Advanced Text Mining and Information Extraction Training Course

Professional Development Training Course Outline

Category	Research and Data Analysis	**Duration**	5 Days
Base Fee	$1,500 USD	**Delivery Mode**	Online / On-site Hybrid

Course Introduction

In today’s data-driven world, unstructured data comprises over 80% of all data generated, creating a growing demand for professionals skilled in Advanced Text Mining and Information Extraction (IE) techniques. Advanced Text Mining and Information Extraction Training Course is designed to equip participants with cutting-edge skills in natural language processing (NLP), semantic analysis, machine learning, and automated knowledge discovery, ensuring the ability to extract actionable insights from massive textual datasets. Whether dealing with academic literature, social media content, or enterprise documents, this course enables you to master the tools and techniques necessary for accurate, scalable, and intelligent text analytics.

Participants will explore trending technologies like deep learning for NLP, entity recognition, relationship extraction, sentiment analysis, and topic modeling. With real-world case studies across sectors such as finance, healthcare, and governance, learners will gain hands-on experience in deploying advanced text mining models using Python, R, and open-source NLP frameworks. The training integrates theory and practical skills for transforming raw text into structured, meaningful knowledge, aligning with current demands in data science and AI-driven analytics.

Programme Curriculum

Advanced Text Mining and Information Extraction Training Course

Introduction

In today’s data-driven world, unstructured data comprises over 80% of all data generated, creating a growing demand for professionals skilled in Advanced Text Mining and Information Extraction (IE) techniques. Advanced Text Mining and Information Extraction Training Course is designed to equip participants with cutting-edge skills in natural language processing (NLP), semantic analysis, machine learning, and automated knowledge discovery, ensuring the ability to extract actionable insights from massive textual datasets. Whether dealing with academic literature, social media content, or

enterprise documents, this course enables you to master the tools and techniques necessary for accurate, scalable, and intelligent text analytics.

Participants will explore trending technologies like deep learning for NLP, entity recognition, relationship extraction, sentiment analysis, and topic modeling. With real-world case studies across sectors such as finance, healthcare, and governance, learners will gain hands-on experience in deploying advanced text mining models using Python, R, and open-source NLP frameworks. The training integrates theory and practical skills for transforming raw text into structured, meaningful knowledge, aligning with current demands in data science and AI-driven analytics.

Course Objectives

1. Understand the fundamentals of text mining and natural language processing.
2. Apply machine learning techniques for intelligent text classification.
3. Perform named entity recognition (NER) and relationship extraction.
4. Utilize sentiment analysis to derive opinion-based insights.
5. Conduct topic modeling using LDA and other algorithms.
6. Implement deep learning models for advanced NLP tasks.
7. Leverage Python and R libraries for scalable text processing.
8. Extract structured information from unstructured sources.
9. Analyze social media and real-time text streams.
10. Evaluate text mining model performance and accuracy.
11. Deploy information extraction pipelines in real-world settings.
12. Interpret results for business intelligence and decision-making.
13. Build domain-specific applications using text analytics.

Target Audiences

1. Data Scientists
2. AI/Machine Learning Engineers
3. Business Intelligence Analysts
4. Academic Researchers
5. Policy Analysts
6. Software Developers
7. Journalists & Media Analysts
8. Public Sector & NGO Researchers

Course Duration: 5 days

Course Modules

Module 1: Introduction to Text Mining and NLP

- Overview of unstructured data challenges
- Core concepts of NLP and linguistic preprocessing
- Text normalization, tokenization, stemming, and lemmatization
- Introduction to key Python and R libraries
- Exploratory text analysis and frequency-based techniques
- **Case Study:** Analyzing customer feedback from e-commerce platforms

Module 2: Text Classification and Clustering

- Supervised vs. unsupervised learning
- Algorithms for text classification (Naïve Bayes, SVM)
- Document clustering using k-means and hierarchical clustering
- Feature engineering and vectorization (TF-IDF, word embeddings)

- Evaluating classifier performance (precision, recall, F1)
- **Case Study:** Spam email detection using classification models

Module 3: Named Entity Recognition and Relationship Extraction

- NER fundamentals and linguistic patterns
- Rule-based vs. statistical models
- Extracting entities, attributes, and relations
- Dependency parsing and semantic role labeling
- Tools: spaCy, Stanford NLP, OpenNLP
- **Case Study:** Extracting financial entities from SEC filings

Module 4: Sentiment Analysis and Opinion Mining

- Lexicon-based and machine learning-based approaches
- Sentiment scoring and polarity detection
- Handling sarcasm, negation, and domain-specific vocabularies
- Emotion detection from texts
- Visualizing sentiment trends
- **Case Study:** Sentiment analysis of tweets during election campaigns

Module 5: Topic Modeling and Semantic Analysis

- Introduction to Latent Dirichlet Allocation (LDA)
- Non-negative Matrix Factorization (NMF)
- Word embeddings: Word2Vec, GloVe
- Visualization of topics using pyLDAvis
- Interpreting and labeling latent topics
- **Case Study:** Topic extraction from academic research papers

Module 6: Deep Learning for NLP

- Introduction to deep learning concepts
- Word embeddings and sequence modeling
- RNNs, LSTMs, and Transformers (BERT, GPT)
- Fine-tuning pre-trained models for domain tasks
- Frameworks: TensorFlow, Hugging Face Transformers
- **Case Study:** Automating summarization of legal documents

Module 7: Social Media Mining and Real-Time Analysis

- Text mining on Twitter, Facebook, Reddit
- APIs for data collection and preprocessing
- Trend detection and event extraction
- Real-time sentiment tracking dashboards
- Ethics in social media mining
- **Case Study:** Real-time crisis monitoring using Twitter feeds

Module 8: Information Extraction Pipelines and Applications

- Building scalable IE workflows
- Integration with business intelligence systems
- Use of knowledge graphs and ontologies
- Deployment and monitoring of text mining systems
- Translating insights into business value
- **Case Study:** Implementing an IE pipeline for healthcare policy analysis

Training Methodology

- Instructor-led lectures and real-time demonstrations
- Practical hands-on coding sessions using Python and R
- Group-based project work and collaborative tasks
- Peer review and feedback sessions on case studies
- Assessments and quizzes to evaluate comprehension
- Certification upon successful course completion

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	27 Nov 2026	Online	$1,000
23 Nov 2026	27 Nov 2026	Physical	$1,500
23 Nov 2026	27 Nov 2026	Physical	$3,000
23 Nov 2026	27 Nov 2026	Physical	$3,000
23 Nov 2026	27 Nov 2026	Physical	$6,000

Start Date	End Date	Location	Price
23 Nov 2026	27 Nov 2026	Physical	$4,500
23 Nov 2026	27 Nov 2026	Physical	$6,000
23 Nov 2026	27 Nov 2026	Physical	$6,000

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

Register Online Now

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