AIM and OBJECTIVES

The aim of this project is to create a web scraping program and compare its performance on serial implementation and parallel implementation through the following parameters:

1. Speedup
2. Efficiency

The deliverables of this project are:

1. Source code
2. Data scraped from web presented in csv format
3. Final presentation video
4. Final report

IMPLEMENTATION

This project was programmed in python language using its SELENIUM library for web scraping and MPI4PY for parallel implementation.

**Selenium** is a powerful tool for controlling web browsers through programs and performing browser automation. Through its array of functions, selenium allows users to create controlled step by step interactions with a webpage and access information from it through user commands. It is compatible on compatible with all major OS and browsers. Even though selenium allows for automation, for this project manual testing has been chosen.

**Mpi4py or MPI** for Python provides Python bindings for the Message Passing Interface (MPI) standard, allowing Python application to exploit multiple processors on workstations, clusters and supercomputers.

**REFERENCES**

1. <https://www.browserstack.com/guide/python-selenium-to-run-web-automation-test>
2. <https://www.geeksforgeeks.org/selenium-python-tutorial/amp/>
3. <https://mpi4py.readthedocs.io/en/stable>