**23-813-0507: Lab 10 – Data Analytics Lab**

**List of Experiments**

1. Install Apache Hadoop
2. Develop a Map Reduce program to calculate the frequency of a given word in a given file.
3. Develop a MapReduce program to find the maximum temperature in each year
4. Develop a MapReduce program to implement matrix multiplication.
5. Develop a program to sort a given input set in PigLatin.
6. Develop a program to calculate the maximum recorded temperature by year wise for the weather dataset in PigLatin.
7. Write queries to create, alter and drop tables using HiveQL
8. Write queries to sort and aggregate the data in a table using HiveQL
9. Install Apache spark.
10. Perform Data cleaning and Data Preprocessing on a standard dataset using pyspark.