Chhatrapati Shahu Ji Maharaj University

Rise from Darkness to Light

Shaping Minds

— Since 1966 —

DEPARTMENT OF INFORMATION TECHNOLOGY
UNIVERSITY INSTITUTE OF ENGINEERING& TECHNOLOGY

To make Data Scientists, Data Engineers and Data Analysts, as well as provide hands-on experience for solving real world data science issues.

Admissions Ongoing

ABOUT COURSE

Data Science (DS) and Machine Learning (ML) are revolutionizing the way everyday world works. As per the World Economic Forum, by 2025, it is estimated that 463 hexabytes of data will be created every day globally. To put this in perspective, this is equivalent to 212,765,957 DVDs per day. This amount of data being generated is helping businesses take informed decisions and optimize the processes. Given the increasing adoption of DS and ML, there is a growing demand for DS talent.

LEVEL OF AWARD

"POST GRADUATE DIPLOMA IN DATA SCIENCE AND MACHINE LEARNING"

PROCEDURE

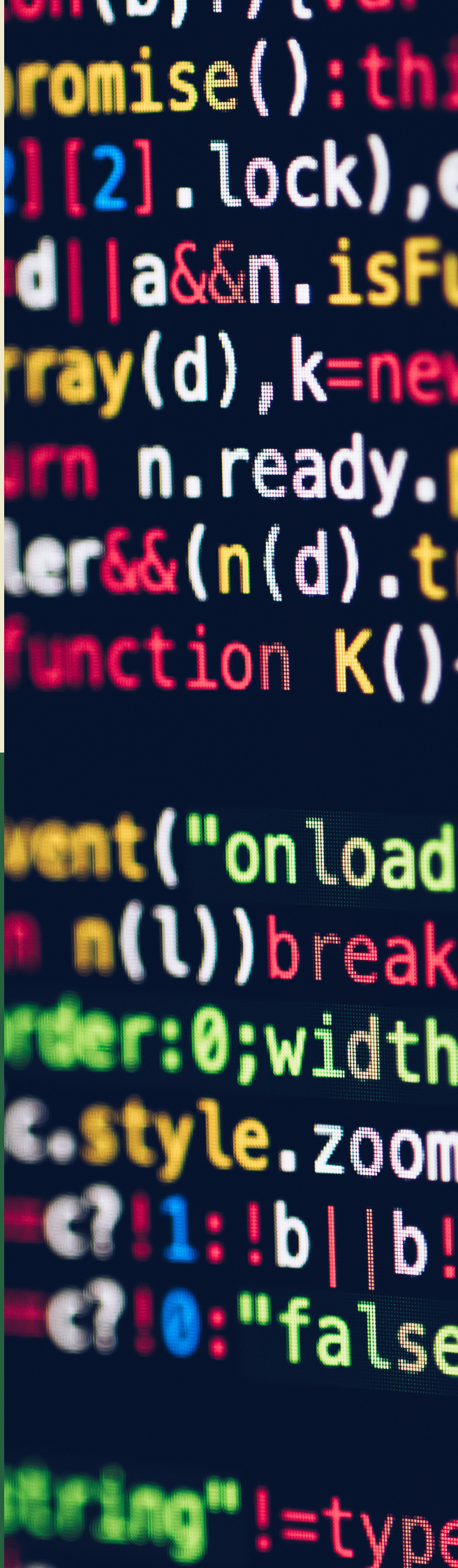
- Admission will be made through entrance exam, as per the rules prescribed by the CSJM University, Kanpur.
- Direct admission will be given to NRI candidates. However, a separate merit list may be prepared in case the number of applicants is more than the number of seats available.

ELIGIBILITY

Bachelor's Degree in Science/Engineering with minimum 50% score

PROJECTS

Includes CAPSTONE Project



CURRICULUM

The curriculum includes lectures, tutorials and practicals. The Curriculum includes one Capstone project, Curriculum incorporates activities like NSS, Sports, eco club etc. as optional.

COMPLETION OF COURSE

The maximum time allowed for completing the program will be 2 years.

COURSE FEE

Semester 1 = Rs. 19000/-
Semester 2 = Rs. 13000/-
Grand Total = Rs. 32,000/-

Syllabus

1 CREDIT = 25 MARKS
TOTAL MARKS = 66 CREDITS = 1650 MARKS

SEMESTER 1 (30 CREDITS)

1. Introduction to Data Science (3-0-4) 6 CREDITS

- Emerging Technologies on AI
- Understanding Data Science and AI
- Fundamentals of Programming
- Foundations of Statistics

Learning Outcome: In this module, you will learn about foundations of Data Science. It is important that you are aware about the tools and technologies that are used for handling data.

2. Mathematics for Data Science and Machine Learning (3-1-0) 4 CREDITS

- Linear Algebra
- Calculus
- Numerical Optimization

Learning Outcome: In this module, you will learn the essential concepts of mathematics which will help you grasp the working of algorithms as you proceed in the programme.

3. Python for DS (0-0-8) 4 CREDITS

- Getting Started with Python
- Hands-on Linear Algebra with NumPy
- Hands-on Data Pre-Processing using Pandas
- Preparation of Times Stamp Intel Distribution of Python

Learning Outcome: Get started with Python Programming for Data Science. You will be learning about the basics of Python and also get introduced to popular Data Science libraries like NumPy, Pandas, and Matplotlib.

4. Introduction to Statistics (3-1-0) 4 CREDITS

- Descriptive Statistics
- Foundations of Probability
- Probability Distributions Inferential Statistics
- Unsupervised Learning: Clusterin
- Case Studies/Hands-on Practice

Learning Outcome: In this module, you will learn about different statistical techniques which can be used to summarise the data. Also, you will learn about generating hypothesis from the data and testing the same.

5. Data Analysis and Visualization (3-0-6) 6 CREDITS

- Sourcing Data from Different Sources
- Data Wrangling
- Working with SQL Primer
- Designing your Own Data for Business Problem
- Exploratory Data Analysis
- Data Visualisation using Tableau

Learning Outcome: Exploring the data and understanding that it is a really important step even before we apply any Machine Learning models. Visualising the data helps to share the insight with all the stakeholders and helps in getting greater insights. You will be using Tableau for visualising the data.

6. Machine Learning (3-0-6) 6 CREDITS

- Python ML Library - Scikit Learn
- Introduction to ML- Types of Learning
- Linear Regression
- Logistic Regression
- k Nearest Neighbors
- Unsupervised Learning : Clustering & Dimensionality Reduction
- Decision Trees
- Support Vector Machines
- Recommender System
- Hands-on Case Studies for ML

Learning Outcome: In this module, you will learn about different Machine Learning algorithms and their applications. You will get hands-on practice with your very first predictions using Machine Learning models.

Summer Training for 6 weeks

SEMESTER 2 (36CREDITS)

7. Text Analytics (3-0-6) 6 CREDITS

- Text Analytics Overview
- Sentiment Analysis on Text Data
- Naïve-Bayes Model for Sentiment Classification
- Document Summarisation
- Topic Modelling
- Hands-on Practice: Text Analytics - Sentiment Analysis

Learning Outcome: Learn to work with the unstructured text data. Learn about how to extract meaningful insights from text data and prepare the data for Machine Learning models.

8. Advance Machine Learning (3-0-6) 6 CREDITS

- Model Tuning
- Overfitting and Regularisation
- Ensemble Models
- Gradient Descent and Stochastic Gradient Descent Algorithms
- Gradient Boosting Machines
- Feature Engineering & Feature Selection Techniques
- Time Series Forecasting

Learning Outcome: Optimising the Machine Learning models is one of the key steps. You will learn about how to tune the hyper parameters of a model and get the desired outcome for a particular business problem.

9. Tensor Flow (0-0-8) 4 CREDITS

- TensorFlow Overview
- TensorFlow 1.X Programming Model
- Tensors
- Computational Graphs
- Sessions
- Linear Algebra with TensorFlow
- TensorFlow 2.0
- Hands-on Exercises with TF 1.x & TF2.0

Learning Outcome: Tensorflow is an end-to-end open source platform for Machine Learning. It provides advantages like easy model building, robust ML production and powerful experimentation. In this module, you will get hands-on experience of programming with Tensorflow

10. Neural Networks (3-0-6) 6 CREDITS

- Introduction to Perceptron
- Perceptron Training
- Deep Neural Networks
- Keras API

Learning Outcome: Get introduced to the concept of a neuron and how multiple neurons can be used to construct an Artificial Neural Network. Deep Learning is a class of Machine Learning algorithms that progressively extract features for better understanding of the problem. You will learn about various Deep Learning models built using Artificial Neural Networks.

11. Elective (3-0-6) 6 CREDITS

12. Capstone Project (0-0-8) 4 CREDITS

1. Creating a Hierarchical Classification Tool for COVID-19 Literature
2. Patch Classification from Image Labels – Healthcare
3. Traffic Sign Recognition

Learning Outcome: End-to-end project to apply the concepts taught in the program.

13. Summer Training (6 weeks from previous semester) 4 CREDITS

List of Electives:

1. Computer Vision and Image Recognition (3-0-6) 6 CREDITS

- Computer Vision with Open CV
- Convolutional Neural Networks (CNN)
- Pre-trained CNN Models
- Image Classification with KERAS
- Object Detection
- Hands-on Practice on Healthcare, Automobile and Retail Analytics

Learning Outcome: One of the popular applications of Deep Learning is in image recognition. You will learn how to build complex image recognition and object detection models and apply them to solve business use cases

2. Speech Recognition (3-0-6) 6 CREDITS

- Overview of Speech Recognition and Basic APIs
- Advanced NLP - using Word Embeddings
- Word2Vec, GLOVE
- Sequence Models to Audio Applications
- Recurrent Neural Networks – RNN
- RNN for Sequence Modelling
- Time Series Forecasting with RNN
- Hands-on Practice on Chatbot Architecture, Building a Chatbot with Dialogflow and Building Alexa Skills

Learning Outcome: Processing the naturally spoken language is one of the complex tasks faced by researchers. In this module, you will learn about Natural Language Processing and how Deep Learning models can be used to build speech recognition applications.

3. Data Engineering (3-0-6) 6 CREDITS

- Introduction to Data Engineering & Big Data
- Working with Data Base
- Connecting 3rd Party Applications to the DBMS i.e., SQL to Python
- Big Data & Bigdata ecosystems
- Hive- ETL
- Hive Pig HBase
- Spark
- Big Data Cluster on Cloud
- Big Data Visualisation

Learning Outcome: Building the data pipelines and deploying the Machine Learning models are some of the important steps in implementing the DS and ML solutions in production. This module will help you learn these tools and techniques

The Campus

Kalyanpur, Kanpur Uttar Pradesh,
India-208024

For Queries:

DR. Rashi Agrawal
HOD, Department of Information Technology,
UIET, CSJM University, Kanpur
Phone: +91 9839223922
Email: Dr.rashiagrawal@gmail.com
csjmu@kanpuruniversity.org