IIM
SHILLONG

Certificate Course in Data Science

Information Booklet

About IIM Shillong

Indian Institute of Management Shillong (IIM Shillong) has been set up with a vision of providing good quality management education and research in the North-Eastern region of India. The decision of coming up with an Indian Institute of Management in the North East was unanimously taken by the Union Minister of Human Resource Development along with the Chief Ministers of the North Eastern States in the Review Meeting held at Shillong during June 2004. Shillong was drafted as the permanent location for the institute after consultation between the Ministry of Development of North Eastern Region (DONER) and the Chief Ministers of States of the region.

Located in the green contours of North Eastern part of the country, IIM Shillong which commenced operations in 2008, remains committed to its goal of excellence in management education and research so as to evolve into a nationally and internationally recognized educational institution. Besides high-quality standards and academic rigor, which are considered as the hallmark of an IIM, in IIM Shillong, there is also an emphasis on the ever-growing significance of sustainable development and business practices. The institute strives to impart to its student's ethical values, compassionate behaviour, and concern for society.

To ensure a holistic management education and to help the graduates grow into innovative leaders of the future, the institute provides opportunities outside the classrooms as well, be it industrial interactions, sports competitions, cultural activities or entrepreneurial pursuits. The quality of instructions, rigorous course curriculum and exposures that the students get at IIM Shillong when coupled with their Summer Internship experience makes them adequately prepared and confident to face the challenges of the corporate world.

Vision: "To become an internationally recognized management institute with a global outlook grounded in Indian values".

Mission: "To generate and disseminate knowledge in all aspects of management education for sustainable development and to develop innovative leaders with strong ethical values".

Core Values:

- Openness to new ideas and experiences
- Intellectual freedom
- Self - Experimentation and Creative pursuit
- Adherence to fair, just and ethical practices
- Compassion

"It is an immense pleasure and pride in introducing to you the **Certificate Course in Data Science** of IIM Shillong. Today, everyone is drowning in the data. But, there is a lack of skilled data scientists who study data in scientific manner to retrieve insights and develop data products that create lasting impact. The Certificate Course in Data Science tries to bridge that gap by producing data scientists. It lays a foundation to an aspiring Data Scientist."

Prof. D.P. Goyal, Director, IIM Shillong

"Data is a very important asset for any industry. Every industry maintains silos of data. Data Science is a scientific study of data to create meaning. Data Scientists are professionals who use scientific methods to create meaning from raw data. Their findings can help businesses grow. The **Certificate Course in Data Science**, a 12-week rigorous training program, is built with the aim of producing such data scientists. "

Dr. Pradeep Kumar Dadabada, Program Lead, Certificate Course in Data Science

About the Program

The Certificate Course in Data Science is designed to prepare aspiring data scientists for entry into the industry. The course has been created after thorough research of industry expectations from the aspirants. This is a rigorous 12-weeks program with over 120+ hours of intensive learning, designed to make the participants industry-ready. The program also aims at giving hands-on experience with toolboxes that any data scientist needs in industry.

Program Outcomes

After completion of this course, the participants are expected to have:

- Ability to translate a business problem to a data science problem, and solve it subsequently
- Ability to suitably apply Machine Learning models in case to case basis
- Ability to handle structured and unstructured data
- Deep knowledge of Machine Learning / Deep Learning / NLP concepts
- Readiness to face data science interviews in industries

Key Highlights

- ❖ Certification from IIM Shillong
- ❖ Weekend classroom session in Guwahati campus of IIM Shillong
- ❖ Case-based interactive sessions
- ❖ India's top faculties from academics and industry
- ❖ Weekend class schedules (4 Sessions on Saturday and 4 Sessions on Sunday)
- ❖ Projects, Lab sessions, Assignments
- ❖ Webinars / meetups with industry experts

Mode of Delivery

The course will be delivered in form of classroom sessions on weekends in IIM Shillong Guwahati campus.

Who should attend?

- Professionals who aspire to join a career in the field of Data Sciences
- Researchers interested to upgrade their skill-sets in the field of Data Science
- Candidates pursuing their graduation and post-graduation in engineering, commerce, management, mathematical sciences, computer application, economics, and physics
- Graduates or equivalent degree holders from any recognized Universities or Institutes

Course Pedagogy

The course will be delivered in the form of lecture sessions, hands-on lab sessions, DIY sessions, and case based discussions. There will be four sessions for each Saturday and Sunday, and each lecture session will be of 1 hour 30 minutes duration.

Assessment

The course will have periodic evaluations. The evaluations will be in form of quizzes, assignments, presentations, mini projects, and other objective/subjective assessments. The period evaluations are aimed at creating a suitable environment for student engagement and for boosting experiential learning process. IIM Shillong will awards certificates to those course participants who successfully complete all the academic requirements and have the necessary attendance criteria (80% attendance).

Program Curriculum

Module 1: Statistics for Data Science (15 hrs)

- Data summaries and descriptive statistics, central tendency, variance, covariance, correlation
- Basic probability: basic idea, expectation, probability calculus, Bayes theorem, conditional probability
- Probability distribution functions—uniform, normal, binomial, chi-square, student's t-distribution, Central limit theorem
- Sampling, measurement, error, random number generation
- Hypothesis testing, A/B testing, confidence intervals, p-values
- ANOVA, t-test
- Linear regression, regularization
- Linear Programming

Module 2: Spreadsheet Modeling and Simulation (15 hrs)

- Introduction to formatting in Excel
- Writing formula
- Writing functions: text, date/time, finance, math and statistics, logical, lookup and reference

- Excel Chart, conditional formatting, data presentation and visualization
- Modeling business scenarios
- What-if analysis, scenario analysis, two-way table, goal seeking
- Monte Carlo Simulation
- Pivot table, pivot chart

Module 3: Data Manipulation (7.5 hrs)

- Basic database terminology
- Loading database in SQL
- SELECT Statement
- INSERT Statement
- UPDATE Statement
- DELETE Statement
- Aggregate functions
- Joins
- Subqueries
- Views

Module 4: Data science Toolkit: R (15 hrs)

- Basic Computations in R
- Data Types in R
- Control Structures in R
- Useful R Packages
- Exploratory Data Analysis in R
- Data Manipulation in R

Module 5: Data science Toolkit: Python (15 hrs)

- Python data types
- Control structures and functions
- Introduction to Numpy
- Operation on Numpy arrays
- Data Visualization with Matplotlib and Seaborn
- Time Series data Analysis
- Forecasting methods

Module 6: Machine Learning (15 hrs)

- Classification methods
- Clustering methods
- Association Mining
- Recommendation Engines

Module 7: Artificial Neural Networks and Deep Learning (15 hrs)

- Neural networks Basics
- Convolutional Neural Networks
- Recurrent Neural Networks
- HyperParameter Tuning
- Transfer learning and Fine-Tuning
- AutoEncoders and Embeddings
- LSTM

Module 8: Natural Language Processing (15 hrs)

- Basic NLP Pipeline
- Text Pre-processing
- Feature engineering on Text data
- POS Tagging
- Named Entity Recognition
- Text classification
- Sentiment analysis

Module 9: Analytics in Business (Non-evaluative) (7.5 hrs)

Capstone project

Course Instructors

	**Dr. Achyanta Kumar Sarmah** is an Assistant Professor in Information Systems and Analytics area. He has more than twenty years of experience in academics. He is graduate in Pure Mathematics and an MCA from Jorhat Engineering College Assam. His teaching and research interests include workflow, FCA, Temporal Algebra and Sustainable Information Systems. He has to his credit many international and national publications.

	**Dr. Basav Roychoudhury** is an Associate Professor in Information Systems and Analytics area. He has more than two decades of experience in academics. He is teaching at IIM Shillong since its very inception. He has earned his Ph.D. in Computer Science and Engineering from Tezpur University. His areas of interest include mobile computing – protocols and security, enterprise systems, and e-Governance. He is also very passionate about business analytics and data modelling. He has been designing and delivering courses on data modelling and analytics for more than a decade. He has also published in acclaimed journals his research work in the area of analytics.
	**Dr. Biplab Bhattacharjee** is an Asst. Professor in Information Systems and Analytics area. He has a cumulative experience of seven years in diverse roles. His teaching interest lies in Management Information Systems, Business Intelligence, Business Analytics, Advanced Tools for Decision Support, Statistical computing in R, Digital Analytics, and Social Media & Web Analytics. He has earned his Ph.D. in Systems and Business Analytics from National Institute of Technology (NIT), Calicut. He has completed his M.Sc. Engineering by Research in the year 2015 and Bachelor of Engineering in the year 2006. He has worked in different data science roles in several companies in Kerala and Bangalore. He is an interdisciplinary researcher and has published extensively in the area of system sciences and data sciences.
	**Dr. Nitin** is an Associate Professor in Information Systems and Analytics area. He worked as Professor Educator in the Department of EECS at University of Cincinnati, OH, and as First Tier Bank Professor in the Peter Kiewit Institute at University of Nebraska at Omaha, NE, USA. He earned his Ph.D. in Computer Science & Engineering from Jaypee University of Information Technology, INDIA and University of Florida (UF), Gainesville, FL, USA under student exchange program. He received B.Engg. in Computer Science & Engineering and M.Engg. in Software Engineering. He is an IBM certified engineer, a Life Member of IAENG, Senior Member of IEEE, ACM & IACSIT and Member of SIAM and ACIS.
	**Dr. Pradeep Kumar Dadabada** is an Asst. Professor in Information Systems and Analytics area. He obtained his Ph.D. in Computer Science from University of Hyderabad and Institute for Development and Research in Banking Technology, Hyderabad, an R&D Center established by RBI. He has publications in peer-reviewed journals and international conference proceedings. Prior to joining IIM, he also worked as Senior Data Scientist at Innominds Software SEZ India Pvt. Ltd., Hyderabad. He also mentored many junior data scientists and analytics aspirants in Hyderabad. For his mentorship, he is recognized as one "The most prominent data science academicians in India for 2019" by Analytics Magazine India. He also worked as a lecturer in RVR&JC College of Engineering, Guntur for 3.5 years. His research interests include data/business analytics, time series forecasting, deep learning and unstructured data mining.

	**Dr. Rohit Joshi** is currently working as an Associate Professor in the Operations Management and Quantitative Techniques area at IIM Shillong. He is a Fulbright fellow. He has done his Postdoctoral Research work at UCLA, the University of California, Los Angeles, USA. He has obtained his Ph.D from Department of Management Studies, Indian Institute of Technology Delhi. He has received Master in Technology from Malviya National Institute of Technology, Jaipur, in Industrial Engineering and his bachelor degree in Mechanical Engineering. He is a Gold Medalist and the University topper in his M.Tech and B.Tech respectively. His areas of interest in consulting and teaching assignments include Supply Chain Management, Operations Management, Six Sigma and Lean Thinking, Business Statistics, Project Management, Value-engineering, Design Thinking and Creative Problem Solving.

Program Fees

Candidates selected for the program needs to pay the fees in one installment.

Tenure	**Deadline**	**Amount**
1st installment	On selection	INR 60,000 (including GST)

Payment Details: The course fees has to be deposited in the following account:

Account Name	RGIIM SHILLONG
Account Number	30276148008
Name of the Bank	State Bank of India
Branch	Laitumkhrah
Branch Address	Lummawrie, Laitumkhrah, Shillong, East Khasi Hills, Meghalaya, Pin-793003
IFSC Code	SBIN0002081

The candidates have to send the fees payment proof along with the scanned application or Filled-in PDF application to the ccd@iimshillong.ac.in

Eligibility and Admission

Click here to Download Application Form

Interested participants can fill-in the details in the application form and send the scanned application or Filled-in PDF to ccd@iimshillong.ac.in on or before 15th May 2020. There are limited seats for the program. Selection is based on fulfilling the course eligibility criteria. Participant's selection will be informed by email.

For any queries contact:

Dr. Pradeep Kumar Dadabada
Program Lead- Certificate Course in Data Science
Indian Institute of Management Shillong
Mayurbhanj Complex, Nongthymmai,
Shillong – 793014, East Khasi Hills District,
Meghalaya, India
Email: pkd@iimshillong.ac.in
Tel- +91 364 2308027, +91 9849555496