



ANDHRA PRADESH STATE COUNCIL OF HIGHER EDUCATION

SKILL COURSE

w.e.f. AY 2023-24

SEMESTER-III

DATA ANALYTICS

Credits: 2

2 hrs/week

Learning Outcomes:

Upon successful completion of the course, the students will be able to

- Understand the framework of big data environment.
- Apply pre-processing techniques that aid in feature selection.
- Classify the data for better understanding.

UNIT - I:

8hrs

Introduction: Overview, Data Science, Big Data Characteristics, Architecture – Core Layers, service layers; roles in data science team, life cycle of data-centric projects, big data life cycle.

UNIT-II:

10hrs

Pre-processing: Introduction, Measures of Central tendency-Mean, Median, Mode, sampling distributions, inferential statistics, ANOVA, feature selection-PCA.

UNIT-III:

12hrs

Methods: Association rules, Apriori algorithm, overview of clustering, k-means algorithm, Regression- Linear, Logistic, Support Vector Machines, Classification- Decision Tree classification, Attribute selection, Naïve Bayes Classification.

Text Books:

1. G. Sudha Sadasivam, R. Thirumahal, “Big Data Analytics”, Oxford University Press.

Reference Books:

1. Paul Zikopoulos, Chris Eaton, “Understanding Big Data Analytics for Enterprise Class Hadoop and Streaming Data”, 1st edition, TMH.

Activities Planned:

1. Identify the roles played by different persons in the team.
2. Understand the phases of big data life cycle.
3. Calculate the central tendency for a given data.
4. Apply Apriori algorithm for generating association rules on a given data.
5. Construct decision tree on a given data for classification.