

6COSC023C.Y - Final Year Project

**TLDR – Machine learning based lecture
note summarizer**

Isuru Weerakoon (2017041 / w1673730)

Supervisors – Ms. Shyani Siriwardene &

Ms. Salitha Dinushika

This report is submitted in partial fulfillment of the requirements for the

BSc (Hons) Business Information Systems

at the University of Westminster.

School of Computer Science & Engineering

University of Westminster

Date of submission – 5th May, 2022

Abstract

TLDR is a machine learning and natural language processing based text summarizer that helps undergraduate students and lecturers summarize study material, thus saving them valuable time. During the research done for this project, it was found students neglect reading their allocated study material for various reasons ranging from losing focus quickly to learning disorders such as ADHD, all these problems had the root cause of students not having enough time to read. Although there are a lot of text summarizers available online, they just offer basic summarization often with limitations, the features they provide are not very useful for students and lecturers. TLDR introduces many features including three summarization techniques (GenSim, NLTK & spaCy) that students and lectures can use to better summarize their material and use it for their academic purposes. Research was conducted into what problems students face and what features they expected from the solution and a prototype was designed. Based on the prototype, the application was implemented using Python. The summarize function of GenSim was used along with two algorithms made for spaCy and NLTK, Tkinter, which is a Python framework was used to give the user a graphical user interface which resulted in better user experience. Stakeholder reviews mentioned that the application was excellent in summarizing texts that were a few pages long but had issues when processing very large files. Overall, the application resulted in a positive experience for the stakeholders in which a majority expressed that they would be interested in utilizing TLDR for their academic purposes.