

Postgraduate Seminar Series

Venue: Graduate Seminar Room

Date & Time: September 14, 2024 at 2:30 PM

Speaker Information

Shamima Naznin (Std No. 0419052090) is a M.Sc. student in the department of CSE, BUET. She completed her undergraduate studies from Khulna University (KU). Her research interest lies in the field(s) of Bioinformatics, Machine Learning, and Software Engineering. She is currently doing her postgraduate thesis under the supervision of Dr. Muhammad Saifur Rahman. She will be speaking about her ongoing research in this talk.

IDENTIFICATION OF CANCER-CAUSING GENES FROM MICROARRAY DATA USING METAHEURISTICS AND MACHINE LEARNING.

In oncogenetics, the intricate relationship between gene expression and cancer development is still an enigma. Identification of the genetic driver for cancer is essential for an effective treatment innovation. Microarray technology holds immense potential for uncovering the epigenetic nature of cancer-causing genes. However, the high dimensionality of microarray gene expression data always poses a challenge in identifying cancer-causing genes. In this study, we have considered this challenge by proposing a framework that integrates metaheuristics, machine learning, and an enrichment analysis approach to find the most relevant genes associated with a particular cancer from the vast amount of expression data. Our methodology involves data preprocessing and ranking, followed by the integration of metaheuristic algorithms and machine learning models within a wrapper framework. A weighted ranking mechanism is used to identify the most significant genes with the highest classification accuracy. Finally, enrichment analysis validates the biological relevance of these top genes in cancer biology.

We evaluated the effectiveness of our method on three microarray cancer gene expression datasets. Our approach achieved superior performance in terms of both classification accuracy and the ability to identify biologically relevant genes. Through rigorous enrichment analysis, we validate the biological significance of the selected genes. Notably, our approach outperforms other state-of-the-art methods in accuracy and biological relevance, suggesting the identified genes are promising candidates for future therapeutic and research targets.