Project 04

Data Science with Python — Real World Project

After learning about Data Science in depth, it is now time to implement the knowledge gained through this course in real-life scenarios. We will provide you with four scenarios where you need to implement data science solutions. To perform these tasks, you can use the different Python libraries such as NumPy, SciPy, Pandas, scikit-learn, matplotlib, BeautifulSoup, and so on.

Here, we ask you to perform the analysis using the Exploratory Data Analysis technique.

The details of these projects and the scope of each project are listed in the sections below.

- > Data acquisition of the movielens dataset
 - users dataset
 - rating dataset
 - movies dataset
- > Perform the Exploratory Data Analysis (EDA) for the users dataset
 - Visualize user age distribution
 - Visualize overall rating by users
 - Find and visualize the user rating of the movie "Toy Story"
 - Find and visualize the viewership of the movie "Toy Story" by age group
 - Find and visualize the top 25 movies by viewership rating
 - Find the rating for a particular user of user id = 2696
 - \circ Visualize the rating data by user of user id = 2696
- ➤ Perform machine learning on first 500 extracted records
 - Use the following features:
 - o movie id
 - o age
 - o occupation
- > Use rating as label

Good Luck!