**VISVESVARAYA TECHNOLOGICAL UNIVERSITY BELAGAVI – 590 018**



**MINI PROJECT REPORT ON**

**“URL Scraping”**

Submitted In Partial Fulfillment of 6th semester WTA laboratory with Mini Project (18CS63)

**Bachelor of Engineering**

In

**Computer Science & Engineering**

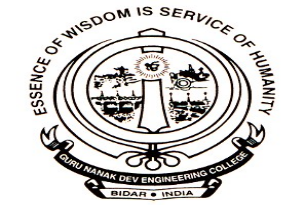
**Submitted By:**

**Abhishek(3GN20CS003)**

Under the guidance of

**Mr. SANDEEP IDDALGAVE**

**Asst. Professor**



**GURU NANAK DEV ENGINEERING COLLEGE**

**BIDAR-585401**

**Department of Computer Science & Engineering**

**Accredited by NBA, New Delhi**

**GURU NANAK DEV ENGINEERING COLLEGE**

**BIDAR-585401, KARNATAKA**

Approved by AICTE, New Delhi & Affiliated to VTU, Belagavi

Phone: 08482- 226949

Website:www.gndecb.ac.in



**DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING**

**CERTIFICATE**

This is to certify that the Mini Project Report entitled “**URL Scraping**” is a bonafide work carried out by **Abhishek (3GN20CS003)** in partial fulfillment of the WTA Laboratory with Mini Project(18CS63) for 6th semester Bachelor of Engineering in Computer Science & Engineering during the year 2022-23. It is certified that all the corrections/suggestions indicated for Internal Assessment have been incorporated in the report deposited in the department library. The mini project has been approved as it satisfies the academic requirements in respect of laboratory work prescribed for the said degree.

**Signature of Guide**

**Prof. Sandeep Iddalgave**

**URL Scraping**

**HTML Code**

<!DOCTYPE html>

<html>

<head>

<link rel="stylesheet" type="text/css" href="styles.css">

<link rel="stylesheet" type="text/css" href="https://cdnjs.cloudflare.com/ajax/libs/twitter-bootstrap/5.2.0/css/bootstrap.min.css">

</head>

<body>

<nav class="navbar navbar-expand-lg navbar-dark bg-dark">

<div class="container-fluid">

<a class="navbar-brand" href="#home">URL Scraper</a>

<button class="navbar-toggler" type="button" data-bs-toggle="collapse" data-bs-target="#navbarNav" aria-controls="navbarNav" aria-expanded="false" aria-label="Toggle navigation">

<span class="navbar-toggler-icon"></span>

</button>

<div class="collapse navbar-collapse" id="navbarNav">

<ul class="navbar-nav ms-auto">

<li class="nav-item">

<a class="nav-link" href="#home">Home</a>

</li>

<li class="nav-item">

<a class="nav-link" href="#about">About Us</a>

</li>

<li class="nav-item">

<a class="nav-link" href="#contact">Contact Us</a>

</li> <li class="nav-item">

<a class="nav-link" href="/Img.html">Screenshots</a>

</li>

</ul>

</div>

</div>

</nav>

<header class="bg-light text-center py-5">

<h1>URL Scraper</h1>

<p class="tagline">Scrape web content easily!</p>

</header>

<main>

<section id="home" class="py-5">

<div class="container">

<div class="row justify-content-center">

<div class="col-md-6">

<input type="text" id="urlInput" class="form-control mb-3" placeholder="Enter URL">

<button onclick="scrapeUrl()" class="btn btn-primary">Scrape the Web</button>

<div id="scrapedContent"></div>

</div>

</div>

</div>

</section>

<section id="about" class="py-5">

<div class="container">

<div class="row">

<div class="col-md-6 offset-md-3">

<div class="card shadow p-3 bg-light">

<div class="card-body">

<h2 class="card-title">About Us</h2>

<p class="card-text">

Web URL scraping is the process of extracting data from websites by programmatically accessing and parsing their HTML content. It involves sending HTTP requests to a specific URL, retrieving the HTML response, and then extracting the desired information from the HTML structure. URL scraping can be used to gather data such as text, images, links, or any other relevant content from web pages.

This web page provides information about the purpose and background of the URL Scraper. It aims to give visitors a better understanding of what the tool does and its benefits. The About Us page will include a pop-up card that provides a brief description of web URL scraping, explaining how it can be used to extract data from websites efficiently. The content will highlight the advantages of URL scraping and its potential applications in various domains such as data analysis, research, and automation.

</p>

</div>

</div>

</div>

</div>

</div>

</section>

<section id="contact" class="py-5">

<div class="container">

<div class="row">

<div class="col-md-6 offset-md-3">

<h2>Contact Us</h2>

<section id="contact" class="py-5">

<div class="container">

<div class="row">

<div class="col-md-6 offset-md-3">

<h2>Contact Us</h2>

<form id="contactForm">

<div class="mb-3">

<label for="nameInput" class="form-label">Name</label>

<input type="text" class="form-control" id="nameInput" placeholder="Enter your name" required>

</div>

<div class="mb-3">

<label for="numberInput" class="form-label">Number</label>

<input type="tel" class="form-control" id="numberInput" placeholder="Enter your phone number" required>

</div>

<div class="mb-3">

<label for="suggestionInput" class="form-label">Suggestion</label>

<textarea class="form-control" id="suggestionInput" rows="5" placeholder="Enter your suggestion" required></textarea>

</div>

<button type="submit" class="btn btn-primary">Submit</button>

</form>

<div id="contactResponse" class="mt-3"></div>

</div>

</div>

</div>

</section>

</main>

</div>

</div>

</div>

</section>

</main>

<script src="script.js"></script>

<script src="https://cdnjs.cloudflare.com/ajax/libs/twitter-bootstrap/5.2.0/js/bootstrap.bundle.min.js"></script>

</body>

</html>

**CSS CODE**

body {

  font-family: Arial, sans-serif;

  margin: 0;

  padding: 0;

  background: linear-gradient(to bottom, #ffafbd, #ffc3a0);

}

header {

  padding: 20px;

}

h1 {

  margin: 0;

}

.tagline {

  margin-top: 10px;

  color: #777;

}

input[type="text"] {

  width: 100%;

  padding: 10px;

  margin-bottom: 10px;

}

button {

  padding: 10px 20px;

  background-color: #333;

  color: #fff;

  border: none;

  cursor: pointer;

  border-radius: 5px;

  width: 100%;

  box-sizing: border-box;

}

#scrapedContent {

  margin-top: 20px;

  border: 1px solid #ccc;

  padding: 10px;

}

@media only screen and (min-width: 600px) {

  input[type="text"] {

    width: 300px;

  }

  button {

    width: auto;

  }

}

*/\* Responsive Styles for Menu \*/*

.navbar-collapse {

  justify-content: flex-end;

}

.navbar-nav {

  flex-direction: column;

}

.navbar-nav .nav-item {

  margin-bottom: 10px;

}

@media only screen and (min-width: 768px) {

  .navbar-collapse {

    justify-content: flex-end;

  }

  .navbar-nav {

    flex-direction: row;

  }

  .navbar-nav .nav-item {

    margin-bottom: 0;

    margin-right: 10px;

  }

}

*/\* Responsive Styles for Input and Button \*/*

@media only screen and (max-width: 600px) {

  input[type="text"] {

    width: 100%;

  }

  button {

    width: 100%;

  }

}

*/\* Gradient Background \*/*

body {

  background: linear-gradient(to bottom, #ffafbd, #ffc3a0);

}

*/\* Responsive Styles for Input and Button \*/*

@media only screen and (max-width: 600px) {

  input[type="text"] {

    width: 100%;

  }

  button {

    width: 100%;

  }

}

*/\* Responsive Styles for Menu \*/*

.navbar-collapse {

  justify-content: flex-end;

}

.navbar-nav {

  flex-direction: column;

}

.navbar-nav .nav-item {

  margin-bottom: 10px;

}

@media only screen and (min-width: 768px) {

  .navbar-collapse {

    justify-content: flex-end;

  }

  .navbar-nav {

    flex-direction: row;

  }

  .navbar-nav .nav-item {

    margin-bottom: 0;

    margin-right: 10px;

  }

}

**JAVASCRIPT CODE**

function scrapeUrl() {

  var url = document.getElementById("urlInput").value;

  if (!url) {

    alert("Please enter a URL.");

    return;

  }

  fetch("https://api.allorigins.win/get?url=" + encodeURIComponent(url))

    .then(*response* => response.json())

    .then(*data* => {

      var content = data.contents;

      document.getElementById("scrapedContent").innerHTML = content;

      alert("Scraping completed!");

    })

    .catch(*error* => {

      console.log("Error:", error);

      alert("Failed to scrape the URL. Please try again.");

    });

}

document.getElementById("contactForm").addEventListener("submit", function (*event*) {

  event.preventDefault();

*// Get form values*

  var name = document.getElementById("nameInput").value;

  var number = document.getElementById("numberInput").value;

  var suggestion = document.getElementById("suggestionInput").value;

*// Display submitted data*

  var contactResponse = document.getElementById("contactResponse");

  contactResponse.innerHTML = `

    <h4>Thank you for your submission!</h4>

    <p>Name: ${name}</p>

    <p>Number: ${number}</p>

    <p>Suggestion: ${suggestion}</p>

  `;

*// Clear form inputs*

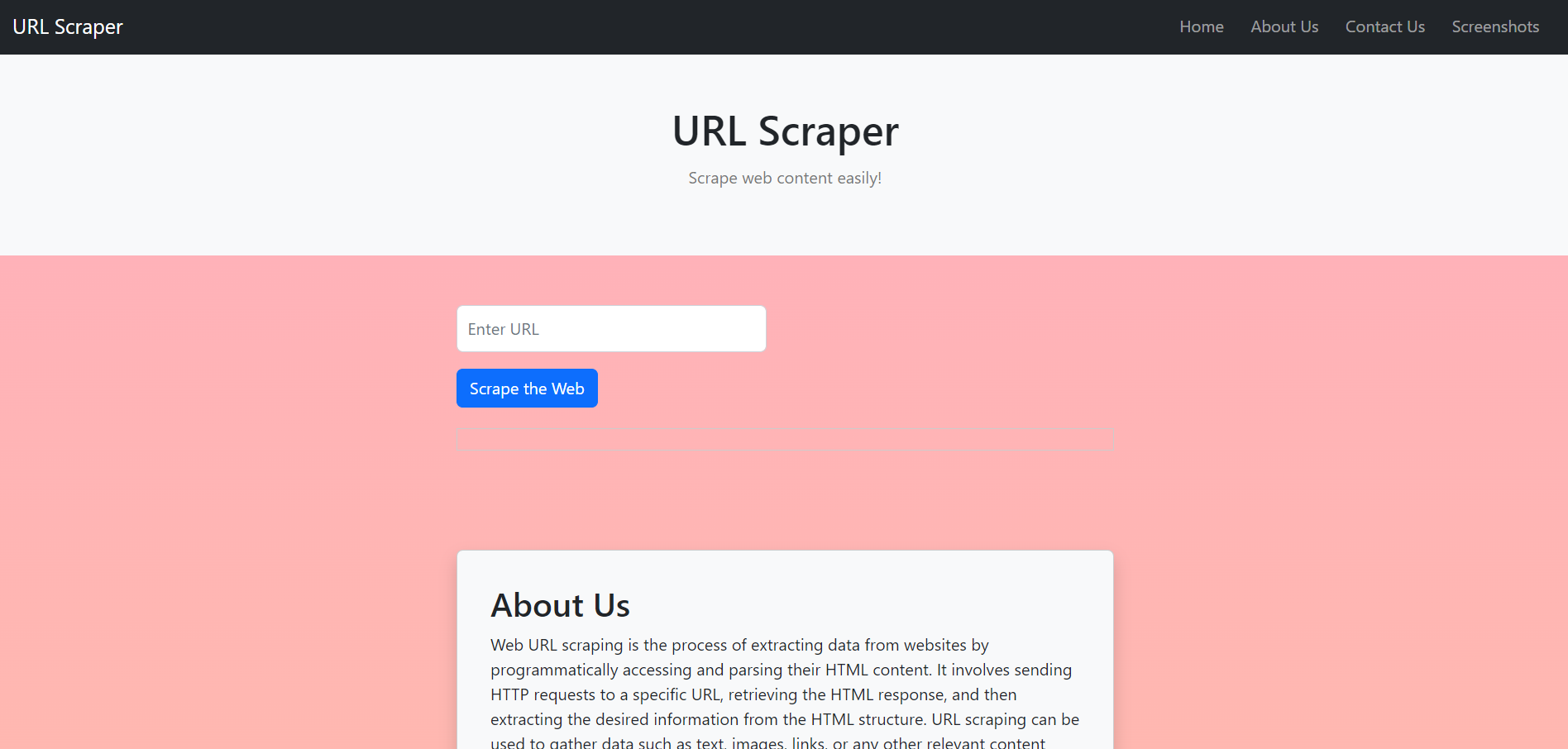
  document.getElementById("nameInput").value = "";

  document.getElementById("numberInput").value = "";

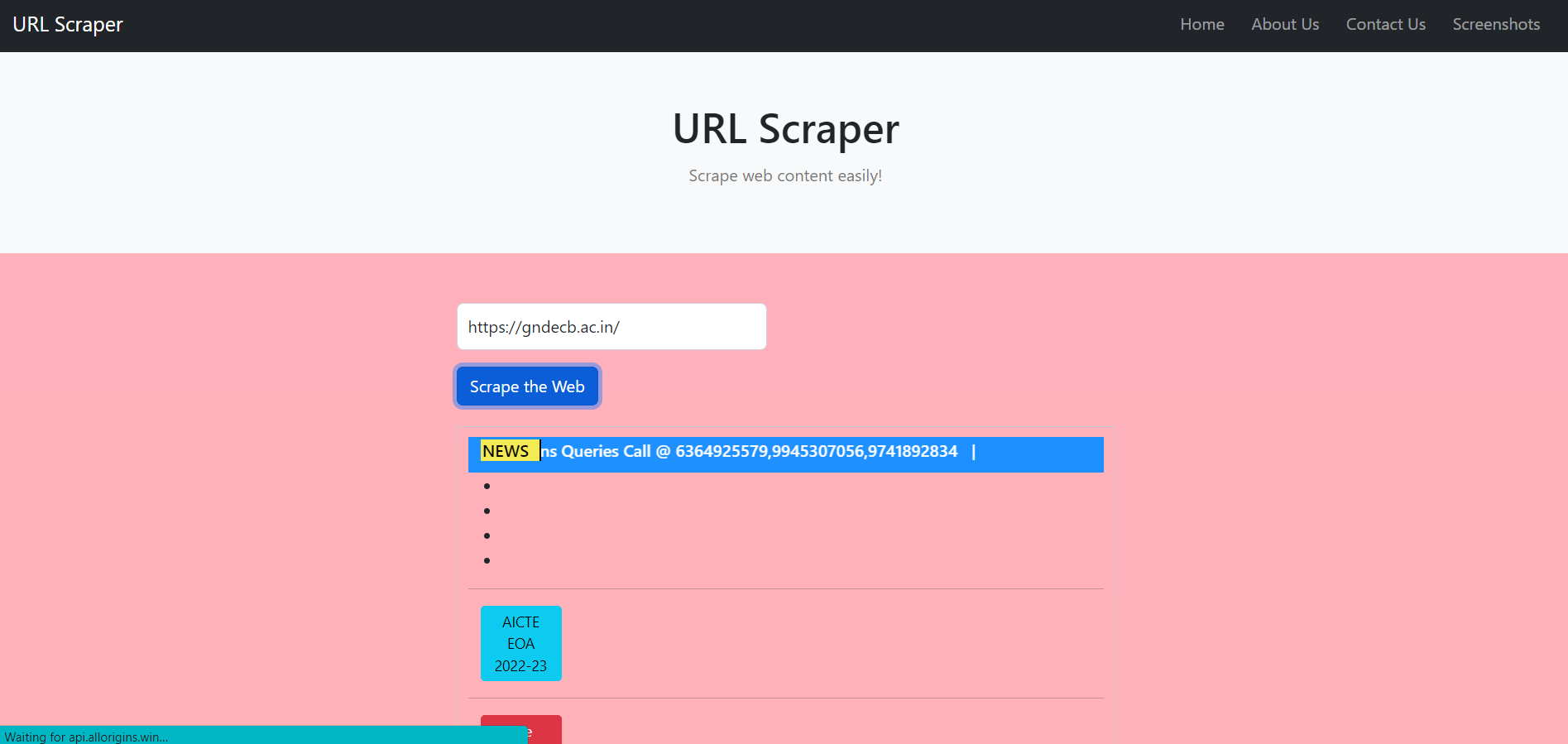
  document.getElementById("suggestionInput").value = "";

});

OUTPUT



*Fig: URL FORM*



*Fig: URL SCRAPPED DATA*