***A Project Report on***

**WEB SCRAPING**



Submitted in partial fulfillment of the Requirements for the award of the degree of

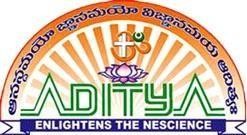
#### MASTER OF COMPUTER APPLICATIONS

**SUBMITTED BY**

**K.LALITHADEVI**

#### (19A91F0008)

**Under the Esteemed Guidance of Mrs.G.Satya, MCA, M.Tech Assistant Professor**



**DEPARTMENT OF MCA**

**ADITYA ENGINEERING COLLEGE**

**An Autonomous Institution**

Approved by AICTE, Affiliated to JNTUK &Accredited by NBA, NAAC with ‘A’ Grade Recognized by UGC under the sections 2(f) and 12(B) of the UGC act 1956 SURAMPALEM-533437, E.G.Dt, ANDHRA PRADESH.

**2019-2022**

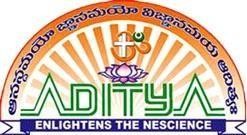
**ADITYA ENGINEERING COLLEGE**

#### An Autonomous Institution

Approved by AICTE, Affiliated to JNTUK &Accredited by NBA, NAAC with ‘A’ GradeRecognized by UGC under the sections 2(f) and 12(B) of the UGC act 1956

SURAMPALEM-533437, E.G.Dt, ANDHRA PRADESH.

**DEPARTMENT OF MCA**



**CERTIFICATE**

This is to certify that the project work entitled, **“Web Scraping”** is a bonafide work of **K.LALITHADEVI** bearing Regd.No:**19A91F0008** submitted to the faculty of Master of Computer Applications in partial fulfillment of the requirements for the award of the degree of **MASTER OF COMPUTER APPLICATIONS** from Aditya Engineering College, Surampalem for the academic year 2019-2022.

##### Project Guide Head of the Department

**Mrs.G.Satya, MCA, M.Tech Mrs.D.Beulah, M.Tech, (Ph.D)**

Assistant Professor, Associate Professor,

Department of MCA, Department of MCA,

Aditya Engineering College, Aditya Engineering College,

Surampalem-533437. Surampalem-533437.

**External Examiner**

## DECLARATION

I hereby declare that the project entitled **“Web Scraping”** submitted to **Aditya Engineering College, Surampalem** has been carried out by me alone under the guidance of **Mrs.G.Satya.**

Place: Surampalem Date:

**(K.LALITHADEVI) Reg No: 19A91F0008**

## ACKNOWLEDGEMENT

The satisfaction and euphoria that accompany the successful completion of any task would be incomplete without the mention of people who made it possible, whose constant guidance and encouragement crowned our efforts with success. It is a pleasant aspect that I have now the opportunity to express my gratitude for all of them.

The first person I would like to thank is my guide **Mrs.G.Satya, MCA, M.Tech, Assistant Professor, Department of MCA**, **Aditya Engineering College, Surampalem.** Her wide knowledge and logical way of thinking have made a deep impression on me. Her understanding, encouragement and personal guidance have provided the basis for this project. She is a source of inspiration for innovative ideas and her kind support is well known to all her students and colleagues.

I wish to thank **Mrs. D.Beulah, M.Tech, (Ph.D), Head of the Department, Aditya Engineering College, Surampalem,** who has extended her support for the success of this project.

I also wish to thank **Sri. M.Sreenivasa Reddy**, **Principal, Aditya Engineering College, Surampalem,** who has extended his support for the success of this project.

I would like to thank all the **Management &Technical Supporting Staff** for their timely help throughout the project.

# ABSTRACT

**ABSTRACT**

Web Scraping is a technique used to automatically extract the data from any websites and save it to a structured format such as spread sheet, database or csv file. Most websites do not allow you to save or download this data. If you need the data, the only option is to manually copy and paste the data - a very tedious job which can take many hours or days to complete. So that instead of manually copying the data from websites, the web scraping software will perform the same task with in fraction of the time.

This project uses the Django web framework to show how scraping can be automated by using the periodic task with celery tasks and workers. Recently, various advanced web scraping techniques and methods have been developed like text processing, artificial intelligence.