

# **Automated Text Classification of Library Books into the Dewey Decimal Classification (DDC)**

**A.W.A.P.L Wanigasooriya**  
*University of Kelaniya, Sri Lanka*

**Wajira Prasad De Silva**  
*Borough Markets, Australia*

Classifying books effectively is vitally essential in any library. Therefore, librarians provide bottom priority to the classification of sources available in their libraries or at the information centers. Classification refers conscious and calculated approach to arrange books in a most desired manner having a permanent sequence with easy access and retrieval. Dewey Decimal Classification (DDC) is a commonly used classification system in libraries, devised by Melvil Dewey in 1876. Currently, the 23rd edition of the full DDC is available through Web Dewey. From the beginning number assigning has been done by librarians. It is not an easy task since classification requires qualified librarians to determine the subject, the disciplinary focus, and the approach or form of the book. Although they have learned classification systems to assign exact no, human-generated classification numbers reported errors such as assigning wrong or complicated numbers and taking pretty much time to generate. Related literature confirmed that misclassification has led to minimizing the easy retrieval from the huge collection of a library. Therefore, this study seeks the possibility to apply Artificial Intelligence to support librarians to apply high accurate fully automate book classification numbers to save their time and effort. The study used text classification using Deep Learning for the purpose. Normally, human reads details of the book before assign an appropriate number for a book and cluster them into main and subcategories. A machine-readable method was used to perform the same task. Firstly, in this proposed system, the machine has scanned the title page of the book and then extract text from the image. The extracted text has been categorized according to DDC subdivisions by the implemented system. The Convolutional Neural Network (CNN), Deep Learning method was employed to develop the system mentioned above using a book depository public dataset. More than, 10,000 book details were train and test through the model. The implemented model reported that this method could perform high-precision (98%) and faster numbering than humans. Additionally, the system developed by this study can be recommended for the easy DDC number generating in any library and it can be improved for the better subject content analysis in any Online Public Access Catalogues (OPACs) for full-text searching as well.

**Keywords:** *Book Classification, CNN, Deep Learning, Dewey Decimal Classification, OPAC, Text Classification*