Mid Term Report SOC'25

Name: Sayam

Roll No.: 24b2240

Track: Summer of Code '25 – Python & Face Detection

Project: Face Detection Model

Reference Course: CS50P – Introduction to Programming with Python

Institution: HarvardX (via edX)

Course Link: https://learning.edx.org/course/course-v1:HarvardX+CS50P+Python/home

1. Introduction

This report highlights my learning progress in Python and its associated modules, completed as part of my Summer of Code (SOC) project under SoS'25. The core objective of the project is to build a face detection model, and to support this, I studied Python intensively over a short span of 4 weeks using the HarvardX CS50P course as the primary resource.

Although the CS50P course is designed as a 10-week curriculum, I successfully completed it in just 4 weeks and simultaneously explored practical applications in data science, debugging, and computer vision foundations.

2. Course Overview – CS50P (HarvardX)

What is CS50P?
It is a beginner-friendly introduction to Python programming by Harvard University via edX. The course covers fundamental programming topics including:

- Variables and Data Types
- Conditionals and Loops
- Functions
- Object-Oriented Programming
- File I/O
- Exceptions and Debugging
- Regular Expressions
- Unit Testing and Test-Driven Development

Progress:

I completed approximately 45 problems out of the 50 assignments offered. These practical exercises helped reinforce each week's concepts.

3. Python Concepts Covered

During this learning phase, I gained practical and theoretical knowledge of:

- Data types – strings, integers, floats, lists, dictionaries
- Control flow – if-else, for, while
- Functions – user-defined and built-in
- OOP – classes, constructors, inheritance
- File Handling – reading and writing files
- Exception Handling – try-except blocks
- Testing – writing unit tests using pytest
- Debugging – using breakpoints and tracebacks

4. Modules & Libraries

NumPy:
What it is: A powerful library used for numerical computations and handling large multi-dimensional arrays and matrices.
What I learned:
- NumPy arrays and differences from Python lists
- Array slicing, indexing, broadcasting
- Matrix operations

Resources:
- https://numpy.org/doc/stable/user/absolute_beginners.html
- https://numpy.org/doc/stable/user/basics.broadcasting.html
- https://www.youtube.com/watch?v=awP79Yb3NaU
- https://www.youtube.com/watch?v=QUT1VHiLmmI

Pandas:
What it is: A data analysis and manipulation library that makes handling structured data easy and efficient.
What I learned:
- DataFrame and Series structures
- Reading/writing CSV files
- Data filtering, grouping, and statistics
- Handling missing data

Resources:
- https://www.w3schools.com/python/pandas/default.asp

- https://www.youtube.com/watch?v=JjuLJ3Sb_9U

Matplotlib:
What it is: A visualization library for creating static, animated, and interactive plots in Python.
What I learned:
- Line plots, bar graphs, and scatter plots
- Plot styling and labeling
- Multiple plots in one figure

Resources:
- https://www.w3schools.com/python/matplotlib_intro.asp
- https://www.youtube.com/watch?v=9GvnrQv138s

5. GitHub Repository

A dedicated GitHub repository has been created to showcase:

- The Python codes and assignments I practiced during the CS50P course
- Code snippets for NumPy, Pandas, and Matplotlib
- The report document and all reference materials
- YouTube links and module documentation

Note: The README.md file has been updated to reflect the contents and structure of each folder.

6. Key Takeaways

- Accelerated Learning: Completed a 10-week course in 4 weeks by maintaining a disciplined schedule.
- Confidence in Python: Built confidence in writing clean, structured, and optimized code.
- Problem Solving: Enhanced logical thinking through weekly coding challenges.
- Tool Proficiency: Gained experience in using NumPy, Pandas, and Matplotlib – critical for data science and machine learning.

7. Future Plans

- Apply learning to real-world mini projects including face detection, object tracking, and data dashboards
- Explore advanced libraries such as OpenCV (for computer vision), Scikit-learn (for ML), and TensorFlow (for deep learning)
- Contribute to open source and further build my GitHub profile

8. References

- CS50P Course: https://learning.edx.org/course/course-v1:HarvardX+CS50P+Python/home
- MODULES.docx: Learning plan for NumPy, Pandas, Matplotlib
- YouTube Tutorials (Hindi & English)
- Official documentation for all three modules
- GitHub Repository (includes assignments, report, code)