HOMEWORK – 3

Data Intensive Computing

Nikhil Saji

50592329

# Task 1.1

Time taken for each set of sentences

* 10 Sentences:
  + Spark –
  + Multi-process – 0.820 secs
  + For-loop – 0.134 secs
* 50 Sentences:
  + Spark –
  + Multi-process – 2.253 secs
  + For-loop – 5.961 secs
* 100 Sentences:
  + Spark –
  + Multi-process – 7.122 secs
  + For-loop – 25.584 secs
* 250 Sentences:
  + Spark –
  + Multi-process – 38.814 secs
  + For-loop – 152.319 secs
* 500 Sentences:
  + Spark –
  + Multi-process – 149.91 secs
  + For-loop – 598.98 secs
* 1000 Sentences:
  + Spark –
  + Multi-process – 298.2 secs
  + For-loop – 1197 secs

# Task 1.2

Time taken for n\_inputs

* 1000 Inputs:
  + Spark – 0.49 seconds
  + Non-Spark – 0.18 seconds
* 5000 Inputs:
  + Spark – 0.35 seconds
  + Non-Spark – 0.96 seconds
* 10000 Inputs:
  + Spark – 0.31 seconds
  + Non-Spark – 2.8 seconds
* 50000 Inputs:
  + Spark – 0.38 seconds
  + Non-Spark – 350.29 seconds
* 100000 Inputs:
  + Spark – 0.4 seconds
  + Non-Spark – 847.21 seconds