

CONSTRUCTION OF A WORD BAG MODEL TO DETECT PLAGIARISM IN THE MATLAB LANGUAGE

SIR-ALEXCI SUAREZ¹, THOMAS EDISON GUERRERO² & ALBERT MIYER SUAREZ³

^{1,2} *Engineering Faculty, University Francisco of Paula Santander Ocaña, Colombia.*

³ *Faculty of Engineering and Architecture. University of Pamplona, Colombia.*

ABSTRACT

A method is presented for building a word bag by extracting tokens from Matlab-coded programs, where each token is labeled according to a value that is assigned as the file is read, and its value varies according to the number of times it is in the program. The levels of plagiarism type 1, 2 and 3 are mainly analyzed, since they are the most used in the academic field. The results show that plagiarism can be detected with programs when using Level 2.

KEYWORDS: *Plagiarism, Matlab, bag of words, histogram of words*

Received: Dec 08, 2020; **Accepted:** Dec 28, 2020; **Published:** Jan 07, 2020; **Paper Id.:** IJMPERDDEC202074