**Pyspark : Enhanced**

**PySpark Coding Environment:**

**1.With Google Colab:**

* Python Colab configurations for pyspark environment:
* Log in to python colab, and create a notebook.
* To install pyspark in it, run the following command.

“!pip3 install pyspark”

* from pyspark import SparkContext, SparkConf
* #initialize the spark.

Conf=SparkConf().setAppName(“test-app-1”).setMaster(“local[3]”)

*(you can put \* instead of no. of cores, this means use all cores that are available.)*

* sc=SparkContext(conf=conf)

sc.defaultParallelism

*(this will show the no. of cores to process the task in parallel)*

So this was about the initial , basic setup for pyspark.

**2.With Desktop Platform:**

This will done using VMware.

(skipping this step here.)

---------------new changes-------------