Cluster Analysis and Classification Techniques Training Course

Professional Development Training Course Outline

Category	Research and Data Analysis	**Duration**	10 Days
Base Fee	$3,000 USD	**Delivery Mode**	Online / On-site Hybrid

Course Introduction

In today's data-driven world, understanding complex datasets is essential for businesses, researchers, and data scientists. Cluster Analysis and Classification Techniques Training Course is meticulously designed to equip participants with the latest data segmentation methods, machine learning classification models, and unsupervised learning approaches. With the growing reliance on predictive analytics, data visualization, and AI-powered decision-making, mastering these techniques allows professionals to derive actionable insights, improve customer segmentation, optimize marketing strategies, and enhance operational efficiency.

This course dives deep into hierarchical clustering, k-means clustering, decision trees, random forests, logistic regression, and support vector machines, among others. Participants will gain hands-on experience through real-world case studies and data applications across various industries. Whether you're a business analyst, data scientist, or researcher, this course provides the essential skill set to classify, group, and analyze large datasets effectively using trending tools like Python, R, and machine learning libraries.

Programme Curriculum

Cluster Analysis and Classification Techniques Training Course

Introduction

In today's data-driven world, understanding complex datasets is essential for businesses, researchers, and data scientists. Cluster Analysis and Classification Techniques Training Course is meticulously designed to equip participants with the latest data segmentation methods, machine learning classification models, and unsupervised learning approaches. With the growing reliance on predictive analytics, data visualization, and AI-powered decision-making, mastering these techniques allows professionals to derive actionable insights, improve customer segmentation, optimize marketing strategies, and enhance operational efficiency.

This course dives deep into hierarchical clustering, k-means clustering, decision trees, random forests, logistic regression, and support vector machines, among others. Participants will gain hands-on experience through real-world case studies and data applications across various industries. Whether you're a business analyst, data scientist, or researcher, this course provides the essential skill set to classify, group, and analyze large datasets effectively using trending tools like Python, R, and machine learning libraries.

Course Objectives

1. Understand the fundamentals of cluster analysis and classification algorithms.
2. Apply k-means, hierarchical clustering, and DBSCAN in real-world data.
3. Analyze data using supervised and unsupervised learning methods.
4. Evaluate model performance using confusion matrices, ROC curves, and accuracy metrics.
5. Explore dimensionality reduction techniques such as PCA.
6. Learn machine learning tools like Scikit-learn, R, and Tableau.
7. Master data preprocessing and feature engineering techniques.
8. Understand class imbalance and resampling strategies.
9. Deploy models for customer segmentation and behavior prediction.
10. Integrate AI and ML algorithms for advanced data classification.
11. Use visualization techniques to interpret clusters and classification outputs.
12. Interpret and validate results using statistical and graphical methods.
13. Gain proficiency in automated clustering/classification pipelines.

Target Audiences

1. Data Scientists and Machine Learning Engineers
2. Business and Marketing Analysts
3. Academic Researchers
4. Statisticians and Mathematicians
5. AI and Data Engineering Professionals
6. Government and NGO Data Officers
7. Graduate Students in STEM fields
8. Corporate Strategy and Product Development Teams

Course Duration: 10 days

Course Modules

Module 1: Introduction to Cluster Analysis

- What is clustering?
- Types of clustering algorithms
- Key concepts: centroids, distance metrics
- Overview of unsupervised learning
- Tools used in clustering
- *Case Study: Customer segmentation for a retail chain*

Module 2: K-Means Clustering

- K-means algorithm steps
- Choosing the number of clusters (elbow method)
- Limitations and enhancements
- Practical coding in Python
- K-means++ initialization
- *Case Study: Clustering telecom customers based on usage data*

Module 3: Hierarchical Clustering

- Agglomerative vs. divisive clustering
- Dendrograms and their interpretation
- Linkage methods: single, complete, average
- Clustering with Scikit-learn
- When to use hierarchical clustering
- *Case Study: Gene expression data clustering in bioinformatics*

Module 4: DBSCAN and Density-Based Clustering

- Concept of density-based clustering
- Parameters: eps and minPts
- Noise and outlier detection
- Comparison with K-means
- Practical use cases in anomaly detection
- *Case Study: Detecting fraudulent transactions in financial data*

Module 5: Classification Fundamentals

- Difference between classification and clustering
- Types of classification: binary, multiclass
- Evaluation metrics: precision, recall, F1-score
- Cross-validation and data splitting
- Overview of supervised learning
- *Case Study: Email spam detection system*

Module 6: Logistic Regression

- Introduction to logistic regression
- Sigmoid function and interpretation
- Binary vs. multinomial logistic regression
- Model training and testing
- ROC curve and AUC
- *Case Study: Predicting diabetes occurrence from health data*

Module 7: Decision Trees

- Structure and components
- Gini index and entropy
- Pruning and overfitting control
- Interpreting decision paths
- Implementing in Python and R
- *Case Study: Loan approval prediction in banking*

Module 8: Random Forests

- Ensemble learning concept
- How random forests work
- Feature importance
- Model tuning and optimization
- Benefits over decision trees
- *Case Study: Customer churn classification*

Module 9: Support Vector Machines (SVM)

- Understanding hyperplanes and margins

- Kernel functions
- SVM in high-dimensional spaces
- Practical coding examples
- Pros and cons of SVM
- *Case Study: Image classification in medical diagnostics*

Module 10: Naïve Bayes Classifier

- Bayes' Theorem fundamentals
- Types of Naïve Bayes models
- Application in text classification
- Handling categorical data
- Real-life applications
- *Case Study: Sentiment analysis on product reviews*

Module 11: Neural Networks for Classification

- Perceptron and multilayer networks
- Activation functions
- Backpropagation basics
- Deep learning vs traditional ML
- Implementing with TensorFlow/Keras
- *Case Study: Predicting cancer types from biopsy images*

Module 12: Dimensionality Reduction

- Importance in clustering/classification
- Principal Component Analysis (PCA)
- t-SNE for visualization
- Choosing optimal number of dimensions
- Integrating with ML pipelines
- *Case Study: Reducing noise in financial market analysis*

Module 13: Evaluation and Model Optimization

- Bias-variance tradeoff
- Grid search and hyperparameter tuning
- Cross-validation strategies
- Interpreting results and pitfalls
- Model deployment checklist
- *Case Study: Improving accuracy of a medical diagnosis model*

Module 14: Real-World Applications

- Business intelligence
- Healthcare analytics
- Marketing and customer insight
- Cybersecurity and anomaly detection
- Industry use cases
- *Case Study: Predictive maintenance in manufacturing systems*

Module 15: Capstone Project

- Overview of end-to-end workflow
- Project selection guidelines
- Data cleaning and preparation

- Model building and evaluation
- Report writing and presentation
- *Case Study: Final project on public dataset (e.g., UCI ML repository)*

Training Methodology

- Interactive lectures with live coding and visual aids
- Guided hands-on labs and simulation exercises
- Case-based learning using real-world industry scenarios
- Group-based problem-solving sessions
- Self-paced quizzes and weekly assignments
- Capstone project for practical experience and evaluation

Register as a group from 3 participants for a Discount

Send us an email: info@fineskilltrainingcenter.org or call *+254769199797*

Certification

Upon successful completion of this training, participants will be issued with a globally- recognized certificate.

Tailor-Made Course

We also offer tailor-made courses based on your needs.

Key Notes

a. The participant must be conversant with English.

b. Upon completion of training the participant will be issued with an Authorized Training Certificate

c. Course duration is flexible and the contents can be modified to fit any number of days.

d. The course fee includes facilitation training materials, 2 coffee breaks, buffet lunch and A Certificate upon successful completion of Training.

e. One-year post-training support Consultation and Coaching provided after the course.

f. Payment should be done at least a week before commence of the training, to Fineskill Training Center account, as indicated in the invoice so as to enable us prepare better for you.

Upcoming Scheduled Sessions

Start Date	End Date	Location	Price
23 Nov 2026	04 Dec 2026	Online	$2,000
23 Nov 2026	04 Dec 2026	Physical	$3,000
23 Nov 2026	04 Dec 2026	Physical	$6,000

Start Date	End Date	Location	Price
23 Nov 2026	04 Dec 2026	Physical	$6,000
23 Nov 2026	04 Dec 2026	Physical	$12,000
23 Nov 2026	04 Dec 2026	Physical	$9,000
23 Nov 2026	04 Dec 2026	Physical	$12,000
23 Nov 2026	04 Dec 2026	Physical	$12,000

Ready to enroll or request a customized program? Book online or contact us:

Register Online Now

Email: info@fineskilltrainingcenter.com | Website: www.fineskilltrainingcenter.com