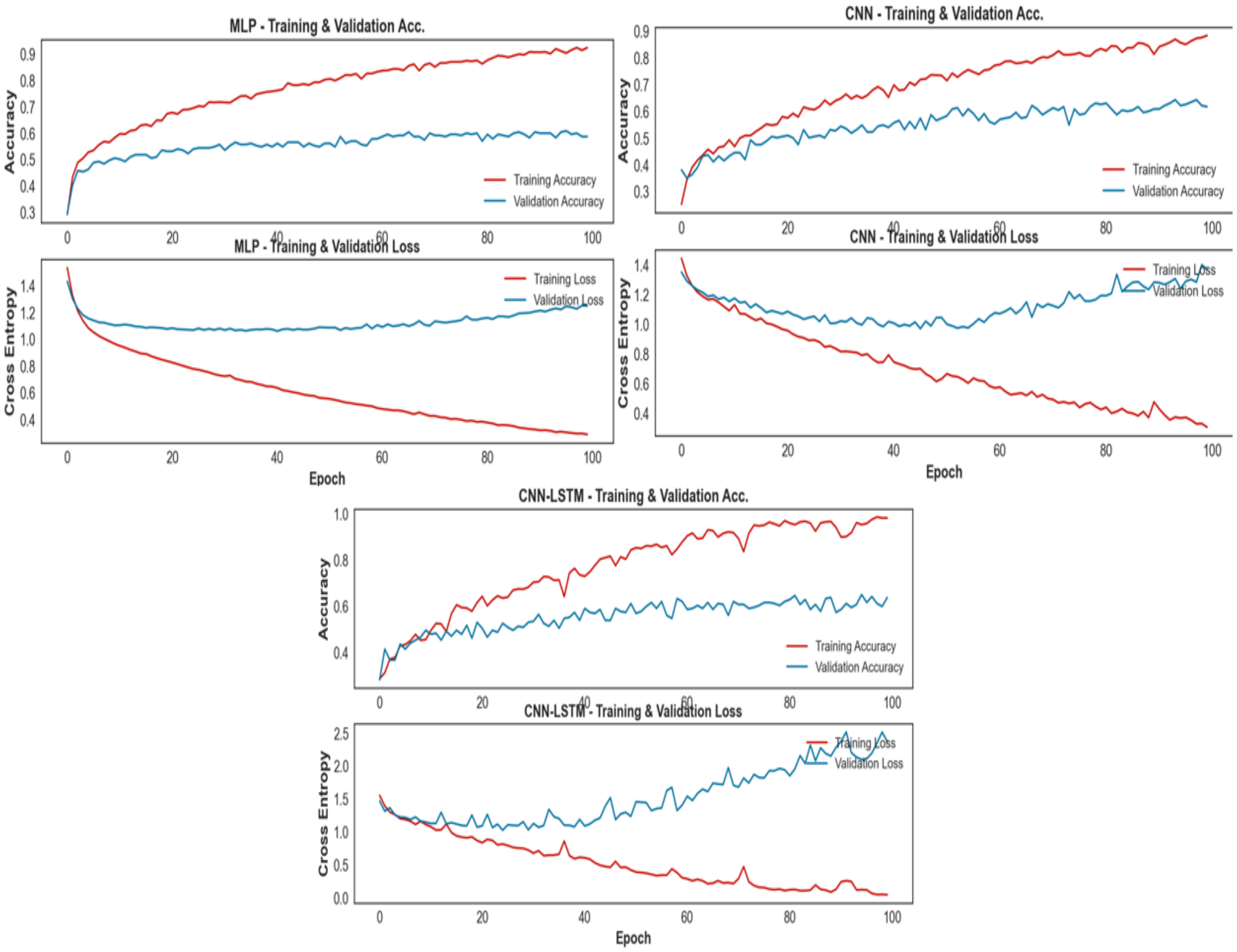

Fig. 11 Accuracy and Loss for Synthetic Dataset (TransCTGAN) showing training and validation for MLP, CNN, and CNN-LSTM

C. Comparison of Model Performance

Below is a table comparing the performance of both conventional machine learning models (Random Forest, and Logistic Regression) and deep learning models (MLP, CNN, and CNN-LSTM) on real vs synthetic datasets, evaluated across various performance metrics

Table I ROC curves for Real and Synthetic dataset

Models	**Real Dataset**	**Synthetic Dataset**		
		ECTGAN	**SCTGAN**	**TransCTGAN**
RF	0.86	0.93	0.98	0.87
LR	0.83	0.75	0.74	0.74

Based on the results in Table I, it is evident that, for the real dataset with gender imbalance, Random Forest (RF) performed the best. However, when synthetic data was introduced using hybrid models of ECTGAN, SCTGAN, and TransCTGAN, RF showed significant improvement, particularly after balancing the dataset using SCTGAN. This approach outperformed ECTGAN and TransCTGAN–based balancing, with RF still showing superior performance.

Table II Accuracy at 100th epoch for Real and Synthetic dataset

Models	**Real Dataset**		**Synthetic Dataset**					
			ECTGAN		**SCTGAN**		**TransCTGAN**	
	Training	**Validation**	**Training**	**Validation**	**Training**	**Validation**	**Training**	**Validation**
MLP	0.96	0.66	0.95	0.64	0.89	0.66	0.91	0.58
CNN	0.93	0.72	0.93	0.70	0.92	0.65	0.89	0.61
CNN-LSTM	0.98	0.70	0.97	0.66	0.96	0.67	0.97	0.63
RF	1.0	0.62	1.0	0.75	1.0	0.72	1.0	0.64
LR	0.59	0.51	0.53	0.49	0.45	0.44	0.5	0.46

Table II summarizes training and validation accuracy at the 100th epoch for five models trained on real and synthetic datasets. While real data yielded the highest accuracy overall, TransCTGAN showed strong performance, particularly with deep learning models.

The CNN-LSTM model achieved 0.97 training and 0.63 validation accuracy with TransCTGAN, close to 0.98/0.70 on real data. Similarly, CNN reached 0.89/0.61 (training/validation), compared to 0.93/0.72 on real data, indicating that TransCTGAN captures spatial and temporal patterns well. MLP showed a larger drop with 0.91/0.58, versus 0.96/0.66 on real data, suggesting simpler models may be more affected by distribution shifts.

Random Forest maintained perfect training accuracy (1.0), and slightly improved validation on TransCTGAN (0.64) compared to real data (0.62), though still below ECTGAN and SCTGAN. Logistic Regression remained the weakest, with 0.50/0.46 on TransCTGAN and 0.59/0.51 on real data.

Overall, TransCTGAN produces reliable synthetic data, especially for deep models, with minimal performance loss. Additionally, training for 100 epochs helped reduce gender imbalance.

The introduction of synthetic data allowed these models to generalize better across both genders, eliminating the bias that was previously present in the real dataset, which had an overrepresentation of males. This ensures that predictions are now more balanced and accurate for both male and female patients, offering a more equitable approach to cirrhosis prediction than relying solely on the imbalanced real dataset. The use of SCTGAN proved effective in addressing gender imbalances, providing a more robust and fairer predictive model.

Table III Confusion Matrix for Real and Synthetic dataset for stages 1,2,3 and 4

Models	**Real Dataset**				**Synthetic Dataset**											
					ECTGAN				**SCTGAN**				**TransCTGAN**			
	1	**2**	**3**	**4**	**1**	**2**	**3**	**4**	**1**	**2**	**3**	**4**	**1**	**2**	**3**	**4**
MLP	0.92	0.42	0.46	0.76	0.92	0.62	0.47	0.73	0.91	0.62	0.50	0.66	0.84	0.56	0.42	0.50
CNN	0.89	0.45	0.5	0.76	0.94	0.78	0.45	0.79	0.96	0.73	0.36	0.71	0.89	0.63	0.50	0.44
CNN-LSTM	0.84	0.39	0.46	0.62	0.90	0.73	0.33	0.71	0.93	0.56	0.49	0.65	0.89	0.61	0.58	0.47

Table III presents the confusion matrix for all stages of cirrhosis (Stages 1, 2, 3, and 4) using deep learning models. From the table, it can be observed that for Stage 1 and Stage 2, SCTGAN outperformed ECTGAN and TransCTGAN, showing better performance in correctly classifying these stages. However, for Stage 3 and Stage 4, the models exhibited similar performance on both real and synthetic datasets, with no significant differences observed in accuracy or classification results between the two. This suggests that while synthetic data enhances prediction accuracy for earlier stages, it has a lesser impact on the more advanced stages of cirrhosis. These findings suggest that TransCTGAN is effective for early-stage cirrhosis prediction but may require further tuning to better capture patterns in later-stage data. Overall, GAN-based synthetic data helps improve stage-wise classification and mitigate imbalances in real-world medical datasets.

Table IV Friedman test to detect differences in model performance across real vs synthetic datasets (ECTGAN, SCTGAN, TransCTGAN)

Models	**p-value**
MLP	0.0022
CNN	0.0159
CNN-LSTM	0.0962
RF	0.3916
LR	0.0020

To assess whether the performance of machine learning models varied significantly when trained on real versus synthetic datasets (generated by ECTGAN, SCTGAN, and TransCTGAN), the Friedman test was conducted as shown in Table IV.. A p-value less than 0.05 indicates statistically significant differences across the datasets. MLP, CNN, and LR demonstrated statistically significant differences in performance when trained on real versus synthetic datasets, indicating greater sensitivity to the quality and distribution of synthetic data. In contrast, CNN-LSTM and Random Forest (RF) did not exhibit significant differences, suggesting these models are more robust across varying data sources.

V. Conclusion

This study demonstrates that cirrhosis can be effectively predicted using a variety of machine learning models, including Random Forest, Logistic Regression, CNN, MLP, and CNN-LSTM. We used three variations of Conditional Tabular GANs (ECTGAN, SCTGAN, and TransCTGAN) to solve gender imbalance in the dataset, which is a prevalent problem in clinical data. Especially for the underrepresented female class, our models were able to produce high-quality synthetic samples that closely resemble the actual data distribution.

Our tests demonstrate that adding synthetic data based on GANs increases prediction accuracy in several models, particularly improving performance on the minority class. The models avoided bias towards the male-dominated data by maintaining balanced performance across genders, even in situations where synthetic augmentation did not result in statistically significant increases.

These results demonstrate how generative AI may be used in conjunction with deep learning and classical models to address data imbalances in the healthcare industry. To verify generalizability, additional validation with bigger and more varied datasets is required. Furthermore, even though synthetic data has potential, its clinical dependability needs to be thoroughly evaluated before being implemented in actual diagnostic systems.

VI. Future Work

Future studies can investigate the application of additional generative models for data augmentation to increase the accuracy of the prediction of cirrhosis. Incorporating a wider range of datasets and investigating the models' interpretability will also be essential for practical implementation in clinical settings.

References

[1] A Utku, A. (2023). Deep Learning Based Cirrhosis Detection. Operational Research in Engineering Sciences: Theory and Applications, 6(1).

[2] Guo, A., Mazumder, N. R., Ladner, D. P., & Foraker, R. E. (2021). Predicting mortality among patients with liver cirrhosis in electronic health records with machine learning. PloS one, 16(8), e0256428.

[3] Abdar, M., Yen, N. Y., & Hung, J. C. S. (2018). Improving the diagnosis of liver disease using multilayer perceptron neural network and boosted decision trees. Journal of Medical and Biological Engineering, 38(6), 953-965.

[4] Lanjewar, M. G., Parab, J. S., Shaikh, A. Y., & Sequeira, M. (2023). CNN with machine learning approaches using ExtraTreesClassifier and MRMR feature selection techniques to detect liver diseases on cloud. Cluster Computing, 26(6), 3657-3672.

[5] Zhang, B., Dong, X., Hu, Y., Jiang, X., & Li, G. (2023). Classification and prediction of spinal disease based on the SMOTE-RFE-XGBoost model. PeerJ Computer Science, 9, e1280.

[6] fedesoriano. (August 2021). Cirrhosis Prediction Dataset. Retrieved [March 2025] from https://www.kaggle.com/fedesoriano/cirrhosis-prediction-dataset.

[7] Shi, Z., Huang, L., & Wang, H. (2024, November). Predicting Complications of Cirrhosis using Synthetic Data Generation Enhanced Dynamic Classifier Selection. In 2024 IEEE International Conference on Medical Artificial Intelligence (MedAI) (pp. 260-267). IEEE.

[8] Xu, L., Skoularidou, M., Cuesta-Infante, A., & Veeramachaneni, K. (2019). Modeling tabular data using conditional.GAN Advances in Neural Information Processing Systems, 32.

[9] Alqulaity, M., & Yang, P. (2024). Enhanced Conditional GAN for High-Quality Synthetic Tabular Data Generation in Mobile-Based Cardiovascular Healthcare. Sensors, 24(23), 7673.

[10] Majeed, A., & Hwang, S. O. (2023). CTGAN-MOS: Conditional generative adversarial network based minority-class-augmented oversampling scheme for imbalanced problems. IEEE Access, 11, 85878-85899.

[11] Hupont, I., & Fernández, C. (2019, May). Demogpairs: Quantifying the impact of demographic imbalance in deep face recognition. In 2019 14th IEEE International Conference on Automatic Face & Gesture Recognition (FG 2019) (pp. 1-7). IEEE.

[12] Straw, I., & Wu, H. (2022). Investigating for bias in healthcare algorithms: a sex-stratified analysis of supervised machine learning models in liver disease prediction. BMJ Health & Care Informatics, 29(1), e100457.

[13] Alvi, R. H., Rahman, M. H., Khan, A. A. S., & Rahman, R. M. (2021). Deep learning approach on tabular data to predict early-onset neonatal sepsis. *Journal of Information and Telecommunication*, *5*(2), 226-246.

[14] Kumar, M., Shelke, N. A., Singh, J., Sharma, K. N., Sharma, R., & Kumar, R. (2024, November). A Multimodal Deep Learning Approach for Advancing Liver Disease Diagnosis and Prognosis Prediction. In *2024 International Conference on IoT, Communication and Automation Technology (ICICAT)* (pp. 838-843). IEEE.

[15] Agyemang, E. F., Mensah, J. A., Nyarko, E., Arku, D., Mbeah-Baiden, B., Opoku, E., & Noye Nortey, E. N. (2025). Addressing Class Imbalance Problem in Health Data Classification: Practical Application from an Oversampling Viewpoint. *Applied Computational Intelligence and Soft Computing*, *2025*(1), 1013769.

[16] Ahmed, H. A., Nepomuceno, J. A., Vega-Márquez, B., & Nepomuceno-Chamorro, I. A. (2025). Synthetic Data Generation for Healthcare: Exploring Generative Adversarial Networks Variants for Medical Tabular Data. *International Journal of Data Science and Analytics*, 1-16.

[17] Musa, A., Prasad, R., & Hernandez, M. (2025). Addressing cross-population domain shift in chest X-ray classification through supervised adversarial domain adaptation. *Scientific Reports*, *15*(1), 11383.

[18] Zhou, B., Zhou, Q., & Li, Z. (2024). Addressing data imbalance in crash data: evaluating Generative Adversarial Network's efficacy against conventional methods. *IEEE Access*.