

M.Tech.

**DATA SCIENCE
& ENGINEERING**



BITS Pilani
Pilani | Dubai | Goa | Hyderabad



Prepare for a career in Data Science with the most comprehensive masters degree programme in Data Science & Engineering without taking a break from your career.

M.Tech. Data Science and Engineering is a four-semester programme designed to help working professionals build mathematical and engineering skills required to advance their career as a Data Scientist or Data Engineer.

The programme features live online lectures conducted on weekends by BITS Pilani faculty.



[Apply Now](#)



WHO SHOULD APPLY?

Software and IT professionals working as Software Engineer, Programmer, Software Test Engineer, Support Engineer, Data Analyst, Business Analyst, who wish to transition to roles such as Data Scientist or Data Engineer should consider applying to this programme.

WHAT ARE THE MAIN HIGHLIGHTS OF THE PROGRAMME?

- ▶ Learn without a career break with live online lectures conducted on weekends by BITS Pilani faculty*
- ▶ The curriculum covers areas that prepare you for the most lucrative careers in the space of Data Science, Data Engineering and Advanced Analytics. It helps learners master critical skills such as Mathematical modeling, Machine learning, Artificial Intelligence, Product development and scripting languages.
- ▶ Benefit from Case Studies, Simulations, Virtual Labs & Remote Labs that allow learners to apply concepts to simulated and real-world situations. Tools and Technologies covered include Apache Spark, Apache Storm for Big Data Systems/ Real time Processing, Tableau for data visualization, Tensorflow for Deep Learning and various packages within Python for data processing, machine learning, data visualization.
- ▶ In addition to live weekly online lectures, supplementary online contact sessions comprising of tutorials, doubt-clearing sessions, and industry talks will also be conducted periodically.
- ▶ The Dissertation (Project Work) in the final semester enables students to apply concepts and techniques learnt during the programme.
- ▶ Semesters 1st, 2nd, and 3rd cover four courses each. The 4th semester covers Dissertation/ Project Work.
- ▶ The programme uses a Continuous Evaluation System that assesses the learners over convenient and regular intervals. Such a system provides timely and frequent feedback and helps busy working professionals stay on course with the programme.
- ▶ Submit fee using easy-EMI option.

[Apply Now](#)

*Classes are conducted by a pool of faculty members comprising of academicians from BITS Pilani, and guest faculty who are experienced industry professionals.



WHAT ARE THE PROGRAMME OBJECTIVES?

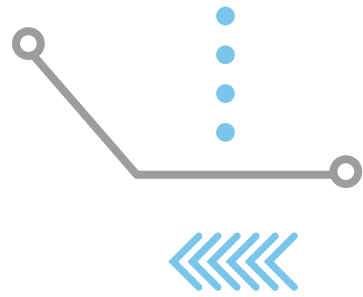
The most lucrative jobs in the areas of Data Science, Data Engineering and Advanced Analytics go to professionals who have mastered a combinations of critical skills such as Mathematical modeling, Machine learning, Artificial Intelligence, Product development and Scripting languages.

The programme aims to help participants build a solid foundation in these areas by developing skills in:

- 
- 
- ▶ Mathematical and Statistical modelling using concepts such as linear algebra and probability to model and solve physical problems
 - ▶ Data structures and algorithms and managing time and space-related complexities
 - ▶ Computer organisation, architecture and Operating systems and advanced techniques for data processing
 - ▶ Data Mining aspects including pre-processing, cleaning & classification, and Data engineering & processing through distributed computing and cloud computing
 - ▶ Advanced computing and analytical skills in areas such as Machine Learning, Artificial Intelligence, Deep Learning and Natural Learning Processing

[Apply Now](#)

WHAT IS THE EDUCATION DELIVERY METHODOLOGY?



EXPERIENTIAL LEARNING

The programme emphasises on Experiential Learning that allows learners to apply concepts learnt in classroom in simulated and real work situations. This is achieved through:



Tools & Technologies: Apache Spark, Apache Storm for Big Data Systems/ Real time Processing; Tableau for data visualisation; Tensorflow for Deep Learning; Various Packages within 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 programme



Dissertation/ Project Work: The fourth semester offers an opportunity for learners to apply their knowledge gained during the programme to a real world like complex project. The learner is expected to demonstrate understanding of vital principles learnt across semesters and their ability to successfully apply these concepts.



ATTEND ONLINE LECTURES OVER WEEKENDS

Lectures are conducted live via online classes. These lectures can be attended via the internet using a computer from any location. These online classrooms offer similar levels of interactivity as regular classrooms at the BITS Pilani campus.

Classes for students admitted during the period Jan - Mar 2020 will begin in Mar 2020. The class schedule is announced within 1 week of completion of the admission process.

The online lectures are conducted usually over weekends for a total of 7-8 hours per week. If you miss a lecture, you can also access the recorded lecture on the internet.

Lectures are conducted on Sat/Sun as per Indian Standard Time.

[Apply Now](#)



CONTINUOUS ASSESSMENT

The learners' performance is assessed continuously throughout the semester using various tools such as quiz, assignments, mid-semester and comprehensive exams. The assessment results are shared with the learners to improve their performance. Each course will entail a minimum of 1 Assignment/ Quiz, a Mid-semester exam and a final Comprehensive exam. Your semester calendar will clearly indicate the dates of the Mid-semester and Comprehensive exam. Typically, a Mid-semester or Comprehensive examination for a course is for 2-3 hours duration. The examinations are typically conducted over a weekend, i.e. Saturday and Sunday.



EXAMINATION CENTERS

Mid-Semester and Comprehensive Exams are conducted at Bengaluru, Chennai, Delhi, Goa, Hyderabad, Kolkata, Mumbai, Pilani, Pune, and Dubai. Participants must appear in-person to take these exams.



SUPPLEMENTARY LEARNING

In addition to live weekly online lectures, supplementary live online sessions will be organised periodically comprising of tutorials, doubt-clearing interactions, and industry talks (18-20 hours per semester).

[Apply Now](#)



WHAT IS THE ELIGIBILITY CRITERIA?

Minimum eligibility to apply - Employed professionals holding B.E. / B.Tech. / MCA / M.Sc. or equivalent with at least 60% aggregate marks or more in their qualifying exam, and minimum one year relevant work experience are eligible to apply.

Applicants should possess basic programming knowledge and adequate background in Mathematics.

FEE STRUCTURE

Application Fees
(one time) :INR 1,500

Admission Fees
(one time) :INR 16,500

Semester Fees
(per semester) :INR 55,000

EMI OPTION

Instant EMI option is now available that allows you to pay programme fee in an easy and convenient way.

- Instant online approval in seconds
- No Credit Cards/ CIBIL score required
- Easy & Secure online process using Aadhaar and PAN number
- Anyone with a Salary Account with Netbanking can apply
- Special interest rate for BITS Pilani WILP applicants
- Pay fee in easy installments of INR 12,298 p.m. for the First Semester, and only INR 9,460 p.m. for subsequent Semesters

[Click here](#) to learn more about EMI option

Apply Now

SEMESTER-WISE PATTERN

The programme features 12 courses between Semester 1-3, and a Dissertation in Semester 4. All the courses will be offered using live online mode.

First Semester

- Data Mining
- Mathematical Foundations for Data Science
- Data Structures and Algorithms Design
- Computer Organization and Systems Software

Second Semester

- Introduction to Statistical Methods
- Introduction to Data Science
- Machine Learning
- Elective 1

Third Semester

- Elective 2
- Elective 3
- Elective 4
- Elective 5

Fourth Semester

- Dissertation

Electives

- Deep Learning
- Natural Language Processing
- Artificial-Intelligence
- Real time-Analytics
- Data Visualisation
- Graphs - Algorithms & Mining
- Optimization Methods for Analytics
- Big Data Systems
- Advanced Topics in Data Processing
- Information Retrieval
- Probabilistic Graphical Models
- Data Warehousing
- Systems for Data Analytics
- Ethics for Data Science

Apply Now

Choice of Electives is made available to enrolled students at the beginning of each semester. Students' choice will be taken as one of the factors while deciding on the Electives offered. However, Electives finally offered will be at the discretion of the Institute



COURSE WISE SYLLABUS

Introduction to Data Science

Learn about:

- The need for data science, with emphasis on data
- Visualization and ethics aspects involved in data science and engineering processes
- Various applications of data science

Topics

- Introduction to data science - definition, motivation; Data science process; Data aspects - types (structured/unstructured/semi structured), sources, formats (csv,tsv,txt etc.), representation; Data pre-processing; Data visualization; Data Science Applications; Ethics - privacy/legal aspects/rightful use of data science, processes, data scientists etc.

Mathematical Foundations for Data Science

Learn about:

- Concepts in linear algebra and use it as a platform to model physical problems
- Analytical and numerical solutions of linear equations
- Mathematical structures, concepts and notations used in discrete mathematics

Topics

- Vector and matrix algebra, systems of linear algebraic equations and their solutions; Eigenvalues, eigenvectors and diagonalization of matrices, Sets and relations; Graphs and digraphs Trees, lists and their uses; Partially ordered sets and lattices

Introduction to Statistical Methods

Learn about:

- Basic and some advanced concepts of probability and statistics
- Concepts of statistics in solving problems arising in data science

Topics

- Different types of data; Data Visualization; Data summarization methods; Tables, Graphs, Charts, Histograms, Frequency distributions, Relative frequency measures of central tendency and dispersion; Box Plot; Chebychev's Inequality on relationship between the mean and the standard deviation of a probability distribution; Basic probability concepts, Conditional probability, Bayes Theorem, Probability distributions, Continuous and discrete distributions, Transformation of random variables, Moments, Correlation and Covariance, Parameter Estimation, Hypothesis Testing

[Apply Now](#)

Data Structures and Algorithms Design

Learn about:

- Applications of basic and advanced data structures & algorithms
- How to determine the space and time complexities of various algorithms
- Identifying and choosing the relevant data structures and algorithms for a given problem and justifying the time and space complexities involved

Topics

- Stack, Queue, Hash Tables, Graphs, Bloom Filters, Trees, Sets, Dictionary, Dynamic Graphs; Divide-and-conquer, Dynamic Programming, Graph Algorithms, Greedy Algorithms - Spanning Tree, Amortized Analysis, Huffman encoding; Page Rank, Map Reduce (map, fork, join etc); Complexity analysis

Computer Organization & Software Systems

Learn about:

- Computer organization, architecture aspects and operating system concepts
- Advanced systems and techniques used for data processing

Topics

- Introduction to computer organization, architecture; Introduction to operating systems; ISA - RISC etc. Processes and Threads; Scheduling, Concurrency; Memory Management



Systems for Data Analytics

Learn about:

- The fundamentals of data engineering; Basics of systems and techniques for data processing - comprising of relevant database, cloud computing and distributed computing concepts

Topics

- Fundamentals of data engineering - data engineering vs data science; Data processing concepts - partitioning, replication, grouping and sorting, data locality; Flynn's taxonomy; Task vs Data Parallelism; Databases, parallel vs distributed databases, architecture - performance; Distributed computing architecture; Processing frameworks - batch, map-reduce, stream processing, parallel processing, real time processing; Cloud fundamentals - virtualization, batch-transactional-continuous workloads, execution models and examples- AWS, Azure etc.



[Apply Now](#)



Data Mining

Learn about:

- Various data pre-processing, cleaning
- Association rule mining, classification, clustering techniques

Topics

- Introduction (Data Mining definitions, Data Mining activities, DM process, DM challenges); Understanding Data (Various Data Types/ Attribute Types, Basic Statistical Descriptions of Data); Measuring Data Similarity and Dissimilarity; Data Pre-processing - Overview; Association Rule Mining - (Apriori Algorithm - FP Growth - generating association rules - efficiency concerns - evaluating patterns for interestingness - Applications; Classification, Linear/Non Linear Regression - Decision Trees, Naive Bayes; Clustering Algorithms: Partitioning Methods, Density Based Approaches to classification, Hierarchical Clustering

Machine Learning

Learn about:

- Basic concepts and techniques of Machine Learning.
- Using recent machine learning software for solving practical problems.
- How to do independent study and research in the field of Machine Learning

Topics

- Introduction to Machine Learning
Supervised Learning - Parametric Methods
Maximum Likelihood Estimation - Evaluate the Estimators - Bias - Variance;
Parametric Classification, Parametric Regression -Tuning Model Complexity and Regularization - Multivariate Methods;
Clustering -Mixture Model -K-Means
Clustering -Expectation- Maximization Algorithm - Applications I; Non - Parametric Methods - Applications; Decision Trees - Linear Discrimination; Linear Discriminants, Basis Functions - Two Classes - Multiple Classes - Pairwise Separation - Gradient Descent Parameter Learning - Logistic Discrimination; Multilayer Perceptrons; Kernel Methods - SVM; Hidden Markov Models

Data Visualization

Learn about:

- Design principles, human perception and effective story telling with data
- Modern visualization tools and techniques

[Apply Now](#)

Topics

- Information overload and issues in decision making; Design of visual encoding schemes to improve comprehension of data and their use in decision making; Presentation and visualization of data for effective communication; Elementary graphics programming, charts, graphs, animations, user interactivity, hierarchical layouts, and techniques for visualization of high dimensional data & discovered patterns

Ethics for Data Science

Learn about:

- Challenges of data privacy
- Data policies for maintaining the privacy of data
- Data Privacy

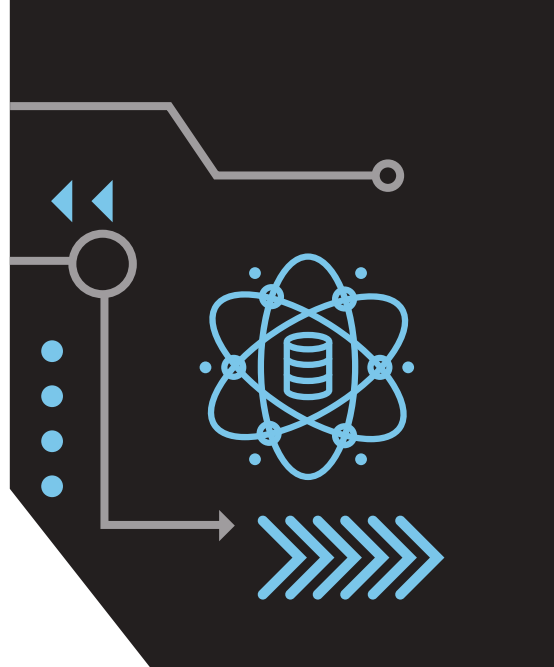
Topics

- Introduction to data ethics, perils of big data, foundations of data privacy, challenges of privacy in the digital age, data policies, consent and fair usage

Graphs - Algorithms and Mining

Learn about:

- Concepts of graph theory so as to understand
- How graph theory concepts are used in different contexts, ranging from puzzles and games to social sciences/ engineering/ computer science
- Model problems in real world using graphs
- Applying mining algorithms to get information from graph structures



Topics

- Basic concepts of graphs and digraphs connectivity, reachability and vulnerability; Trees, tournaments and matroids; Planarity Routing and matching problems; Representations; Various algorithms applications, Introduction to graph mining, Graph Pattern Mining, Graph Classification, Graph Compression, graph model, graph dynamics, social network analysis, visualization, summarization, graph clustering, link analysis, applications of graph patterns

Optimization Methods for Analytics

Learn about:

- Applying linear programming techniques to complex business problems across various functional areas including finance, economics, operations, marketing and decision making
- Implementing optimization techniques to business and industrial problems

[Apply Now](#)

Topics

- Solving various problems related to planning, production, transportation, Microeconomics, etc. using LP models; Decision making in the context of multi stage LP models; Application of Goal Programming (GP) and Analytic Hierarchy Process (AHP) for decisions relating to large teams and complex problems with long term implications; Quadratic Programming, convex Optimizations

Big Data Systems

Learn about:

- Concepts related to big data and its processing
- Applying the concepts of storage, retrieval, interfaces and processing frameworks to a given problem and design solutions for the same by choosing the relevant ones

Topics

- What is big data - are existing systems sufficient; Data Warehouse v/s Data Lakes; Hadoop - Components; Storage - Relational DBs/ NoSQL dbs / HDFS / HBase / Object Data stores - S3; Serialization; Interfaces - Hive/ Pig; Stream Processing; Spark; Mahout

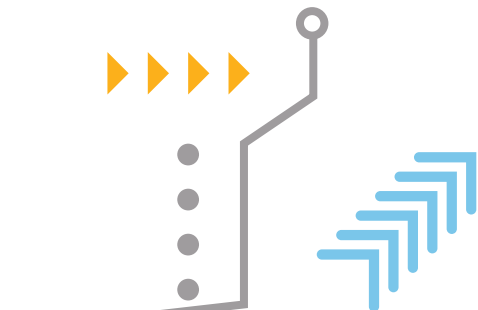
Advanced Topics in Data Processing

Learn about:

- Advanced strategies for data processing
- The relationship between the scale of data and the systems used to process it
- The importance of scalability of algorithms as the size of datasets increase

Topics

- Distributed software architectures, runtime and storage strategies; Consistency in Large Distributed Systems: Paxos consensus, CAP theorem; Massive Parallel Programming models; HPC, wsc (work station clusters), mpp"; Algorithms for Large Datasets: Time-accuracy trade-offs, scalable algorithms; Case Study: Online Social Networks (Distributed social network data storage, Influence diffusion and maximization, Rumor blocking, Information source detection)



Apply Now



Information Retrieval

Learn about:

- Structure and organization of various components of an IR system
- Information representation models, term scoring mechanisms, etc. in the complete search system
- Architecture of search engines, crawlers and the web search
- Cross lingual retrieval and multimedia information retrieval

Topics

- Basic IR concepts; Vector Space model; Text mining; Web search; Cross lingual retrieval; Multimedia IR; Recommender systems

Deep Learning

Learn about:

- Learn about deep learning techniques, constructing deep network structures specific to applications and tuning for parameters

Topics

- Common Architectural Principles of Deep Networks; Building Blocks of Deep Networks; Convolutional Neural Networks (CNNs); Recurrent Neural Networks; Recursive Neural Networks; Building Deep Networks with ND4J; Applications to Sequence Data; Anomaly Detection; Tuning Deep Networks; Vectorization; Data Mining (Pre-requisites)

Natural Language Processing

Learn about:

- Learn about natural language processing techniques such Parts-of-Speech tagging, syntactic and semantic modelling of languages

Topics

- Language modelling with N-gram; Spelling correction; Neural networks and neural language models; Parts-of-Speech tagging; Syntactic parsing; Language semantics; Computational semantics



[Apply Now](#)



Artificial Intelligence

Learn about:

- Classic AI Techniques

Topics

- Intelligent Agents; Search Techniques; Adversarial Search; Logical Agents; CSP; Inference on FOL; Classical Planning; Knowledge Representation (Logics; Probabilistic); Exact and Approximate Inference; Reasoning over time
- Introduction to Learning

Real time analytics

Learn about:

- Processing frameworks for real time analytics, and Analytics techniques for real time streaming data

Topics

- Real Time Streaming Architecture; Service Configuration and Coordination; Data Flow Management, Storing and Processing Streaming Data; Visualization Techniques for Real Time Streaming Data; Aggregation (Timed Counting, Multi Resolution Time Series Aggregation); Statistical Approximation; Approximating with sketches

Probabilistic Graphical Models

Learn about:

- Representation, learning and reasoning techniques for graphical models

Topics

- HMM; Markov Random Field; Bayesian networks; Representation; Learning Inference; Dynamic Bayesian Networks and Temporal Bayesian networks applications

Data Warehousing – Elective

Learn about:

- Concepts needed to design, develop, and maintain a data warehouse
- End user access tools like OLAP and reporting

Topics

- Introduction to Data Warehousing; Introduction to Dimensional Modelling; Architectural Components of a Data Warehouse; Extraction; Transformation & Loading; Advanced Dimensional Modelling; Online Analytical Processing (OLAP) & Multidimensional Databases (MDDB); Query Performance Enhancing Techniques; Metadata, Support for Data Warehousing in RDBMS/SQL; Real-time Data Warehousing; Current Trends in Data Warehousing

[Apply Now](#)

HOW TO APPLY

- ▶ [Click here](#) to visit the BITS Pilani Online Application Center. Create your login at the Application Center by entering your unique Email id and create a password of your choice. Once your login has been created, you can anytime access the online Application Center using your email id and password.
- ▶ Once you have logged in, you will see a screen showing 4 essential steps to be completed to apply for the programme of your choice
- ▶ Begin by clicking on Step 1 - 'Fill/ Edit and Submit Application Form'. This will enable you to select the programme of your choice. After you have chosen your programme, you will be asked to fill your details in an online form. You must fill all details and press 'Submit' button given at the bottom of the form
- ▶ Take the next step by clicking on Step 2 - 'Download Application PDF Copy'. This will download a pdf copy of the application form on your computer
- ▶ Now, click on Step 3 - 'Pay Application Fee' to pay INR 1,500/- using Net banking/ Debit Card/ Credit Card
- ▶ Take a printout of the downloaded Application Form and note down the Application Form Number that appear on the top-right corner of the first page. This Application Form Number should be referred in all future correspondence with BITS Pilani
- ▶ In the printout of the downloaded Application Form, you will notice on page no. 3 a section called the Employer Consent Form. Complete the Employer Consent Form. This form needs to be signed and stamped by your organisation's HR or any other authorised signatory of the company
- ▶ Further on page no. 4 of the printed Application Form is a section called the Mentor Consent Form. The Mentor Consent Form needs to be signed by the Mentor.

WHO IS A MENTOR:

Candidates applying to Work Integrated Learning Programmes must choose a mentor, who will monitor the academic progress of the candidate, and act as an advisor & coach for successful completion of the programme.

Apply Now

Candidates should ideally choose the immediate supervisor or another senior person from the same organisation. In case a suitable mentor is not available in the same organisation, a candidate could approach a senior person in another organisation who has the required qualifications. Wherever the proposed Mentor is not from the same employing organization as that of the candidate, a supporting document giving justification for the same should be provided by the candidate's employer.

Candidates applying to M.Tech. Data Science & Engineering programme should choose a Mentor who is an employed professional with:

- a. B.E./ M.Sc./ M.B.A./ M.C.A. etc. and with a minimum of five years of relevant work experience
- OR
- b. M.E./ M.S./ M.Tech./ Higher Degree of BITS or equivalent

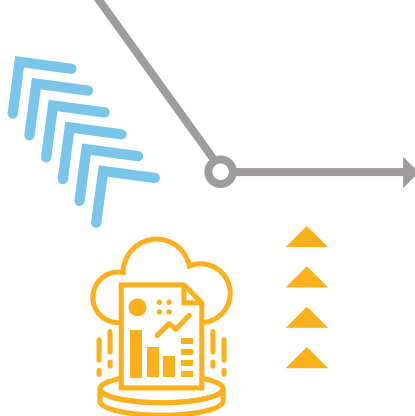
► Further on page no. 5 of the downloaded Application Form, is a Checklist of Enclosures/ Attachments.

- a. Make photocopies of the documents mentioned in this Checklist
- b. Applicants are required to self-attest all academic mark sheets and certificates

► Finally, click on Step 4 - 'Upload & Submit All Required Documents'.

This will allow you to upload one-by-one the printed Application Form, Mentor Consent Form, Employer Consent Form, and all mandatory supporting documents and complete the application process. Acceptable file formats for uploading these documents are .DOC, .DOCX, .PDF, .ZIP and .JPEG.

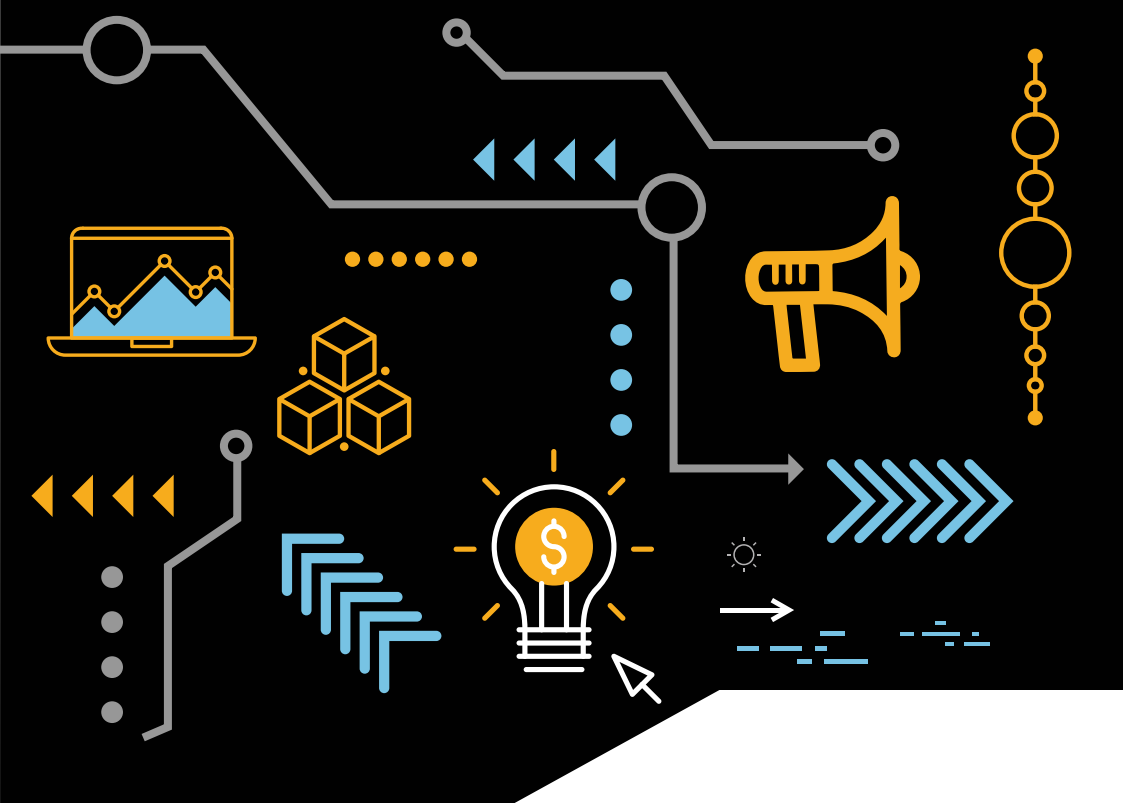
Note: It is also possible that some candidates may receive a link to take an Online Learning Readiness Evaluation from the Admission Cell, after filling up the Online Application Form. This is a one-hour objective-type exercise which will ascertain minimum mathematical and programming acumen needed to pursue the programme. You will be given typically 48-hours to complete this online evaluation. A sample model paper will also be provided to help you understand the format of Online Learning Readiness Evaluation. After the Online Learning Readiness Evaluation is completed, the Admission Cell will intimate selected candidates by email within one week.



► Upon receipt of your Application Form and all other enclosures, the Admissions Cell will scrutinise them for completeness, accuracy and eligibility.

► Admission Cell will intimate selected candidates by email within two weeks of submission of application with all supporting documents. The selection status can also be checked by logging in to the Online Application Centre. Candidates will need to accept the Admission Offer Letter, and submit the Admission Fee and Semester 1 Fee to enrol for the programme.

Apply Now



DISCLAIMER

Ever since it was declared as a Deemed to be University in 1964, BITS Pilani has been offering higher education programmes in science and technology, and has earned an enviable reputation for its innovations in this sphere. The Work Integrated Learning Programmes (WILP) of BITS Pilani constitutes a unique set of educational offerings for working professionals. These programmes, which BITS began to offer in 1979, have, over the years, evolved along the lines envisaged in the National Policy on Education, 1986.

The WILP are rigorous higher education programmes in technology areas, designed keeping the evolving needs of industry in view, and meant for working professionals in their respective domains. The very intent is to deliver the education at the workplace, in order that the greatest degree

of work integration of the education is achieved, and thus the WILP are very distinct in philosophy and pedagogy from open, distance learning programmes. Though it is incorrect and improper, at times the WILP are compared to ODL programmes. Accordingly, it has been our constant endeavor to engage with the regulator, and provide all necessary information about these programmes.

The WILP have been well received, and accepted by industry, because of the high quality of the programmes in terms of the curriculum and the instruction, and also because of the high degree of work integration, which results not only in up gradation of knowledge, but also in up skilling, and productivity increase.