Practical 2 Preparation:

A software project is not just about writing a few hundred lines of source code to achieve a particular objective. The scope of a software project is comparatively quite large, and such a project could take several years to complete. However, the phrase "quite large" could only give some (possibly vague) qualitative information. As in any other science and engineering discipline, one would be interested to measure how complex a project is. One of the major activities of the project planning phase, therefore, is to estimate various project parameters in order to take proper decisions. Some important project parameters that are estimated include:

* **Project size:** What would be the size of the code written say, in number of lines, files, modules?
* **Cost:** How much would it cost to develop a software? A software may be just pieces of code, but one has to pay to the managers, developers, and other project personnel.
* **Duration:** How long would it be before the software is delivered to the clients?
* **Effort:** How much effort from the team members would be required to create the software?

**1. Project Size**

* **Lines of Code (LOC):** ~8,000–25,000 (depending on deep learning usage)
* **Files:** ~20–50 files
  + Data Collection (data\_collection.py)
  + Image Preprocessing (image\_preprocessing.py)
  + Feature Extraction (feature\_extraction.py)
  + Model Training (model\_training.py)
  + Model Evaluation (model\_evaluation.py)
  + GUI or Web API (app.py for Flask/FastAPI)
  + Jupyter Notebooks for testing
* **Modules/Libraries:**
  + **Image Processing:** OpenCV, PIL
  + **Machine Learning:** sklearn, xgboost
  + **Deep Learning (if CNN is used):** TensorFlow, PyTorch
  + **Handwriting Feature Extraction:** tesseract-ocr, Handwriting Libraries

**2. Cost Estimation**

* **Team Composition:**
  + 1–2 ML Engineers
  + 1 Data Engineer
  + 1 Backend Developer
  + 1 Researcher (for handwriting traits)
* **Total Cost:**
  + **Small-scale (Student/Academic project):** $2,000–$5,000
  + **Startup-level:** $10,000–$30,000
  + **Enterprise-scale:** $50,000+

**3. Duration**

* **Data Collection & Annotation:** **1–3 months** (handwriting dataset is hard to obtain)
* **Preprocessing & Feature Engineering:** **1–2 months**
* **Model Training & Optimization:** **2–3 months**
* **Deployment & Testing:** **1–2 months**
* **Total Duration:** **~5–10 months** (depending on dataset size & complexity)

**4. Effort Estimation (Person-Months)**

* **Small-scale (4–5 members)** → **8–15 person-months**
* **Medium-scale (6–8 members)** → **12–20 person-months**
* **Large-scale (enterprise)** → **25+ person-months**

Would you like help finding handwriting datasets, feature extraction techniques, or model selection?

COCOMO (Constructive Cost Model) was proposed by Boehm. According to him, there could be three categories of software projects: organic, semidetached, and embedded. The classification is done considering the characteristics of the software, the development team and environment. These product classes typically correspond to application, utility and system programs, respectively. Data processing programs could be considered as application programs. Compilers, linkers, are examples of utility programs. Operating systems, real-time system programs are examples of system programs. One could easily apprehend that it would take much more time and effort to develop an OS than an attendance management system.

The concept of organic, semidetached, and embedded systems are described below.

* **Organic:** A development project is said to be of organic type, if
  + The project deals with developing a well understood application
  + The development team is small
  + The team members have prior experience in working with similar types of projects
* **Semidetached:** A development project can be categorized as semidetached type, if
  + The team consists of some experienced as well as inexperienced staff
  + Team members may have some experience on the type of system to be developed
* **Embedded:** Embedded type of development project are those, which
  + Aims to develop a software strongly related to machine hardware
  + Team size is usually large