MSCS305 BIG DATA ANALYTICS (3-0-0)

MODULE-I (9Hrs.)
Introduction to Big Data: Analytics, Nuances of big data, Value, Issues, Case for Big data, Big data sources, Acquisition, Nuts and Bolts of Big data. Features of Big Data, Security, Compliance, auditing and protection, Evolution of Big data, Best Practices for Big data Analytics, Big data characteristics, Volume, Veracity, Velocity, Variety, Data Appliance and Integration tools.

Module – II (9Hrs.)
Data Analysis : Evolution of analytic scalability, Convergence, parallel processing systems, Cloud computing, grid computing, map reduce, enterprise analytic sand box, analytic data sets, Analytic methods, analytic tools, Cognos, Microstrategy, Pentaho. Analysis approaches, Statistical significance, business approaches, Analytic innovation, Traditional approaches.

Module – III (8 Hrs.)

Stream Computing : Introduction to Streams Concepts, Stream data model and architecture, Stream Computing, Sampling data in a stream, Filtering streams, Counting distinct elements in a stream, Estimating moments, Counting oneness in a window, Decaying window, Real-time Analytics Platform(RTAP) applications, IBM Infosphere, Big data at rest, Infosphere streams, Data stage, Statistical analysis, Intelligent scheduler, Infosphere Streams

Module – IV (10Hrs)
Predictive Analytics and Visualization : Predictive Analytics, Supervised, Unsupervised learning, Neural networks, Kohonen models, Normal, Deviations from normal patterns, Normal behaviours, Expert options ,Variable entry, Mining Frequent itemsets, Market based model, Apriori Algorithm, Handling large data sets in Main memory, Limited Pass algorithm, Counting frequent itemsets in a stream,
Clustering Techniques, Hierarchical, K- Means, Clustering high dimensional data Visualizations, Visual data analysis techniques, interaction techniques; Systems and applications

Text Books:
1. Frank J Ohlhorst, "Big Data Analytics: Turning Big Data into Big Money", Wiley and SAS Business Series
2. Colleen Mccue, "Data Mining and Predictive Analysis: Intelligence Gathering and Crime Analysis", Elsevier
Reference Books:
1. AnandRajaraman and Jeffrey David Ullman, "Mining of Massive Datasets", Cambridge University Press.
2. Jiawei Han, MichelineKamber "Data Mining Concepts and Techniques", Second Edition, Elsevier