California State University
Los Angeles

Machine Learning

Machine Learning Libraries

Nikhil Dhiman
Computational Molecular Biology (COMB) Lab

Agenda Overview

01 PyTorch

02 Why PyTorch?

03 What is a Tensor?

04 What is TensorFlow?

05 PyTorch vs TensorFlow

Important Machine Learning Libraries

PyTorch

- PyTorch is an open-source deep learning framework developed by Facebook AI Research (FAIR).
- It allows researchers and developers to build and train neural networks with ease, flexibility, and speed.

What Can You Do with PyTorch?

- Build and train deep learning models (CNNs, RNNs, Transformers)
- Perform image and text classification
- Apply transfer learning and fine-tuning
- Develop research prototypes or production-ready solutions

Why PyTorch?

1. Pythonic & Intuitive
   - Feels like writing regular Python code
   - Simple syntax, easy to debug
   - Great for beginners and researchers

2. Dynamic Computation Graph (Eager Execution)
   - Build and modify your model on-the-fly
   - Immediate feedback at each line (ideal for experimentation)
   - No need to compile the full model before running

3. Autograd for Differentiation
   - Automatically computes gradients during training
   - No manual math or derivative rules needed
   - Perfect for backpropagation in neural networks

4. GPU Acceleration

- Run tensor operations on NVIDIA GPUs using CUDA
- Huge speedup for training deep models
- Easy transfer: .cuda() or .to(device)

5. Strong Ecosystem

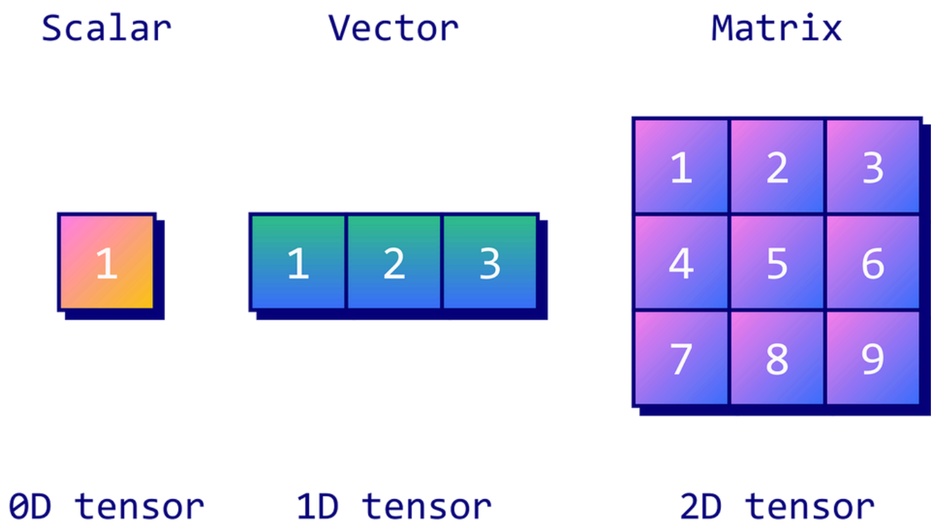
- TorchVision (images), TorchText (NLP), TorchAudio (speech)
- PyTorch Lightning for easier model training
- Integrates with tools like HuggingFace, ONNX, FastAI

6. Preferred in Research

- Dominates academic papers and ML research projects
- Rapid prototyping and easy model customization

What is a Tensor?

- A tensor is a multi-dimensional array — like a generalization of numbers, vectors, and matrices.
- It's the core data structure used in PyTorch, TensorFlow, and many ML frameworks.



Why Arrays (like NumPy) Fall Short in ML:

- No GPU Support:
  - NumPy arrays only run on the CPU. For big models and datasets, that's too slow.

- No Autograd (No Gradient Calculation):
  - Training neural networks requires gradients (calculus!). NumPy doesn't do this.

- Not ML-Aware:
  - NumPy doesn't know how to handle backpropagation, layers, or loss functions.
  - You'd have to code everything yourself.

Why Tensors Rock (Especially in PyTorch):

- They look and behave like arrays — so they're familiar!
- But they come with superpowers:
  - Keep track of operations (so you can compute gradients)
  - Move to GPU easily for speed
  - Work directly with neural network layers and training code

What is TensorFlow?

- TensorFlow is an open-source machine learning and deep learning framework developed by Google Brain.
- It provides a flexible ecosystem of tools for building, training, and deploying ML models at scale.

Core Features

- Computational Graphs: Originally static (TF1), now supports eager execution (TF2)
- Cross-platform: Works on desktop, mobile, web, and embedded devices
- GPU & TPU support: Accelerates model training and inference
- Scalable: Designed for production deployment
- Integrated with Keras: High-level API for building and training models

PyTorch vs TensorFlow

- **Execution Mode:** PyTorch uses dynamic graphs (eager execution), TensorFlow used static graphs (TF1) but supports eager in TF2
- **Syntax**: PyTorch is Pythonic and intuitive, TensorFlow is more verbose and abstract
- **Debugging**: PyTorch allows easy debugging with native Python, TensorFlow requires specialized tools
- **Use Case**: PyTorch is popular in research, TensorFlow dominates industry/production
- **Deployment**: PyTorch supports TorchScript/ONNX, TensorFlow supports TFLite, TF Serving, TF.js
- **Hardware** Support: PyTorch supports GPU (CUDA), TensorFlow supports GPU and TPU
- **Ecosystem**: PyTorch has TorchVision, Lightning, TensorFlow has Keras, TensorBoard, TF Hub
- **Learning Curv**e: PyTorch is easier to learn, TensorFlow has a steeper learning curve
- **Community**: PyTorch has a fast-growing research community, TensorFlow has a large, mature ecosystem

California State University
Los Angeles

Thank You

Computational Molecular Biology (COMB) Lab