The main goal of the task is to conduct the basic mining software repository process to obtain target data and, furthermore, to extract useful information via basic statistical analysis and machine learning.

* **Task 1**. Please use the *PyDriller* library to obtain the list of commits and issues of the Kubernetes project on GitHub (https://github.com/kubernetes/kubernetes) for the last 24 months. Each commit should contain at least the basic information, e.g., committer id, commit time, changed files, changed LOC, etc. Similarly, each issue should contain information, e.g., issue content, issue type, issue status, issue open time, issue closed time, issue creator, issue closer, etc. In addition, please clean the data and save them in CSV form.
* **Task 2**. Please use a proper tool to visualize the data retrieved from the previous task. Here we would like to see 1) the distribution of contributors’ daily contributions, 2) the most frequent types of issues, and 3) the monthly number of commits, and (open and closed) issues for the last 5 years.
* **Task 3**. Please use a proper method, together with the data obtained previously, to predict the number of commits of the project in December 2024 and reason how accurate your prediction is.