**Steps to generate balanced ETFD dataset:**

1. Run the ‘cluster\_undersampling.py’ code to generate the balanced ETFD dataset. The ‘cluster\_undersampling.py’ code performs the following operations:
   1. It reads the ‘ETFD\_extracted\_feature.txt’ file from ‘1. Feature extraction’ folder
   2. It performs clustered undersampling and saves the balanced dataset in the ‘ETFD\_balanced.txt’ file