

Mohammad Idrees Bhat

Tech Skills Trainer | AI/ML Consultant

# IMDB Reviews

Scraping & Cleaning the movie ratings

Mohammad Idrees Bhat

## PROJECT DESCRIPTION

Scrape movie reviews from IMDB and clean data to prepare it for analysis. Students will use web scrapping tools and data cleaning techniques to create a high quality database ready for futher analysis.

## PROJECT DELIVERABLES

1. Cleaned dataset
2. Jupyter notebook with code and documentation
3. Report on data acquisition and cleaning processess

## FURTHER INSTRUCTIONS

### IMDb Movie Reviews Project Instructions:

#### 1. Web Scraping:

- Use a web scraping tool (like BeautifulSoup or Scrapy) to extract IMDb movie reviews.
- Ensure you gather relevant fields (e.g., review, name, date).

(Choose the relevant fields based on your descretion)

#### 2. Data Cleaning:

- Remove irrelevant data, empty reviews, and duplicates.
- Handle missing data appropriately

#### 3. Data Transformation: (Optional/Advanced)

- Convert all text to lowercase.
- Optionally, perform text tokenization or stemming for future analysis.

#### 4. **Prepare Data for Analysis:**

- Store cleaned data in a structured format (CSV).
- Create a Jupyter notebook documenting your process with code, explanations, and necessary visualizations.

#### 5. **Report:**

- Summarize the steps taken in data acquisition, cleaning, and transformation.

Students can use the IMDb website for reviews: [IMDb Movie Reviews](#)

\_\_\_\_\_ BEST OF LUCK ... :)

\_\_\_\_\_

THANK YOU!

**Mohammad Idrees Bhat**

Tech Skills Trainer | AI/ML Consultant