

**GHOSTBUSTER: A WRITING STYLE DETECTION  
SYSTEM TO IDENTIFY GHOST-WRITTEN  
PLAGIARISM**

**Namila Dammalage**

A dissertation submitted in partial fulfillment of the requirements  
for the BSc (Hons) in Business Information Systems degree at the  
University of Westminster.

**Department of Computing**

**Informatics Institute of Technology, Sri Lanka**

**in collaboration with**

**University of Westminster, UK**

**2020**

## Abstract

Students at any level of academics go through a variety of problems through their university life. No matter what the problem might be, no student would like their academic grading or qualification to be of a lower level than level they desire. Gone are the days where a simple copy-paste off the internet will give students an edge over their colleagues. This is mainly due to the superior level of plagiarism detectors available in the market today. The dilemma arises when these world class plagiarism detectors are unable to detect when a specific report has been outsourced because there are no direct indicators available. Due to this reason a new market has opened by people that are experts in various fields where they complete these difficult assignments for the students for a price. This almost guarantees a good grade for the students with absolutely no effort. These writers are called ghostwriters and thanks to advancements in technology, they are extremely easy to find and are available for very low prices across the world.

The project aims to identify the reasons for why these plagiarism detectors fail to identify ghost-written reports and to design, develop and evaluate a solution that is capable of identifying the differences in writing patterns and styles between the student that submitted the report versus the writing patterns of the ghost-writer that actually wrote the report.

**GhostBuster** would provide an in-depth analysis of the text characteristics of a known report of the student and the report that is under suspicion of ghost-written plagiarism. It will be able to provide quantitative evidence on how similar the 2 reports are thereby providing the markers of reports with enough evidence to apprehend the students that resort to ghost-written plagiarism. The solution will further provide graphical analysis of all the deductions that are made to enable users to understand with ease.

The solution has been tested and evaluated by expert markers and graders along with non-expert stakeholders to evaluate the success of the project and the value that the solution will bring upon was appreciated.