

Developing a Machine Learning Model for Handwriting Recognition in the Context of Grading Assignment

Gunawardhana, M.U.K.¹, Shalika, K.B.¹, Weerasooriya, W.M.S.M.M.¹, Kariyapperuma, K.A.D.M.H.¹, Dissanayake, D.M.Y.U.B.¹, Rathnayake, R.M.L.M.¹, Padmasiri, K.M.P.U.¹

¹*Department of Information and Communication Technology,
Faculty of Technology, Rajarata University of Sri Lanka.*

*itt1718030@tec.rjt.ac.lk

A common technique used to assess student achievement in the realm of education is the written assignment. In the study, the focus is on the Machine Learning (ML) model for MCQ handwriting identification. The amount of human labor needed for the assessment is quite significant and is dependent on several variables, including the teacher's expertise, grading criteria, and time provided. Traditional evaluation procedures are extremely time-consuming and expensive in terms of the amount of work required to complete the evaluation, verify and publish the results. This study aims to develop a ML model that can accurately and efficiently grade handwritten responses. System's scope includes improving efficiency in answer evaluation, enhancing accessibility for students with disabilities, and streamlining the grading process in educational institutions by using Word2Vec, Support Vector Machines (SVM), and evaluation metrics of precision, recall, and F1-score. A system has been developed to grade handwritten assignments in picture format, extract text from scanned answer scripts using an optical character recognition tool, and grade answer sheets using ML and natural language processing techniques. The model interprets and classifies answers as A, B, C, D, or E. Two datasets were used: image datasets for preprocessing images and identifying words, and the question-answer dataset for building the ML model. With the pre-trained models, the highest accuracy of 78.95% was attained in the experiments. The scores are based on cosine set similarity measures, with each word in the evaluated answer paper carrying a corresponding mark.

Keywords: *Natural language processing; Optical character recognition; Educational assessment; WordToVec; Image preprocessing*