

DEPARTMENT OF DATA SCIENCE & CYBER SECURITY

II CSD-B Mini Project

A.Y. 2023-2024

Date:04-04-2024

Domain / Areas	DATA SCIENCE
Title of the Project	WEB DATA EXTRACTION FROM WEBSITES
Team Leader Name	1. K.VENKATA ADITHYA 22R21A6792
Team Members Name with Roll No	2. K.TAUFEEQ UMAR 22R21A6786
	3. B.VARUN 22R21A6768
Guide Name	Mrs.Srinija
Guide Signature	
<u>ABSTRACT:</u> Web scraping, a dynamic technique, has emerged as an indispensable tool for gathering structured information from websites and online sources. Web scraping encompasses a wide range of automated data extraction methods that enable users to access and collect data from websites and web applications. These techniques involve parsing the HTML, XML, or other structured data formats of web pages to extract specific information, such as text, images, links, and more. Web scraping plays a pivotal role in the digital age, offering unprecedented access to valuable data for research, business, and decision-making purposes. As the digital frontier continues to expand, web scraping remains a powerful tool for unlocking insights and opportunities from the vast online ecosystem. Web scraping should be conducted ethically and within legal boundaries. Some websites explicitly prohibit web scraping in their terms of service and scraping sensitive or personal data without consent is illegal in many jurisdictions. Web scraping is a powerful tool for collecting and utilizing data from the web.	
Software/Hardware Needs	Python, Javascript, HTML, CSS, VS Code.

Signature of the Incharge

Signature of the HOD(CS&DS)