AI & Data Masters

A comprehensive journey through the complete AI and data science ecosystem, from foundational programming to cutting-edge generative AI. This programme covers nine essential sections designed to transform you into a complete AI and data professional.

Python for AI & Data

Master Python fundamentals, data structures, and advanced programming concepts essential for AI development.

SQL for AI & Data

Learn database management, complex queries, and data manipulation with PostgreSQL.

PowerBI for Data Analysis

Create compelling visualisations and interactive dashboards for business intelligence.

Data Engineering with MS Fabric

Build scalable data pipelines and manage enterprise data infrastructure.

Maths & Stats for AI & Data

Develop the mathematical foundation required for machine learning algorithms.

Python Libraries for AI & Data

Master NumPy, Pandas, Matplotlib, Seaborn, and Plotly for data manipulation and visualisation.

Machine Learning

Implement supervised and unsupervised learning algorithms from scratch to production.

Deep Learning & NLP

Build neural networks, CNNs, RNNs, and natural language processing systems.

Generative AI & Agentic AI

Create intelligent agents and leverage large language models for real-world applications.

Digital Edify

India's First AI-Native Training Institute

Learn AI. Build Agents. Lead Future.

hello@digitaledify.ai

+91 6304982304

About Digital Edify

India's #1 Training Institute for the AI Era

Established: 2016 **Headquarters:** Hyderabad, Telangana **Reach:** Global (Online + Offline)

The Transformation Narrative

Digital Edify has evolved from a premium training institute in the Automation Era to an AI-first organisation leading the Agentic AI revolution. Since 2016, we've transformed over 100,000 professionals and built partnerships with more than 1,000 industry leaders. Our journey reflects the technological evolution of our time—from traditional job placement to career transformation, and now to building AI-native professionals who will shape the future of work.

1

Automation Era (2016-2023)

Premium Training Institute focused on job placement with 100K+ students trained

2

AI Revolution (2024-2025)

AI-Powered Training with industry-AI integration and career transformation focus

3

Agentic AI Leadership (2026+)

AI First Institute building AI-Native Professionals with 1 Million AI-Native Vision

> "We started in the Automation Era. We evolved through the AI Revolution. Now, we're leading the Agentic AI Future—with 100,000+ professionals already transformed and 1,000+ industry partners trusting our graduates."

Vision & Mission

Vision

"To Create 1 Million AI-Native Professionals Who Will Build the Agentic Future of Work"

Mission

"We transform learners into AI-native professionals through industry-aligned programmes that integrate Agentic AI into every discipline—from development to data science to enterprise platforms."

Course Highlights

Section 1: Python for AI & Data

Build strong Python programming foundations for data processing, automation, and AI applications.

Section 2: SQL for AI & Data

Design, query, optimize, and manage relational databases using advanced PostgreSQL concepts.

Section 3: Power BI for Data Analysis

Create interactive dashboards and enterprise reports using Power BI, DAX, and data modeling.

Section 4: Data Engineering with MS Fabric

Build scalable data pipelines, lakehouses, real-time analytics, and AI-enabled data platforms using Microsoft Fabric.

Section 5: Math & Stats for AI & Data

Develop mathematical, probabilistic, and statistical foundations required for data science and machine learning.

Section 6: Python Libraries for AI & Data

Analyze and visualize data efficiently using NumPy, Pandas, and modern visualization libraries.

Section 7: Machine Learning

Build, evaluate, and deploy machine learning models across supervised, unsupervised, and reinforcement learning.

Section 8: Deep Learning & NLP

Create intelligent models using neural networks and language systems.

Section 9: Generative AI & Agentic AI

Build advanced AI systems using LLMs and agent workflows.

Section 1: Python for AI & Data

MODULE 1

Python Fundamentals

Begin your Python journey with comprehensive fundamentals covering everything from environment setup to control flow. This module establishes the essential programming skills needed for AI and data science applications.

Environment & Syntax

- Python interpreter installation for Windows and Mac
- Visual Studio Code IDE configuration
- 35 essential Python keywords
- Identifiers and naming conventions

Data Types & Operations

- Variables and memory management
- Simple and complex data types
- Type conversion and casting
- Arithmetic, comparison, and logical operators

Control Flow

- Conditional statements (if, elif, else, match-case)
- While and for loops with range() function
- Break, continue, and pass statements
- User input with input() function

MODULE 2

String Manipulation

Master Python's powerful string manipulation capabilities, essential for text processing in AI and data applications. Learn indexing, slicing, formatting, and the comprehensive suite of string methods.

Core String Operations

- Positive and negative indexing techniques
- Advanced slicing with start:end:step syntax
- String concatenation and repetition
- Modern f-strings and format() methods

String Methods

- Case conversion (upper, lower, title, capitalize)
- Search operations (find, index, count)
- Validation methods (isalpha, isdigit, isalnum)
- Trimming (strip, lstrip, rstrip)
- Replacement and split/join operations

String Immutability: Understanding that strings cannot be modified in place is crucial for efficient Python programming and memory management.

MODULE 3

Data Structures: Lists & Tuples

Explore Python's fundamental collection types that form the backbone of data manipulation. Lists provide flexible, mutable sequences whilst tuples offer immutable alternatives for data integrity.

Lists

Mutable sequences supporting dynamic operations

- Creation, indexing, and slicing
- Adding elements (append, insert, extend)
- Removing elements (remove, pop, clear)
- Sorting and reversing
- List comprehensions for elegant code

Tuples

Immutable sequences for data protection

- Creation and basic operations
- Tuple packing and unpacking
- Performance advantages
- Use cases for immutability
- Comparison with lists

> List comprehensions provide a concise way to create lists, combining the power of loops and conditionals in a single elegant expression.

MODULE 4

Data Structures: Dictionaries & Sets

Dictionaries

Key-value pairs providing fast lookups and flexible data organisation. Essential for representing structured data in AI applications.

- Creation and access patterns
- Keys, values, and items methods
- Dictionary comprehensions
- Nested dictionary structures
- Practical applications in data processing

Sets

Unordered collections with unique elements, perfect for mathematical operations and data deduplication.

- Set creation and UUU properties (Unordered, Unique, Unindexed)
- Union, intersection, and difference operations
- Subset and superset relationships
- Frozen sets for immutability
- Real-world applications

MODULE 5

Advanced Collections & Iterators

Elevate your Python skills with advanced collection types and iteration patterns. Master memory-efficient techniques essential for processing large datasets in AI applications.

Collections Module

Specialized container datatypes: namedtuple for readable tuples, Counter for counting, defaultdict for default values, and deque for efficient queue operations.

Iterators & Generators

Master iteration protocols, create custom iterators, and leverage generators with yield statements for memory-efficient data processing pipelines.

Functional Programming

Lambda functions, higher-order functions (map, filter, reduce), and functional programming concepts for elegant, concise code.

MODULE 6

Functions & Scope

Functions are the building blocks of reusable code. Master function definition, parameter handling, scope management, and advanced concepts like recursion and lambda expressions.

1

Function Basics

Function definition, calling conventions, parameters versus arguments, and return statements including multiple return values.

2

Parameter Types

Positional arguments, keyword arguments, default arguments, arbitrary positional arguments (*args), and arbitrary keyword arguments (**kwargs).

3

Scope Management

Local and global scope, global keyword usage, built-in versus user-defined functions, and proper documentation with docstrings.

4

Advanced Concepts

Lambda functions for inline operations, immediately invoked function expressions (IIFE), and recursive functions for elegant problem-solving.

MODULE 7

Modules & Packages

Organise and reuse code effectively through Python's module and package system. Learn to create, import, and manage code libraries for scalable AI projects.

Module Types

- Built-in modules (math, random, datetime, os, sys)
- User-defined modules
- External packages via pip

Package Management

- Package structure and __init__.py files
- Nested packages for complex projects
- pip package manager and PyPI
- requirements.txt for dependency management
- Popular packages: requests, pandas, numpy

Proper module organisation is crucial for maintaining large AI projects and collaborating with teams.

MODULE 8

Working with Data Formats

Data persistence and exchange are fundamental to AI applications. Master file operations, CSV processing, and JSON handling for seamless data workflows.

File Operations

CRUD operations with open() function, reading methods (read, readline, readlines), writing and appending, plus directory management with os and shutil modules.

CSV Processing

Working with tabular data using csv.reader and csv.writer for basic operations, plus csv.DictReader and csv.DictWriter for dictionary-based workflows.

JSON Handling

Data serialisation and deserialisation with JSON operations: dump and dumps for writing, load and loads for reading structured data.

MODULE 9

Advanced Python Concepts

Elevate your Python expertise with exception handling, decorators, generators, and context managers. These advanced patterns are essential for writing robust, production-ready AI code.

1 Exception Handling

Master try-except-else-finally blocks, catch specific exceptions, raise and re-raise exceptions, create custom exception classes, and understand built-in exception types.

2 Decorators

Function decorators for code enhancement, decorators with arguments, applying multiple decorators, class decorators, and practical applications in logging and timing.

3 Generators Deep Dive

Generator expressions for memory efficiency, infinite generators for streaming data, and advanced iteration patterns for large datasets.

4 Context Managers

Understanding context managers for resource management, creating custom context managers with __enter__ and __exit__ methods.

MODULE 10

Object-Oriented Programming

Object-oriented programming is the cornerstone of modern software development. Master classes, objects, and the four pillars of OOP to build scalable AI systems.

Encapsulation

Data hiding through access modifiers (public, protected, private), controlling access to class attributes and methods for data integrity.

Polymorphism

Method overriding for different behaviours, duck typing for flexible interfaces, and runtime polymorphism for dynamic systems.

Inheritance

Code reuse through single, multi-level, and multiple inheritance patterns. Method overriding and super() function for extending functionality.

Abstraction

Abstract classes and methods using ABC module, defining interfaces and contracts for consistent implementation across classes.

Core Concepts

- Classes and objects fundamentals
- __init__ constructor method
- Instance, class, and static methods
- Special methods (__str__, __repr__, __len__)

Advanced Topics

- Instance versus class variables
- @classmethod and @staticmethod decorators
- Abstract base classes
- Real-world OOP applications

Section 2: SQL for AI & Data

MODULE 1

Foundations of Databases & PostgreSQL

Build a solid foundation in relational databases and PostgreSQL. Understand ACID properties, database objects, and constraints essential for data integrity in AI applications.

Database Fundamentals

- Introduction to DBMS and RDBMS
- ACID properties: Atomicity, Consistency, Isolation, Durability
- PostgreSQL installation across platforms
- Tools: psql command-line and pgAdmin 4

Database Objects

- Databases, schemas, and tables
- Data types: numeric, character, date/time, boolean, special
- Creating and managing database structures
- INSERT operations and data population

Constraints & Integrity

- PRIMARY KEY and FOREIGN KEY
- UNIQUE, NOT NULL, CHECK
- DEFAULT values
- Referential integrity enforcement

MODULE 2

Querying and Analysing Data

Master the art of data retrieval and analysis with comprehensive SQL querying techniques. From basic SELECT statements to advanced window functions and complex joins.

Query Fundamentals

- SELECT statements and column aliases
- WHERE clause filtering
- Comparison and logical operators
- BETWEEN, IN, LIKE operators
- NULL handling (IS NULL, IS NOT NULL)
- ORDER BY sorting
- DISTINCT for removing duplicates
- LIMIT and OFFSET for pagination

Advanced Techniques

- String, numeric, and date/time functions
- Aggregate functions (COUNT, SUM, AVG, MIN, MAX)
- GROUP BY and HAVING clauses
- Window functions (ROW_NUMBER, RANK, LAG, LEAD)

1

INNER JOIN

Returns matching rows from both tables

2

LEFT/RIGHT JOIN

Includes all rows from one table with matches from another

3

FULL OUTER JOIN

Combines all rows from both tables

4

CROSS JOIN

Cartesian product of both tables

MODULE 3

Advanced Queries & Data Manipulation

Advance your SQL expertise with subqueries, CTEs, set operators, and transaction management. Essential skills for complex data manipulation in AI pipelines.

Subqueries & CTEs

Subqueries in WHERE, SELECT, and FROM clauses. Correlated subqueries, EXISTS/NOT EXISTS. Common Table Expressions (CTEs) and recursive CTEs for hierarchical data.

Set Operations

UNION and UNION ALL for combining results, INTERSECT for common rows, EXCEPT for difference operations. Understanding set theory in SQL.

Data Modification

UPDATE statements with expressions and JOINs, DELETE operations with subqueries, TRUNCATE versus DELETE comparison for data removal.

Transaction Management

BEGIN, COMMIT, ROLLBACK operations, savepoints for partial rollbacks, transaction isolation levels, and concurrency control mechanisms.

MODULE 4

Database Programming & Automation

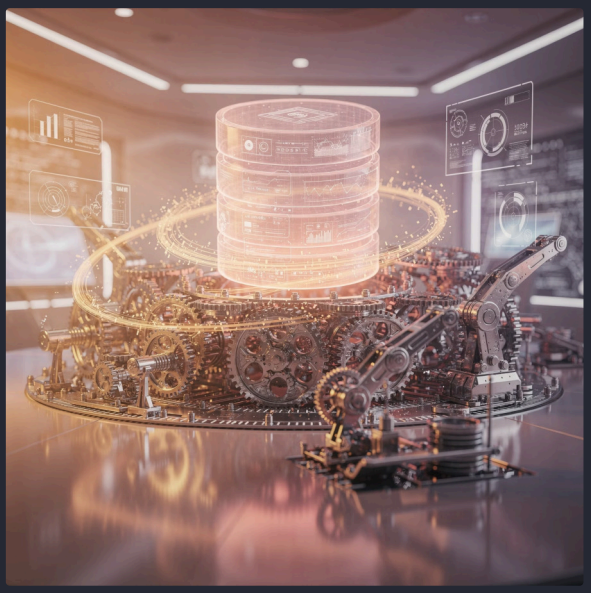
Transform your database skills with programming capabilities. Master stored functions, procedures, triggers, and PL/pgSQL for automated, efficient database operations.

Schema Management

- ALTER TABLE operations
- Adding, modifying, and dropping columns
- Managing constraints dynamically
- Index types and creation strategies

Views & Abstraction

- Creating and managing views
- Updatable views
- Materialized views for performance
- Refreshing materialized views



PL/pgSQL: PostgreSQL's procedural language enables complex logic, control structures, and exception handling directly in the database.

Stored Functions

PL/pgSQL programming language, function parameters and return types, control structures (IF, CASE, LOOP), functions returning tables.

Stored Procedures

Procedure versus function differences, exception handling in PL/pgSQL, transaction control within procedures.

Triggers

BEFORE, AFTER, INSTEAD OF triggers, trigger functions, audit logging, data validation, and advisory locks.

MODULE 5

Database Design & Optimisation

Design robust, efficient databases with proper modelling, normalisation, and optimisation techniques. Essential knowledge for scalable AI data infrastructure.

ER Modelling

Entities, attributes, and relationships. Relationship types (1:1, 1:M, M:N). Creating comprehensive ER diagrams.

Best Practices

Naming conventions, data type selection, primary and foreign key strategies, design patterns.

Normalisation

First, Second, and Third Normal Forms (1NF, 2NF, 3NF). Benefits and trade-offs. When to denormalise.

Optimisation

EXPLAIN and EXPLAIN ANALYZE, execution plans, index strategies, query rewriting, performance tuning.

Performance Tuning

- Database statistics with ANALYZE
- VACUUM and maintenance operations
- Connection pooling strategies
- Table partitioning techniques

Query Optimisation

- Reading execution plans
- Index selection strategies
- Query rewriting techniques
- Performance monitoring

Section 3: PowerBI for Data Analysis

Module 1: Introduction to Business Intelligence & Power BI

Business intelligence fundamentals and modern analytics. Power BI components, architecture, interface navigation, and first report creation. Understanding Desktop versus Service capabilities.

Module 2: Connecting to Data Sources

File, database, cloud, and web source connectivity. Import versus DirectQuery versus Live Connection modes. Data source settings, credential management, and performance considerations.

MODULES 3-4

Data Preparation & Modelling

Module 3: Power Query

Master data transformation with Power Query's intuitive interface and applied steps methodology.

- Power Query interface and applied steps
- Data profiling and quality assessment
- Essential transformations: filtering, splitting, merging
- Reshaping: pivot, unpivot, grouping
- Combining queries: append and merge operations

Module 4: Data Modelling

Build robust data models with proper relationships and hierarchies for optimal performance.

- Star schema versus snowflake schema design
- Creating and managing table relationships
- Primary and foreign keys
- Hierarchies and date dimension tables
- Data model optimisation strategies

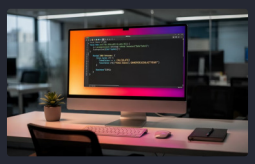
MODULES 5-6

Visualisation & DAX Fundamentals

Create compelling visual stories and implement powerful calculations with DAX. Transform raw data into actionable insights through strategic visualisation and business logic.

Module 5: Visual Reports

Data visualisation principles and chart selection. Core visualisations: charts, tables, maps, KPIs. Interactive elements: slicers, filters, bookmarks, drill-through. Dashboard layout, mobile optimisation, and storytelling with data.



Module 6: Introduction to DAX

DAX syntax and structure. Calculated columns versus measures. Essential functions: aggregation, logical, text, date/time. CALCULATE and FILTER functions. Creating KPIs and business metrics.

MODULES 7-8

Advanced Analytics & Custom Visuals

Module 7: Time Intelligence

- Time intelligence functions: YTD, MTD, QTD
- Prior period comparisons and growth rates
- Filter context versus row context
- Variables and iterator functions
- DAX performance optimisation

Module 8: Advanced Visualisations

Extend Power BI's capabilities with custom visuals and AI-powered analytics.

- Custom visuals from AppSource
- Advanced chart types: waterfall, funnel, decomposition tree
- R and Python integration
- AI visuals: Key Influencers, Q&A, Smart Narratives
- Dynamic visuals with parameters

MODULES 9-10

Publishing & Enterprise Deployment

01

Publishing & Sharing

Publishing and workspace management. Dashboards versus reports. Data refresh and gateway configuration. Sharing strategies and Power BI apps. Integration with Teams, SharePoint, Excel, PowerPoint.

02

Enterprise Governance

Power BI admin portal and tenant settings. Row-Level Security (RLS) and Object-Level Security (OLS). Incremental refresh and aggregations. Dataflows and deployment pipelines.

03

Administration

Performance optimisation and capacity management. Enterprise licensing models. APIs and embedded analytics for custom applications.

> Enterprise Power BI deployment requires careful consideration of security, governance, and performance to ensure scalable, reliable analytics across the organisation.

Section 4: Data Engineering with MS Fabric

Module 1: Introduction

Microsoft Fabric fundamentals and architecture. Fabric workloads: Data Factory, Data Engineering, Warehouse, Science, Real-Time Intelligence, Power BI. Licensing (F2-F2048), workspaces, tenant structure.

Module 2: OneLake

OneLake architecture and unified data lake concepts. Delta Lake and Parquet formats. ACID transactions and versioning. OneLake shortcuts, catalog, security, and access control.

MODULES 3-4

Lakehouse & Data Integration

Module 3: Lakehouse Architecture

Comprehensive lakehouse fundamentals combining data lake flexibility with data warehouse structure.

- Creating and managing Lakehouses
- Medallion architecture (Bronze, Silver, Gold layers)
- SQL Analytics Endpoint
- Delta table operations and optimisation
- Time travel queries

Module 4: Data Factory

Build robust data integration pipelines with comprehensive connectivity and orchestration.

- Data Factory capabilities and connectors
- Data pipelines creation and configuration
- Dataflows Gen2 and Power Query
- Database mirroring (Azure SQL, Cosmos DB, PostgreSQL)
- Pipeline orchestration and CI/CD

MODULES 5-6

Spark Engineering & Data Warehouse

Harness the power of Apache Spark for large-scale data processing and build enterprise data warehouses with T-SQL support.

1

Module 5: Apache Spark

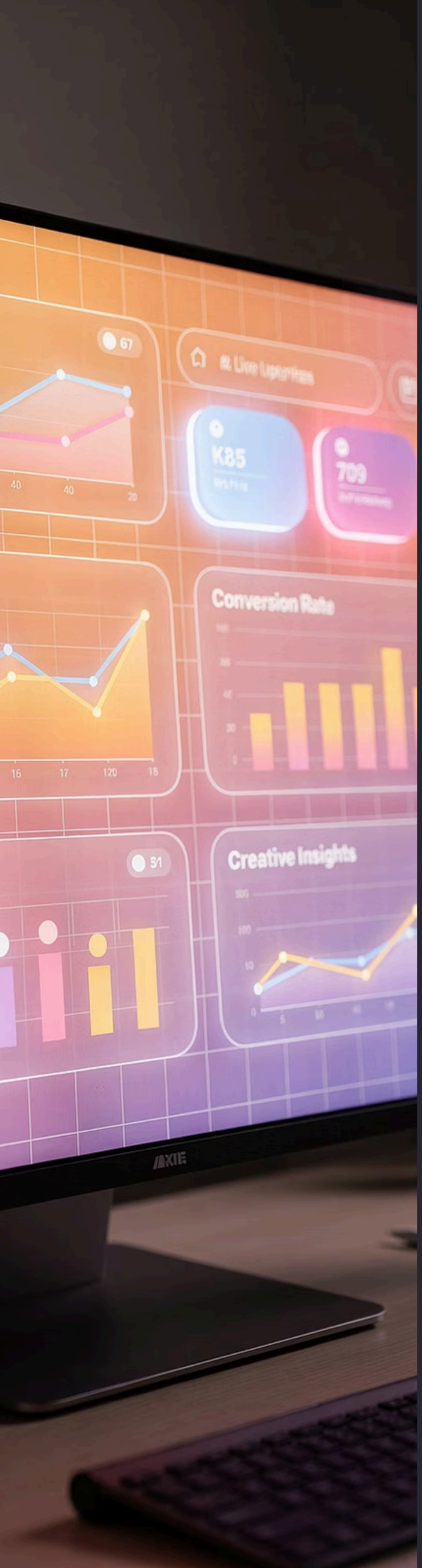
Apache Spark in Fabric, Fabric Notebooks with Copilot, PySpark DataFrames and transformations, Spark SQL queries and optimisation, Spark job definitions and scheduling, AI functions (Summarisation, Classification, PII obfuscation).

2

Module 6: Data Warehouse

Fabric Data Warehouse overview, T-SQL support and user-defined functions, schema design and table management, star schema and dimensional modelling, Slowly Changing Dimensions (SCD), SQL Database in Fabric, performance optimisation.

hello@digitaledify.ai



MODULES 7-8

Real-Time Intelligence & Power BI Integration

Module 7: Real-Time Intelligence

Process and analyse streaming data with real-time capabilities.

- Real-Time Intelligence overview
- Eventstreams and streaming sources
- Kusto Query Language (KQL)
- Eventhouse and KQL databases
- Graph in Fabric for relationship modelling
- Maps for geospatial analytics
- Real-time dashboards and alerting

Module 8: Power BI & Semantic Models

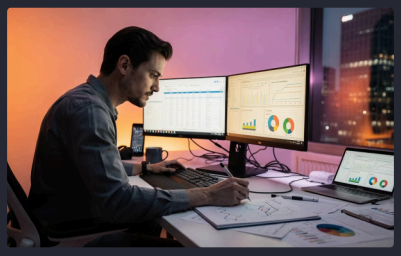
Integrate Power BI seamlessly with Fabric for advanced analytics.

- Power BI integration with Fabric
- Direct Lake mode and storage modes
- Semantic models and relationships
- DAX fundamentals and syntax
- Row-level and object-level security
- Incremental refresh and aggregations
- Report development with Copilot

hello@digitaledify.ai

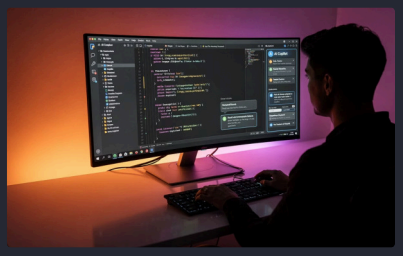
MODULES 9-10

Data Science & AI Integration



Module 9: Data Science

Data Science experience in Fabric. Exploratory data analysis. ML model training and versioning. MLflow and experiment tracking. Semantic Link (SemPy) integration. Batch scoring and predictions.



Module 10: AI & Copilot

Copilot across workloads (Notebooks, SQL, KQL, Pipelines, Reports). Data Agents for conversational AI. Operations Agents for monitoring. Fabric IQ and ontology models. AI Functions and Azure AI Foundry integration.

MODULES 11-12

User Functions & Security Governance

Module 11: User Data Functions

Create custom serverless functions for extended capabilities.

- User Data Functions overview
- Python-based serverless functions
- VS Code extension for development
- Integration with Notebooks, Pipelines, and SQL
- Testing and deployment workflows

Module 12: Security & Governance

Implement comprehensive security and governance frameworks.

- Fabric security model
- Authentication and authorisation
- Row-level and column-level security
- Dynamic data masking
- Microsoft Purview integration
- Data lineage, catalog, compliance, and auditing

MODULES 13-14

Administration & Advanced Topics

1

Module 13: Administration & Monitoring

Fabric admin portal and tenant settings. Capacity management and SKUs. Monitoring Hub and performance dashboards. Query and pipeline monitoring. Git integration and deployment pipelines. CI/CD patterns for Fabric.

2

Module 14: Advanced Topics & Best Practices

Query and performance optimisation. Partition strategies and caching. Enterprise architecture patterns. Data mesh implementation. Migration strategies. Developer tools and extensibility (Fabric CLI, APIs). Integration with Azure and third-party services.

hello@digitaledify.ai

MODULE 15

Certification Preparation

1

DP-600 Exam Structure

Comprehensive coverage of exam structure and skills measured for the Microsoft Fabric Analytics Engineer certification.

2

DP-700 Certification

Fabric Data Engineer certification path, requirements, and preparation strategies for success.

3

Practice Tests

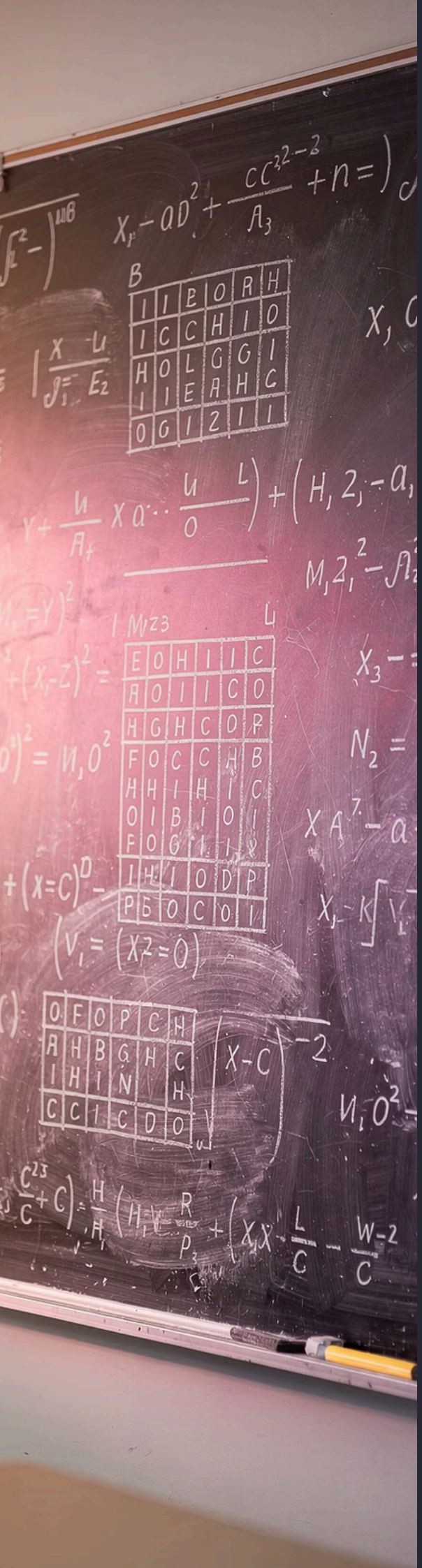
Strategic approach to practice tests, time management, and exam-taking techniques.

4

Hands-On Labs

Alignment of practical labs with exam objectives, ensuring comprehensive skill development.

Certification validates your expertise and demonstrates your ability to design, implement, and manage enterprise-scale data solutions with Microsoft Fabric.



Section 5: Maths & Stats for AI & Data

MODULE 1

Mathematical Foundations

Build the mathematical foundation essential for understanding machine learning algorithms. Master linear algebra, vectors, matrices, and eigenvalue decomposition.

Functions & Graphs

Set theory and logical operations. Functions and graphs: linear, polynomial, exponential, logarithmic. Understanding mathematical relationships.

Vectors & Matrices

Vector operations, dot product, magnitude. Matrix creation, operations, transpose, inverse, determinant. Matrix algebra and applications.

Linear Systems

Systems of linear equations. Gaussian elimination method. Eigenvalues and eigenvectors. Diagonalisation and matrix decomposition. Introduction to PCA intuition.

MODULE 2

Probability Fundamentals

Random Variables

Understanding discrete versus continuous random variables is fundamental to statistical modelling.

- Random variables (discrete vs continuous)
- Probability Mass Function (PMF)
- Probability Density Function (PDF)
- Monte Carlo simulations
- Law of Large Numbers

Core Concepts

- Basics of probability and sample spaces
- Events and probability axioms
- Conditional probability and independence
- Bayes' theorem and applications

MODULE 3

Probability Distributions

Master the probability distributions that underpin statistical inference and machine learning. From discrete distributions to the Central Limit Theorem.

Discrete Distributions

Binomial distribution for binary outcomes, Poisson distribution for rare events, Geometric distribution for waiting times. Applications in classification and counting problems.

Continuous Distributions

Normal (Gaussian) distribution as the foundation of statistics, Exponential distribution for time-based events, Uniform distribution for random sampling.

Statistical Measures

Expectation (mean) and variance. Standard deviation and higher moments. Law of Large Numbers demonstration. Central Limit Theorem. Sampling distributions and standard error.

MODULE 4

Statistical Analysis & Visualisation

Descriptive Statistics

- Mean, median, mode
- Standard deviation and variance
- Measures of spread (range, IQR)
- Outlier detection methods
- Data visualisation (histograms, boxplots, scatterplots)

Hypothesis Testing

- Hypothesis testing framework
- Null and alternative hypotheses
- p-values and statistical significance
- t-tests, chi-square test, ANOVA
- Confidence intervals
- Type I and Type II errors

MODULE 5

Applied Statistics for Machine Learning

Apply statistical concepts directly to machine learning. Master correlation, regression, and gradient descent —the optimisation algorithm powering modern AI.

Correlation Analysis

Covariance and correlation (Pearson, Spearman). Understanding correlation versus causation. Identifying relationships in data.

Linear Regression

Simple and multiple linear regression. Least squares method. R^2 and Adjusted R^2. Residual analysis for model validation.

Logistic Regression

Logistic regression for binary classification. Sigmoid function and log-odds. Probability estimation for classification tasks.

Optimisation

Cost functions and loss functions (MSE, Cross-entropy). Gradient descent algorithm. Learning rate and convergence. Batch versus stochastic gradient descent.

Section 6: Python Libraries for AI & Data

MODULE 1

NumPy: Numerical Computing

NumPy is the foundation of scientific computing in Python. Master array operations, broadcasting, and linear algebra for efficient numerical computations in AI.

Array Fundamentals

- NumPy arrays versus Python lists
- Array creation methods (zeros, ones, arange, linspace, random)
- Array indexing, slicing, and fancy indexing
- Array shapes, reshaping, and transposing

Operations & Functions

- Broadcasting rules and operations
- Universal functions (ufuncs)
- Boolean indexing and conditional operations
- Sorting and searching in arrays

Advanced Computations

- Linear algebra operations (dot products, matrix multiplication)
- Statistical operations (mean, median, std, variance, percentiles)
- Efficient numerical algorithms

MODULE 2

Pandas: Data Manipulation & Analysis

Pandas is the essential tool for data manipulation in Python. Master DataFrames, data cleaning, transformation, and analysis for AI-ready datasets.

1 Data Structures

Introduction to Series and DataFrame. Creating and reading DataFrames (CSV, Excel, JSON). Data exploration (head, tail, info, describe).

2 Data Selection

Indexing and selecting data (loc, iloc, at, iat). Boolean indexing and conditional selection. Advanced filtering techniques.

3 Data Cleaning

Handling missing data (isna, dropna, fillna). Detecting and removing duplicates. Data type conversions and string operations.

4 Transformation

Apply, map, and lambda functions. GroupBy operations and aggregations. Pivot tables and cross-tabulations. Merging, joining, and concatenating DataFrames.

Advanced Operations

- Reshaping data (stack, unstack, melt, pivot)
- Time series analysis and date/time handling
- Data input/output operations

Pandas' powerful GroupBy functionality enables complex aggregations and transformations essential for feature engineering in machine learning.

MODULE 3

Matplotlib: Static Data Visualisation

Matplotlib is Python's foundational plotting library. Master static visualisations from basic plots to complex multi-panel figures for publication-quality graphics.

Core Concepts

- Matplotlib architecture and components
- Figure and Axes objects
- Pyplot interface versus object-oriented approach
- Basic plot types (line, scatter, bar, histogram, pie)

Customisation

- Colours, markers, and line styles
- Labels, titles, and legends
- Ticks, grid lines, and axis formatting
- Annotations and text elements

Advanced Features

- Multiple subplots and grid layouts
- Multiple y-axes and error bars
- Box plots and violin plots
- Matplotlib styles and themes
- Saving figures in various formats

MODULE 4

Seaborn: Statistical Data Visualisation

Distribution Plots

Visualise data distributions with histplot, kdeplot, boxplot, and violinplot. Understand data spread and identify outliers.

Categorical Plots

Categorical data visualisation with barplot, countplot, and pointplot. Compare categories and show relationships.

Relational Plots

Explore relationships with scatterplot, lineplot, and relplot. Identify correlations and trends in data.

Matrix & Regression

Heatmaps and clustermaps for matrix visualisation. Regression plots (regplot, lmplot, residplot) for trend analysis.

Advanced Features

Pair plots and joint plots for multi-dimensional data. Facet grids for complex visualisations. Colour palettes and customisation.

MODULE 5

Plotly: Interactive Data Visualisation

Plotly brings interactivity to data visualisation with hover effects, zooming, and animations. Create engaging, web-ready visualisations for presentations and dashboards.

Plotly Fundamentals

Introduction to Plotly ecosystem (Express, Graph Objects, Dash). Plotly Express fundamentals. Interactive features (hover, zoom, pan, selection).

Chart Types

Line charts, scatter plots, and bar charts. Box plots, violin plots, and histograms. 3D scatter plots and surfaces. Geographic maps and choropleths.

Advanced Features

Animated visualisations for temporal data. Candlestick charts for financial data. Layout templates and customisation. Subplots and multi-plot dashboards. Exporting interactive plots.

Interactive visualisations transform static data into engaging stories, allowing viewers to explore insights at their own pace and discover patterns through interaction.

Section 7: Machine Learning

MODULE 1

Introduction to Machine Learning

Begin your machine learning journey with foundational concepts. Understand how machines learn from data, the types of learning, and the role of models, parameters, and optimisers.

ML Fundamentals

What is Machine Learning? AI versus ML versus Deep Learning. Understanding how machines learn from data.

Applications

Real-world applications across industries. ML environment setup (Python, Jupyter, scikit-learn).

Learning Types

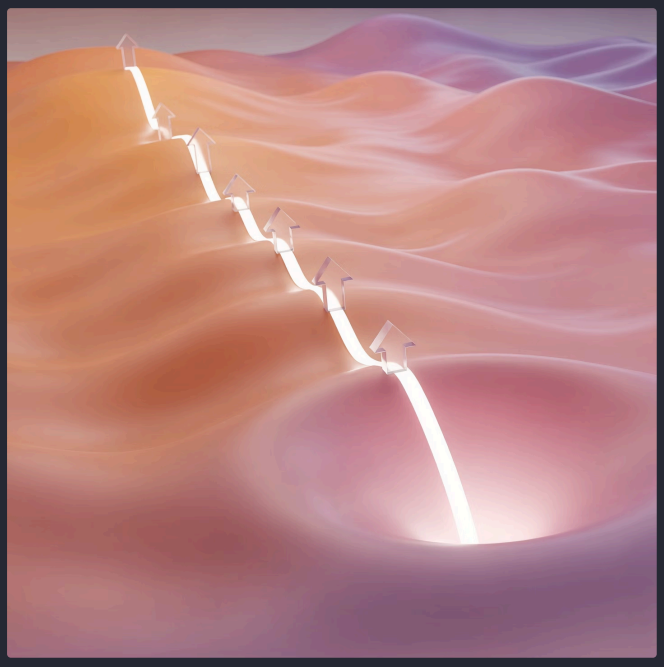
Supervised learning with labelled data. Unsupervised learning for pattern discovery. Reinforcement learning through interaction.

Model Architecture

Models, parameters, and hyperparameters. Role of optimisers in training. Model architecture and complexity.

MODULE 2

Mathematical Foundations & Optimisation



Model Learning

- Weights, biases, and feature representations
- Loss functions (MSE, MAE, Cross-Entropy, Huber Loss)
- Gradient descent fundamentals
- Learning rate and convergence criteria

Optimisation Algorithms

Advanced optimisers power modern machine learning by efficiently navigating complex loss landscapes.

- Batch versus Stochastic versus Mini-batch Gradient Descent
- Advanced optimisers (Adam, RMSprop, AdaGrad, Momentum)
- Local minima versus global minima
- Optimisation landscape visualisation

MODULE 3

ML Life Cycle & Data Preparation

Master the complete machine learning workflow from problem definition to model deployment. Data preparation is crucial—quality data leads to quality models.

01

Project Workflow

End-to-end ML project workflow (CRISP-DM). Problem definition and success metrics. Data collection and quality assessment.

02

Exploratory Analysis

Exploratory Data Analysis (EDA). Statistical analysis and data visualisation. Understanding data distributions and relationships.

03

Data Cleaning

Handling missing data and outlier detection. Data quality assessment and validation. Ensuring data integrity.

04

Feature Engineering

Feature engineering and feature selection. Data preprocessing (scaling, normalisation, standardisation). Encoding categorical variables (One-Hot, Label, Ordinal).

05

Data Splitting

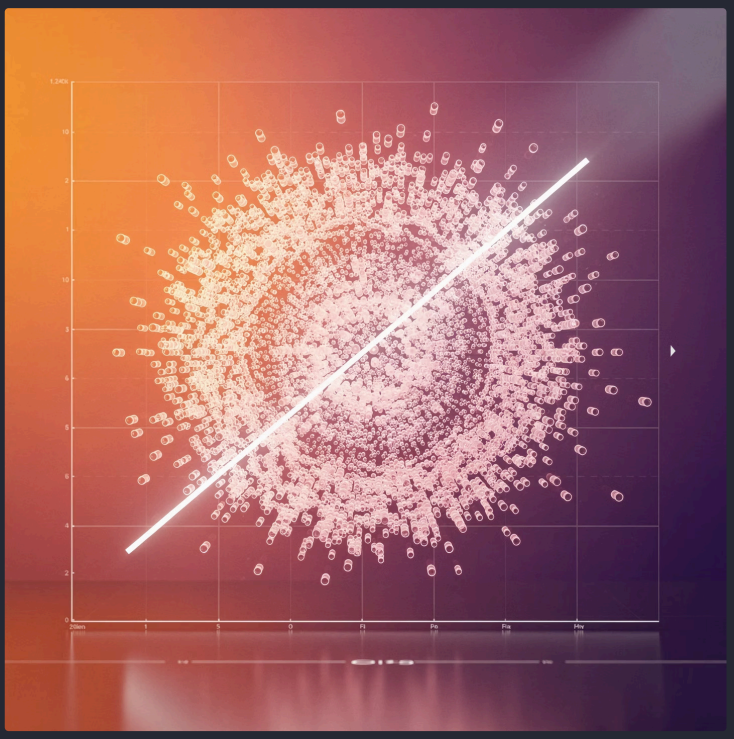
Train-test-validation split strategies. Cross-validation techniques for robust evaluation.

MODULE 4

Linear Regression & Regression Techniques

Regression Fundamentals

- Regression versus Classification
- Simple and multiple linear regression
- Cost function (MSE) and Ordinary Least Squares (OLS)
- Implementing Linear Regression with Gradient Descent



Model Evaluation

- R^2, Adjusted R^2, RMSE, MAE, MAPE
- Assumptions of linear regression
- Residual analysis and diagnostics
- Multicollinearity and VIF
- Feature scaling impact on regression

MODULE 5

Polynomial Regression & Regularisation

1

Polynomial Regression

Polynomial regression and feature transformation. Capturing non-linear relationships. Bias-Variance Tradeoff fundamentals.

2

Overfitting & Underfitting

Understanding overfitting versus underfitting. Model complexity and generalisation. Learning curves for model diagnosis.

3

Regularisation

Ridge Regression (L2) and Lasso Regression (L1). Elastic Net combining both. Regularisation parameter tuning.

4

Hyperparameter Tuning

Cross-validation for hyperparameter selection. GridSearchCV and RandomizedSearchCV for optimal parameters.

MODULE 6

Classification Algorithms I: Probabilistic Methods

Logistic Regression

Binary versus multi-class classification. Logistic regression and sigmoid function transforming linear outputs to probabilities.

- Log Loss (Binary Cross-Entropy)
- Decision boundary visualisation
- Multi-class strategies (One-vs-Rest, One-vs-One, Softmax)

Naive Bayes Classifier

Probabilistic classification based on Bayes' Theorem. Fast and effective for text classification.

- Types of Naive Bayes (Gaussian, Multinomial, Bernoulli)
- Laplace smoothing for zero probabilities
- Applications in spam detection and sentiment analysis

MODULE 7

Classification Algorithms II: Tree-Based Methods

Tree-based methods are powerful, interpretable algorithms. Master decision trees, random forests, and gradient boosting for robust classification.

1

Decision Trees

Decision tree fundamentals and structure. Entropy, Information Gain, and Gini Impurity for splitting. Tree construction algorithms (ID3, C4.5, CART). Tree pruning strategies to prevent overfitting.

2

Random Forest

Bootstrap Aggregating (Bagging) for ensemble learning. Random Forest combining multiple trees. Out-of-Bag (OOB) error estimation. Feature importance interpretation.

3

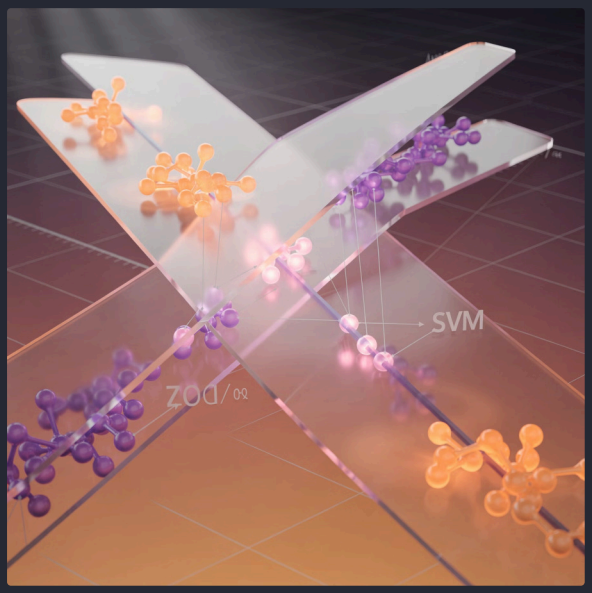
Gradient Boosting

Gradient Boosting fundamentals and sequential learning. XGBoost, LightGBM, and CatBoost implementations. Boosting versus Bagging comparison. Hyperparameter tuning for optimal performance.

hello@digitaledify.ai

MODULE 8

Support Vector Machines & Model Evaluation



Support Vector Machines

- SVM fundamentals and margin maximisation
- Support vectors and decision boundaries
- Hard margin versus Soft margin SVM
- Kernel Trick and kernel types (Linear, RBF, Polynomial, Sigmoid)

Model Evaluation

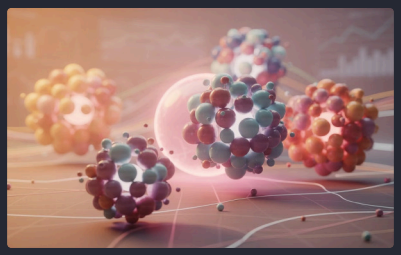
Comprehensive evaluation metrics for classification models.

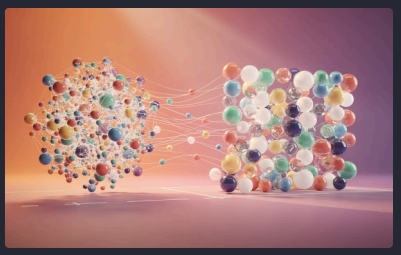
- Confusion Matrix interpretation
- Accuracy, Precision, Recall, F1-Score
- ROC curves and AUC-ROC score
- Handling imbalanced datasets
- SMOTE and sampling techniques

MODULE 9

Unsupervised Learning: Clustering & Dimensionality Reduction

Discover hidden patterns in unlabelled data through clustering and dimensionality reduction. Essential techniques for exploratory analysis and data compression.





Clustering Algorithms

K-Means Clustering and elbow method. Hierarchical Clustering and dendrograms. DBSCAN (density-based clustering). Gaussian Mixture Models (GMM). Clustering evaluation metrics.

Dimensionality Reduction

Principal Component Analysis (PCA) for feature extraction. t-SNE for visualisation. Curse of dimensionality. Feature extraction versus feature selection techniques.

MODULE 10

Reinforcement Learning & Model Deployment

Reinforcement Learning

Learn through interaction with the environment.

- RL fundamentals and paradigm
- Agent, Environment, State, Action, Reward
- Markov Decision Process (MDP)
- Q-Learning Algorithm and Bellman equation
- Exploration versus Exploitation tradeoff

Model Deployment

Take models from development to production.

- Model persistence (pickle, joblib, ONNX)
- Creating prediction APIs (Flask/FastAPI)
- Model versioning and A/B testing
- Model monitoring and drift detection
- ML best practices and ethics
- Responsible AI principles

Section 8: Deep Learning & NLP

MODULES 1-2

Neural Networks & Deep Learning Frameworks

Module 1: Neural Networks Fundamentals

Introduction to artificial neural networks and perceptrons. Activation functions (sigmoid, tanh, ReLU, and variants). Forward propagation and backpropagation algorithms. Gradient descent and optimisation techniques (SGD, Adam, RMSprop). Loss functions and performance metrics. Building ANNs from scratch in NumPy. Overfitting, underfitting, and regularisation.

Module 2: Deep Learning Frameworks

Introduction to PyTorch and TensorFlow. Tensors, computational graphs, and automatic differentiation. Building neural networks with PyTorch nn.Module. Data loading, batching, and augmentation. Training loops, validation, and checkpointing. GPU acceleration and performance optimisation. Debugging neural networks.

MODULES 3-4

Convolutional Neural Networks & Computer Vision

Module 3: CNNs

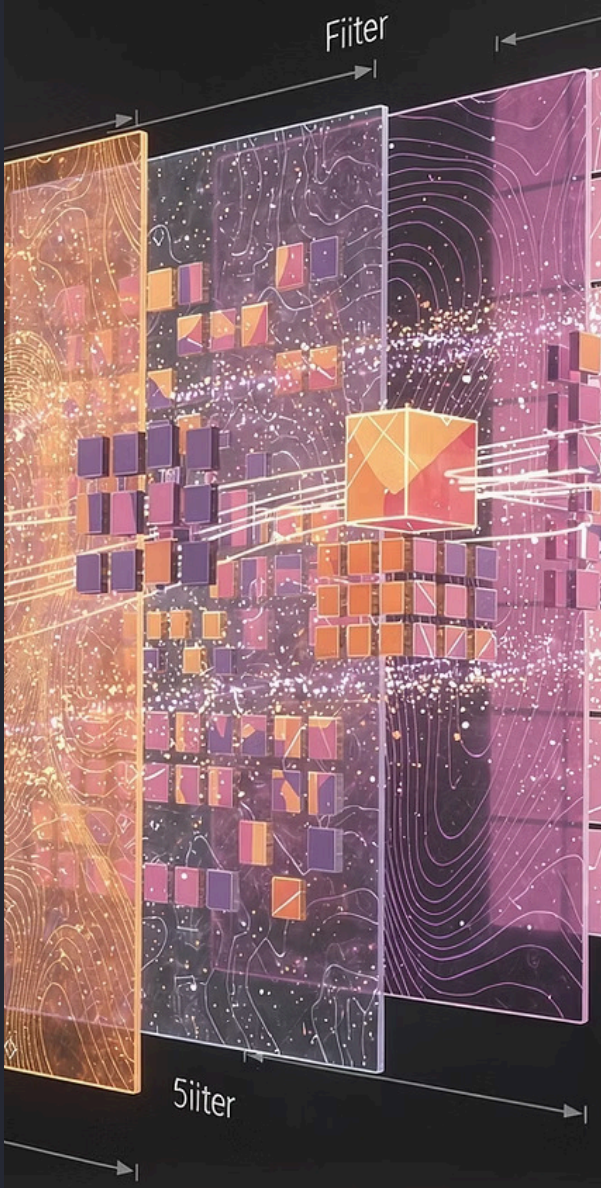
Master the architecture powering computer vision.

- Convolution operation (filters, strides, padding)
- Pooling layers and feature extraction
- Classic architectures (LeNet, AlexNet, VGG)
- Building CNNs from scratch
- Visualising CNN features and activations
- Understanding receptive fields

Module 4: Advanced CNNs

State-of-the-art architectures and techniques.

- Advanced architectures (ResNet, Inception, DenseNet)
- Batch normalisation and dropout
- Transfer learning and fine-tuning
- Data augmentation techniques
- Object detection basics (R-CNN concepts)
- Image segmentation fundamentals
- Model optimisation and compression

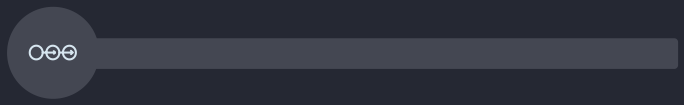


hello@digitaledify.ai

MODULES 5-6

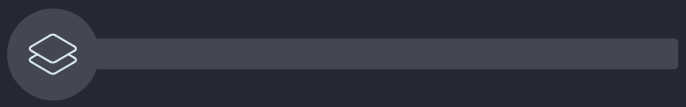
Recurrent Neural Networks & Advanced Architectures

Master sequential data processing with RNNs and LSTMs. Essential for time series, natural language, and any data with temporal dependencies.



Module 5: RNN Fundamentals

Sequential data and temporal dependencies. Vanilla RNNs and vanishing gradient problem. Backpropagation Through Time (BPTT). Long Short-Term Memory (LSTM) networks. Gated Recurrent Units (GRU). LSTM gates and cell states. Training strategies for RNNs.



Module 6: Advanced RNNs

Bidirectional RNNs for context from both directions. Stacked and deep RNN architectures. Sequence-to-sequence models. Encoder-decoder architectures. Teacher forcing and inference strategies. Handling variable-length sequences. Advanced training techniques.

MODULES 7-8

Natural Language Processing Foundations

Module 7: NLP Foundations

- Text preprocessing and tokenisation
- Bag of Words and TF-IDF
- Word embeddings (Word2Vec, GloVe, FastText)
- Training custom word embeddings
- Language modelling fundamentals
- N-gram models and neural language models
- Evaluation metrics for NLP

Module 8: Text Classification

- Text classification pipeline design
- Feature extraction for text data
- CNN for text classification
- RNN and LSTM for text classification
- Sentiment analysis techniques
- Multi-class and multi-label classification
- Handling imbalanced text datasets
- Model evaluation and error analysis

MODULES 9-10

Sequence Models & Information Extraction

Module 9: Seq2Seq Models

Sequence-to-sequence architecture deep dive. Machine translation fundamentals. Encoder-decoder models. Handling long sequences. Beam search and decoding strategies. Evaluation metrics (BLEU, ROUGE). Training strategies.

1

2

Module 10: NER & Extraction

Named Entity Recognition (NER) fundamentals. Part-of-speech tagging. Sequence labelling with RNNs and LSTMs. BiLSTM-CRF for NER. Information extraction. Relation extraction. Domain-specific NER systems. Evaluation metrics.

Section 9: Generative AI & Agentic AI

MODULES 1-2

Foundations of Generative AI

Enter the frontier of AI with large language models and generative systems. Master the fundamentals of LLMs, prompt engineering, and context design.

Module 1: LLM Foundations

- Large Language Models fundamentals
- Transformer architecture
- Comparing major LLMs (GPT, Claude, Gemini, DeepSeek)
- Evolution from GPT-1 to 2026 frontier models
- LLM architecture and tokenisation
- Model selection for different use cases
- Cost optimisation strategies

Module 2: Prompt Engineering

- Advanced prompt engineering techniques
- Context engineering and design
- Reasoning mode optimisation
- Reducing hallucinations
- Zero-shot, few-shot, and chain-of-thought prompting
- Multimodal prompting (text, image, audio)
- Domain-specific prompt design

MODULES 3-4

LLM APIs & RAG Systems

Module 3: LLM APIs & LangChain

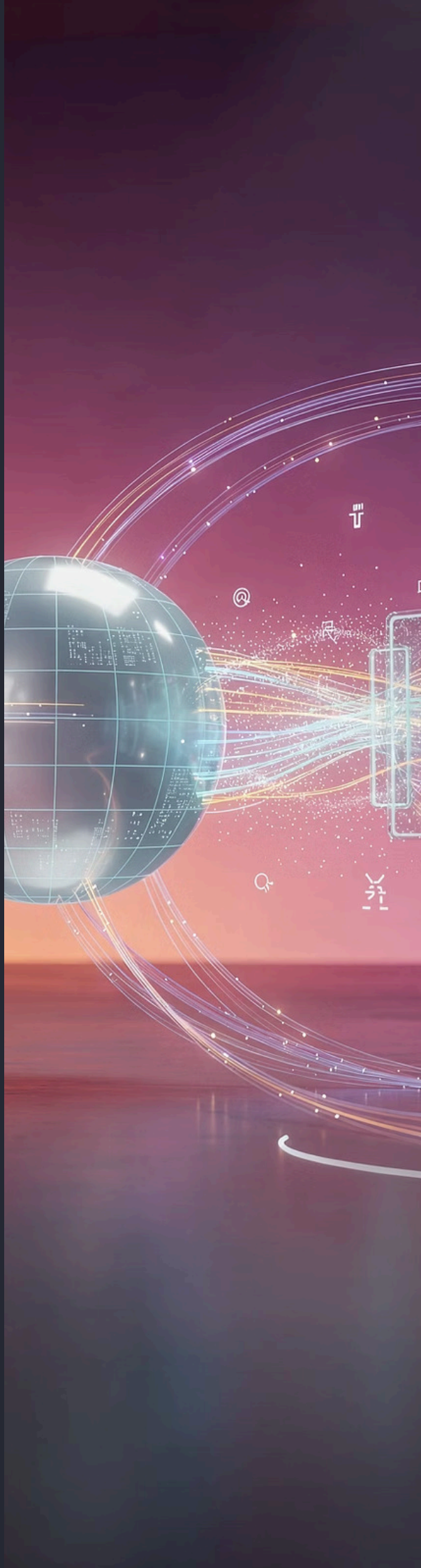
Integrate multiple LLM providers seamlessly.

- OpenAI, Anthropic, Google, and DeepSeek APIs
- LangChain 1.0 fundamentals
- Create_agent abstraction
- Middleware systems for customisation
- Multi-provider integration
- Streaming and batching
- Function calling and structured outputs
- Cost-optimised pipelines

Module 4: RAG & Vector Databases

Build production-ready retrieval-augmented generation systems.

- Vector databases (ChromaDB, Pinecone, Qdrant)
- Building production RAG pipelines
- Agentic RAG and self-improving retrieval
- MCP-Enhanced RAG
- Embedding strategies
- Hybrid search (semantic and keyword)
- Document processing at scale
- Hallucination reduction techniques



hello@digitaledify.ai

MODULES 5-6

Production Deployment & Agentic AI

1

Module 5: Production Deployment

Streamlit and Gradio interfaces. LangGraph Platform deployment. Cost optimisation strategies. AI governance and EU AI Act compliance. API security and rate limiting. Monitoring and observability. Scaling strategies. Integration with enterprise tools.

2

Module 6: Introduction to Agentic AI

Agentic AI fundamentals (plan, reason, act). LangChain 1.0 Agents with middleware. Model Context Protocol (MCP). Tool integration patterns. Enterprise adoption and use cases. Agent architectures and design patterns.

MODULES 7-8

LangGraph & Advanced Workflows

Module 7: LangGraph Fundamentals

- LangGraph 1.0 architecture
- State management and graph-based logic
- Node caching for development
- Pre/Post hooks for guardrails
- Building AI workflows
- Production use cases

Module 8: Advanced Patterns

- Parallel execution with deferred nodes
- Conditional routing and decision trees
- Iterative refinement loops
- Type-safe streaming
- Essay evaluation systems
- Customer feedback routing
- Multi-stage approval workflows
- Quality-gated content generation

MODULES 9-10

Persistence & Production Agentic Systems

Build robust, production-ready agentic systems with persistence, human-in-the-loop workflows, and comprehensive security measures.

1

Module 9: Persistence & Human-in-the-Loop

Durable state management for long-running workflows. Built-in persistence with PostgreSQL and Redis. Human-in-the-loop (HITL) implementations for critical decisions. Multi-day workflow support. Enterprise compliance and audit trails. Restart and failure recovery mechanisms.

2

Module 10: Production Agentic Systems

LangGraph Platform deployment at scale. Multi-agent system design and orchestration. Google A2A Protocol for agent-to-agent communication. LangSmith observability and monitoring. MCP security model. Prompt injection prevention. Compliance and audit trails. Agent guardrails and safety measures.

> Production agentic systems require careful consideration of security, reliability, and human oversight to ensure safe, effective autonomous operation in enterprise environments.