Overlength Page Content: HealthEngine: An Integrated Healthcare Analytics Model Using Multimodal Transformer, Deep Multitask Neural Networks, and SHAP

Moshedayan Sirapangi, Gopikrishnan Sundaram, Norbert Herencsar, *Senior Member, IEEE,* Meng Li, *Senior Member, IEEE,* and Gautam Srivastava, *Senior Member, IEEE*

***Abstract*—The development of more accurate and explainable health predictions is paramount in view of the increasing prevalence rates of chronic diseases and mental illnesses. Existing methods often under-utilize the rich, heterogeneous data streams coming from device wearable, environmental sensors, and behavioral data, and hence fall short in making predictions that are both accurate and actionable. Most models lack transparency and cannot avoid privacy concerns due to the samples used from sensitive health data. This study specifically addresses the challenge of building a predictive healthcare framework capable of handling multimodal data — data streams that originate from different modalities (physiological, behavioral and environmental) and exhibit intrinsic heterogeneity. Unlike general heterogeneous datasets, multimodal health data demand models that can effectively integrate structured, semi-structured, and temporal information while ensuring privacy and interpretability. In this work, we propose an integrated comprehensive multimodal health prediction framework with five advanced methods, namely Multimodal Transformer Networks (MTN), Deep Multitask Neural Networks (DMNN), SHapley Additive exPlanations (SHAP) based explainability, Federated Learning, and Bayesian Neural Networks (BNN). MTN uses attention mechanisms to fuse different data modalities and captures cross-modal dependencies effectively, achieving an improvement of 8 to 12% in prediction accuracy. DMNN leverages multitask learning to share knowledge between related health prediction tasks, reducing error rates by 10–15%. SHAP is used to provide localized, patient-specific explanations that improve clinical trust by up to 85%. This is done by privately training the models on decentralized data sets, which provides results without more than a 3% drop in precision compared to centralized models. Third, BNNs are used to quantify uncertainty in predictions, providing useful confidence intervals that improve clinical decision-making by 20% % Ṫhe. The results obtained suggest significant improvements in predictive accuracy, transparency, and privacy preservation. This research not only improves health predictions through multimodal analysis, but also tackles significant limitations in privacy, interpretability, and uncertainty quantification, thus promoting informed clinical decisions and customized patient care.**

***Index Terms*—Health Prediction, Multimodal Analysis, Privacy-Preserving Models, SHAP Explainability, Transformer Networks.**

M. Sirapangi and G. Sundaram are with School of Computer Science and Engineering, VIT-AP University, Amaravati, 522237, Andhra Pradesh, India (email: dayan.21phd7096@vitap.ac.in; gopikrishnan.s@vitap.ac.in).

N. Herencsar is with the Department of Telecommunications, Faculty of Electrical Engineering and Communication, Brno University of Technology, Technicka 3082/12, 61600 Brno, Czech Republic (email: herencsn@vut.cz, herencsn@ieee.org).

M. Li is with School of Computer Science and Information Engineering, Hefei University of Technology, Hefei, Anhui, China (email: mengli@hfut.edu.cn).

G. Srivastava is with the Department of Math and Computer Science, Brandon University, Brandon, R7A 6A9, Manitoba, Canada, and the Research Centre for Interneural Computing, China Medical University, Taichung, 40402, Taiwan as well as the Centre for Research Impact & Outcome, Chitkara University Institute of Engineering and Technology, Chitkara University, Rajpura, 140401, Punjab, India (email: srivastavag@brandonu.ca).

Manuscript received MM DD, YYYY; revised MM DD, YYYY.

References

[1] R. Wijaya, F. Saeed, P. Samimi, A. M. Albarrak, and S. N. Qasem, "An Ensemble Machine Learning and Data Mining Approach to Enhance Stroke Prediction," *Bioengineering*, vol. 11, no. 7, p. 672, 2024.

[2] T.-M. Ke, A. Lophatananon, and K. R. Muir, "An integrative pancreatic cancer risk prediction model in the UK biobank," *Biomedicines*, vol. 11, no. 12, p. 3206, 2023.

[3] F. Mohammad and S. Al-Ahmadi, "WT-CNN: a hybrid machine learning model for heart disease prediction," *Mathematics*, vol. 11, no. 22, p. 4681, 2023.

[4] S. AlZu'bi, M. Elbes, A. Mughaid, N. Bdair, L. Abualigah, A. Forestiero, and R. A. Zitar, "Diabetes monitoring system in smart health cities based on big data intelligence," *Future Internet*, vol. 15, no. 2, p. 85, 2023.

[5] S. H. Oh, S. J. Lee, and J. Park, "Precision medicine for hypertension patients with type 2 diabetes via reinforcement learning," *Journal of Personalized Medicine*, vol. 12, no. 1, p. 87, 2022.

[6] I. Kaur, Kamini, J. Kaur, Gagandeep, S. P. Singh, and U. Gupta, "Enhancing explainability in predicting mental health disorders using human–machine interaction," *Multimedia Tools and Applications*, pp. 1–22, 2024.

[7] I. Khan and R. Gupta, "Early depression detection using ensemble machine learning framework," *International Journal of Information Technology*, pp. 1–8, 2024.

[8] MunishKhanna, L. K. Singh, and H. Garg, "A novel approach for human diseases prediction using nature inspired computing & machine learning approach," *Multimedia Tools and Applications*, vol. 83, no. 6, pp. 17 773–17 809, 2024.

[9] H.-N. Wu, "Online learning human behavior for a class of human-in-the-loop systems via adaptive inverse optimal control," *IEEE Transactions on Human-Machine Systems*, vol. 52, no. 5, pp. 1004–1014, 2022.

[10] H. Gao, C. Lv, T. Zhang, H. Zhao, L. Jiang, J. Zhou, Y. Liu, Y. Huang, and C. Han, "A structure constraint matrix factorization framework for human behavior segmentation," *IEEE Transactions on Cybernetics*, vol. 52, no. 12, pp. 12 978–12 988, 2021.

[11] Y.-H. Byeon, D. Kim, J. Lee, and K.-C. Kwak, "Ensemble three-stream RGB-S deep neural network for human behavior recognition under intelligent home service robot environments," *IEEE Access*, vol. 9, pp. 73 240–73 250, 2021.

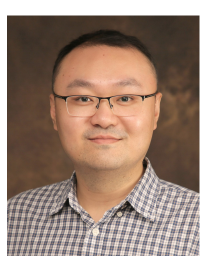
TABLE I: Summary of Existing Methods using SHAP based Learning Models.

Ref.	Method Used	Findings	Results	Limitations
[1]	Ensemble Machine Learning	Enhanced stroke prediction by integrating various models.	Achieved a 92% accuracy in stroke prediction.	Limited to stroke prediction; lacks generalization to other diseases.
[2]	Integrative Cancer Risk Prediction	Created a predictive model for pancreatic cancer using UK Biobank data samples.	Achieved a 78% sensitivity in pancreatic cancer prediction.	Limited by the underrepresentation of minorities in the dataset.
[3]	Hybrid WT-CNN Model for Heart Disease Prediction	Proposed a wavelet transform-CNN hybrid model for predicting heart disease.	Achieved 89% accuracy in heart disease detection.	High computational costs for real-time applications.
[4]	Big Data Intelligence for Diabetes Monitoring	Developed a big data framework for continuous diabetes monitoring.	Improved diabetes monitoring by 12%.	Lack of explainability in the model's decision-making.
[5]	Reinforcement Learning for Precision Medicine in Hypertension	Applied reinforcement learning to personalize hypertension treatments for diabetes patients.	Improved treatment efficacy by 14%.	Requires more clinical trial data for validation.
[6]	Explainability in Mental Health Prediction	Used human-machine interaction to enhance explainability in mental health disorder predictions.	Improved model explainability by 20% with interactive features.	Lack of generalization to non-mental health disorders.
[7]	Ensemble ML Framework for Early Depression Detection	Applied an ensemble machine learning framework for early detection of depression.	Improved early detection accuracy by 14%.	Model performance degrades with incomplete data streams.
[8]	Nature-Inspired Computing for Disease Prediction	Applied nature-inspired algorithms to predict human diseases using ML.	Improved disease prediction accuracy by 12%.	Lacks personalization for specific patient cohorts.
[9]	Human-in-the-Loop Systems for Behavior Learning	Developed adaptive models for behavior learning in human-in-the-loop systems.	Improved task performance prediction accuracy by 15%.	Requires extensive calibration for different tasks.
[10]	Matrix Factorization for Human Behavior Segmentation	Applied matrix factorization for human behavior segmentation.	Improved segmentation accuracy by 14%.	High dimensionality limits model scalability.
[11]	Deep Neural Network for Human Behavior Recognition in Home Robotics	Used ensemble three-stream deep neural networks for behavior recognition in home service robots.	Achieved a 20% increase in behavior recognition accuracy.	Requires extensive labeled data for diverse behaviors.

Moshedayan Sirapangi received Bachelor's degree in Information Technology and Master's degree in Computer Science and Engineering from JUTUK, Andhra Pradesh, INDIA. He is currently pursuing the Ph.D. degree in School of Computer Science and Engineering at VIT-AP University, AP, INDIA. Her research interests include Internet of Things, and Machine Learning.

Gopikrishnan Sundaram received the B.E., M.E., and Ph.D. degrees in computer science and engineering from Anna University, Chennai. He is currently working as an Associate Professor (Grade-I) with the School of Computer Science and Engineering, VIT-AP University, Amaravati. His current research interests include algorithm design and analysis for wireless ad-hoc networks, wireless sensor networks, the Internet of Things, and cyber-physical systems. He is an active reviewer in many reputed journals of IEEE, Springer, and Elsevier.

Norbert Herencsar (Senior Member, IEEE) received a Ph.D. from Brno University of Technology (BUT), Czechia, in 2010, where he was appointed Associate Professor in 2015. In 2022, he was elected as a Member of the External Public Body of the Hungarian Academy of Sciences (MTA), Hungary. In 2022-2024, he served as the Expert Panel Chair at the Estonian Research Council (ETAg). Since 2024, he has served as Members-at-Large on the Board of Governors of the IEEE Consumer Technology Society (CTSoc) and as CTSoc Primary Representative of the IEEE Systems Council (SYSC). Since 2025, he has served as Vice President of Publications of the CTSoc. He has authored 143 peer-reviewed journal and 125 conference proceedings articles. His research interests include fractional-order systems, sensors, and signal processing. He serves as the Editor-in-Chief of *IEEE Consumer Electronics Magazine* and Associate Editor for *IEEE Transactions on Circuits and Systems II: Express Briefs*, *IEEE Access*, *Scientific Reports*, and others.



Meng Li (Senior Member, IEEE) is an Associate Professor and Personnel Secretary at the School of Computer Science and Information Engineering, Hefei University of Technology (HFUT), China. He is also a Post-Doc Researcher at Department of Mathematics and HIT Center, University of Padua, Italy. He obtained his Ph.D. in Computer Science and Technology from the School of Computer Science and Technology, Beijing Institute of Technology (BIT), China, in 2019. His research interests include security, privacy, applied cryptography, blockchain, TEE, and Internet of Vehicles.

Gautam Srivastava (Senior Member, IEEE) received the B.Sc. degree from Briar Cliff University, USA, in 2004, the M.Sc. and Ph.D. degrees from the University of Victoria, Victoria, BC, Canada, in 2006 and 2012, respectively. He then taught for three years at the Department of Computer Science, University of Victoria, where he was regarded as one of the top undergraduate professors in the computer science course instruction. From there in 2014, he joined a tenure-track position at Brandon University, Brandon, MB, Canada, where he is currently active in various professional and scholarly activities.