# EDL-PCC: AN ENSEMBLE METHOD FOR PAN-CANCER CLASSIFICATION USING GENOMIC DATA

**ABSTRACT**

Cancer is a leading cause of death worldwide and early detection is crucial for successful treatment. Pan-cancer classification, also known as multi-cancer classification, refers to the use of deep learning techniques to classify and diagnose different types of cancer. This approach is important because it enables the detection and diagnosis of cancer across multiple organs and tissue types, rather than being limited to a specific type of cancer. Pan-cancer classification using deep learning has some limitations, such as data scarcity, data imbalance, overfitting, and limited interpretability. These limitations can affect the performance and robustness of deep learning models, making it challenging to achieve accurate and reliable results in pan-cancer classification. In this work, we present an ensemble method, EDL-PCC, for classification of 33 cancer types based on gene expression data to overcome the limitations of deep learning methods by combining the predictions of multiple models to improve performance and robustness, and increase interpretability. The method combines the efficiency of VGG 16, ResNet 50, Inception V3 models to capture high-level features from structured input data. Additionally, a class-weighting mechanism is used to overcome data imbalance issues. Experiments were conducted on a gene expression dataset consisting of 10,267 cancer samples from 33 cancer types.

**Team Members Internal Guide**

K. Bhavitha Sri – 20WH1A1272 Mr. Mukhtar Ahmad Sofi

P. Ruchitha Sree – 20WH1A1277 Assistant Professor

Ch. Akhila – 20WH1A1288