Research and evaluation of methods for handling data imbalance in the classification of tea leaf diseases

Do Tu Duong

Abstract: This is a study aimed at identifying and evaluating the effectiveness of data imbalance handling methods in improving the accuracy of tea leaf disease classification systems. The practical context shows that data imbalance is a major challenge for machine learning-based applications, especially in the agricultural field such as tea leaf disease classification. The research objective is to address this problem by evaluating and comparing the performance of various data processing methods. Through this, an optimized tea leaf disease classification system is built. The expected result is to provide important insights into approaches for handling data imbalance, while contributing to improving the quality of automatic classification systems in the agricultural industry.

Keywords: Data imbalance, tea leaf disease classification, machine learning