

Abstract No: MP-09

Development of rapid detection strip for amines from other organic functional groups

Ravindu M. A. Y.^{1*}, Maheshani Y. K. D. C.², Perera L. R. V. D.¹, Kumarika B. M. T.² and Wanniarachchi D. D. C. de S.¹

¹Department of Chemistry, Faculty of Science, University of Kelaniya, Sri Lanka.

²Department of Statistics and Computer Science, Faculty of Science, University of Kelaniya, Sri Lanka.
dakshikacw@kln.ac.lk*

Identifying organic compounds in a laboratory requires a lot of chemicals and hence, the process is expensive. To address the challenges of controlling costs and reducing chemical waste, an investigation into the integration of chemistry with computer science techniques has been initiated. This approach emphasizes the significance and innovative aspects of the research. The research focuses on predicting Organic Compounds using both color strips and machine learning methods. A disposable strip was designed with ten separate holes, each serving as a colorimetric indicator. The first hole does not contain any chemical, from the second hole FeCl₃, Chromic Acid, CuCl₂, FCP, Methyl Orange, Phenol Red, Bromophenol Blue, Thymol Blue, Bromocresol Green were in holes respectively. These sensor indicators react with Functional group, causing distinctive color changes. RGB values from colorimetric strips were extracted as the dataset using ImageJ, an image analysis software, which analyzed photos of the sensor strip to obtain the RGB values for each hole. Two methods were used to classify compounds. Initially, the dataset containing RGB values of every compound was subjected to Principal Component Analysis (PCA) to evaluate the sensor array's intrinsic capacity for distinguishing between distinct categories of organic compounds. Second, specific chemicals were categorized using their RGB profiles because of the development of machine learning algorithms. It was shown that alcohol, ester, aldehyde, ketone, carboxylic acid could not be effectively separated using a single-color value (red, green, or blue) using PCA. But in the green value PCA plots, amines frequently formed unique clusters that allowed for their independent identification. Using PCA-derived green values, the K-Nearest Neighbors (KNN) model proved to be the most effective among all models for classifying chemicals as amines or non-amines, with an accuracy of 94%, recall of 95%, and precision of 95%. The KNN model achieved 99% training accuracy by adding additional amine and non-amine chemicals (27 and 26 respectively) to the training dataset. This study demonstrates the potential of RGB data for chemical identification, particularly for amines, suggesting that the colorimetric sensor array can be used as an identification strip for amine compounds in environmental samples and for educational purposes. Clustering mixtures like Carboxylic Acid-Aldehyde, Alcohol-Ketone into different categories was shown to be a substantial issue. Mixtures' color patterns frequently matched the dominating component's (amine in an amine-alcohol-ester combination, for example). This shows that in complex samples, clear categorization is made difficult by solvent effects or inter-component interactions.

Keywords: Colorimetric Sensor Array, Machine Learning, Organic Compound, Principal Component Analysis, RGB Values