

CPGET-2023

M.Sc. Data Science

Eligibility Criteria:

Must have obtained B.Sc./B.E./B. Tech. degree with Data Science subject as optional with minimum of 24 Credits.

OR

Must have obtained B.Sc. with Mathematics, Statistics and Computer Science as optional subjects.

78. M.Sc. Data Science Syllabus for Entrance Test - 2023

- 1. Fundamentals of Information Technology:** Data and Information, Acquisition of Numbers and Textual Data, Central Processing Unit, Computer Networks, Input Output Devices, Computer Software, The Software Problem. Programming Principles and Guidelines.
- 2. Problem Solving and Python Programming:** Introduction to Computing and Problem Solving. Introduction to Python Programming. Control Flow Statements. Functions. Strings. Files and exception. Object-Oriented Programming. Functional Programming.
- 3. Data Engineering with Python:** Data Science; Files and Working with Text Data, Working with Text Data. Working with Text Data, Regular Expression Operations, Working with Databases, Working with Tabular Numeric Data (Numpy with Python), Working with Data Series and Frames, Plotting (Plotting with Pandas).
- 4. Machine Learning:** Introduction, Limits of Learning, Geometry and Nearest Neighbours, The Perceptron, Practical Issues, Linear Models, Probabilistic Modeling, Neural Networks, Unsupervised Learning, Association Rules.
- 5. Natural Language Processing:** Language Processing and Python, Accessing Text Corpora and Lexical Resources, Processing Raw Text, Categorizing and Tagging Words, Learning to Classify Text, Deep Learning for NLP, Extracting Information from Text, Analysing Sentence Structure.
- 6. NoSQL Data Bases:** Why NoSQL, Aggregate Data Models, More Details on Data Models, Distribution Models, Consistency, Version Stamps, Map-Reduce, Key-Value Databases, Document Databases, Column-Family Stores, Graph Databases.
- 7. Big Data:** Getting an overview of Big Data, Introducing Technologies for Handling Big Data, Understanding Hadoop Ecosystem, Hadoop Distributed File System, Introducing HBase, Understanding MapReduce Fundamentals and HBase, Understanding Big Data Technology Foundations, Storing Data in Databases and Data Warehouses, NoSQL Data Management.
- 8. Deep Learning:** Introduction to deep learning Algorithms, Neural Networks, Scalars (0D tensors), Vectors (1D tensors), Matrices (2D tensors), 3D tensors and higher-dimensional tensors, Key attributes, Manipulating tensors in Numpy, The notion of data batches, Real-world examples of data tensors, Vector data, Timeseries data or sequence data, Image data, Video data. Tensor operations: Element-wise operations, Broadcasting, Tensor dot. Tensor reshaping, Geometric interpretation of tensor operations, A geometric interpretation of deep learning, Gradient-based optimization, Derivative of a tensor operation, Stochastic gradient descent. Chaining derivatives: the Backpropagation algorithm Neural networks: Anatomy, Layers, Models, Loss functions and optimizers. Introduction to Keras, Keras, TensorFlow, Theano, and CNTK, Recurrent neural networks: A recurrent layer in Keras, Understanding the LSTM and GRU Layers.