



# Master of Technology

Data Science and Analytics

**Indian Institute of Information Technology Allahabad**

An Institute of National Importance by Act of Parliament

# 6 REASONS TO CHOOSE IIIT ALLAHABAD



**High quality education**



**Vibrant student life**



**Excellent placement opportunities**



**Affordable tuition fee**



**State of the art classrooms**



**Rich scientific work**

# ABOUT THE PROGRAM

M. Tech Data Science and Analytics at the Indian Institute of Information Technology - Allahabad, is a four semesters program designed to develop the skills to analyze, interpret and infer from datasets complemented by a solid foundation of mathematical, statistical and information technology skills. This multidisciplinary program, also the first inter-departmental program at IIIT–Allahabad is steered by the Dept. of Applied Sciences and driven by a collaborative effort of faculties from the Dept. of Information Technology and Dept. of Management Studies. The curriculum designed is mindful of the diversity of students, industry demands and recent technology available for data scientists while emphasizing the foundational concepts for data science. The coursework is designed to include a balance of functional knowledge as well as practical learning spread over four semesters covering Mathematics, Statistics, Computer programming and emerging topics such as Big Data Analytics, Deep Learning, and Data Modeling tools to deal with large scale data.



## WHO SHOULD APPLY?

Researchers and professionals who wish to shape their careers as Data Scientists to solve data driven decision making and problem solving in the fields of Information Technology, Applied Mathematics, Statistics, Biological Information System, Medical Informatics, Business Analytics, Data Science and Applied Physics, etc. should consider applying to this two years Master of Technology Program in Data Science and Analytics.

# WHY DATA SCIENCE?

The explosive growth of data coupled with the increasing demands on data management systems has made effective data engineering a necessity for several diverse domains such as healthcare, fundamental sciences, data intensive technology, business and finance. Given the increasing complexity of data, the need for efficient data engineering techniques for mining useful (pertinent) information from the data is paramount. Information is now regarded as an asset, which facilitates decision-making processes in any organization, thereby saving the organization both time as well as money. Consequently, there is an increasing industry demand for specialized professionals in the area of Data Engineering. The current chasm between the number of opportunities and available resources in the domain of data science and analytics is undeniable.

## WHY AT IIIT ALLAHABAD?

As a premier institute of Information Technology and allied areas, IIIT –Allahabad will certainly bring an incredible amount of value addition while offering the above mentioned coursework for training and creating the data analysts and data scientist of future generations. Development of high-throughput data analysis tools is an integral part of current IT trends, where such a course can provide critical contribution towards the need of the nation. Additionally, the development of high performance computational tools which are associated with biomedical engineering domains, biophysics, biomaterial, bio-imaging, market and financial analysis, astrophysics etc. require a thorough understanding of computer science concepts. Further, with the current focus of the Government of India and its funding agencies on innovations in Artificial Intelligence and Big Data Analytics in health, process automation, forecasting, predictive analysis and other ancillary areas, this program within the aegis of an IT institution is an excellent fit.

Currently, very few institutions offer data science as a degree program that spans two years including almost one year research exposure in this field. Data science encompasses acquisition, storage, retrieval, processing and finally the conversion of data into knowledge where the quantum of data is very large. Data Science relies heavily on the continuous aspects of mathematics & statistics. It is different from statistics insofar as there is an intense concentration on efficient algorithms to handle very large data sets, in some cases streamed data, and their realization via different architectures and platforms. There is a growing realization that most of the knowledge lies in the data. For example, lot of cutting edge science (LHC, radio astronomy, genetic medicine, genomics and proteomics, molecular simulation, organic synthesis), commerce (supply chains, e-commerce, and finance), medicine (epidemiology, drug discovery) would be impossible without data science. Consequently, there is a large amount of ongoing work in data science that cuts across disciplines. Given its importance, the need to bring the different disciplinary participants together on a common academic platform is just plain common sense. IIIT-Allahabad with its specialized computational resources and multidisciplinary faculty expertise is uniquely placed to offer this program.



## OBJECTIVE OF THIS PROGRAM

With advent new high throughput technologies across the domain and large scale experimental setups, the volume of data being generated and being made available for researchers is growing exponentially. These “big-data” sets create a challenge for systematic exploration, mining and analysis due to its large volume, variety and high velocity. These datasets need to be analyzed and mined for the (Knowledge Discovery in Databases) KDD and inference making processes, which has led to the need of Data Scientists/ Data Analyst in almost every domain of science and technology. The main objective of this program is to develop manpower for data and information management opportunities. This will be achieved by an approach involving rigorous and comprehensive academic course work covering theory, fundamentals and hands-on experience with real-world applications through Industrial collaboration. Data Science is gaining prominence in academia and industry. Introduction to this program opens up several new opportunities in science and engineering research and Data Scientist positions are new roles where industry is looking for trained resources. This area offers immense potential and demands for trained manpower with solid foundation of Mathematics, Statistics, and Computer programming and skill sets of Big Data Analytics, Machine Learning and Data Modeling. To address this growing demand, the Dept of Applied Sciences is offering this M. Tech Program in Data Science and Analytics. The program is also backed by the active support from Faculty Members associated with Centre of Cognitive Computing at IIIT Allahabad. This is the first Inter-Departmental Program (IDP), involving Dept. of Management Studies and Dept. of Information Technology.

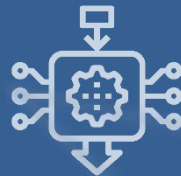
# WHAT IS THE EDUCATION DELIVERY METHODOLOGY?

Each subject is divided into two components each carrying two units of L: Lecture, T: Tutorials and P: Practice (LTP).

The program emphasizes on Experiential Learning that allows learners to apply classroom concepts in simulated and real world situations. This is achieved through:

## TOOLS & TECHNOLOGIES:

Apache Spark, Apache Storm for Big Data Systems/ Real time Processing, TensorFlow for Deep Learning; Python for data processing, machine learning, data visualization etc.



## CASE STUDIES AND ASSIGNMENTS:

Carefully chosen real-world cases & assignments are both discussed and used as problem-solving exercises during the program.

## CAPSTONE PROJECT:

Various capstone projects floated by Industrial partners will help the learner to utilize their skill sets into problem solving and analytical thinking, while using the tools and techniques to interpret and infer from the data.



## DISSERTATION/ PROJECT WORK:

The third semester and fourth semester offers an opportunity for learners to apply their knowledge gained to a real world like complex project. The learner is expected to demonstrate understanding of vital principles and their ability to successfully apply these concepts.





## CONTINUOUS ASSESSMENT

IIIT Allahabad follows Credit Based Continuous Assessment and Award System. (Choice & Credit Linked Continuous Assessment and Award: CCLCAA). The Continuous assessment takes into account all academic activities undertaken by the student during each academic semester, in which a student is enrolled. This entails activities ensuring continuous assessments such as conduct of Review Test(s), Quizzes, Tutorials, Assignments, Practical assessments, Quizzes, Project Work, Term Work, Seminars / Discussions, etc.

Every coursework is divided into two Components: C1 and C2, each with two units. While the Review Tests would be held during the Semester at the end of said two component C1 & C2 periods respectively, C3 the End Semester Exams shall be scheduled by the (Assessment And Awards) AAA Section of the Institute. Faculty Member(s) shall declare the schedule of Quizzes, Tutorials, Assignments, Seminars, discussions, etc. Certain unannounced assignments may also make a part of the overall assessment. However, students shall be made aware of the assessment modalities that are going to be followed in a course by the faculty.

**More at <https://aaa.iiita.ac.in/>**

Mode of teaching will be offline, online, and in few cases as hybrid model. This with ensure that the learners get the best inputs from not only the highly qualifies faculty members at IIIT Allahabad but also the other national and international experts from Industry and Academia.



# LEARNING OUTCOMES

By the end of the course, candidates are expected to be proficient in Programming and Statistical Techniques for Data Science along with Data scraping and Data wrangling.

It will also make them competent in Big Data Technologies, Cloud Computing, Deep Learning and Advance AI Applications and able to handle diverse Data Science domains.

Four specializations basket that combine to provide a wide learning and application spectrum shall be provided to facilitate the candidate in selecting their dominant interest:

1. Financial Services Analytics/ Marketing Analytics/ Supply Chain Analytics.
2. Biological Data Analytics, Systems Modeling and forecasting of biological systems
3. Unstructured Data Analysis/ Process Automation
4. Cognitive Computing

The output of this M. Tech. Data Science and Analytics program will be skilled manpower which can cater to the needs of the country in various data intensive areas like health informatics, business informatics, forecasting and automation. This program will definitely boost the development and creation of future knowledge creators and serve as a valuable pipeline between the Institute and existing data intensive industries and operation sector.

Government of India and private sectors are providing lots of endorsement and thrust in the areas of Data Science and Analytics. At the end of the program, the student will find themselves identifying cutting edge research opportunities at various research laboratories as well as several industries which are currently involved in this field. Almost 2.5 quintillion bytes of data is getting generated every day. Additionally, about 90 % of traceable data has been generated in just last 5 years! With a growth rate of 40 % compounded annually, data science and analytics has become one of the most demanding technologies in recent years. Almost all IT, Pharmaceutical, Financial forecasting firms are currently working in the area of analytics and the market of Data Analytics is anticipated to grow at an exponential rate. This gives the prospective students of the program a very bright future and promising placement scenario.



# PAST RECRUITERS

<b>A</b>	Aakash Accenture Accolite Adobe AirTight Amazon AMD Amdocs American Express Apigee Aptech Aquamarine ArachnoMesh Arcesium Arista Networks Avanti Axtria	<b>D</b>	D.E. Shaw DevFactory Digital Green Direct-i DRDL	<b>I</b>	IBM Icon Resources ION Trading Incture Technology Info Edge Infosys Infurnia Furnishings Inmodi Ittiam IVP
<b>B</b>	Barkley BDL Belzabar Bharat Bhasha Bidgely BlackRock BlueJeans Bluestacks BTU	<b>E</b>	Edelweiss Expedia	<b>J</b>	JDA Jigserv Juniper
<b>C</b>	Capgemini Cisco Systems India CitiCorp CMU Corlax Coupon Dunia Curofy CXC Codenation	<b>F</b>	FactSet Flaberry Flipkart Fractal Freescale Frrole Futures First	<b>K</b>	KPMG Kritikal Kronos Kuliza
		<b>G</b>	Geek Shastra Germin8 GE Digital Google Goldman Sachs Grofers	<b>L</b>	LeafAir LinkedIn
		<b>H</b>	HackerRank HashedIn HCL-CDC HealthKart HeroMoto Hike Holidify Housing HT Media	<b>M</b>	MakeMyTrip Mahindra Comviva MAQ Software Microsoft MoonFrog Morgan Stanley Mu Sigma Must See India Myntra

<b>N</b>	Nagarra Softwares Naukri.com NEC Tech NeoGrowth NewGen National Instruments Niki.ai (Techbins) Nucleas Software Nvidia	<b>S</b>	Samsung SAP Labs Sapient Nitro Simplify Shopclues Smartpix Snapdeal SportsWave Streamoid Success Factors Sutra Analysis Sutra Swiggy	<b>W</b>	Walmart Labs WhistleApp Wipro WorksApp Wrigg
<b>O</b>	OFSS OkuTech Online Tyari Oxygen	<b>T</b>	TCS Tesco Tech Racers Thorogood Times Internet Tiny Owl Toppr	<b>Y</b>	Yodlee
<b>P</b>	PagalGuy Pay U PayTM Polaris Power2SME Practo Proptiger	<b>U</b>	Unbxd Urbanresto	<b>Z</b>	Zendrive Zomato Zoomcar ZopHop ZopNow ZS ZScalar
<b>Q</b>	Qualcomm Quantum Qubole	<b>V</b>	Vassar Labs Verizon Vitara VIZ		
<b>R</b>	RB RBS Routofy				

# ADMISSION PROCEDURE

True to the multi-disciplinary nature of Data Science and Analytics, the potential students for this graduate program will be those with engineering undergraduate or equivalent degree.

Candidates with a Four Years Bachelor's degree in any branch of Engineering/Technology from educational Institutions approved by AICTE/Government, Four year B.Sc. or equivalent in appropriate discipline (with first class or 60% marks (55% for SC/ST) or a minimum Cumulative Grade Point Average (CGPA)/Cumulative Performance Index (CPI) of 6.5 (6.0 for SC/ST) on the scale of 10 in qualifying degrees).

**OR**

Master's degree in Physics/Mathematics/Life Sciences and related subjects from educational Institutions approved by UGC/Government may apply. Students having MBBS and MBA degree can also apply for admission to this program. (With first class or 60% marks (55% for SC/ST) or a minimum Cumulative Grade Point Average (CGPA)/Cumulative Performance Index (CPI) of 6.5 (6.0 for SC/ST) on the scale of 10 in qualifying degrees).

The initial intake of this program is 20 students and split between two channels:

1. **Funded Applicants:** Candidates who wish to avail a fellowship from the institute must possess a valid GATE<sup>1</sup> score and also need to successfully qualify the entrance test<sup>2</sup> conducted by IIIT-Allahabad.
2. **Self-Financed Candidates:** Candidates will need to successfully qualify the entrance test<sup>2</sup> conducted by IIIT-Allahabad.

## <sup>1</sup>**GATE**

A valid Graduate Aptitude Test in Engineering (GATE) score in CS, EC, EE, IN, MA, ST, BM, PH or XE is a must. {Instrumentation Engineering(IN), Mathematics(MA), Computer Science and Information Technology(CS), Electronics and Communication Engineering(EC), Electrical Engineering(EI), Engineering Sciences(XE), Statistics(ST), Biomedical(BM), Physics(PH)}

## <sup>2</sup>**Entrance Test**

An entrance test will be conducted through a proctored ONLINE test. The candidate needs to have their own laptop/desktop/tab with a working webcam for the same. This test will be proctored and the candidates will be intimated well in advance about the modalities for the test. Selected candidates will be called for an online personal interactive session. The list of successful candidates will be put up on the program website. (<https://dsa.iiita.ac.in>).



## FEE STRUCTURE

<b>Application Fee:</b> INR 500.00 (One Time) OR INR 300.00 (ST/SC)
<b>Admission Fee, ID Card Charges, Enrollment Fee:</b> 1000.00 (One Time)
<b>Semester Tuition Fee:</b> INR 75000.00 (Per semester)
<b>Semester Exam fee:</b> INR 7000.00 (Per Semester)

## HOW TO APPLY

Interested candidates will have to apply online through the application portal of IIIT-Allahabad (<https://apply.iiita.ac.in/application/authenticate/mtech/>). They must register at the portal by providing a valid email address and a valid cellular number. Once registered – they must also pay an application fee (through the payment gateway) and record the payment transaction number while completing their online application. Upon successful submission of their application, they shall receive an electronic copy of their application at the registered email for their reference and records. Additional details for the application procedure are available at the program webpage (<https://dsa.iiita.ac.in>).

Please check the Dates and other details about Admission at <https://dsa.iiita.ac.in>

# COURSE CURRICULUM

2 Year M. Tech. Data Science and Analytics Course Curriculum (Total Credit = 64)

Sl. No.	First Semester	Credits
1	Statistical Foundations for Data Science (L2,T1,P1)	4 Credit Hours each 20
2	Mathematical Foundation for Data Science (L2,T1,P1)	
3	Advance Programming (L2,T1,P1)	
4	Fundamental of Data Analytics (L2,T1,P1)	
5	Optimization and Bio-inspired Computing (L2,T1,P1)	
Sl. No.	Second Semester	Credits
1	Information Systems (L2,T1,P1)	4 Credit Hours each 20
2	Big Data for Data Science (L2,T1,P1)	
3	Machine Learning with Large Dataset (L2,T1,P1)	
4	Exploratory Data Analysis (L2,T1,P1)	
5	Knowledge Engineering (L2,T1,P1)	
Sl. No.	Third Semester	Credits
Elective 1 & Elective 2	<b>Two Electives from any of the three Pools with hands-on with specializations</b> <b>Pool-1</b> -Biological Information Processing <b>Pool-2</b> -Social Networks/e-commerce and processes <b>Pool-3</b> -Finance, Strategy and Planning <b>Pool-4</b> Cognitive Computing	4 Credit Hours each + 4 Credit Hours Minor Project 12
Minor Project	Mini-Project / Capstone Project	
Sl. No.	Final Semester	Credits
Major Project	Major Project with Industrial Collaboration	12 Credit Hours 12

## LIST OF ELECTIVES:

IOT, Information Retrieval, Natural Language Processing, Deep Learning, Time series analysis and forecasting, Data Warehousing, Probabilistic Graphical Models, Astroinformatics, Systems Biology, Medical Imaging, Material Informatics, Climate Informatics, Social Network, Human Inspired Computing, Business Analytics, Marketing Intelligence.

# Other electives will be floated with inputs from Industries.



# Indian Institute of Information Technology Allahabad

An Institute of National Importance by Act of Parliament

Devghat, Jhalwa, Prayagraj-211015, U. P. INDIA



<https://dsa.iiita.ac.in>



[dsa@iiita.ac.in](mailto:dsa@iiita.ac.in)



+91-532-2922197