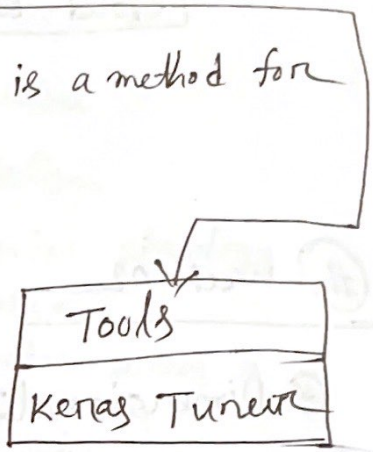
Specialization - 03

"Machine Learning Modeling pipelines in production"

Week-01

① Hyperparameter Tuning

- Neural architecture search (NAS) is a method for automating the design of ANN.
- Can tune parameters
  - Architecture choice
  - activation functions
  - weight initial strategy
  - Learning rate, stop condition

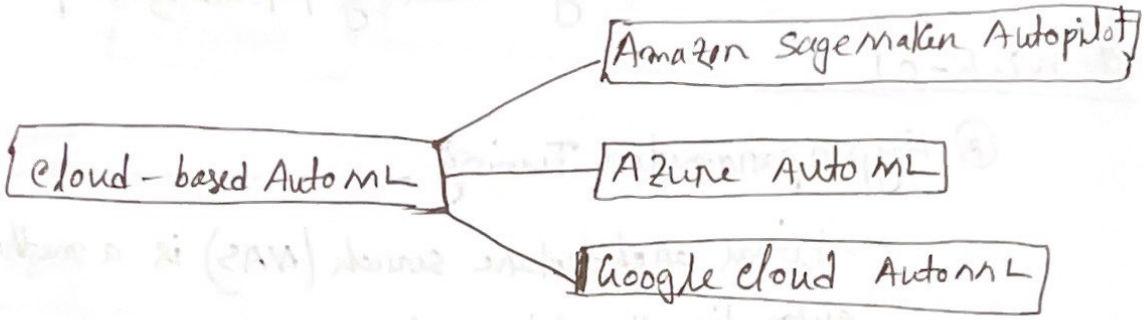


② AutoML → applies machine learning and search techniques to the process of create machine learning models and pipelines. It covers the complete pipeline from the raw dataset to the deployable.

AutoML automates the entire ML workflow.

- NAS: a few search strategies
  - Grid search • Random search • Bayesian optimization
  - Evolutionary Algorithms • Reinforcement Learning

⊛ AutoML on the cloud



QwiK Labs
↳ Free cloud use for demo.

Week 02

⊛ Dimensionality

- dimensionality reduction
  - Linear dimensionality reduction
  - principal component Analysis (PCA)
- classification: LDA
- Regression: PLS (partial Least squares)
- Unsupervised: PCA
- others: SVD, ICA, NMF (Non-negative Matrix Factor)
  ↳ deemes non-square matrices.

⊛ Quantization & Pruning (Model deployment on Edge devices)

↓

- Shrinking model file size
- post training quantization
- Quantization Aware training (QAT)

- ML Kit
- Core ML
- TF Lite

- pruning

$$P_n = 1 - (1-p)^n$$

Week - 03

High performance Modeling

- Distributed training → Data parallelism
  → Model parallelism
- tf. distribute. strategy → one device strategy
  → mirrored "
  → parameter server "
- Tools → GPipe → distributed parallelism
- Knowledge Distillation (KD)

Week - 04

Model performance Analysis

Tools
→ Tenson board
→ TFMa / tfma
Model performance Analysis.

Model Debugging

- Benchmark models

Sensitivity Analysis & Adversial Attacks.

- partial dependence plots.
- PDP box and pyCEbox

Tool
→ what - if

Adversial Attacks
↓
Tools

elevenhang
Foolbox

→ Adversial attack is to confuse the Model.

~~Ris~~ Residual Analysis

Model Remediation

Statistical process control

$$\mu = np_t \qquad \sigma = \sqrt{\frac{p_t(1-p_t)}{n}}$$

Sequential Analysis

$$p_t + \sigma_t \geq p_{min} + 3\sigma_{min}$$

clustering

- Algorithm = OLINDDA, MINAS, ECSMiner, GC3

* google cloud AI continuous Evaluation

Week 05

Ⓐ Explainable AI (XAI)

Ⓐ Responsible AI (RAI)

Ⓐ Interpretability

Ⓐ Model specific or Model Agnostic

→ Interpretable Models

- Linear · Logistic · Decision Tree · RuleFit, K-nearest Neighbors
- TF Lattice

→ Model Agnostic methods

- These methods separate explanations from the machine learning model.

  - partial Dependence plots (PDP)
  - permutation Feature
  - Shapley values.
  - Testing concept Activition
  - LIME
  - TCAV & LIME
  - AI Explanations.