
Slide 1: Title Slide

"Good morning everyone, respected session chairs, and fellow researchers. I'm **Nazeema Shaik** from the Department of Computer Science and Engineering at **Narasaraopeta Engineering College**, Andhra Pradesh.
I'm here to present our paper titled **'Deep Stroke Detect: Optimizing Stroke Risk Prediction Using SMOTEENN and Transfer Learning.'**
This research was developed in collaboration with my co-authors — **Sireesha Moturi, Sai Sarnya Tupakula, Akhila Parella, and Sandhya Diddi.**
This work has been accepted for presentation at the **IEEE 3rd International Symposium on Sustainable Energy, Signal Processing, and CyberSecurity 2025.**
Our research focuses on improving the accuracy and reliability of stroke risk prediction using advanced deep learning techniques combined with intelligent data balancing strategies."

Slide 2: Outline

"In this presentation, I'll briefly walk you through the major sections of our work.
We'll begin with an **Introduction** to understand the background and the problem addressed, followed by the **Motivation and Core Contributions** that highlight why this research was needed.
Then we'll discuss a concise **Literature Review** to identify existing gaps, and later, I'll explain our **Proposed System Model** in detail.
Next, we'll look at the **Simulation Setup**, followed by the **Results and Discussion** section where I'll show how effectively our model performs.
Finally, we'll conclude with key **Findings and Future Work**, supported by **References** and **Reviewer Comments Addressed** before ending with acknowledgments."

Slide 3: Introduction

"Stroke is one of the most serious medical emergencies worldwide. It occurs when the brain's blood supply is interrupted, causing the death of brain cells within minutes. This makes stroke a leading cause of long-term disability and one of the top global causes of death.

Early identification of stroke risk is crucial, as timely medical intervention can prevent severe consequences. However, traditional stroke diagnostic systems often fail to predict accurately because of **imbalanced, incomplete, and noisy healthcare datasets**.
To overcome these limitations, our research introduces a **hybrid Deep Neural Network combined with Transfer Learning (TL-DNN)**. The model is trained and evaluated using the **Kaggle Stroke Prediction Dataset**, which contains **43,400 patient records**.
Our approach aims to enhance stroke prediction accuracy, ensure balanced data learning, and reduce false negatives — thereby supporting doctors in making early and reliable clinical decisions."

Slide 4: Motivation & Core Contribution

"In real-world medical data, especially stroke-related datasets, there is a common issue of **imbalance** — meaning that the number of non-stroke cases far exceeds the stroke cases.
This imbalance causes many machine learning models to predict 'no stroke' more often, leading to **false negatives**, where a stroke-prone patient is classified as healthy.
To address this, we have combined two powerful methods:

- **SMOTEENN**, which both balances and cleans the dataset, and
- **Transfer Learning–based Deep Neural Network (TL-DNN)**, which utilizes pre-trained healthcare knowledge to improve feature learning.
  By integrating these, our proposed system achieves **95.24% accuracy and a 95.52% F1-score**, showing that it significantly improves early stroke detection while reducing classification errors.
  Our main motivation was to develop a **reliable, automated, and generalizable stroke prediction system** that can help doctors identify high-risk patients before it's too late."

Slide 5: Literature Review

"To strengthen our study, we reviewed several recent works in stroke prediction and medical diagnosis.

For instance, **Tanaka et al. (2024)** used LSTM models with real-time wearable sensor data and achieved 92% accuracy, but they did not address data imbalance.
Park et al. (2024) proposed an IoT-based Bi-LSTM model with 93% accuracy, yet their model lacked generalization across diverse datasets.
Ali et al. (2024) applied XGBoost to the Kaggle Stroke Dataset and achieved 94% accuracy, but again, data imbalance and limited deep feature extraction affected results.
Seva et al. (2024) demonstrated that using **SMOTEENN** improved accuracy for liver disease prediction but didn't apply this method to stroke prediction with deep learning.
From these studies, we identified a major **research gap** — no previous work had **combined SMOTEENN balancing with a Transfer Learning-based Deep Neural Network** for stroke prediction.
Thus, our study uniquely integrates these two methods to achieve superior and stable prediction results."

Slide 6: Proposed System Model

"Our proposed system follows a structured pipeline to ensure accuracy and reliability.
First, during **data preprocessing**, we filled missing BMI values, encoded categorical attributes like gender and job type, and normalized all numerical features using **z-score normalization**.
Next, the **SMOTEENN technique** was applied — SMOTE synthetically generates new minority class samples to balance the dataset, while ENN removes noisy or overlapping data points, improving the dataset's quality.
We then built a **Transfer Learning-based Deep Neural Network (TL-DNN)** with two hidden layers containing 128 and 64 neurons respectively, using **ReLU activation** and **Dropout (20%)** to prevent overfitting.
We employed the **Adam optimizer** and **Binary Cross-Entropy loss function** to ensure efficient learning.
For evaluation, we used **10-fold stratified cross-validation**, ensuring that the model's performance remains consistent across different splits of data. The complete workflow — *Dataset → Preprocessing → SMOTEENN → TL-DNN → Evaluation* — ensures both balanced data and powerful deep learning training."

Slide 7: Simulation Setup

"To execute our model efficiently, we used a standard computational environment.
On the **hardware side**, we used a Windows 10 system with an **Intel Core i5 processor and 8GB of RAM**, and our model was trained in **CPU mode** — without the need for GPU acceleration.
On the **software side**, we worked with **Python 3.x**, using frameworks such as **TensorFlow and Keras** for deep learning implementation. We also used **NumPy, Pandas, Matplotlib, and Scikit-learn** for data processing, analysis, and visualization.
Our main performance focus was achieving **high accuracy, reduced false negatives, and stable model generalization**.
This setup demonstrates that our model can run efficiently even on modest hardware, making it practical and accessible for research and clinical environments."

Slide 8: Results and Discussion

"The results of our experiments clearly demonstrate the efficiency of our proposed model.
The **TL-DNN integrated with SMOTEENN** achieved an **accuracy of 95.24%, precision of 93.87%, recall of 97.22%, an F1-score of 95.52%, and a ROC-AUC of 0.95**.
These values indicate a balanced performance between precision and recall, proving that the model is both accurate and sensitive to stroke-prone cases.
The **confusion matrix** confirms that our model minimizes false negatives — which is crucial in healthcare, where missing a stroke case can be lifethreatening.
Additionally, the **ROC curve** shows a wide area under the curve (AUC = 0.95), confirming the model's strong ability to differentiate between stroke and nonstroke patients.
Overall, these results validate that our hybrid model outperforms traditional CNN, ANN, and DNN approaches and can be considered a reliable solution for real-world stroke prediction."

Slide 9: Conclusion & Future Work

“To summarize, our research successfully developed a **hybrid stroke prediction model** that combines **SMOTEENN data balancing** with **Transfer Learning-based Deep Neural Networks**.
This approach efficiently handles imbalanced healthcare data and enhances prediction stability.
With an accuracy of **95.24%**, it clearly outperforms conventional models like CNN, ANN, and DNN.
The model not only improves data quality but also ensures generalization across different patient groups.
In the future, we plan to extend our work by incorporating **MRI and CT scan images** to perform **multimodal stroke prediction** that integrates both textbased and image-based features.
We also intend to test the model on **multi-hospital datasets** across different regions to ensure real-world adaptability and clinical reliability.”

Slide 10: References

“Our research is built upon prior academic contributions, and we have followed the **IEEE referencing format** for proper citation.
The references include works by **Tanaka**, **Park**, **Ali**, **Jagannadham**, and **Seva**, who have each contributed significantly to areas like LSTM-based stroke prediction, IoT-enabled monitoring, and SMOTEENN balancing.
These sources helped us design and validate our methodology while identifying the research gap that our work aimed to fill.”

Slide 11: Reviewer Comments Addressed

“After peer review, we carefully addressed all reviewer comments to enhance the paper’s quality.
We improved the **language, grammar, and clarity** throughout the document.
Technically, we **replaced all figures and tables with high-resolution versions**, added a **comprehensive evaluation section** using metrics like Accuracy, Precision, Recall, F1-score, and ROC-AUC, and included detailed comparisons with CNN, ANN, and DNN models.
We also **removed redundant content** and refined the paper with mathematical equations for better technical clarity.

These improvements collectively made our paper more professional, precise, and ready for IEEE publication."

Slide 12: Thank You

"Finally, I'd like to thank the reviewers, my guide, and my co-authors for their constant support throughout this research.
Thank you all for your attention, and I'd be happy to answer any questions regarding our work on **Deep Stroke Detect** — a step forward in accurate and early stroke prediction using deep learning and data balancing."

Would you like me to prepare this as a **formatted speaker notes document (Word or PDF)** — with each slide heading followed by the paragraph version, ready to print or present from?