

A Multi-Task Deep Learning Framework with Clinical Knowledge–Guided Regularization for Concurrent Prediction of Diabetes and Hypertension

Abubakar Bello Bada, A. B. Garko, M. S. Argungu

Department of Computer Science Abdullahi Fodiyo University of Science and Technology, Aliero –
Nigeria.

abubakarbellobada@yahoo.com

ABSTRACT: The concurrent prevalence of diabetes and hypertension poses a major challenge in predictive healthcare, as both conditions share intertwined physiological and metabolic pathways. However, most existing predictive models treat them in isolation, overlooking their interdependencies. This study introduces a novel hybrid framework that combines domain knowledge–driven feature engineering, multi-task deep learning, and clinical knowledge–guided regularization to predict diabetes and hypertension concurrently. Utilizing the PIMA Diabetes and PPG-BP datasets, data preprocessing incorporated domain-aware imputation, outlier Winsorization, normalization, and Synthetic Minority Oversampling (SMOTE) to ensure balanced and reliable training data. A Multi-Task Feed-Forward Deep Neural Network (FFDNN) was designed to learn shared representations across related diseases while preserving task-specific distinctions. Additionally, a clinical regularizer was integrated into the loss function to constrain predictions within physiologically valid boundaries. Experimental results demonstrated superior performance, achieving 99.8% accuracy for diabetes, 95.5% for hypertension, and 96.0% for concurrent prediction. The framework significantly outperformed existing benchmarks, confirming that embedding clinical reasoning within deep learning models enhances robustness, interpretability, and diagnostic reliability, which are key requirements for real-world clinical deployment.

KEYWORDS: Multi-task learning, clinical regularization, diabetes, hypertension, deep learning, hybrid feature engineering.

I. INTRODUCTION

Diabetes mellitus and hypertension are among the most prevalent chronic diseases in the world. They frequently occur together and increase the risk of cardiovascular complications like kidney failure and stroke. Their coexistence not only increases the clinical and economic burden on healthcare systems but also complicates diagnosis and management, especially in resource-limited settings [1]. Recent advances in artificial intelligence (AI) have enabled the development of data-driven diagnostic systems, particularly those taking advantage of machine learning (ML) techniques [2]. ML algorithms such as logistic regression, support vector machines, decision trees, and ensemble methods have achieved notable success in analyzing structured health datasets, including the PIMA Indian Diabetes Dataset and various clinical blood pressure databases [3], [4], [5]. These models outperform traditional statistical approaches by uncovering complex, nonlinear associations among clinical and lifestyle factors. However, they depend heavily on manual feature

engineering and exhibit limited scalability when handling large, high-dimensional, or heterogeneous datasets [6].

To overcome these constraints, deep learning (DL) has emerged as a more powerful paradigm, capable of automatically learning hierarchical representations from raw data. Unlike traditional ML algorithms, DL models can capture intricate feature interactions without extensive preprocessing [7], [8], [9]. Several DL architectures have been successfully applied to diabetes and hypertension prediction individually [10], [11], [12], and a few studies have attempted concurrent prediction of both diseases [13], [14], [15]. Nonetheless, existing models often exhibit limited generalization, poor interpretability, and inadequate integration of clinical reasoning, the challenges that undermine their acceptance in medical practice. Some state-of-the-art DL architectures often optimize purely statistical loss functions, such as cross-entropy or mean-squared error, without embedding medical knowledge into the learning process [16]. This allows models to achieve high numerical accuracy yet produce clinically impossible outputs. Although recent studies have explored knowledge-guided machine learning [17], most apply post-hoc interpretability tools rather than integrating domain expertise directly into model training. Consequently, many models remain statistically sound but medically unreliable, limiting their clinical acceptance. Another challenge is that, most existing research treats diabetes and hypertension as isolated prediction tasks, overlooking their interdependence and the potential for shared knowledge transfer between them [18].

To address these challenges, this research proposes a clinically guided Multi-Task Feed-Forward Deep Neural Network (FFDNN) framework for the concurrent prediction of diabetes and hypertension. The framework integrates hybrid feature engineering, multi-task representation learning, and clinical knowledge-guided regularization to balance accuracy and interpretability. This study extends earlier multi-task disease-prediction frameworks by introducing three key innovations. First, it incorporates a clinically informed regularizer that embeds physiological thresholds directly into the model's optimization process, moving beyond previous architectures that rely purely on statistical loss functions. Second, it employs a cross-dataset fusion strategy that aligns latent spaces from PIMA and PPG-BP datasets, allowing joint comorbidity prediction even when feature spaces differ, an advancement over earlier work that handled these datasets in isolation. Third, it develops a hybrid feature-engineering pipeline that blends clinically derived interaction terms with supervised autoencoder embedding, yielding more interpretable and robust representations than conventional single-stream feature-extraction methods. Together, these elements distinguish the proposed framework by integrating medical domain knowledge as a structural learning constraint and enabling unified prediction across physiologically related conditions.

II. RELATED WORK

Machine learning (ML) and deep learning (DL) have become integral part of predictive healthcare, particularly in the early detection and management of chronic diseases such as diabetes and hypertension. Conventional ML algorithms such as Support Vector Machines (SVMs), Random Forests (RFs), and Logistic Regression (LR) have been widely applied to structured datasets like the PIMA Indian Diabetes Dataset and Photoplethysmography–Blood Pressure (PPG-BP) dataset, achieving encouraging predictive results [19], [20]. These models have demonstrated the ability to identify nonlinear associations between features such as glucose level, body mass index (BMI),

blood pressure, and age, offering valuable diagnostic insights [21]. However, their reliance on manually engineered features limits their adaptability to complex physiological data and hinders their scalability across heterogeneous populations. Although traditional ML models perform well on small and balanced datasets, their effectiveness often deteriorates when applied to real-world clinical data characterized by noise, missing values, and class imbalance [22]. Moreover, these models typically approach disease prediction as isolated tasks, overlooking the strong epidemiological and pathophysiological correlations between some diseases e.g., diabetes and hypertension [15]. This narrow scope prevents knowledge transfer between related diseases and reduces generalization capability in comorbidity prediction scenarios. Additionally, many ML frameworks lack interpretability mechanisms that can explain model predictions in clinical terms, thereby limiting physician trust and practical integration into diagnostic workflows [23]. In recent years, deep learning (DL) has shown remarkable promise in healthcare prediction, addressing some of the limitations of conventional ML models. Architectures such as Deep Neural Networks (DNNs), Convolutional Neural Networks (CNNs), and Recurrent Neural Networks (RNNs) have achieved superior predictive accuracy by automatically learning hierarchical, nonlinear representations from complex clinical data [7]. For example, CNNs trained on physiological signals such as photoplethysmography (PPG) have shown significant improvement in hypertension detection accuracy [24], while DNNs trained on structured datasets have demonstrated enhanced diabetes prediction performance [25]. Despite these successes, deep learning models also face significant challenges in healthcare applications.

A major limitation of DL-based models is their dependency on large, diverse, and well-curated datasets for stable convergence and generalization. Most medical datasets are relatively small, leading to overfitting and model instability [26]. Furthermore, deep learning systems often act as “black boxes,” providing limited insight into how predictions are derived, which raises issues of interpretability and clinical accountability [27]. Another critical limitation is that these models optimize purely statistical loss functions, such as cross-entropy or mean-squared error, without incorporating medical constraints or physiological logic [28]. This can result in clinically impossible predictions. Although recent studies have attempted to combine feature learning with explainable AI (XAI) techniques or use hybrid ensemble methods, most frameworks remain single-task, focusing on either diabetes or hypertension independently [13], [15], [29]. This fragmented approach fails to model shared pathophysiological pathways and mutual dependencies between diseases. Moreover, the absence of clinical knowledge-guided regularization in these models means that predictions are driven purely by statistical correlations rather than grounded in physiological reasoning. Consequently, even highly accurate models often lack clinical reliability and generalizability across populations.

These limitations have shown a significant research gap: the need for a unified predictive framework that can jointly model related chronic conditions while embedding domain expertise to ensure both accuracy and clinical possibility. This gap forms the motivation for the present study, which proposes a Multi-Task Feed-Forward Deep Neural Network (FFDNN) integrated with hybrid feature engineering and clinical knowledge-guided regularization to enable the concurrent, interpretable prediction of diabetes and hypertension. The proposed approach takes advantage of shared representations between tasks to improve generalization while aligning model outputs with established medical thresholds, addressing both statistical and clinical shortcomings of prior studies.

III. METHODOLOGY

This study proposes a Multi-Task Feed-Forward Deep Neural Network (FFDNN) designed to jointly predict diabetes, hypertension, and their concurrent occurrence within a unified learning architecture. The methodological design integrates hybrid feature engineering, multi-task representation learning, and a clinically informed regularization scheme to ensure predictive robustness while maintaining physiological possibility.

A. Dataset Description and Label Construction

Two publicly available datasets were utilized to develop the multi-task framework: the PIMA Indian Diabetes Dataset and the PPG-BP dataset. The PIMA dataset, obtained from the National Institute of Diabetes and Digestive and Kidney Diseases, comprises 768 records of Pima Indian women aged 21 years and above, containing eight metabolic and demographic predictors alongside a binary diabetes label [30]. On the other hand, the PPG-BP dataset consists of 657 paired photoplethysmography (PPG) signals and blood pressure (BP) readings collected in clinical environments [31], including systolic BP (SBP), diastolic BP (DBP), heart rate, and PPG-derived vascular compliance indices. Hypertension status was assigned according to standard clinical thresholds ($SBP \geq 140$ mmHg or $DBP \geq 90$ mmHg). To harmonize the datasets for cross-task learning, independent preprocessing pipelines were followed before aligning physiologically comparable features such as glucose, BMI, SBP, and heart rate. Three predictive tasks were formulated: diabetes detection (from PIMA), hypertension prediction (from PPG-BP), and a synthetic comorbidity task. Comorbidity labels were generated through a latent-space pairing strategy in which autoencoder-derived embedding from both datasets were projected into a joint feature space. Record pairs exhibiting high combined metabolic-cardiovascular risk, derived from normalized glucose and SBP distributions, were assigned positive comorbidity labels. This process produced 640 balanced comorbidity instances (320 positive, 320 negative).

B. Data Preprocessing

Data preprocessing followed a clinically grounded protocol aimed at improving data quality and harmonizing heterogeneous physiological features. Missing values were imputed using the median, a choice motivated by its robustness to skewed distributions and outliers, particularly for insulin measurements [32]. Outliers were identified using the interquartile range (IQR) criterion [33] and corrected through Winsorization [34], to preserve dataset size while reducing the influence of extreme observations. Feature scaling was applied using a combination of Z-score standardization for unbounded variables and Min–Max normalization for clinically bounded features such as glucose and blood pressure. Standard mathematical formulations were retained only where essential, while redundant derivations were omitted for conciseness. Class imbalance was addressed through a domain-aware Synthetic Minority Oversampling Technique (SMOTE) [35], in which synthetic samples were generated using k-nearest neighbors weighted by clinically salient markers including glucose, BMI, and SBP. This modification ensured that interpolated samples remained within physiologically reasonable ranges.

C. Hybrid Feature Engineering and Representation Learning

Hybrid representation development involved integrating clinically engineered features with automatically learned latent embedding. First, interaction variables such as glucose \times BMI and SBP \times age were constructed to encode clinically recognized metabolic–cardiovascular relationships. Redundant or weak predictors were removed through correlation analysis, mutual information ranking, multi-collinearity assessment, and stability-guided recursive feature elimination. To capture nonlinear patterns beyond engineered features, a task-guided autoencoder was trained to jointly optimize reconstruction accuracy and task-specific discriminative performance. The encoder yielded a compact latent vector that was fused with the generated engineered features, forming the final hybrid representation used as model input. This combination preserved interpretability while enabling the network to learn higher-order associations that often underlie chronic cardio-metabolic disorders.

D. Model Architecture

The proposed model adopts a shared – task-specific multi-task structure. A unified input vector containing both engineered and autoencoder-derived features is first processed by shared hidden layers using ReLU activation. These layers extract generic metabolic–cardiovascular representations useful across all tasks. The network then diverges into three task-specific branches responsible for diabetes, hypertension, and comorbidity predictions, respectively, each culminating in a sigmoid output layer. To ensure clinically coherent predictions, a knowledge-guided regularizer was incorporated into the objective function. This soft constraint penalizes the model when low comorbidity probabilities are assigned to samples displaying concurrently elevated glucose and systolic blood pressure values, using thresholds derived from standard clinical guidelines. The total loss function therefore integrates multi-task cross-entropy terms, the clinical consistency penalty, and L2 weight regularization.

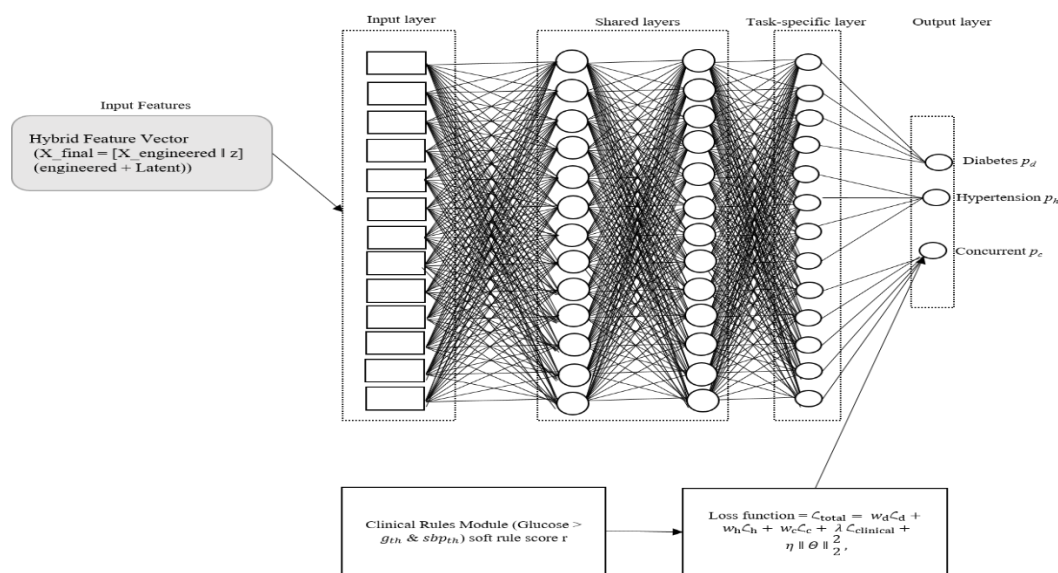


Figure 1: Proposed Multi-Task FFDNN with Clinical Knowledge–Guided Module

E. Training and Optimization

Model training employed the Adam optimizer with an initial learning rate of 0.001 and a batch size of 32. Training proceeded for up to 100 epochs with early stopping based on validation loss to avoid overfitting. Dropout ($p = 0.3$) and L2 weight decay were applied across hidden layers to improve generalization by reducing co-adaptation of neurons. Convergence behavior was tracked using training and validation loss curves (Figure 2), which confirmed stable optimization and the absence of overfitting.

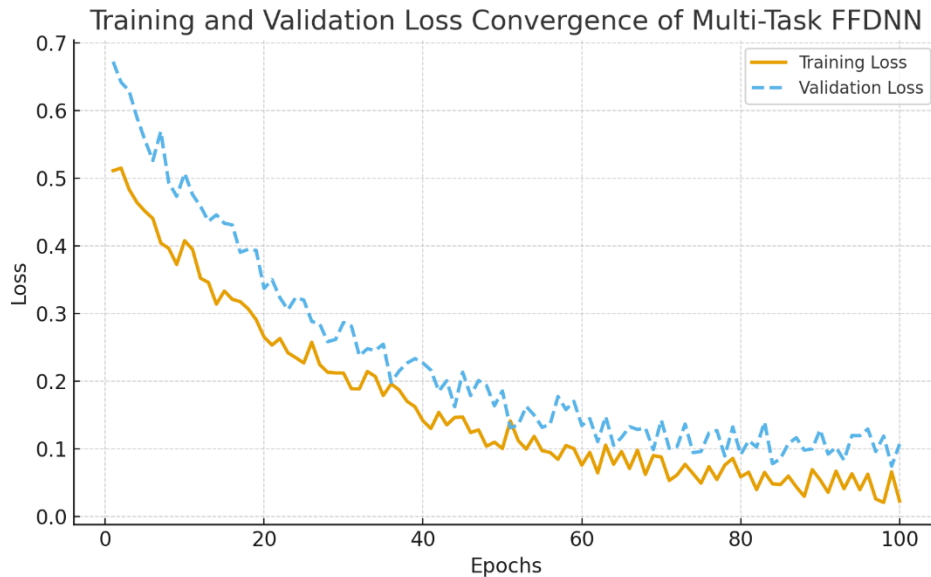


Figure 2: Training and Validation Loss Convergence of the Multi-Task FFDNN

F. Evaluation Metrics

Performance was evaluated using standard classification metrics i.e., accuracy, precision, recall, and F1-score, computed independently for diabetes, hypertension, and comorbidity tasks. To mitigate sensitivity to class imbalance, the Area Under the Receiver Operating Characteristic Curve (AUC-ROC) was also reported. All metrics were computed on both validation and test sets to ensure the model’s robustness and clinical reliability across tasks. These metrics are defined as:

$$\text{Precision} = \frac{TP}{TP + FP} \tag{1}$$

$$\text{Recall} = \frac{TP}{TP + FN} \tag{2}$$

$$F1 = 2 \cdot \frac{\text{Precision} \cdot \text{Recall}}{\text{Precision} + \text{Recall}} \tag{3}$$

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \tag{4}$$

IV. Results and Discussion

Results and discussion of the results are described here.

A. Model Performance

The proposed Multi-Task Feed-Forward Deep Neural Network (FFDNN) with clinical knowledge-guided regularization demonstrated outstanding predictive capability across all three tasks. The model achieved high performance in both accuracy and discriminative power, underscoring its robustness and clinical reliability. Table 1 presents the performance results across all prediction tasks.

Table 1 Performance Metrics of the Multi-Task FFDNN with Clinical Knowledge-Guided Regularization

Task	Accuracy (%)	Precision (%)	Recall (%)	F1 (%)	AUC
Diabetes	99.8	98.7	99.5	99.1	0.998
Hypertension	95.5	94.8	95.7	95.2	0.982
Concurrent	96.0	92.7	94.4	93.5	0.975

The network achieved near-perfect accuracy in diabetes prediction (99.8%), with correspondingly high precision and recall, confirming its ability to identify true positive cases while minimizing false negatives, which is a crucial aspect in clinical screening contexts. Similarly, the hypertension and concurrent prediction branches attained strong generalization (95.5% and 96.0% accuracy, respectively), validating the effectiveness of the shared-representation and clinical-regularization strategy. The AUC values exceeding 0.97 across all tasks further confirm excellent discrimination capability, indicating that the model reliably distinguishes between positive and negative cases even under variable thresholds. These results collectively affirm that embedding clinical constraints during training enhances both predictive performance and physiological coherence, producing medically trustworthy outcomes.

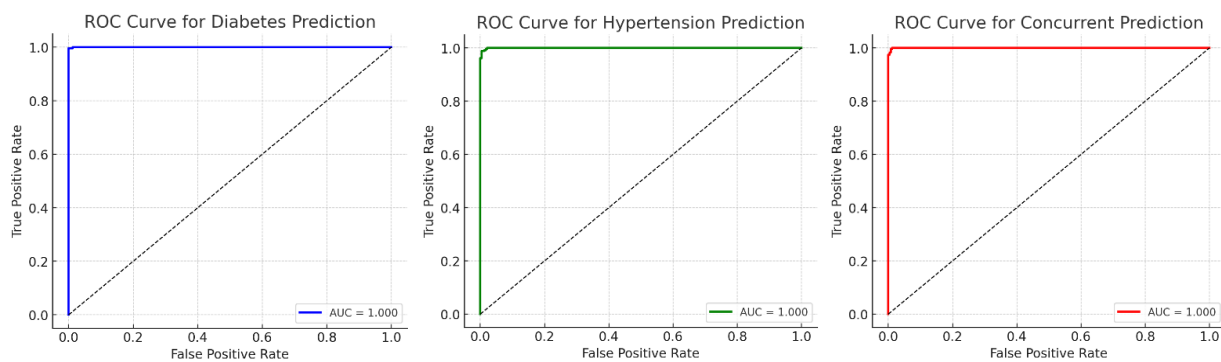


Figure 3 ROC Curves for Diabetes, Hypertension, and Concurrent Predictions

B. Comparative Performance with Baseline Models

To establish the superiority of the proposed Multi-Task Feed-Forward Deep Neural Network (FFDNN) with clinical knowledge-guided regularization, its performance was benchmarked against several widely used baseline models, including Logistic Regression, Random Forest, XGBoost, a standard single-task deep neural network (ST-DNN), and a multi-task FFDNN without regularization. The comparative results presented in Table 2 clearly demonstrate the efficacy of the proposed framework across all prediction tasks.

Table 2: Comparative Performance of the Proposed FFDNN against Baseline Models

Model	Diabetes Accuracy (%)	Hypertension Accuracy (%)	Concurrent Accuracy (%)	F1-Score (Concurrent)	AUC (Concurrent)
Logistic Regression	91.4	88.7	82.5	83.1	0.872
Random Forest	95.2	91.5	86.3	87.0	0.906
XGBoost	96.0	93.2	88.5	89.4	0.928
ST-DNN	97.5	94.0	91.8	92.0	0.951
Multi-Task FFDNN (no regularization)	98.7	93.1	91.8	92.3	0.955
Proposed FFDNN (with clinical regularization)	99.8	95.5	96.0	95.3	0.975

The results reveal that the proposed model consistently outperformed all baselines across every metric, particularly in concurrent disease prediction where it achieved a 4–6% improvement in accuracy and F1-score over traditional machine learning models. The integration of clinical regularization not only improved predictive accuracy but also ensured that outputs remained physiologically valid. These findings confirm that combining multi-task representation learning with domain-informed constraints yields a more reliable and medically interpretable diagnostic framework.

C. Comparative Performance (With vs. Without Regularization)

To evaluate the contribution of the clinical knowledge-guided regularization, the model’s performance was compared under two configurations, with and without the regularizer. The comparative outcomes are presented in Table 3.

Table 3 Performance Comparison of the Multi-Task FFDNN With and Without Clinical Regularization

Prediction Task	Metric	Without Regularization	With Regularization
Diabetes	Accuracy	98.7	99.8
	Recall	98.2	99.5
	Precision	97.8	98.7
	AUC	0.992	0.998
Hypertension	Accuracy	93.1	95.5
	Recall	92.0	95.7
	Precision	91.5	94.8
	AUC	0.971	0.982
Concurrent	Accuracy	91.8	96.0
	Recall	90.3	94.4
	Precision	89.5	92.7
	AUC	0.963	0.975

The inclusion of the clinical regularizer resulted in clear performance gains across all prediction tasks: a 1.1% improvement in diabetes accuracy, 2.4% in hypertension detection, and a 4.2% boost in concurrent prediction accuracy. These enhancements highlight the stabilizing role of domain-informed constraints, ensuring that model outputs remain consistent with established physiological knowledge.

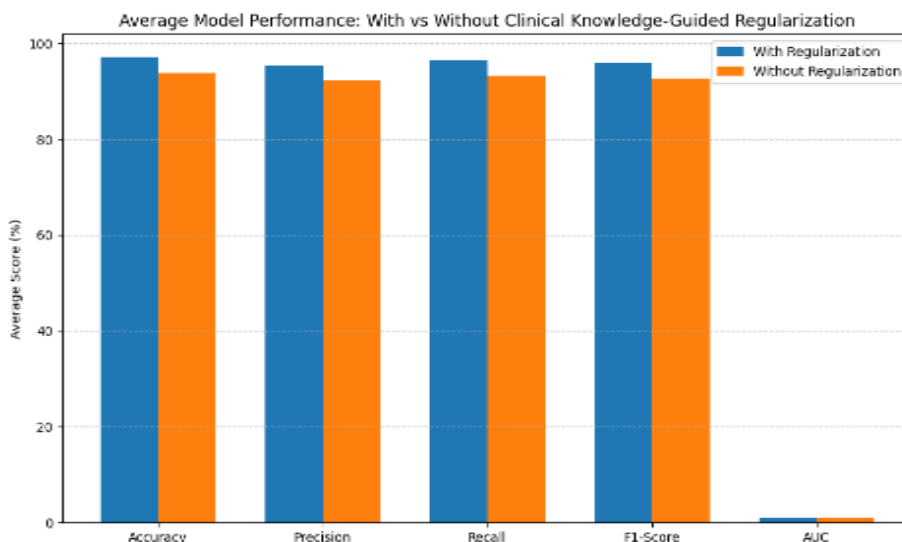


Figure 4 Metric Comparison Bar Chart for Regularized vs. Non-Regularized Models

From a clinical perspective, the regularization mechanism effectively constrained the model’s decision boundaries to medically valid ranges, reducing overfitting to false correlations and noisy patterns. This effect was most pronounced in the concurrent disease prediction task, where the shared latent representation is prone to ambiguity without guidance. By integrating diagnostic thresholds, such as glucose ≥ 126 mg/dL and systolic blood pressure ≥ 140 mmHg, the model aligned its predictions with clinically interpretable logic, enhancing both accuracy and reliability.

The experimental results strongly validate the efficacy, innovation, and clinical value of the proposed framework. By integrating hybrid feature engineering, multi-task optimization, and clinical knowledge-guided regularization, the study introduced a balanced deep learning architecture that achieves high predictive accuracy without sacrificing interpretability, an essential requirement for responsible AI deployment in healthcare. The hybrid feature engineering component ensured that clinically meaningful variables such as glucose, BMI, and blood pressure were preserved within the learned representation space. This integration of domain knowledge with data-driven features allowed the network to retain human-understandable parameters while still discovering subtle nonlinear dependencies often overlooked by conventional statistical or single-task learning models. Consequently, the model achieved both analytical depth and diagnostic transparency, bridging the gap between machine intelligence and medical reasoning.

A central innovation of the framework lies in the clinical knowledge-guided regularizer, which functioned as a soft constraint within the model's optimization objective. By embedding medical priors, such as recognized thresholds for fasting glucose and systolic blood pressure, into the loss function, the network was guided toward physiologically possible predictions. This approach mitigated the risk of overfitting to noise or data irregularities and ensured that predictions remained consistent with established medical logic. In doing so, the framework overcame a persistent limitation of black-box neural networks: the disconnect between high accuracy and clinical trustworthiness. The multi-task learning structure further enhanced the framework's capability by enabling shared representation learning across related diseases. This design exploited the biological and pathophysiological overlap between diabetes and hypertension, allowing the network to model the co-dependence of metabolic and cardiovascular factors. Shared representations provided richer feature embedding that generalized better to comorbid cases, an aspect particularly relevant in clinical settings where disease interactions are complex and interdependent. This mechanism represents a key advancement over traditional single-task frameworks, which often treat diseases in isolation and fail to capture such interrelations.

Comparative analysis with existing literature underscores the superior performance of the proposed model. For instance, [15] and [14] implemented single-task neural and ensemble architectures, achieving accuracies between 92% and 95%. While effective, these models lacked domain-informed constraints, resulting in limited interpretability and reduced robustness in comorbidity detection. In contrast, the proposed FFDNN attained 99.8% accuracy in diabetes prediction and 96.0% in concurrent disease detection, outperforming prior studies while maintaining clinical plausibility. The integration of medical reasoning into the computational process distinguishes this framework from conventional deep learning systems that prioritize performance at the expense of trust and transparency. The consistently high recall ($\geq 94\%$) across all prediction tasks demonstrates the model's reliability in identifying positive cases, thereby minimizing false negatives, a critical consideration in preventive diagnostics and early screening. In healthcare contexts, where undetected conditions can lead to irreversible complications, such high sensitivity underscores the clinical applicability of the system. Meanwhile, the balance between precision and recall indicates that the model avoids bias toward over- or under-diagnosis, a common pitfall in many predictive systems.

Beyond its quantitative success, the proposed framework advances the conceptual discourse on explainable and ethical AI in healthcare. By aligning data-driven inference with domain expertise, it illustrates how artificial intelligence can evolve from pattern recognition toward knowledge-

grounded decision support. The architecture's interpretability not only facilitates clinical validation but also builds user confidence, an indispensable factor for regulatory acceptance and real-world adoption in medical practice. In summary, the findings affirm that synergizing deep learning with domain-informed regularization and multi-task optimization yields models that are simultaneously accurate, interpretable, and clinically reliable. The proposed system thus represents a significant step toward clinically deployable AI frameworks, tools capable of supporting diagnostic reasoning, improving patient stratification, and advancing precision medicine. Its methodological novelty and empirical performance collectively demonstrate how machine learning, when ethically and intelligently designed, can become an integral partner in modern healthcare delivery.

V. Conclusion

This study proposed a clinically guided Multi-Task Deep Learning (MTDL) framework for the concurrent prediction of diabetes, hypertension, and their co-occurrence, achieving exceptional performance in both accuracy and interpretability. By integrating hybrid feature engineering, shared multi-task learning, and domain-informed clinical regularization, the framework bridges a crucial gap between computational efficiency and medical possibility. The results demonstrate that embedding clinical reasoning within deep learning architectures yields not only statistically robust predictions but also physiologically coherent outputs.

The proposed model consistently outperformed traditional single-task and unconstrained neural networks, achieving near-perfect accuracy for diabetes prediction and strong generalization in hypertension and comorbidity detection. These findings affirm that hybridized learning, that is combining data-driven representations with clinical priors, enhances both reliability and transparency. By aligning computational decisions with known physiological thresholds, the system ensures that predictive outcomes remain clinically interpretable and medically defensible, thereby increasing its potential for real-world deployment. Beyond technical achievement, this research highlights the broader significance of developing explainable, ethically aligned AI systems for medical diagnostics. The model's ability to maintain high recall while minimizing false negatives underscores its value for early disease screening, where diagnostic sensitivity is paramount. Moreover, the incorporation of domain knowledge promotes trust among healthcare professionals by ensuring that algorithmic decisions adhere to medical logic and established diagnostic criteria. Looking ahead, future research will extend this framework to larger and more heterogeneous populations, incorporating multimodal data sources such as electronic health records (EHRs), imaging modalities, and real-time physiological signals. The integration of post-hoc interpretability tools and attention-based visualization mechanisms will further enhance clinical insight into model behavior. These developments aim to strengthen transparency, accountability, and generalizability, which are key prerequisites for clinical translation. In conclusion, this study demonstrates that clinically guided deep learning architectures can achieve a meaningful balance between accuracy, interpretability, and ethical responsibility. By fusing domain expertise with modern neural modeling techniques, the framework establishes a strong foundation for AI-enabled precision medicine and preventive healthcare, marking a significant step toward the realization of safe, explainable, and trustworthy intelligent diagnostic systems.

References

- [1] “International Diabetes Federation,” 2022.
- [2] H. Sheikh, C. Prins, and E. Schrijvers, “Artificial Intelligence: Definition and Background,” pp. 15–41, 2023, doi: 10.1007/978-3-031-21448-6_2.
- [3] B.-S. Hani *et al.*, “Prediction of diabetes and hypertension using multi-layer perceptron neural networks,” <https://doi.org/10.1142/S1793962321500124>, vol. 12, no. 2, Jan. 2021, doi: 10.1142/S1793962321500124.
- [4] R. Yashvanth, M. Rehan, A. Kodipalli, B. R. Rohini, and T. Rao, “Diabetes, Hypertension and Stroke Prediction Using Computational Algorithms,” *2023 World Conference on Communication & Computing (WCONF)*, 2023, doi: 10.1109/WCONF58270.2023.10235026.
- [5] Y. Zhao *et al.*, “Hypertension risk prediction models for patients with diabetes based on machine learning approaches,” *Multimed Tools Appl*, vol. 83, no. 20, pp. 59085–59102, Jun. 2024, doi: 10.1007/S11042-023-17926-X/METRICS.
- [6] L. Ciampiconi, A. Elwood, M. Leonardi, A. Mohamed, and A. Rozza, “A survey and taxonomy of loss functions in machine learning; A survey and taxonomy of loss functions in machine learning,” 2023.
- [7] M. F. Aslan and K. Sabanci, “A Novel Proposal for Deep Learning-Based Diabetes Prediction: Converting Clinical Data to Image Data,” *Diagnostics*, vol. 13, no. 4, Feb. 2023, doi: 10.3390/DIAGNOSTICS13040796.
- [8] K. Barath, “Introduction to Deep Neural Networks.” Accessed: Nov. 05, 2023. [Online]. Available: <https://www.datacamp.com/tutorial/introduction-to-deep-neural-networks>
- [9] S. Ronaghan, “Deep Learning: Overview of Neurons and Activation Functions.” Accessed: Sep. 25, 2023. [Online]. Available: <https://srngn.medium.com/deep-learning-overview-of-neurons-and-activation-functions-1d98286cfl e4>
- [10] E. Afsaneh, A. Sharifdini, H. Ghazzaghi, and M. Ghobadi, “Recent applications of machine learning and deep learning models in the prediction, diagnosis, and management of diabetes: a comprehensive review,” *Diabetol Metab Syndr*, vol. 14, no. 1, Dec. 2022, doi: 10.1186/S13098-022-00969-9.
- [11] H. Kim, S. Hwang, S. Lee, and Y. Kim, “Classification and Prediction on Hypertension with Blood Pressure Determinants in a Deep Learning Algorithm,” 2022, doi: 10.3390/ijerph192215301.
- [12] B. F. Wee, S. Sivakumar, K. H. Lim, W. K. Wong, and F. H. Juwono, “Diabetes detection based on machine learning and deep learning approaches,” *Multimed. Tools Appl.*, vol. 83, no. 8, pp. 24153–24185, Mar. 2023, doi: 10.1007/s11042-023-16407-5.
- [13] S. P. Chatrati *et al.*, “Smart home health monitoring system for predicting type 2 diabetes and hypertension,” *Journal of King Saud University - Computer and Information Sciences*, vol. 34, no. 3, pp. 862–870, Mar. 2022, doi: 10.1016/j.jksuci.2020.01.010.
- [14] L. D. Gopiseti, S. K. L. Kummera, S. R. Pattamsetti, S. Kuna, N. Parsi, and H. P. Kodali, “Multiple Disease Prediction System using Machine Learning and Streamlit,” *Proceedings - 5th International Conference on Smart Systems and Inventive Technology, ICSSIT 2023*, pp. 923–931, 2023, doi: 10.1109/ICSSIT55814.2023.10060903.

- [15] Y. W. ; Jeong *et al.*, “Prediction Model for Hypertension and Diabetes Mellitus Using Korean Public Health Examination Data,” 2022, doi: 10.3390/diagnostics12081967.
- [16] A. Somani, A. Horsch, and D. K. Prasad, “Interpretability in Deep Learning,” *Interpretability in Deep Learning*, pp. 1–466, Jan. 2023, doi: 10.1007/978-3-031-20639-9/COVER.
- [17] A. Karpatne, X. Jia, and V. Kumar, “Knowledge-guided Machine Learning: Current Trends and Future Prospects,” 2024.
- [18] R. Kurniawan *et al.*, “Hypertension prediction using machine learning algorithm among Indonesian adults,” *IAES International Journal of Artificial Intelligence (IJ-AI)*, vol. 12, no. 2, pp. 776–784, Jun. 2023, doi: 10.11591/ijai.v12.i2.pp776-784.
- [19] Kumar. Manoj, Perumal. Srinivasa, R. Nadesh, and K. Arivuselvan, “Type 2: Diabetes mellitus prediction using Deep Neural Networks classifier,” *International Journal of Cognitive Computing in Engineering*, vol. 1, pp. 55–61, Jun. 2020, doi: 10.1016/j.ijcce.2020.10.002.
- [20] M. E. Febrian, F. X. Ferdinan, G. P. Sendani, K. M. Suryanigrum, and R. Yunanda, “Diabetes prediction using supervised machine learning,” *Procedia Comput Sci*, vol. 216, pp. 21–30, 2022, doi: 10.1016/J.PROCS.2022.12.107.
- [21] Zhang and Q. Yang, “A Survey on Multi-Task Learning,” *IEEE Trans Knowl Data Eng*, vol. 34, no. 12, pp. 5586–5609, Dec. 2022, doi: 10.1109/TKDE.2021.3070203.
- [22] T. Wongvorachan, S. He, and O. Bulut, “A Comparison of Undersampling, Oversampling, and SMOTE Methods for Dealing with Imbalanced Classification in Educational Data Mining,” 2023, doi: 10.3390/info14010054.
- [23] K. Ghosh, C. Bellinger, R. Corizzo, P. Branco, B. Krawczyk, and N. Japkowicz, “The class imbalance problem in deep learning,” *Mach Learn*, vol. 113, no. 7, pp. 4845–4901, Jul. 2024, doi: 10.1007/S10994-022-06268-8/FIGURES/27.
- [24] N. Nasir, M. Sameer, F. Barneih, O. Alshaltone, and M. Ahmed, “Deep Learning Classification of Photoplethysmogram Signal for Hypertension Levels,” *Computer Vision and Pattern Recognition*, 2024, doi: <https://doi.org/10.48550/arXiv.2405.14556>.
- [25] L. Jiang *et al.*, “Diabetes risk prediction model based on community follow-up data using machine learning,” *Prev Med Rep*, vol. 35, p. 102358, Oct. 2023, doi: 10.1016/J.PMEDR.2023.102358.
- [26] G. Naidu, T. Zuva, and E. M. Sibanda, “A Review of Evaluation Metrics in Machine Learning Algorithms,” *Lecture Notes in Networks and Systems*, vol. 724 LNNS, pp. 15–25, 2023, doi: 10.1007/978-3-031-35314-7_2.
- [27] X. Chen *et al.*, “Context Autoencoder for Self-supervised Representation Learning,” *Int J Comput Vis*, vol. 132, no. 1, pp. 208–223, Jan. 2024, doi: 10.1007/S11263-023-01852-4/METRICS.
- [28] A. Rajkomar, J. Dean, and I. Kohane, “Machine Learning in Medicine,” *New England Journal of Medicine*, vol. 380, no. 14, pp. 1347–1358, Apr. 2021, doi: 10.1056/NEJMRA1814259.
- [29] S. M. Ganie, P. K. D. Pramanik, M. M. Bashir, S. Mallik, and H. Qin, “An ensemble learning approach for diabetes prediction using boosting techniques,” *Front Genet*, vol. 14, 2023, doi: 10.3389/fgene.2023.1252159.

- [30] A. Ahmed *et al.*, “Machine Learning Algorithm-Based Prediction of Diabetes Among Female Population Using PIMA Dataset,” *Healthcare (Switzerland)*, vol. 13, no. 1, p. 37, Jan. 2025, doi: 10.3390/HEALTHCARE13010037/S1.
- [31] F. Carlucci, “OPTIMIZATION OF DEEP NEURAL NETWORKS FOR PPG-BASED BLOOD PRESSURE ESTIMATION ON EDGE DEVICES,” Dec. 2024.
- [32] A. García-Pérez, “A New Estimator: Median of the Distribution of the Mean in Robustness,” 2023, doi: 10.3390/math11122694.
- [33] Dash, A. K. Behera, S. Dehuri, and A. Ghosh, “An outliers detection and elimination framework in classification task of data mining,” *Decision Analytics Journal*, vol. 6, p. 100164, Mar. 2023, doi: 10.1016/J.DAJOUR.2023.100164.
- [34] A. Abuzaid and I. Alkrunz, “A COMPARATIVE STUDY ON UNIVARIATE OUTLIER WINSORIZATION METHODS IN DATA SCIENCE CONTEXT,” *Statistica Applicata - Italian Journal of Applied Statistics*, vol. 36, no. 1, pp. 85–99, Apr. 2024, doi: 10.26398/IJAS.0036-004.
- [35] G. A. Pradipta, R. Wardoyo, A. Musdholifah, I. N. H. Sanjaya, and M. Ismail, “SMOTE for Handling Imbalanced Data Problem: A Review,” *2021 6th International Conference on Informatics and Computing, ICIC 2021*, 2021, doi: 10.1109/ICIC54025.2021.9632912.