



VIT[®]
Vellore Institute of Technology
(Deemed to be University under section 3 of UGC Act, 1956)

School of Computer Science and Engineering

J Component report(Comparison)

Programme : MIS (M.Tech Software Engineering)

Course Title : BIG DATA ANALYTICS

Course Code : SWE2011

Slot : E2+TE2

Title:- COVID'19 LOCKDOWN ANALYSIS

Team Members: B DEVI PRASAD | 19MIS1018

CHAITANYA K | 19MIS1099

Faculty: Prof. S.P .Syed Ibrahim

Sign:

Date:

COMPARISON OF RESULTS:

Algorithm	In ML	In PySpark
K-means	2.03070479425353	0.078 sec
Linear and Logistic Regression	1.791163	0.7446808510638298

We have implemented two models in traditional vs big data framework with our dataset containing Countries dataset which contains the Deaths, Recovered , Confirmed cases.

The Algorithms are :

K-means

Logistic Regression

Among the implemented methods in ML and in Bigdata framework the Bigdata is faster than the traditional methods. PySpark is faster than Pandas, and even in the bench marking test, it shows PySpark leading Pandas.