IDM TECHPARK

GUIDE`S FOR PERFECT CAREER PATHWAY

Data Science Course Syllabus

N·S·D·C National Skill Development Corporation

CERTIFIED ISO 9001:2015 COMPANY

IDMTECHPARK
GUIDE'S FOR PERFECT CAREER PATHWAY

1

About Us

IDMTECHPARK global retail & corporate training solutions provider in Coimbatore, Erode, Trichy & Salem that offers a comprehensive range of training and placement services for both fresher's and professionals seeking new opportunities. The company commenced its IT training business in 2016. A pioneer in IT education, over the years, we have trained over 50k students. Idmtechpark has a wide range of courses, maintains education standards & provides placement assistance.

About IDMTechpark Education Quality

IDMTECHPARK is managed and developed by industry specialists with more than 8 years of expertise in the field. IDMTECHPARK offers a staff of highly skilled professional trainers who deliver effective IT training in a friendly setting, concentrating on the needs of each individual to help them succeed in a demanding work world. In the book of career and success, our staff never leaves a page unturned.

IDMTECHPARK's versatile instructor-led training class rooms and lower-class sizes enable people to engage more easily and absorb knowledge, resulting in remarkable results for both themselves and the organizations for which they work. Our training programmes are adaptable and customizable to ensure that each participant gets the most out of their time with us. IDMTECHPARK focuses in providing hands-on IT training in over 30 different courses.

- We teach in-demand courses
- We provide impactful learning material
- Our teachers are well-selected & trained
- We follow world-class teaching methods
- Our courses include E-Projects
- We conduct technical workshops
- Exams are held and based on Exams providing Certification
- Certificates are recognized the world over
- Our course timings are flexible

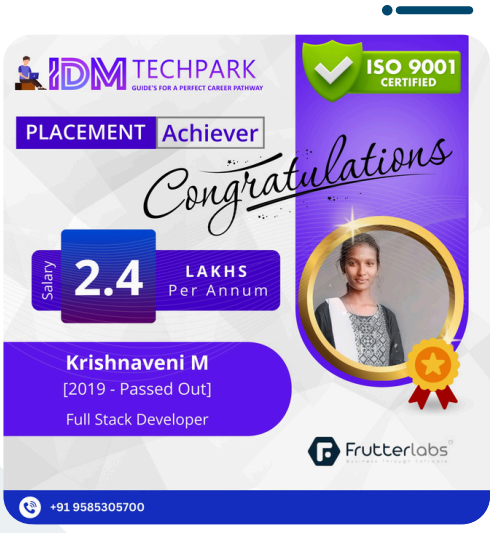
3

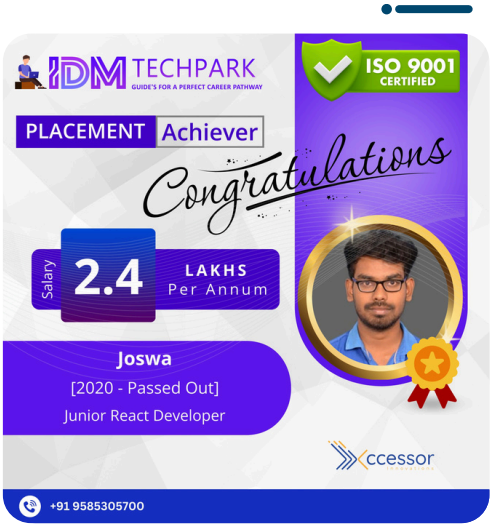
Our Recent Placement

Idmtechpark assists students in getting job placements on successful completion of their courses. Idmtechpark also provides recruitment assistance to organizations. Idmtechpark students are shortlisted based on the organization's requirement. To make students job-ready, Idmtechpark conducts workshops e.g. How to do Group Discussions, how to behave in a Personal Interview. From time to time, job fairs & campus recruitments are conducted. Workplace skills such as time management, making effective presentations and communication skills are also provided. All this helps students find appropriate jobs in the IT industry while also helping save companies recruitment costs.

Krishnaveni M

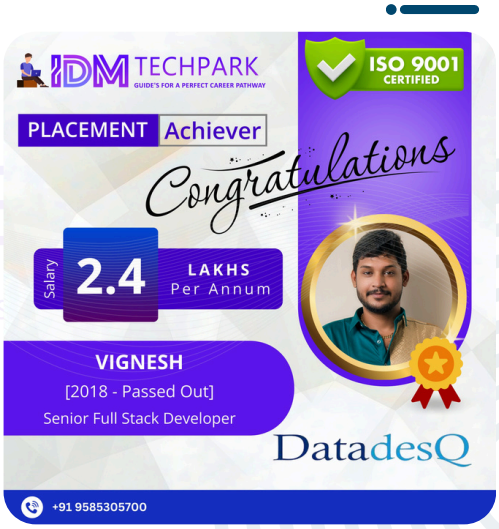
Frutterlabs





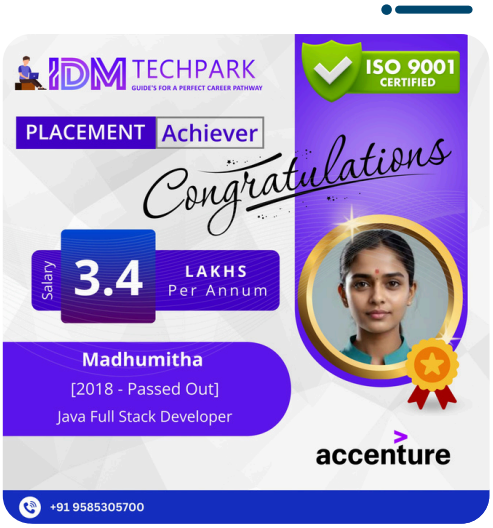
Joshwa

Xccessor



Vignesh

Datadesq

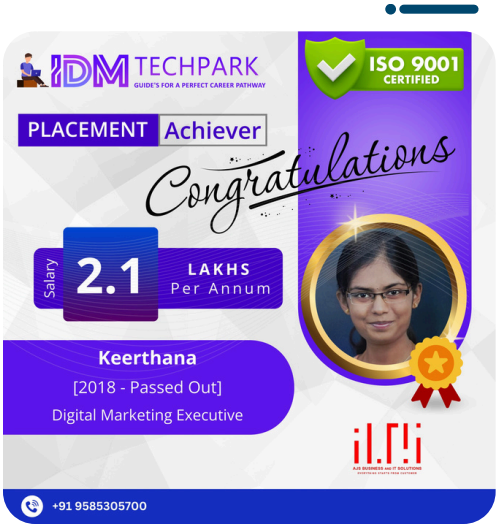


Madhumitha

Accenture

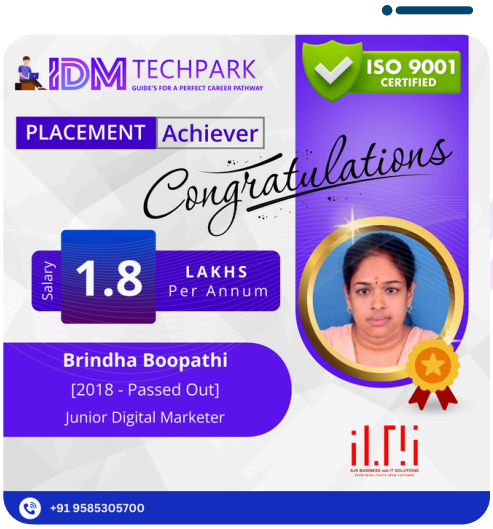
Keerthana

AJS



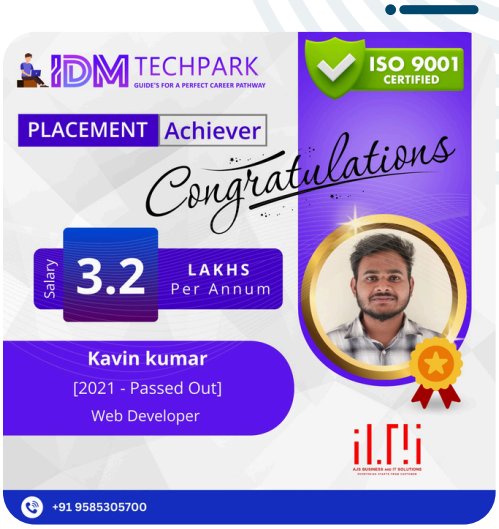
Brindha Boopathi

AJS



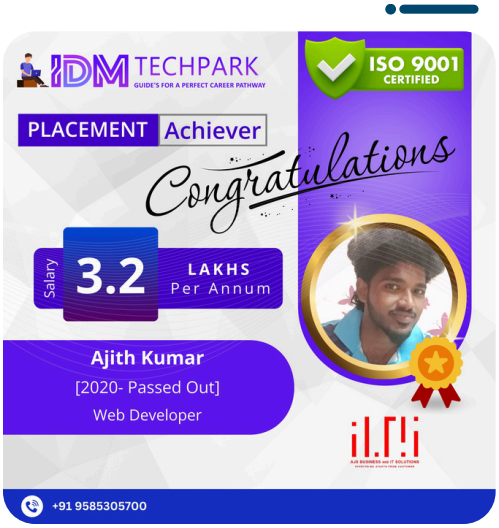
Kavin Kumar

AJS



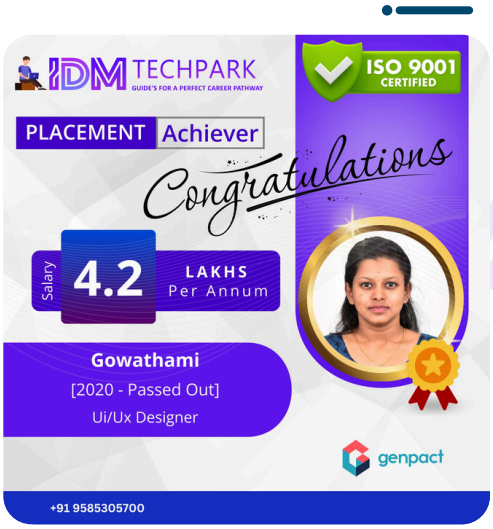
Ajithkumar

AJS



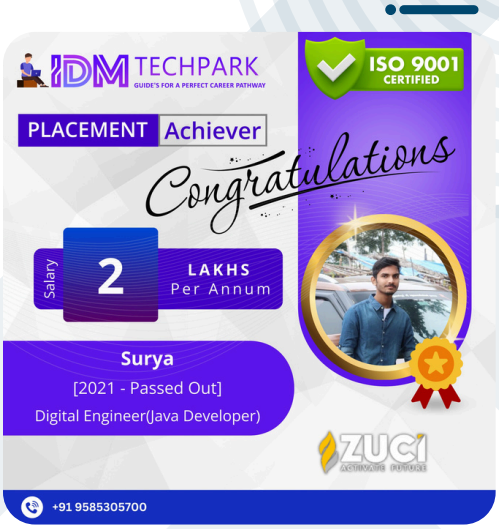
Gowthami

Genpact



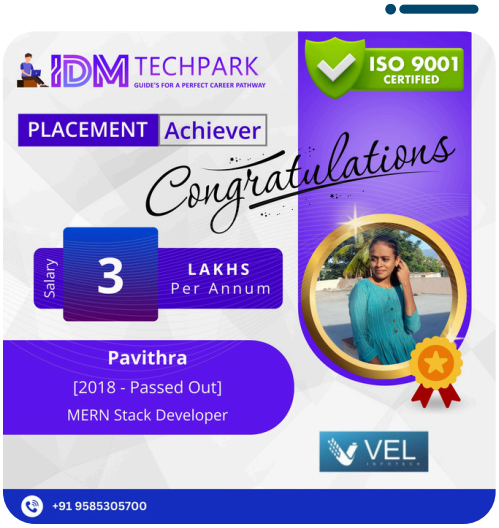
Surya

ZUCI



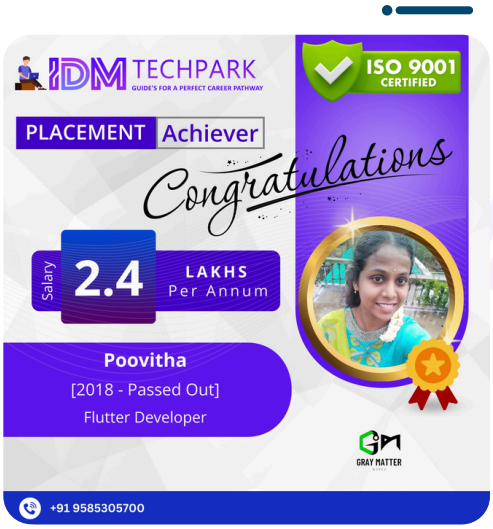
Pavithra

Vel Info Tech



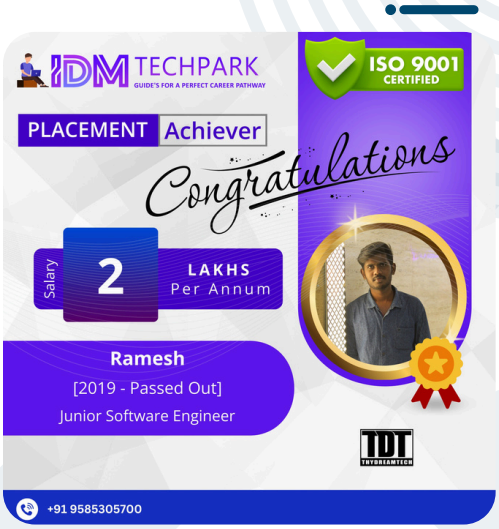
Poovitha

Gray Matter



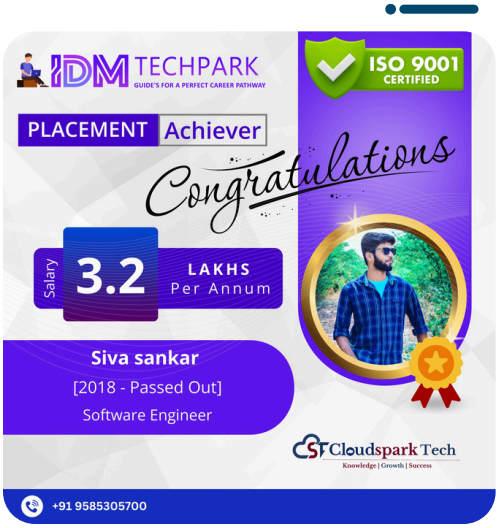
Ramesh

TDT



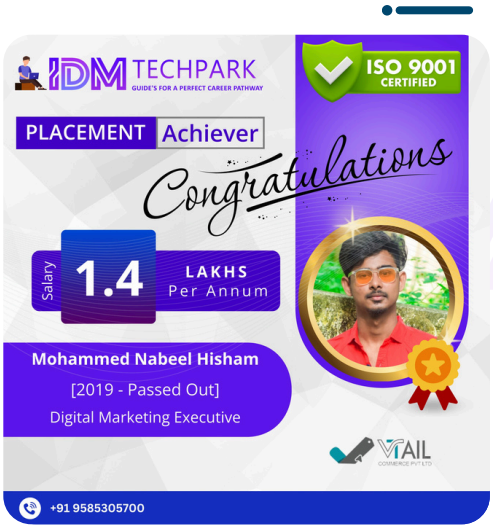
Siva Sankar

ST Cloudspark tech



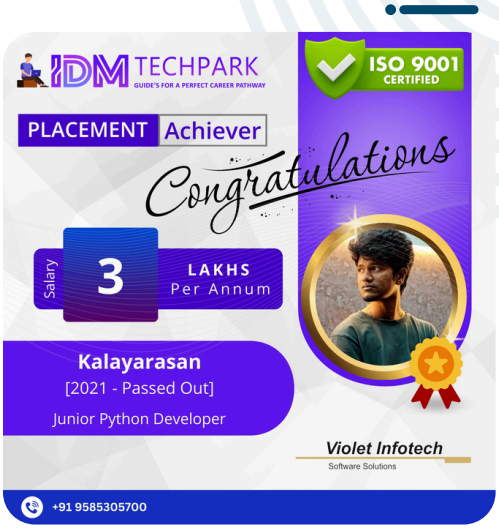
Nabeel Hisham

VTail



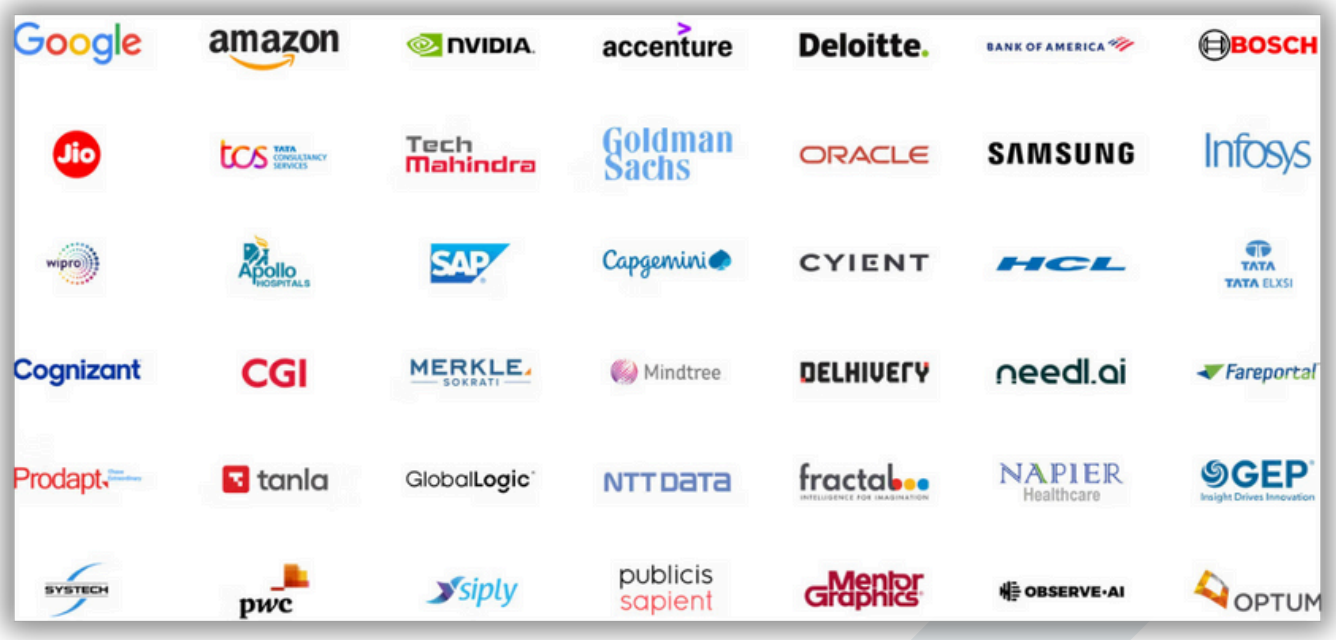
Kalayarasan

Violet Infotech



4

Our Alumini Working At



MODULE 1

Introduction to Data Science

- Overview of Data Science
  - What is Data Science?
  - Data Science vs. Data Analytics vs. Machine Learning vs. Artificial Intelligence
- The Data Science Lifecycle
  - Problem identification, data collection, data cleaning, analysis, modeling, and interpretation
- Tools and Technologies in Data Science
  - Introduction to Python, R, and key libraries (e.g., Pandas, NumPy)

MODULE 2

Python for Data Science

- Python Programming Basics
  - Data types, loops, functions, and conditionals
- Python Libraries for Data Science
  - Pandas for data manipulation
  - NumPy for numerical computations
  - Matplotlib and Seaborn for data visualization

MODULE 3

Data Collection and Data Wrangling

- Data Collection Methods
  - Web scraping, APIs, and databases (SQL, NoSQL)
  - Understanding data formats (CSV, JSON, XML)
- Data Wrangling
  - Handling missing values, outliers, and duplicates
  - Data cleaning and transformation techniques

MODULE 4

Exploratory Data Analysis (EDA)

- Introduction to EDA
  - Descriptive statistics (mean, median, mode, variance)
  - Data distribution and visualization (histograms, box plots, etc.)
- Data Visualization
  - Correlation matrices, pair plots, scatter plots
  - Visualizing categorical and numerical variables

MODULE 5

Probability and Statistics for Data Science

- Fundamentals of Probability
  - Probability distributions (normal, binomial, Poisson)
  - Conditional probability and Bayes' Theorem
- Descriptive and Inferential Statistics
  - Hypothesis testing, p-values, confidence intervals
  - Sampling techniques and bias

MODULE 6

Linear Algebra for Data Science

- Introduction to Linear Algebra
  - Vectors, matrices, and operations
  - Matrix multiplication and dot product
- Applications in Data Science
  - Data transformations and dimensionality reduction

MODULE 7

Introduction to Machine Learning

- Overview of Machine Learning
  - Supervised vs. unsupervised learning
  - Common algorithms (linear regression, decision trees, KNN)
- Model Evaluation
  - Train-test split, cross-validation
  - Metrics: accuracy, precision, recall, F1-score

MODULE 8

Supervised Learning – Regression

- Linear Regression
  - Simple and multiple linear regression
  - Assumptions, cost function, and gradient descent
- Polynomial and Logistic Regression
  - Regularization (L1 and L2)
  - Overfitting and underfitting

MODULE 9

Supervised Learning – Classification

- Classification Algorithms
  - K-Nearest Neighbors (KNN), Decision Trees, Naive Bayes
  - Support Vector Machines (SVM)
- Evaluation Metrics for Classification
  - Confusion matrix, ROC curve, AUC, F1-score

MODULE 10

Unsupervised Learning – Clustering

- Clustering Techniques
  - K-Means clustering
  - Hierarchical clustering, DBSCAN
- Dimensionality Reduction
  - Principal Component Analysis (PCA)
  - t-SNE for data visualization

MODULE 11

Unsupervised Learning – Association

- Association Rule Learning
  - Apriorism algorithm
  - Market Basket Analysis
- Applications of Association Rules
  - Recommender systems and customer behavior analysis

MODULE 12

Neural Networks and Deep Learning

- Introduction to Neural Networks
  - Perceptrons, multi-layer perceptrons (MLP)
  - Activation functions (ReLU, sigmoid)
- Deep Learning Fundamentals
  - Deep neural networks (DNN), backpropagation
  - Introduction to frameworks like TensorFlow and Keras

MODULE 13

Natural Language Processing (NLP)

- Text Preprocessing Techniques
  - Tokenization, stemming, lemmatization
  - Bag-of-Words, TF-IDF, and Word2Vec
- Text Classification and Sentiment Analysis
  - Naive Bayes, LSTM, BERT
  - Applications of NLP in chatbots, sentiment analysis, and translation

MODULE 14

Time Series Analysis

- Time Series Data
  - Understanding time series components: trend, seasonality, noise
  - Autocorrelation and stationarity
- Time Series Forecasting
  - ARIMA, SARIMA, Holt-Winters exponential smoothing
  - Advanced forecasting techniques

MODULE 15

Feature Engineering and Selection

- Feature Engineering
  - Creating new features, handling categorical variables (one-hot encoding, label encoding)
  - Feature scaling (normalization, standardization)
- Feature Selection
  - Importance of feature selection
  - Techniques: Recursive feature elimination, L1 regularization (Lasso)

MODULE 16

Advanced User Research Techniques

- Overfitting and Underfitting
  - Bias-variance trade-off
  - Regularization techniques
- Hyperparameter Tuning
  - Grid search and random search
  - Bayesian optimization

MODULE 17

Advanced Machine Learning Algorithms

- Ensemble Methods
  - Random Forests, Gradient Boosting, XGBoost, LightGBM
  - Stacking, bagging, and boosting techniques
- Support Vector Machines (SVM)
  - Kernel tricks, SVM for classification and regression

MODULE 18

Big Data Technologies

- Introduction to Big Data
  - Characteristics of big data (Volume, Variety, Velocity)
  - Hadoop and Spark ecosystem
- Distributed Computing for Data Science
  - MapReduce, Spark SQL, and PySpark
  - Data storage and retrieval in big data environments

MODULE 19

Data Engineering for Data Science

- Data Pipelines
  - ETL (Extract, Transform, Load) processes
  - Tools like Apache Airflow, Luigi, and Kafka
- Databases and SQL
  - Relational databases, NoSQL databases (MongoDB, Cassandra)
  - Advanced SQL queries and optimization

MODULE 20

Cloud Computing and Data Science

- Cloud Platforms for Data Science
  - AWS, Google Cloud, Microsoft Azure
  - Cloud-based data storage (S3, BigQuery, Data Lake)
- Data Science in the Cloud
  - Training machine learning models on the cloud
  - Cloud-based data processing and analysis tools

MODULE 21

Model Deployment and Monitoring

- Model Deployment Techniques
  - Flask/Django APIs for model deployment
  - Cloud services for deployment (AWS Lambda, Google AI Platform)
- Model Monitoring and Maintenance
  - Continuous monitoring, drift detection, and model retraining
  - A/B testing for deployed models

MODULE 22

Ethics in Data Science

- Ethical Considerations
  - Privacy, data security, and fairness in machine learning
  - Addressing biases in data and models
- Regulations and Policies
  - GDPR, CCPA, and data protection laws

MODULE 23

Data Science Project Management

- Project Lifecycle and Documentation
  - Defining project scope, timeline, and deliverables
  - Documentation best practices for reproducible research
- Collaboration and Communication
  - Working with stakeholders, presenting findings, and insights

MODULE 24

Advanced Topics in Data Science

- Reinforcement Learning
  - Introduction to Q-learning and policy gradients
  - Applications of reinforcement learning (e.g., game AI, robotics)
- Generative Models
  - GANs (Generative Adversarial Networks)
  - Variational Autoencoders (VAEs)

MODULE 25

Capstone Project

- End-to-End Data Science Project
  - Problem formulation, data collection, EDA, feature engineering, model building
  - Presentation of results and insights
- Peer Review and Critique
  - Presenting projects to peers and receiving feedback
  - Final project evaluation

IDM TECHPARK
GUIDE'S FOR PERFECT CAREER PATHWAY

Thank You

+91 9585305700
www.idmtechparkcoimbatore.com
www.idmtechparkerode.com