**List of problems:-**

Install miniconda and create a new environment. Install required packages in it.

Packages: python3, pytorch, hdf5, kafka\_client, numpy

**Problem statement 1:**

Create a kafka producer consumer program which reads data from a binary(int) file and dumps it into kafka. Other side consumer reads that data and stores it into a binary file.

**Kafka producer:** Java technology base code

**Kafka consumer:** Python technology

**Message size:** 512 bytes of data should be included into one message.

Include headers in the kafka message (headers: length of data, count: message counter)

**Sample data file:**

Store uint16\_t integer data into a binary file.

“Requirements: Apache kafka installed locally”

**Extension of above problem:**

Create automation of above problem statement through shell script which takes input from command line parameters and handles pulse length value in one message.

**Problem statement 2:**

Python program which takes input as a file contains a list of student names. Program should sort names alphabetically and store it into another file. Also calculate length of each name and store those values into HDF5 file.

NOTE: don’t use inbuilt function like “sort”

Input format: names.txt

Output : Output\_names.txt (Sorted names) , lengths.hdf5 (HDF5 file contains length of each name.)

**Problem Statement 3:**

Create a python program which reads integer data from .txt file

Applies filter on it: remove odd numbers from input and create remaining numbers size numpy 1D array and stores into .hdf5 file.