Ayutech Ai Labs

AAL-501: Five-Day Technical Programme on Applied Machine Learning

(Customizable as per Client's Specific Requirements)

Key Learning Outcomes & Prerequisites

- **Build ML models from first principles** - Implement standard machine learning models from their underlying algorithmic logic, not only by calling a library.
- **Write clean, professional Python workflows** - Work fluently with NumPy, Pandas and Scikit-learn to move from raw data to a trained model.
- **Train, tune, validate and deploy** - Fit models, tune hyperparameters, cross-validate honestly, and deploy onto a compute target.
- **Work with text and image data** - Construct a basic NLP pipeline and a computer vision classification and segmentation task.
- **Prerequisite** - Basic working knowledge of programming. Prior exposure to Python is helpful but not essential; an engineering or analytical background is preferred.

Day 1: Foundational Mathematics & Mechanics

09:30 AM - 11:00 AM

Session 1: What a Model Actually Learns

- From Data to Decision Rule - The mental model behind every supervised algorithm - inputs, parameters, predictions and error - established before any code is written.
- The Statistics You Genuinely Need - Distributions, expectation, variance and correlation, covered only to the depth that later sessions will demand.
- Linear Models as the Starting Point - Why linear and logistic regression remain the reference point against which every more complex model is judged.

11:00 AM - 11:15 AM

Tea break

11:15 AM - 01:00 PM

Session 2: Loss Functions - Telling the Model It Is Wrong

- What a Loss Function Encodes - How the choice of loss quietly defines what the model treats as a serious mistake and what it is willing to ignore.
- Regression Losses: MSE and MAE - Squared versus absolute error, their differing treatment of outliers, and when each is the honest choice.
- Classification Losses: Cross-Entropy - Log loss and cross-entropy explained through probability, and why accuracy alone is a poor training signal.
- Hands-On: Computing Loss by Hand - Calculating loss manually on a small dataset, then reproducing the same number in Python.

01:00 PM - 02:00 PM

Lunch break

02:00 PM - 03:30 PM

Session 3: Gradients and the Direction of Improvement

- Derivatives, Intuitively - Slope as a measure of sensitivity, and gradient as the direction of steepest change, built up visually rather than symbolically.
- The Chain Rule and Backpropagation - How gradients flow backwards through a composed function, the single idea that makes deep learning possible.
- Hands-On: Deriving Gradients for Linear Regression - Working out the gradient of the loss with respect to each parameter, then verifying it numerically in code.

03:30 PM - 03:45 PM

Tea break

03:45 PM - 05:00 PM

Session 4: Optimization Algorithms

- Gradient Descent, Step by Step - Batch, stochastic and mini-batch gradient descent, and the practical trade-offs between them.
- Learning Rate: The Parameter That Decides Everything - Watching a model diverge, crawl, or converge cleanly purely as a function of step size.
- Momentum, RMSProp and Adam - What the modern optimizers add over plain gradient descent, and when the difference actually matters.
- Hands-On: Train a Model From Scratch - Implementing gradient descent in plain Python and watching the loss curve fall, with no ML library involved.

Day 2: The Python Ecosystem for Machine Learning

09:30 AM - 11:00 AM

Session 1: NumPy and Matrix Manipulation

- Arrays, Shapes and Axes - Building fluency with the array as the fundamental data structure of scientific Python.
- Broadcasting and Vectorization - Replacing loops with vectorized operations, and understanding the order-of-magnitude speed difference this creates.
- Hands-On: Linear Algebra for Models - Matrix multiplication, transposition and inversion, applied directly to the regression equations from Day 1.

11:00 AM - 11:15 AM

Tea break

11:15 AM - 01:00 PM

Session 2: Data Cleaning with Pandas

- Loading, Inspecting and Profiling Data - First contact with a messy dataset - shape, types, ranges and the questions to ask before modelling anything.
- Missing Values, Duplicates and Outliers - Practical strategies for each, and the reasoning required to choose between dropping, imputing and flagging.
- Reshaping, Merging and Grouping - Joins, group-by aggregations and pivots, the operations that consume most of a working day in practice.
- Hands-On: Cleaning a Real Dataset - Taking a deliberately messy file through to an analysis-ready table, documenting every decision made.

01:00 PM - 02:00 PM

Lunch break

02:00 PM - 03:30 PM

Session 3: Scikit-learn Fundamentals

- The Estimator API - The fit, predict and transform pattern that unifies every model in the library.
- Feature Engineering and Encoding - Scaling, categorical encoding and feature construction, and why these choices often outweigh model choice.
- Pipelines: Making Preprocessing Reproducible - Binding preprocessing and model into a single object, and how this prevents the most common source of leakage.

03:30 PM - 03:45 PM

Tea break

03:45 PM - 05:00 PM

Session 4: Lab - From Raw File to Trained Model

- Implementing Logistic Regression From Scratch - Writing the model in NumPy using the loss and gradients derived on Day 1.
- The Same Model in Three Lines of Scikit-learn - Reproducing the result with the library, comparing outputs, and understanding exactly what was abstracted away.
- Train-Test Split and First Evaluation - Holding out data honestly, and reading a confusion matrix, precision, recall and ROC-AUC.

Day 3: Generalization, Validation & Capstone I

09:30 AM - 11:00 AM

Session 1: Overfitting and the Bias-Variance Trade-Off

- Diagnosing Overfitting from Loss Curves - Reading training and validation curves together to distinguish underfitting, overfitting and a genuine data problem.
- The Bias-Variance Decomposition - Where prediction error actually comes from, and why the most accurate model on training data is rarely the best model.
- Learning Curves and Data Sufficiency - Using learning curves to answer a question every sponsor asks - would more data actually help?

11:00 AM - 11:15 AM

Tea break

11:15 AM - 01:00 PM

Session 2: Regularization, Cross-Validation & Tuning

- L1 and L2 Regularization - Ridge and Lasso as penalties on complexity, and the feature-selection side effect of L1.
- K-Fold and Stratified Cross-Validation - Building a validation scheme that survives contact with small, imbalanced or grouped data.
- Hyperparameter Search - Grid search, random search and early stopping, with an honest look at the compute cost of each.
- Data Leakage: The Silent Failure - The ways validation scores get inflated without anyone noticing, and how to design them out.

01:00 PM - 02:00 PM

Lunch break

02:00 PM - 03:45 PM

Session 3: Capstone I - Build, Tune and Validate a Model

- Project Brief and Dataset Handover - A structured tabular prediction problem, issued with a business objective rather than a target metric.
- Build the Pipeline - Constructing a full preprocessing and modelling pipeline, with every transformation inside the cross-validation loop.
- Tune and Validate - Running a disciplined hyperparameter search and reporting a validation score that will not collapse later.
- Defend the Result - Short peer review in which each participant must justify their validation design to the group.

03:45 PM - 04:00 PM

Tea break

04:00 PM - 05:00 PM

Session 4: Professional Development

Client to classify trainees as per criteria mentioned in appendix 1 and to select one of the relevant themes mentioned in appendix 2.

Day 4: Advanced Modalities & Capstone II

09:30 AM - 10:30 AM

Session 1: Professional Development

Client to classify trainees as per criteria mentioned in appendix 1 and to select one of the relevant themes mentioned in appendix 2. This theme should be different from the one selected for day 3.

10:30 AM - 10:45 AM

Tea break

10:45 AM - 12:30 PM

Session 2: Natural Language Processing

- Text as Data: Tokenization - Words, sub-words and tokens, and why tokenization decisions propagate through everything downstream.
- From Counts to Embeddings - Bag-of-words and TF-IDF, then the shift to dense embeddings and what that representation captures.
- Hands-On: A Working Text Classifier - Building a complete classification pipeline over a text dataset and evaluating it properly.
- Where Transformers Fit - A conceptual bridge from these fundamentals to the architecture behind current large language models.

12:30 PM - 01:30 PM

Lunch break

01:30 PM - 03:00 PM

Session 3: Computer Vision

- Images as Arrays - Pixels, channels and normalization, connecting image data back to the NumPy work of Day 2.
- Convolution, Intuitively - What a filter actually does to an image, demonstrated visually before any network is trained.
- Convolutional Networks and Transfer Learning - Layer stacking, pooling, and using a pretrained backbone rather than training from zero.
- Hands-On: Classification and Segmentation Demo - Running an image classification task, then a basic segmentation task, and interpreting the failure cases.

03:00 PM - 03:15 PM

Tea break

03:15 PM - 05:00 PM

Session 4: Capstone II - An Unstructured Data Project

- Choose Your Modality - Each participant selects either a text or an image problem from a small set of provided briefs.
- Build and Train - Constructing the full pipeline from raw text or images through to a trained, evaluated model.
- Evaluate Beyond Accuracy - Examining class-wise performance and inspecting the examples the model gets wrong, not only the headline score.
- Document the Work - Applying the morning's handover format to the project just completed.

Day 5: Deployment, Capstone III, Ethics & Wrap-Up

09:30 AM - 10:30 AM

Session 1: Professional Development

Client to classify trainees as per criteria mentioned in appendix 1 and to select one of the relevant themes mentioned in appendix 2. This theme should be different from the two selected for days 3 and 4.

10:30 AM - 10:45 AM

Tea break

10:45 AM - 01:00 PM

Session 2: Capstone III - Build and Validate an End-to-End Pipeline

- Project Brief: A Complete Operational Pipeline - A single project running from raw data ingestion through to a served prediction, delivered as a working artefact.
- Ingestion, Preprocessing and Feature Pipeline - Assembling the stages into one reproducible pipeline object rather than a sequence of manual steps.
- Train, Tune and Validate - Applying the full Day 3 validation discipline to a project that will actually be deployed.

01:00 PM - 02:00 PM

Lunch break

02:00 PM - 03:30 PM

Session 3: Capstone III - Deploy, Monitor and Critique

- Serialize and Serve the Model - Packaging the trained pipeline and exposing it behind a simple prediction endpoint on a compute target.

- Test the Deployed Model - Sending fresh, unseen records to the running service and confirming the output matches expectations.
- Monitoring and Model Drift - What to measure after go-live, and the signals that indicate a model has quietly stopped working.
- Structured Critique - Each team presents its pipeline and receives structured challenge on its weakest assumption.

03:30 PM - 03:45 PM

Tea break

03:45 PM - 04:30 PM

Session 4: Ethics and Governance

- Bias in Data, Models and Deployment - Where unfairness enters a pipeline, and the practical checks that catch it before a model reaches production.
- Explainability and Contestability - What an affected person is owed when an automated decision goes against them.
- Privacy, Consent and Data Handling - Working responsibly with personal data, and the governance controls that make this auditable.
- Governance in Practice - Model documentation, approval gates and review cycles, sized for a real organization rather than a policy document.

04:30 PM - 05:00 PM

Session 5: Wrap-Up, Q&A and Next Steps

- Open Q&A - An unstructured session addressing whatever remains unresolved from the week.
- What You Can Now Do - A consolidated review of the week against the stated learning outcomes.

- Your First Thirty Days - Each participant leaves with one specific, scoped problem from their own work to take forward.

Appendices

Appendix 1: Classification of Audience

Choices of professional development modules to be included as part of this training programme will depend on the seniority and stated priorities of the cohort in question. Broadly speaking, audiences can be classified into the one of the following categories with curriculum emphasis shifting accordingly:

Cohort	Profile	Curriculum Emphasis
Senior Professionals	Leaders with extensive experience, accustomed to setting direction and accountable for outcomes at scale.	For this cohort, the programme is pitched toward governance and strategic judgement: artificial intelligence as a matter for decision-making rather than daily execution, communication as a vehicle for influence and presence, and marketing as a question of brand and reputation.
Mid-Career Professionals	Individuals managing teams or substantial projects, positioned between execution and people leadership.	Here the emphasis falls on practical application: using artificial intelligence to improve team output, leading difficult conversations, understanding team behaviour, and running a digital marketing campaign from brief to measurement.
Early-Career Professionals	Those in the earlier stages of their careers, still forming professional habits and identity.	The focus here is foundational: responsible everyday use of artificial intelligence, workplace communication and behaviour, an introduction to self-awareness through assessment, and the fundamentals of digital marketing and mindfulness.

Appendix 2: The Seven Themes for Professional Development

The table below sets out seven proposed professional development themes, along with their scopes. The table also suggests the target audience class for which each theme should be considered.

Clients should choose three different themes, one each for days 3, 4 and 5.

Theme	Core Focus	Scope	Recommended Audience Classes
Effective Communication	The structuring of messages, the discernment of audience, and the management of high-stakes or difficult conversations.	The module comprises a structured framework, recorded practice, and peer feedback.	Early-Career Professionals Mid-Career Professionals
Understanding Behaviour: Self and Others	Assessment-based behavioural self-awareness, together with the capacity to read and adapt to others.	This module includes an assessment debrief, an introduction to the Johari Window, and a team-based application exercise.	Mid-Career Professionals
Mindfulness and Resilience	Attention, recovery and stress management, treated as professional performance skills rather than wellness peripherals.	This will be divided into two brief sessions, culminating in a personal sustainability plan.	Senior Professionals
Digital Marketing in the Age of Artificial Intelligence	The principal digital channels, campaign fundamentals, and the evolving influence of artificial intelligence on content and targeting.	This topic covers fundamentals through a hands-on campaign-building exercise.	Early-Career Professionals Mid-Career Professionals

Change Agility and Problem-Solving	A design-thinking approach to navigating ambiguity and resolving genuine organisational problems.	This module is built around a structured problem-solving sprint addressing a live organisational issue.	Early-Career Professionals
Personal Branding and Professional Presence	The manner in which one is perceived, both internally and externally: professional profile, visibility and reputation.	This session comprises a personal brand audit, a workshop on professional profile and online presence, and peer feedback.	Mid-Career Professionals Senior Professionals
Artificial Intelligence Fluency and Application	The judicious and responsible use of generative artificial intelligence and automation in everyday professional practice.	A full-day module encompassing foundational concepts, guided application to an authentic work task, and a concise treatment of ethical and governance considerations.	Early-Career Professionals Mid-Career Professionals