

Data Science with AI

► Foundations of Python for Data Science

- ▶ Python basics: variables, loops, functions, modules
- ▶ Working with files, error handling, virtual environments
- ▶ Popular DS libraries: NumPy, Pandas, Matplotlib, Seaborn
- ▶ Writing clean, readable, reusable code
- ▶ Mini Project: Data exploration on a small CSV (e.g., sales or healthcare)

► Data Handling & Pre-processing

- ▶ Handling missing values, outliers, skewed data
- ▶ Encoding categorical data
- ▶ Scaling & normalization
- ▶ Feature engineering essentials
- ▶ Pandas groupby, merges, joins for real datasets

► Exploratory Data Analysis (EDA)

- ▶ Statistical summaries
- ▶ Correlation, distribution analysis
- ▶ Visualization patterns for business insights
- ▶ Storytelling with graphs
- ▶ Colab Activities: Interactive plots with Plotly
- ▶ Mini Project: EDA report for a retail or healthcare dataset

► Statistics & Probability for Data Science

- ▶ Descriptive & inferential statistics
- ▶ Probability distributions
- ▶ Hypothesis testing (t-test, chi-square, ANOVA)
- ▶ Confidence intervals
- ▶ Sampling methods
- ▶ Mini Project: Statistical analysis of A/B testing dataset

► Machine Learning Fundamentals

- ▶ Train-test split, cross-validation
- ▶ Regression (Linear, Lasso, Ridge)
- ▶ Classification (Logistic, KNN, SVM)
- ▶ Model evaluation metrics
- ▶ Bias-variance, overfitting, underfitting
- ▶ Mini Project: Predict house prices / classify customer churn

► Advanced Machine Learning

- ▶ Ensemble methods (Random Forest, XGBoost, AdaBoost)
- ▶ Feature importance analysis
- ▶ Hyperparameter tuning (GridSearchCV, RandomSearch)
- ▶ Pipelines and automation
- ▶ Mini Project: Build a high-performance ML model for tabular data

► **Unsupervised Learning**

- K-Means, Hierarchical clustering
- DBSCAN
- Dimensionality reduction (PCA, t-SNE)
- Colab Activity: Visual cluster plots
- Mini Project: Customer segmentation analysis

► **Deep Learning & Neural Networks**

- Neural network basics (perceptron, activation, loss)
- Using TensorFlow/Keras in Colab
- CNN basics for image data
- RNN/LSTM basics for sequence data
- Transfer learning
- Mini Project: Image classifier using pre-trained models

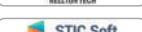
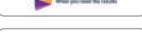
► **Applied AI & Large Language Models**

- Introduction to LLMs (GPT, Gemini, LLaMA)
- Prompt engineering
- Using open-source models in Colab (HuggingFace)
- Vector embeddings & semantic search
- Building simple AI assistants
- Mini Project: Build a text-based chatbot using an open LLM





**Our Students
Are Placed In**

 TATA CONSULTANCY SERVICES	 vsoft Ingenious at work	 Cognizant	 Microsoft	 ENVISION	 SUNSEAZ Technologies Pvt Ltd	 wipro	 iSOFT CONSULTING	 amazon
 PrimeSoft	 statista	 Practera	 PIERSOFT	 HELLTON TECH	 Capgemini	 KENSUIM	 SOLIX Empowering Data Management	 sumtotal TM A Skillsoft Company
 ValueMomentum	 PERVACIO	 Apollo	 LARSEN & TOUBRO	 STIC Soft since you need the results	 Advi Tech Solutions	 NOOR	 servicenow	 ValueLabs
 CyberTech	 BLACK KNIGHT	 Micron	 highradius	 HEXaware	 NextSphere Technologies	 GLOBAL Management Services On the Right Side of Business	 splashBI TM	 archents
 Nendrasy Innovating Business Technologies	 magnasoft TM Digital Learning	 Innominds Powering the Digital Next	 High Noon Consulting People & Everything	 JRD SYSTEMS	 Medtronic	 MAXO	 PTW	 Uber
 SOLUGENIX	 Collabera TM Value. Accelerated.	 iPrism Technologies	 meslova	 actlogica Actuarial Finance Technology	 S&P Global	 FISSION Technologies	 SOLIGENIX	 TechnoFacts
 SAHASYA Global Solutions								