Sathya™
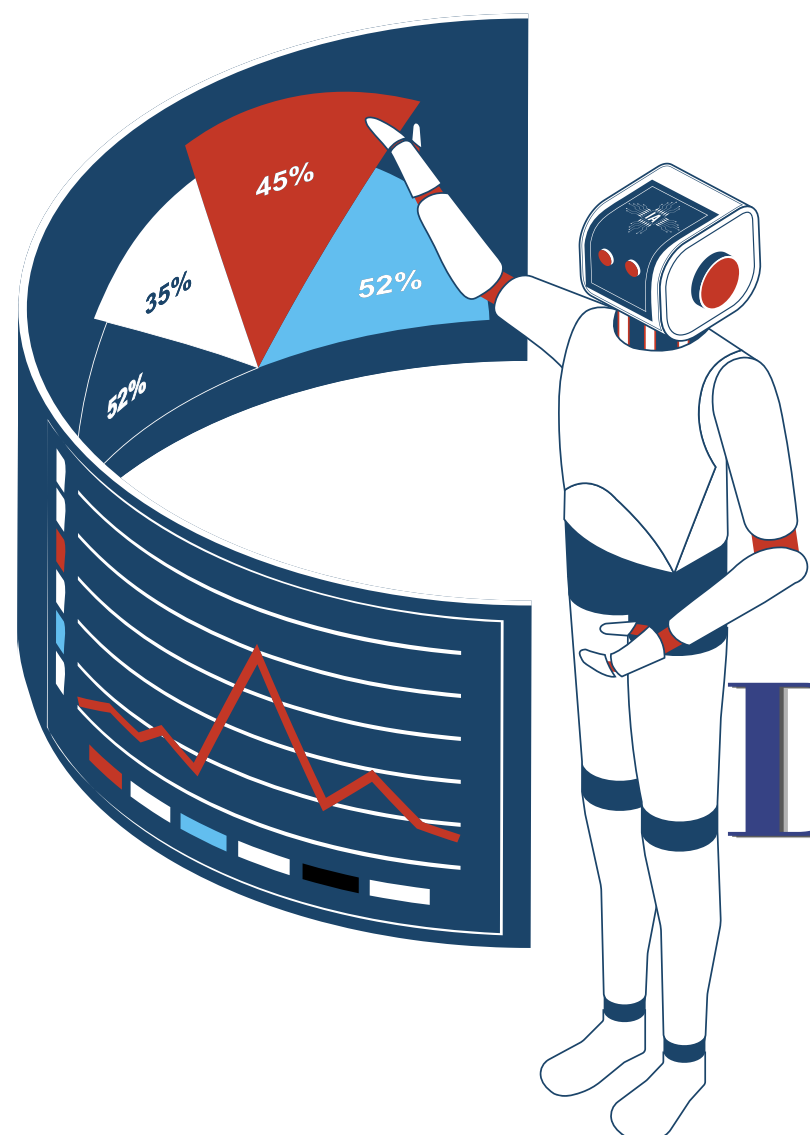## TECHNOLOGIES

DATA SCIENCE
And GenAI

DATA SCIENCE IT IS A SOFTWARE HERE DISTRIBUTING AND PROCESSING THE LARGE SET OF DATA INTO THE CLUSTER OF COMPUTERS. THIS COURSE IS DESIGNED TO MASTER YOURSELF IN THE DATA SCIENCE TECHNIQUES AND UPGRADE YOUR SKILL SET TO THE NEXT LEVEL TO SUSTAIN YOUR CAREER IN EVER CHANGING THE SOFTWARE INDUSTRY.

Sathya™
TECHNOLOGIES

2ND FLOOR, SRI SAI ARCADE, BESIDE ADITYA TRADE CENTER, AMEERPET, HYDERABAD-500038.
TEL: +91-9100920092, 9100940094 WWW.SATHYATECH.COM

Introduction to Data Science

- **Need for Data Scientists**
- **Foundation of Data Science**
- **What is Business Intelligence**
- **What is Data Analysis, Data Mining, and Machine Learning**
- **Analytics vs Data Science**
- **Value Chain**
- **Types of Analytics**
- **Lifecycle Probability**
- **Analytics Project Lifecycle**

Data

- **Basis of Data Categorization**
- **Types of Data**
- **Data Collection Types**
- **Forms of Data and Sources**
- **Data Quality, Changes and Data Quality Issues, Quality Story**
- **What is Data Architecture**
- **Components of Data Architecture**
- **OLTP vs OLAP**
- **How is Data Stored?**

Python

- **Python Installation**
- **Jupyter Notebook Tutorial**
- **Variable**
- **Function**
- **Lambda Expression**
- **Loops**
- **List**
- **Tuple**
- **Set**
- **Dictionary**

Advance Python

- Introduction to Numpy Creating Arrays
- Selection and Indexing
- Basic Operations on Arrays
- Mathematical Operation on Arrays
- Linear Algebra Operation on Arrays
- Stacking Arrays
- Data Types and Type Conversion
- Introduction to Pandas
- Creating Data Frames Reading and Writing Operation
- Selection and Indexing
- Conditional Selection
- Groupby
- Pivot Table
- Merge
- Join
- Concat
- Missing Value Treatment
- Drop Duplicates
- Dealing with Date Time Data
- Introduction to Series
- Series Operation
- Pandas Builtin Functions for Data Visualisation

Data Base (SQL)

- Introduction to Databases
- Basics of SQL
- DML, DDL, DCLand Data Types
- Common SQLcommands using SELECT, FROM and WHERE
- Logical Operators in SQL
- Filtering and Sorting
- Advanced filtering using IN, OR and NOT
- Sorting with GROUPBYand ORDER BY
- SQLJoins
- INNER and OUTER joins to combine data from multiple tables
- RIGHT, LEFTjoins to combine data from multiple tables
- SQLAggregations
- Common Aggregations including COUNT, SUM, MIN and MAX
- CASE and DATE functions as well as work with NULLvalues
- Subqueries and Temp Tables
- Subqueries to run multiple queries together
- Temp tables to access a table with more than one query
- Window Functions
- ROW_NUMBERS(), RANK(), DENSE_RANK(), LAG, LEAD, SUM, COUNT, AVG

Reporting Tool (Power BI)

- **What is Business Intelligence?**
- **Power BI Introduction**
- **Quadrant report**
- **Comparison with other BI tools**
- **Power BI Desktop overview**
- **Power BI workflow**
- **Installation query addressal**
- **Data import options in Power BI**
- **Import from Web (hands on)**
- **Why Visualization?**
- **Visualization types**
- **Categorical data visualization**
- **Trend Data viz**
- **Visuals for Filtering**
- **Slicer details and use**
- **Formatting visuals**
- **KPI visuals**
- **Tables and Matix**

Visualisation

- **Introduction To Plotly**
- **Scatter Plot**
- **Line Plot**
- **Scatter Matrix**
- **Box Plot**
- **Bar Chart**
- **Histogram**
- **Sun Burst Chart**
- **Create Dash Board**

Statistics

- **Central Limit Theorem**
- **Measure of Dispersion**
- **Quartiles**
- **Inter Quartile Ranges**
- **Variance**
- **Standard Deviation**
- **Z Score**
- **Normal Distribution**
- **Pearson Correlation Coefficient- R**
- **R Square**
- **Multi Collinearity Detection Techniques**
- **Multi Collinearity Removal Techniques**
- **Outliers Detection and Removal**

Machine Learning

- Introduction to Machine Learning
- Difference Between Supervised & Unsupervised Learning
- Difference Between Classification and Regression
- Machine Learning Application
- Data Science Project Life Cycle
- Linear Regression
- Theory of Linear Regression
- Cost Function
- Optimization Using Gradient Descent
- Mathematical Interpretation of Gradient Descent
- Model Validation Techniques
- Mean Squared Error
- Root Mean Squared Error
- Mean Absolute Error
- Polynomial Regression
- Understanding Polynomial Regression
- Implementing Polynomial Regression Using Python
- Overfitting, Underfitting, Right Fit Logistic Regression
- Understanding Logistic Regression Step by Step
- Decision Tree and Random Forest
- ID3 Algorithm vs CART
- Entropy
- Information Gain
- Step by Step Understanding of How Decision Tree Work

- How to overcome overfitting in DT
- Cross Validation
- Bootstrap Aggregation/Bagging
- Introduction to Random Forest
- How Random Forest Works
- Feature Selection
- Model Validation Techniques
- Accuracy
- Confusion Matrix
- Classification Report
- Recall
- Precision
- Hyper parameter Tuning
- K-Means Clustering
- What is Euclidian Distance
- Introduction to Unsupervised Learning
- Step By Step Mathematical Derivation
- Pros and Cons Of K Means
- Elbow Method to Find K

Deep Learning

- What is Deep Learning
- Deep Learning VS Machine Learning
- What is a Perceptron
- How Neural Network Learns
- Multi Layer Perceptron
- Activation Function
- Introduction to Keras
- What is Feed Forward Network
- Detail Explanation of ANN
- What is Cost Function
- Optimization Technique
- Vanilla Gradient Descent
- Mini Batch Gradient Descent
- Stochastic Gradient Descent
- Softmax
- Cross Entropy Loss
- MSE vs Cross Entropy

Image Processing, CNN & Computer Vision

- Introduction to Computer Vision
- Challenges in Computer Vision
- Introduction to Open CV Image Basics
- Reading and Writing Images/Videos
- Rescaling/Normalisation
- Colour Mapping
- Thresholding of an Image
- Morphological Transformation
- Image Augmentation Using Keras
- What is Image Filters
- Different Kind of Filters
- Convolution
- What is Convolutional Neural network
- Pooling
- Overfitting In Deep Learning
- Drop Outs

Sathya TECHNOLOGIES

Natural Language Processing-Text Mining

- What is Unstructured Data
- Introduction to NLTK and Spacy
- Tokenization
- Stop Word Removal
- Stemming
- Lemmatization
- N-Grams
- What is Word Embedding
- Count Vectorizer
- Tf-Idf Vectorizer
- Pattern Matching
- Regular Expression

Generative AI on Cloud Topics

- Cloud Computing Foundations AWS S3
- Amazon EC2 Trn1n Amazon EC2 Inf2
- Amazon Code Whisperer Amazon Bedrock
- Azure Open AI

Generative AI and its Industry Applications Topics

- Generative AI Principles Types of Generative Models
- Applications of Generative Models Machine Learning Algorithms with GenAI
- Applications of Generative AI
- Generative AI: Advantages and Disadvantages Ethical Considerations

Working with ChatGPT Topics

- Introduction to ChatGPT
- Leveraging ChatGPT for Productivity Mastering Excel through ChatGPT
- Social Media Marketing using ChatGPT Keyword Search and SEO using ChatGPT
- Generating Content using ChatGPT
- Implementing ChatGPT for Customer Service ChatGPT for Developers
- ChatGPT for Creating Programs ChatGPT for Debugging
- Documenting the Code with ChatGPT
- ChatGPT vs DeepSeek