



CertNexus Certified Artificial Intelligence (AI) Practitioner Course Description: Exam: AIP-110

Course Content

*Orange text indicates activity

Lesson 1: Solving Business Problems Using AI and ML

Topic A: Identify AI and ML Solutions for Business Problems

- The Data Hierarchy—Making Data Useful
- Big Data
- Guidelines for Working with Big Data
- Data Mining
- Examples of Applied AI and ML in Business
- Guidelines to Select Appropriate Business Applications for AI and ML
- Identifying Appropriate Business Applications for AI and ML

Topic B: Follow a Machine Learning Workflow

- Machine Learning Model
- Machine Learning Workflow
- Data Science Skillset
- Traditional IT Skillsets
- Concept Drift
- Transfer Learning
- Guidelines for Following the Machine Learning Workflow
- Planning the Machine Learning Workflow

Topic C: Formulate a Machine Learning Problem

- Problem Formulation
- Framing a Machine Learning Problem
- Differences Between Traditional Programming and Machine Learning
- Differences Between Supervised and Unsupervised Learning
- Randomness in Machine Learning
- Uncertainty
- Random Number Generation
- Machine Learning Outcomes
- Guidelines for Formulating a Machine Learning Outcome
- Selecting a Machine Learning Outcome

Topic D: Select Appropriate Tools

- Open Source AI Tools
- Proprietary AI Tools
- New Tools and Technologies
- Hardware Requirements
- GPUs vs. CPUs
- GPU Platforms
- Cloud Platforms
- Guidelines for Configuring a Machine Learning Toolset
- How to Install Anaconda
- Selecting a Machine Learning Toolset

Lesson 2: Collecting and Refining the Dataset

Topic A: Collect the Dataset

- Machine Learning Datasets

- Structure of Data
- Terms Describing Portions of Data
- Data Quality Issues
- Data Sources
- Open Datasets
- Guidelines for Selecting a Machine Learning Dataset
- Examining the Structure of a Machine Learning Dataset
- Extract, Transform, and Load (ETL)
- Machine Learning Pipeline
- ML Software Environments
- Guidelines for Loading a Dataset
- Loading the Dataset

Topic B: Analyze the Dataset to Gain Insights

- Dataset Structure
- Guidelines for Exploring the Structure of a Dataset
- Exploring the General Structure of the Dataset
- Normal Distribution
- Non-Normal Distributions
- Descriptive Statistical Analysis
- Central Tendency
- When to Use Different Measures of Central Tendency
- Variability
- Range Measures
- Variance and Standard Deviation
- Calculation of Variance
- Variance in a Sample Set

- Calculation of Standard Deviation
- Skewness
- Calculation of Skewness Measures
- Kurtosis
- Calculation of Kurtosis
- Statistical Moments
- Correlation Coefficient
- Calculation of Pearson's Correlation Coefficient
- Guidelines for Analyzing a Dataset
- Analyzing a Dataset Using Statistical Measures

Topic C: Use Visualizations to Analyze Data

- Visualizations
- Histogram
- Box Plot
- Scatterplot
- Geographical Maps
- Heat Maps
- Guidelines for Using Visualizations to Analyze Data
- Analyzing a Dataset Using Visualizations

Topic D: Prepare Data

- Data Preparation
- Data Types
- Operations You Can Perform on Different Types of Data
- Continuous vs. Discrete Variables
- Data Encoding
- Dimensionality Reduction

- Impute Missing Values
- Duplicates
- Normalization and Standardization
- Summarization
- Holdout Method
- Guidelines for Preparing Training and Testing Data
- Splitting the Training and Testing Datasets and Labels

Lesson 3: Setting Up and Training a Model

Topic A: Set Up a Machine Learning Model

- Design of Experiments
- Hypothesis
- Hypothesis Testing
- Hypothesis Testing Methods
- p-value
- Confidence Interval
- Machine Learning Algorithms
- Algorithm Selection
- Guidelines for Setting Up a Machine Learning Model
- Setting Up a Machine Learning Model

Topic B: Train the Model

- Iterative Tuning
- Bias
- Compromises
- Model Generalization
- Cross-Validation

- k-Fold Cross-Validation
- Leave-p-Out Cross-Validation
- Dealing with Outliers
- Feature Transformation
- Transformation Functions
- Scaling and Normalizing Features
- The Bias–Variance Tradeoff
- Parameters
- Regularization
- Models in Combination
- Processing Efficiency
- Guidelines for Training and Tuning the Model
- Refitting and Testing the Model

Lesson 4: Finalizing a Model

Topic A: Translate Results into Business Actions

- Know Your Audience
- Visualization for Presentation
- Guidelines for Presenting Your Findings
- Translating Results into Business Actions

Topic B: Incorporate a Model into a Long-Term Business Solution

- Put a Model into Production
- Production Algorithms
- Pipeline Automation
- Testing and Maintenance
- Consumer-Oriented Applications

- Guidelines for Incorporating Machine Learning into a Long-Term Solution
- Incorporating a Model into a Long-Term Solution

Lesson 5: Building Linear Regression Models

Topic A: Build a Regression Model Using Linear Algebra

- Linear Regression
- Linear Equation
- Linear Equation Data Example
- Straight Line Fit to Example Data
- Linear Equation Shortcomings
- Linear Regression in Machine Learning
- Linear Regression in Machine Learning Example
- Matrices in Linear Regression
- Normal Equation
- Linear Model with Higher Order Fits
- Linear Model with Multiple Parameters
- Cost Function
- Mean Squared Error (MSE)
- Mean Absolute Error (MAE)
- Coefficient of Determination
- Normal Equation Shortcomings
- Guidelines for Building a Regression Model Using Linear Algebra
- Building a Regression Model Using Linear Algebra

Topic B: Build a Regularized Regression Model Using Linear Algebra

- Regularization Techniques
- Ridge Regression

- Lasso Regression
- Elastic Net Regression
- Guidelines for Building a Regularized Linear Regression Model
- Building a Regularized Linear Regression Model

Topic C: Build an Iterative Linear Regression Model

- Iterative Models
- Gradient Descent
- Global Minimum vs. Local Minima
- Learning Rate
- Gradient Descent Techniques
- Guidelines for Building an Iterative Linear Regression Model
- Building an Iterative Linear Regression Model

Lesson 6: Building Classification Models

Topic A: Train Binary Classification Models

- Linear Regression Shortcomings
- Logistic Regression
- Decision Boundary
- Cost Function for Logistic Regression
- A Simpler Alternative for Classification
- k-Nearest Neighbor (k-NN)
- k Determination
- Logistic Regression vs. k-NN
- Guidelines for Training Binary Classification Models
- Training Binary Classification Model

Topic B: Train Multi-Class Classification Models

- Multi-Label Classification
- Multi-Class Classification
- Multinomial Logistic Regression
- Guidelines for Training Multi-Class Classification Models
- Training a Multi-Class Classification Model

Topic C: Evaluate Classification Models

- Model Performance
- Confusion Matrix
- Classifier Performance Measurement
- Accuracy
- Precision
- Recall
- Precision–Recall Tradeoff
- F1 Score
- Receiver Operating Characteristic (ROC) Curve
- Thresholds
- Area Under Curve (AUC)
- Precision–Recall Curve (PRC)
- Guidelines for Evaluating Classification Models
- Evaluating a Classification Model

Topic D: Tune Classification Models

- Hyperparameter Optimization
- Grid Search
- Randomized Search
- Bayesian Optimization
- Genetic Algorithms

- Guidelines for Tuning Classification Models
- Tuning a Classification Model

Lesson 7: Building Clustering Models

Topic A: Build k-Means Clustering Models

- k-Means Clustering
- Global vs. Local Optimization
- k Determination
- Elbow Point
- Cluster Sum of Squares
- Silhouette Analysis
- Additional Cluster Analysis Methods
- Guidelines for Building a k-Means Clustering Model
- Building a k-Means Clustering Model

Topic B: Build Hierarchical Clustering Models

- k-Means Clustering Shortcomings
- Hierarchical Clustering
- Hierarchical Clustering Applied to a Spiral Dataset
- When to Stop Hierarchical Clustering
- Dendrogram
- Guidelines for Building a Hierarchical Clustering Model
- Building a Hierarchical Clustering Model

Lesson 8: Building Advanced Models

Topic A: Build Decision Tree Models

- Decision Tree
- Classification and Regression Tree (CART)

- Gini Index Example
- CART Hyperparameters
- Pruning
- C4.5
- Continuous Variable Discretization
- Bin Determination
- One-Hot Encoding
- Decision Tree Algorithm Comparison
- Decision Trees Compared to Other Algorithms
- Guidelines for Building a Decision Tree Model
- Building a Decision Tree Model

Topic B: Build Random Forest Models

- Ensemble Learning
- Random Forest
- Out-of-Bag Error
- Random Forest Hyperparameters
- Feature Selection Benefits
- Guidelines for Building a Random Forest Model
- Building a Random Forest Model

Lesson 9: Building Support-Vector Machines

Topic A: Build SVM Models for Classification

- Support-Vector Machines (SVMs)
- SVMs for Linear Classification
- Hard-Margin Classification
- Soft-Margin Classification
- SVMs for Non-Linear Classification

- Kernel Trick
- Kernel Trick Example
- Kernel Methods
- Guidelines for Building an SVM Model
- Building an SVM Model

Topic B: Build SVM Models for Regression

- SVMs for Regression
- Guidelines for Building SVM Models for Regression
- Building an SVM Model for Regression

Lesson 10: Building Artificial Neural Networks

Topic A: Build Multi-Layer Perceptrons (MLP)

- Artificial Neural Network (ANN)
- Perceptron
- Multi-Label Classification Perceptron
- Perceptron Training
- Perceptron Shortcomings
- Multi-Layer Perceptron (MLP)
- ANN Layers
- Backpropagation
- Activation Functions
- Guidelines for Building MLPs
- Building an MLP

Topic B: Build Convolutional Neural Networks (CNN)

- Traditional ANN Shortcomings
- Convolutional Neural Network (CNN)

- CNN Filters
- CNN Filter Example
- Padding
- Stride
- Pooling Layer
- CNN Architecture
- Generative Adversarial Network (GAN)
- GAN Architecture
- Guidelines for Building CNNs
- Building a CNN

Lesson 11: Promoting Data Privacy and Ethical Practices

Topic A: Protect Data Privacy

- Protected Data
- Obligation to Protect PII
- Relevant Data Privacy Laws
- Privacy by Design
- Data Privacy Principles at Odds with Machine Learning
- Guidelines for Complying with Data Privacy Laws and Standards
- Complying with Applicable Laws and Standards
- Open Source Data Sharing and Privacy
- Data Anonymization
- Guidelines for Data Anonymization
- The Big Data Challenge
- Guidelines for Protecting Data Privacy
- Protecting Data Privacy

Topic B: Promote Ethical Practices

- Preconceived Notions
- The Black Box Challenge
- Prejudice Bias
- Proxies for Larger Social Discriminations
- Ethics in NLP
- Guidelines for Promoting Ethical Practices
- Promoting Ethical Practices

Topic C: Establish Data Privacy and Ethics Policies

- Privacy and Data Governance for AI and ML
- Intellectual Property
- Humanitarian Principles
- Guidelines for Establishing Policies Covering Data Privacy and Ethics
- Establishing Policies Covering Data Privacy and Ethics

Appendix A: Mapping Course Content to CertNexus® Certified Artificial Intelligence (AI) Practitioner (Exam AIP-100)