

Guest Editorial Preface

Special Issue on Explainable AI in Healthcare Informatics and Clinical Decision Support

Rajeev Kumar Gupta, Computer Science and Engineering, Pandit Deendayal Energy University, Gandhinagar, India

Arti Jain, Computer Science and Engineering, Jaypee Institute of Information Technology, Noida, India

D. Jude Hemanth, Electronic and Communication Engineering, Karunya University, Coimbatore, India

As the availability of structured and unstructured data grows, it is becoming quite difficult to analyse this data in real-time that too with the higher accuracy. Artificial Intelligence (AI) is a tool that can be used to develop a model which analyses data quickly and automatically. Due to the high complexity of AI models, they are black box in nature, so it is impossible to know why the model provides the specific answer. Lack of transparency is one of the main barriers to the adoption of the AI model in the healthcare sector. It has become essential to develop some methods to explore the black box nature of AI and to understand how AI can be designed to operate responsibly. This is where Explainable Artificial Intelligence (XAI) comes into the picture. XAI will assist the developers in building customer trust in the transparency, accountability, and fairness of AI solutions.

This special issue, “Explainable AI in Healthcare Informatics and Clinical Decision Support,” collects contributions that address the applications of AI in the healthcare domain, explainable and interpretable AI as well as recent developments, innovations, and challenges in the healthcare sector.

This special issue of the *International Journal of Business Analytics* (IJBAN), IGI Global, has received a total of 45 papers, out of which six are finally selected that cover a range of aspects of deep learning and XAI in medical domain. Each of these papers has undergone full double blind peer review, prior to being selected for this special issue.

The first paper is “Futuristic Prediction of Missing Value Imputation Methods Using Extended ANN” by Ashok Kumar Tripathi, Hemraj Saini and Geetanjali Rathee. One of the most significant issues facing researchers in artificial intelligence is how to deal with missing values. This article offers a method for predicting missing values using an extended Artificial Neural Network (ANN). It provides a comparative study of four approaches that are used for missing value prediction: K-Nearest Neighbour (KNN), Recurrent Neural Network (RNN), Iterative KNN imputation, and extended ANN.

Mohammad Kamel Daradkeh presents a data analytics framework that aims to analyse topics and sentiments associated with COVID-19 vaccine misinformation in social media in his paper “Analyzing Sentiments and Diffusion Characteristics of COVID-19 Vaccine Misinformation Topics in Social Media: A Data Analytics Framework.” A total of 40,359 tweets related to COVID-19 vaccination are collected between January 2021 and March 2021. Misinformation is detected using multiple predictive machine learning models. Latent Dirichlet Allocation (LDA) topic model is used to identify dominant topics in the COVID-19 vaccine misinformation. Sentiment orientation of misinformation is analyzed using a lexicon-based approach. An independent-samples t-test is performed to compare the number of

replies, retweets, and likes of misinformation with different sentiment orientations. Based on the data sample, the results show that COVID-19 vaccine misinformation included 21 major topics. Across all misinformation topics, the average number of replies, retweets, and likes of tweets with negative sentiment are 2.26, 2.68, and 3.29 times higher, respectively, than those with the positive sentiment.

In “Alleviation of Delay in Tele-Surgical Operations Using Markov Approach-Based Smith Predictor,” Ratish Kumar, Rajeev Kumar, and Madhav Ji Nigam have discussed delay in Tele-Surgical Operations Using Markov Approach-Based Smith Predictor. Networked Control System (NCS) involves the feedback control loop system wherein the control components such as actuators and sensors are controlled and allowed to share their feedback over real time network with distributed users spread geographically. This paper has designed and simulated the functionality of a model-based Smith predictive controller. The model and randomized error estimations are employed through Markov approach and Kalman techniques. The simulation results show a delay of 49.926ms from master controller to slave controller and 79.497ms of delay from sensor to controller results, to a total delay of 129.423ms. This reduced delay improves the surgical accuracy and eliminate the risk factors to criticality of patients’ health.

Siddhartha Kumar Arjaria, Abhishek Singh Rathore, and Gyanendra Chaubey explore the XAI method for Thyroid Disease Prediction in their paper “Developing an Explainable Machine Learning-Based Thyroid Disease Prediction Model: XAI-Based Thyroid Disease Prediction.” The presented work explores a thyroid disease diagnosis system. SHAP, a popular method based on coalition game theory is used for interpretability of results. The work explains the system behaviour both locally and globally and shows how machine learning can be used to ascertain the causality of the disease and support doctors to suggest the most effective treatment of the disease. The work not only demonstrates the results of machine learning algorithms but also explains related features, their importance and model insights.

In “A Framework for Feature Selection using Natural Language Processing for User Profile Learning for Recommendations of Healthcare Related Content,” Mona Tanwar, Sunil Kumar Khatri, and Ravi Pendse have described feature selection methods for healthcare related content. In the proposed framework, user profile learning and modelling has been done by analysing and embedding semantics of terms from the healthcare text referred by the users. Multiple user interest profiles for each user have been constructed by segmenting related terms into non-overlapping semantic clusters. There are three main components in the proposed framework are highlighted as- Extracting relevant terms from the healthcare text-based content referred by users, Semantic analysis of the terms, and Clustering the semantically related terms. The experiments have shown improved precision for the proposed approach as compared to the state-of-the-art technique with a mean reciprocal rank of 0.76.

In “Evaluation of Diagnostic Performance of Machine Learning Algorithms to Classify the Fetal Heart Rate Baseline From Cardiotocograph,” Sahana Das, Obaidullah Md S. K., Kaushik Roy, and Chanchal Kumar Saha have discussed classification of fetal heart rate baseline from cardiotocograph by using machine learning. The CardioTocoGraphy (CTG) is the widely used cost-effective, non-invasive technique to monitor the fetal heart and mother’s uterine contraction pressure to assess the wellbeing of the fetus. Since visual estimation has its limitations, the authors have used various machine learning algorithms to classify the baseline. The 110 CTG traces from CTU-UHB dataset are divided into three subsets using stratified sampling to ensure that the sample is the accurate depiction of the population. The results are analyzed using various statistical methods and compared with the visual estimation by three obstetricians. The FURIA provides the highest accuracy of 98.11%.

Rajeev Kumar Gupta

Arti Jain

D. Jude Hemanth

Guest Editors

IJBAN