Training Course on Feature Engineering for Machine Learning

Professional Development Training Course Outline

Category	Information Technology	**Duration**	10 Days
Base Fee	$3,000 USD	**Delivery Mode**	Online / On-site Hybrid

Course Introduction

In today's data-driven world, the success of **machine learning models** hinges not only on sophisticated algorithms but also, critically, on the quality and relevance of the input data. **Feature engineering**, the art and science of transforming raw data into meaningful and informative features, stands as a cornerstone of building high-performing predictive systems. This comprehensive training course delves deep into the essential techniques and best practices of **feature engineering**, empowering participants to extract maximum value from their datasets and significantly enhance the accuracy, efficiency, and interpretability of their **machine learning models**. By mastering **data preprocessing**, **feature scaling**, **feature selection**, and the creation of novel **feature transformations**, you will gain a competitive edge in the rapidly evolving field of **artificial intelligence** and **data science**.

This intensive program is meticulously designed to equip data scientists, machine learning engineers, and analysts with practical, hands-on experience in tackling real-world data challenges. Through a blend of theoretical foundations, engaging case studies, and practical exercises, you will learn to identify crucial **data patterns**, handle **missing values**, encode **categorical variables**, engineer **time-series features**, and effectively reduce **dimensionality**. You will also explore advanced techniques for generating **polynomial features**, interaction terms, and domain-specific features. Upon completion, you will be proficient in building robust feature pipelines that are essential for deploying successful **machine learning applications** across diverse industries.

Programme Curriculum

Training Course on Feature Engineering for Machine Learning

Introduction

In today's data-driven world, the success of **machine learning models** hinges not only on sophisticated algorithms but also, critically, on the quality and relevance of the input data. **Feature engineering**, the art and science of transforming raw data into meaningful and informative features, stands as a cornerstone of building high-performing predictive systems. This comprehensive training course delves deep into the essential techniques and best practices of **feature engineering**, empowering participants to extract maximum value from their datasets and significantly enhance the accuracy, efficiency, and interpretability of their **machine learning models**. By mastering **data preprocessing**, **feature scaling**, **feature selection**, and the creation of novel **feature transformations**, you will gain a competitive edge in the rapidly evolving field of **artificial intelligence** and **data science**.

This intensive program is meticulously designed to equip data scientists, machine learning engineers, and analysts with practical, hands-on experience in tackling real-world data challenges. Through a blend of theoretical foundations, engaging case studies, and practical exercises, you will learn to identify crucial **data patterns**, handle **missing values**, encode **categorical variables**, engineer **time-series features**, and effectively reduce **dimensionality**. You will also explore advanced techniques for generating **polynomial features**, interaction terms, and domain-specific features. Upon completion, you will be proficient in building robust feature pipelines that are essential for deploying successful **machine learning applications** across diverse industries.

Course Duration

10 days

Course Objectives

This training course aims to equip participants with the following key skills and knowledge:

1. Learn to effectively clean, handle missing data, and prepare raw data for feature engineering.
2. and apply various scaling techniques to optimize model performance.
3. Explore different methods for converting categorical data into numerical representations suitable for machine learning algorithms
4. Identify and select the most relevant features while reducing data dimensionality
5. Create insightful features from temporal data, including lag features, rolling statistics, and trend analysis
6. Discover how to create new features by combining existing ones to capture non-linear relationships.
7. Learn to leverage domain knowledge to create custom features that are highly predictive for specific applications.
8. Understand the importance of creating automated and reproducible feature engineering workflows
9. Learn to assess how different feature engineering techniques affect model performance.
10. Explore techniques for addressing class imbalance during feature engineering
11. Learn basic techniques for transforming text data into numerical features
12. Understand how feature engineering strategies vary for classification, regression, and clustering problems.
13. insights into cutting-edge methods and emerging trends in the field

Organizational Benefits

- Enhanced feature quality directly translates to more accurate and reliable machine learning models.

- Well-engineered features can lead to faster training times and reduced computational costs.
- The process of feature engineering fosters a deeper understanding of the underlying data patterns.
- More accurate models support better-informed business decisions.
- Organizations with strong feature engineering capabilities can build more sophisticated and effective AI solutions.
- By optimizing data input, the need for complex model tuning may be reduced.
- Efficient feature pipelines contribute to quicker deployment of machine learning applications.
- A strong understanding of feature engineering can unlock new possibilities for data-driven innovation.

Target Audience

This training course is ideal for individuals in the following roles:

1. Data Scientists
2. Machine Learning Engineers
3. Data Analysts
4. AI Researchers
5. Software Engineers with an interest in AI/ML
6. Business Analysts seeking to leverage data insights
7. Statisticians looking to apply their skills in machine learning
8. Students and Academics in relevant fields

Course Outline

Module 1: Introduction to Feature Engineering

- The importance of feature engineering in the machine learning lifecycle.
- Understanding the impact of data quality on model performance.
- Different types of features and their characteristics.
- The relationship between feature engineering and model interpretability.
- Setting up the development environment (Python, libraries).

Module 2: Data Exploration and Preprocessing

- Techniques for exploring and understanding datasets.
- Identifying and handling missing values (imputation strategies).
- Detecting and managing outliers in the data.
- Data cleaning and formatting best practices.
- Understanding data distributions and their implications.

Module 3: Feature Scaling and Normalization

- The need for feature scaling in machine learning algorithms.
- Standardization (Z-score normalization) and its applications.
- Min-max scaling and its use cases.
- Robust scaling for data with outliers.
- Choosing the appropriate scaling technique for different scenarios.

Module 4: Encoding Categorical Variables

- Challenges of using categorical data in machine learning.
- One-hot encoding and its advantages and disadvantages.

- Label encoding and ordinal encoding for ordered categories.
- Handling high-cardinality categorical features.
- Advanced encoding techniques (e.g., target encoding).

Module 5: Feature Selection Techniques

- The benefits of reducing the dimensionality of the data.
- Filter methods (e.g., variance thresholding, correlation analysis).
- Wrapper methods (e.g., forward selection, backward elimination).
- Embedded methods (e.g., L1 regularization).
- Choosing the right feature selection approach.

Module 6: Dimensionality Reduction Techniques

- Understanding the concept of dimensionality reduction.
- Principal Component Analysis (PCA) and its applications.
- Linear Discriminant Analysis (LDA) for classification.
- Non-linear dimensionality reduction techniques (e.g., t-SNE).
- Evaluating the impact of dimensionality reduction on model performance.

Module 7: Engineering Time-Series Features

- Working with temporal data in machine learning.
- Creating lag features and their significance.
- Calculating rolling statistics (e.g., moving average, standard deviation).
- Extracting trend and seasonality components.
- Handling time-based cross-validation.

Module 8: Generating Polynomial Features and Interaction Terms

- Capturing non-linear relationships in the data.
- Creating polynomial features of different degrees.
- Generating interaction terms between features.
- The risk of overfitting with polynomial features.
- Strategies for selecting relevant polynomial and interaction terms.

Module 9: Domain-Specific Feature Engineering

- Leveraging domain knowledge to create informative features.
- Examples of domain-specific feature engineering in different industries (e.g., finance, healthcare).
- The importance of collaboration with domain experts.
- Documenting and sharing domain-specific feature engineering insights.
- Ethical considerations in domain-specific feature engineering.

Module 10: Building Feature Pipelines

- The importance of creating automated feature engineering workflows.
- Using tools and libraries for building feature pipelines (e.g., scikit-learn Pipelines).
- Ensuring reproducibility and maintainability of feature pipelines.
- Integrating feature pipelines with machine learning models.
- Deployment considerations for feature pipelines.

Module 11: Feature Engineering for Imbalanced Datasets

- Challenges of working with imbalanced datasets.
- Techniques for oversampling the minority class (e.g., SMOTE).
- Techniques for undersampling the majority class.
- Generating synthetic samples for imbalanced learning.
- Evaluating model performance on imbalanced datasets.

Module 12: Feature Engineering for Text Data

- Introduction to Natural Language Processing (NLP) for feature engineering.
- Basic text preprocessing techniques (e.g., tokenization, stemming, lemmatization).
- Bag-of-Words (BoW) and TF-IDF representations.
- Introduction to word embeddings (e.g., Word2Vec, GloVe).
- Creating features from text data for machine learning tasks.

Module 13: Evaluating and Iterating on Features

- Metrics for evaluating the quality of engineered features.
- Analyzing the impact of features on model performance.
- Iterative feature engineering and experimentation.
- Techniques for visualizing feature importance.
- Best practices for documenting and managing engineered features.

Module 14: Advanced Feature Engineering Techniques

- Introduction to automated feature engineering tools.
- Exploring feature engineering with deep learning.
- Working with graph data for feature engineering.
- Feature engineering for interpretable machine learning.
- Emerging trends and research in feature engineering.

Module 15: Real-World Case Studies and Applications

- In-depth analysis of successful feature engineering applications across various industries.
- Discussion of challenges and lessons learned from real-world projects.
- Hands-on exercises applying feature engineering techniques to real datasets.
- Group projects focused on solving real-world machine learning problems through effective feature engineering.
- Ethical considerations in deploying machine learning models with engineered features.

Training Methodology

This course employs a blended learning approach that combines:

- **Interactive Lectures:** Engaging presentations covering the theoretical concepts and practical applications of feature engineering.
- **Hands-on Lab Sessions:** Practical exercises using Python and relevant libraries (e.g., scikit-learn, pandas) to implement feature engineering techniques on real-world datasets.
- **Case Studies:** In-depth analysis of real-world scenarios showcasing the impact of effective feature engineering.
- **Group Discussions:** Collaborative sessions to foster peer learning and the exchange of ideas.
- **Individual Assignments:** Practical tasks to reinforce learning and assess understanding.

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
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

Start Date	End Date	Location	Price
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