

# Postgraduate Seminar Series

*Venue: Graduate Seminar Room*

*Date & Time: September 28, 2024 at 2:30 PM*

## **Speaker Information**

Joydeb Kumar Sana (Student No. 1018054003) is a part-time Ph.D. student in the Department of Computer Science and Engineering (CSE) at the Bangladesh University of Engineering and Technology (BUET). He earned his undergraduate degree from Rajshahi University of Engineering and Technology (RUET) in 2006 and M.Sc. degree from BUET in 2015. His research interests encompass Applied Machine Learning, Big Data, Financial Informatics, Health Informatics and Data Privacy. He is currently doing his Ph.D. under the guidance of Prof. M Saifur Rahman. In this talk, he will discuss his ongoing research.



## **Privacy-Preserving Customer Churn Prediction Model in the Context of Telecommunication Industry**

Data is the main fuel of a successful machine learning model. A dataset may contain sensitive individual records e.g. personal health records, financial data, industrial information, etc. Training a model using this sensitive data has become a new privacy concern when someone uses third-party cloud computing. Trained models also suffer privacy attacks which leads to the leaking of sensitive information of the training data. This study is conducted to preserve the privacy of training data in the context of customer churn prediction modeling for the telecommunications industry (TCI). In this work, we propose a framework for privacy-preserving customer churn prediction (PPCCP) model in the cloud environment. We have proposed a novel approach which is a combination of Generative Adversarial Networks (GANs) and adaptive Weight-of-Evidence (aWOE). Synthetic data is generated from GANs, and aWOE is applied on the synthetic training dataset before feeding the data to the classification algorithms. Our experiments were carried out using eight different machine learning (ML) classifiers on three openly accessible datasets from the telecommunication sector. We then evaluated the performance using six commonly employed evaluation metrics. In addition to presenting a data privacy analysis, we also performed a statistical significance test. The training and prediction processes achieve data privacy and the prediction classifiers achieve high prediction performance (87.1% in terms of F-Measure for GANs-aWOE based Na<sup>ive</sup> Bayes model). In contrast to earlier studies, our suggested approach demonstrates a prediction enhancement of up to 28.9% and 27.9% in terms of accuracy and F-measure, respectively.