Annexure -II

M. Tech. Data Analytics

Data Analytics involves using computing for reporting, analysis, monitoring, and prediction. International Data Corporation (IDC) forecasts that business analytics market will reach $50.7 billion in 2016. According to Avendus Capital, US alone is expected to have a shortage of 140,000 – 190,000 analytics professionals by 2018. NASSCOM predicts that the analytics market in India could reach $2.3 billion by the end of 2017-18.

In India, a good number of consulting, KPO and IT companies are providing specialised services in the field of data analytics. Some are providing general solutions in all segments whereas some others are specialising in specific verticals like banking, retail, marketing, financial, CRM, risk analysis, web, clinical, telecom, insurance, etc. These companies include- IBM Analytics, Mu-Sigma, LatentView, HCL Technologies, Accenture, Genpact Analytics, Cognizant Analytics, TCS Analytics, Wipro Analytics, McKinsey Analytics Knowledge Centre, Deloitte Analytics, PwC Analytics, AbsolutData, Fractal Analytics, iCreate, Dunhumby, Global Analytics, Manhattan Systems, Capillary Technologies, Nabler, Activecubes, ICRA Technology Services, WNS Analytics, Opera Solutions, Data Monitor, Ipsos, EXL Services, Meritus, Modeanalytics, Bridge i2i Analytics, Cytel, Neural Techsoft (Financial & Risk Analytics), Vehere Interactive, Aegis Global, Datamatics, Marketelligent, TNS Global, NettPositive Analytics, Affine Analytics, EVALUESERVE, Innovacer, Crisp Analytics.

In addition, many other captive companies/groups provide solutions to their parent organisation like. These include - HSBC Analytics, Citi Bank Analytics, American Express, Fidelity Analytics, GE Capital, RBS Business Services, Barclays Shared Services, Target Analytics, Spencer Analytics, Amazon Analytics, Dell Analytics, HP Analytics, eBay / PayPal, Experian India, Fair Isaac India, Dun & Bradstreet.

Further, a large number of new IT ventures are starting in this area to fill the global demand and talent gap in this area. Telangana state government has already announced a knowledge park on Data Analytics.

Program Objectives:

MTech (Data Analytics) is an interdisciplinary program offered by department of CSE & IT and is designed to meet the huge manpower shortage in this area that has been well recognized as one of the fastest growing areas. All business and government organisation working in commerce, policy, insurance, finance, economics, engineering, infrastructure, energy, health care, education, security, sports, media, culture, etc. are increasing relying on computational tools and techniques of data analytics for taking informed decisions.

This program has been designed to develop the ability to apply and develop computational techniques and systems to draw insights from big data in a variety of application domains. It is expected that the graduates of the proposed program will mostly join the fast growing Data Analytics industry as technical experts and few will even launch their own IT start-ups in this field. Some graduates will also choose to go for doctoral studies and/or pursue their career in academia. The available and growing academic opportunities being developed through our ongoing MTech programs in ‘CSE’ and ‘IT and Entrepreneurship’ will also be leveraged to enrich this program with advanced level computer science elective courses in various areas on one hand and design and entrepreneurship related elective courses on the other. Jaypee Centre for Entrepreneurship Development (JCED) will also support the students’ entrepreneurial activity that spins off from this program.

The curriculum exposes students with all aspects of data analytics including research design, data collection, preparation, analysis, integration, visualization, and interpretation. In addition to the CSE & IT department, the department of mathematics as well as business school/department of HSS will also contribute courses for this program. The core courses include statistical data analysis, financial econometrics, data warehousing and data mining, pattern recognition and machine learning, large scale graph analytics, empirical research and laboratories. Students will also be offered several electives on theoretical, systemic, algorithmic, and applied aspects of data analytics. In addition to the CSE & IT department, the department of mathematics as well as business school/department of HSS will also contribute courses for this program. The curriculum is strongly oriented towards industry and emphasises industry internship, projects, and entrepreneurship.

Eligibility:

- BTech (in any discipline) or equivalent
- Masters (in Computer Applications/Computer Science/IT/Maths/Statistics/Operations Research/Physics/Electronics/Instrumentation/Economics/Commerce)/Bioinformatics or equivalent
M. Tech. Data Analytics: Curriculum

**FIRST SEMESTER**

<table>
<thead>
<tr>
<th>S.No</th>
<th>Contact Hours</th>
<th>Credits</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>L</td>
<td>T</td>
</tr>
<tr>
<td>1</td>
<td>3</td>
<td>-</td>
</tr>
<tr>
<td>2</td>
<td>3</td>
<td>-</td>
</tr>
<tr>
<td>3</td>
<td>3</td>
<td>-</td>
</tr>
<tr>
<td>4</td>
<td>3</td>
<td>-</td>
</tr>
<tr>
<td>5</td>
<td>3</td>
<td>-</td>
</tr>
<tr>
<td>6</td>
<td>3</td>
<td>-</td>
</tr>
<tr>
<td>7</td>
<td></td>
<td>-</td>
</tr>
<tr>
<td>TOTAL</td>
<td>18</td>
<td>4</td>
</tr>
</tbody>
</table>

**SECOND SEMESTER**

<table>
<thead>
<tr>
<th>S.No</th>
<th>Contact Hours</th>
<th>Credits</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>L</td>
<td>T</td>
</tr>
<tr>
<td>1</td>
<td>3</td>
<td>-</td>
</tr>
<tr>
<td>2</td>
<td>3</td>
<td>-</td>
</tr>
<tr>
<td>3</td>
<td>2</td>
<td>-</td>
</tr>
<tr>
<td>4</td>
<td>3</td>
<td>-</td>
</tr>
<tr>
<td>5</td>
<td>3</td>
<td>-</td>
</tr>
<tr>
<td>6</td>
<td>2</td>
<td>-</td>
</tr>
<tr>
<td>7</td>
<td>-</td>
<td>4</td>
</tr>
<tr>
<td>TOTAL</td>
<td>16</td>
<td>0</td>
</tr>
</tbody>
</table>

**THIRD SEMESTER**

<table>
<thead>
<tr>
<th>S.No</th>
<th>Contact Hours</th>
<th>Credits</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>L</td>
<td>T</td>
</tr>
<tr>
<td>1</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>TOTAL</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

**FOURTH SEMESTER**

<table>
<thead>
<tr>
<th>S.No</th>
<th>Contact Hours</th>
<th>Credits</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>L</td>
<td>T</td>
</tr>
<tr>
<td>1</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>TOTAL</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>
Total Credits = 70

1\textsuperscript{st} Sem Electives (Elective I and II):
1. Web Algorithms
2. Information Retrieval
3. Intelligent Systems
4. Multimedia Analytics
5. E-Commerce and Social-Web
6. Algorithmic Graph Theory
7. Mobile and Pervasive Computing
8. Distributed Systems  
   (Core of MTech (CSE))
9. High Performance Software Engineering  
   (Core of MTech (CSE))
10. Activity Centred Design and User Experience  
    (Core of MTech (ITE))
11. IT Venture Creation  
    (Core of MTech (ITE))
12. Data Analytics for Bioinformatics  
    (by Biotech)
13. Socio-Political Analysis  
    (by HSS)
14. Political -Economic Analysis  
    (by HSS)
15. Demographic Analysis  
    (by HSS)

2\textsuperscript{nd} Sem Electives (Elective III, IV, and V):
1. Advanced Machine Learning
2. Interactive Data Analysis
3. Ecommerce and Social Media Analytics
4. Mobile and IoT Analytics
5. Forensic Analytics
6. Information Integration and Visualization
7. Computer Vision
8. Machine Translation
9. Predictive Analytics
10. Web Services and Cloud Computing
11. Cryptography and Computer Security
12. Advanced Algorithms  
    (Core of MTech (CSE))
13. Performance Evaluation of Computing systems  
    (Core of MTech (CSE))
14. IT Product Innovation and Creativity  
    (Core of MTech (ITE))
15. Game Theory and Business Intelligence  
    (by JBS)
16. CRM Analytics  
    (by JBS)
17. Financial analysis  
    (by HSS)
18. Ethics & laws in Social Data Analysis  
    (by HSS)

Electives can be added flexibly.
Any other core/ elective course offered to MTech (CSE)/MTech (IT&E)/MTech (ACM) can be opened for this program as well.
Syllabus of Bridge and Core courses:

1. **Econometrics:** (3 credits Bridge Course, not open for students with MA Economics)

2. **Programming Foundations for Data Sciences:** (3 credits Bridge Course, not open for students with CS/IT background)
   HTML5, DHTML, Overview of Python, Operators and Expressions, Flow Control, Sequence Data (list operations, list methods, Strings sets, Dictionaries), Functions, Files, Errors and Exception Handling, Modules, Regular Expressions, Object oriented programming using Python, GUI, Database connectivity: Python Database API to connect MySql Database. Python XML processing: XML parser Architecture, Parsing XML with SAX & DOM APIs, Introduction to cloud environment.

3. **Mathematics for Data Analytics:** (3 credits)

4. **Data Warehousing and Data Mining:** (3 credits)
   Schema integration, Data Mart/Data Warehouse, Metadata, OLTP Vs OLAP, Data Summarization, Data cleaning, de-duplication, Dimensional Modelling, Fact Tables, Star schema, Snowflake schema, OLAP, ROLAP, MOLAP, online aggregation, From Data Warehousing to Data Mining, Data Cube Computation and Data Generalization, MDX, Classification, Clustering, Computational Tools for Integration Management (ETL, EAI, etc), Information Management (DW/Mart/ODS, Olap Servers, etc), Information Delivery (Portal, Dashboard, Analytics/OLAP Client, etc), and data mining (e.g., RapidMiner, Weka, etc.)

5. **Machine Learning:** (3 credits)
   Bayesian reasoning, Parametric methods, Supervised learning, Decision trees, Dimensionality reduction, mixture densities and expectation maximisation, Unsupervised learning, Optimisation and search, ANN-multilayer preceptron and back propagation network, Kernel Machines, Evolutional learning, Reinforcement learning, Deep learning, Computational Tools, e.g., Octave

6. **Empirical Research:** (3 credits)
   Philosophy of Knowledge and Science, Argument and fallacies, Frameworks of Critical Thinking, Research thinking, Formulating a research problem, IT assisted Informative VS Normative research, Case Study, Experiment design, Controlled experiment, Quasi experiment, IT assisted sampling, Confounding variable, Qualitative and quantitative data, IT assisted data collection (direct, indirect, independent), IT assisted data analysis, IT assisted data visualisation, Concept mapping tools, Theory building, Falsifiability, Validation, Validity threats, Web based empirical research, Empirical research in software engineering, information systems, usability, e-commerce, Social media, business, policy making, and social sciences, Introduction to some of the following open source empirical research tools- ODK, QDA Miner Lite, CAT, CATMA, Aquad, Compendium, QCAmap, Octave, PSPP, hackystat, etc.
7. **Big Data Technologies**: (3 credits)
   Big data, Structured data, Unstructured data, Storage and indexing, of massive databases, Algorithmic topics e.g., sketching, hashing, random projections, etc. Massive spatial, multimedia, and real time databases, SQL extensions, NoSQL databases, Massive parallel processing, Big data programming- Hadoop/HDFS, Map-reduce, Apache Spark, event processing, Big Data Visualisation, Big Data Analytics, Applications of Machine learning techniques.

8. **Large Scale Graph Analytics**: (2 credits)
   Large scale graph representation and storage, Social networks, Indexing techniques for large graphs, graph compression, query processing, link analysis, evolving graphs, heterogeneous graphs, integrating graph with non graph data, distributed graph management, large scale graph visualisation, Applications of Machine learning techniques, Computational Tools, e.g., Gephi, HyperGraphDB, Neo4j, etc.