Machine Learning

Introduction to the Course

Nevin L. Zhang
lzhang@cse.ust.hk

Department of Computer Science and Engineering
The Hong Kong University of Science and Technology

This set of notes is based on internet resources.

Course Coverage

Deployment	**Adversarial Attack (Security)**	**XAI (Trust/Fairness)**	Federated Learning(Privacy), Meta-learning (Learn to Learn), Domain Adaption/Generalization, Lifelong learning, Ethics, ...		
	General Issues	**Supervised**	**Self-Supervised**	**Unsupervised**	**Reinforcement**
Deep Learning	Dropout Normalization Optimizers	Feedforward NN Convolutional NN ViT	Recurrent NN Transformer BERT, LLM, CLIP Contrastive Learning	VAE GAN Diffusion	DQN Policy gradient Actor-critic
Machine Learning	Overfitting Bias, variance Regularization Validation	Linear Regression Logistic Regression Generative models SVM, Decision tree, K-NN, Emsemble methods		Finite Mixtures PCA, Bayesian Nets HMMs	Q-learning REINFORCE
Foundation Principles Algorithms	**Probability Theory** Likelihood, Bayes theorem		**Information Theory** Cross entropy Divergence		**Optimization Theory** Gradient Descent Newton Primal-dual

- Machine Learning (ML) has become a vast field, encompassing a wide array of topics.
- We will focus on the topics highlighted in large font, while omitting those in smaller font.

Course Coverage

Deployment	**Adversarial Attack (Security)**	**XAI (Trust/Fairness)**	Federated Learning(Privacy), Meta-learning (Learn to Learn), Domain Adaption/Generalization, Lifelong learning, Ethics, ...		
	General Issues	**Supervised**	**Self-Supervised**	**Unsupervised**	**Reinforcement**
Deep Learning	Dropout Normalization Optimizers	Feedforward NN Convolutional NN ViT	Recurrent NN Transformer BERT, LLM, CLIP Contrastive Learning	VAE GAN Diffusion	DQN Policy gradient Actor-critic
Machine Learning	Overfitting Bias, variance Regularization Validation	Linear Regression Logistic Regression Generative models SVM, Decision tree, K-NN, Emsemble methods		Finite Mixtures PCA, Bayesian Nets HMMs	Q-learning REINFORCE
Foundation Principles Algorithms	**Probability Theory** Likelihood, Bayes theorem		**Information Theory** Cross entropy Divergence	**Optimization Theory** Gradient Descent Newton Primal-dual	

- Three basic things to do in ML: 1). Choose a model, 2). Set up an objective/loss function, and 3). Optimize it.
- We will start in Part 0 with the principles behind the learning objectives, which are from Probability Theory and Information theory.
- Optimization methods, derived from Optimization Theory, will be discussed in the context specific models.

Course Coverage

Deployment	**Adversarial Attack (Security)**	**XAI (Trust/Fairness)**	Federated Learning(Privacy), Meta-learning (Learn to Learn), Domain Adaption/Generalization, Lifelong learning, Ethics, ...		
	General Issues	**Supervised**	**Self-Supervised**	**Unsupervised**	**Reinforcement**
Deep Learning	Dropout Normalization Optimizers	Feedforward NN Convolutional NN ViT	Recurrent NN Transformer BERT, LLM, CLIP Contrastive Learning	VAE GAN Diffusion	DQN Policy gradient Actor-critic
Machine Learning	Overfitting Bias, variance Regularization Validation	Linear Regression Logistic Regression Generative models SVM, Decision tree, K-NN, Emsemble methods		Finite Mixtures PCA, Bayesian Nets HMMs	Q-learning REINFORCE
Foundation Principles Algorithms	**Probability Theory** Likelihood, Bayes theorem	**Information Theory** Cross entropy Divergence	**Optimization Theory** Gradient Descent Newton Primal-dual		

- In Part 1, we will cover the foundation of Machine Learning:
  - Fundamental models for regression and classification
  - Basic optimization algorithms
  - Fundamental issues in Machine learning

Course Coverage

Deployment	**Adversarial Attack (Security)**	**XAI (Trust/Fairness)**	Federated Learning(Privacy), Meta-learning (Learn to Learn), Domain Adaption/Generalization, Lifelong learning, Ethics, ...		
	General Issues	**Supervised**	**Self-Supervised**	**Unsupervised**	**Reinforcement**
Deep Learning	Dropout Normalization Optimizers	Feedforward NN Convolutional NN ViT	Recurrent NN Transformer BERT, LLM, CLIP Contrastive Learning	VAE GAN Diffusion	DQN Policy gradient Actor-critic
Machine Learning	Overfitting Bias, variance Regularization Validation	Linear Regression Logistic Regression Generative models SVM, Decision tree, K-NN, Emsemble methods		Finite Mixtures PCA, Bayesian Nets HMMs	Q-learning REINFORCE
Foundation Principles Algorithms	**Probability Theory** Likelihood, Bayes theorem		**Information Theory** Cross entropy Divergence	**Optimization Theory** Gradient Descent Newton Primal-dual	

- In Part 2, we will cover basic deep learning models, as well as techniques for optimizing them:
  - Feedforward Neural Networks
  - Convolutional Neural Networks
  - Recurrent Neural Networks
  - Basic techniques for optimizing deep models

Course Coverage

	General Issues	Supervised	Self-Supervised	Unsupervised	Reinforcement
Deployment	**Adversarial Attack (Security)**	**XAI (Trust/Fairness)**	Federated Learning(Privacy), Meta-learning (Learn to Learn), Domain Adaption/Generalization, Lifelong learning, Ethics, ...		
Deep Learning	Dropout Normalization Optimizers	Feedforward NN Convolutional NN ViT	Recurrent NN Transformer BERT, LLM, CLIP Contrastive Learning	VAE GAN Diffusion	DQN Policy gradient Actor-critic
Machine Learning	Overfitting Bias, variance Regularization Validation	Linear Regression Logistic Regression Generative models SVM, Decision tree, K-NN, Emsemble methods		Finite Mixtures PCA, Bayesian Nets HMMs	Q-learning REINFORCE
Foundation Principles Algorithms	**Probability Theory** Likelihood, Bayes theorem		**Information Theory** Cross entropy Divergence	**Optimization Theory** Gradient Descent Newton Primal-dual	

- In Part 3, we will cover advanced deep learning models:
  - Transformer, BERT and GPT
  - Vision Transformers (ViT)
  - Vision-Language Models (CLIP)

Course Coverage

Deployment	**Adversarial Attack** **(Security)**	**XAI** **(Trust/Fairness)**	Federated Learning(Privacy), Meta-learning (Learn to Learn), Domain Adaption/Generalization, Lifelong learning, Ethics, ...		
	General Issues	**Supervised**	**Self-Supervised**	**Unsupervised**	**Reinforcement**
Deep Learning	Dropout Normalization Optimizers	Feedforward NN Convolutional NN ViT	Recurrent NN Transformer BERT, LLM, CLIP Contrastive Learning	VAE GAN Diffusion	DQN Policy gradient Actor-critic
Machine Learning	Overfitting Bias, variance Regularization Validation	Linear Regression Logistic Regression Generative models SVM, Decision tree, K-NN, Emsemble methods		Finite Mixtures PCA, Bayesian Nets HMMs	Q-learning REINFORCE
Foundation Principles Algorithms	**Probability Theory** Likelihood, Bayes theorem		**Information Theory** Cross entropy Divergence		**Optimization Theory** Gradient Descent Newton Primal-dual

- In Part 4, we will cover deep learning models for unsupervised learning (genearative AI):
  - Variational autoencoders
  - Generative adversarial networks
  - Diffusion Models

Course Coverage

Deployment	**Adversarial Attack (Security)**	**XAI (Trust/Fairness)**	Federated Learning(Privacy), Meta-learning (Learn to Learn), Domain Adaption/Generalization, Lifelong learning, Ethics, ...		
	General Issues	**Supervised**	**Self-Supervised**	**Unsupervised**	**Reinforcement**
Deep Learning	Dropout Normalization Optimizers	Feedforward NN Convolutional NN ViT	Recurrent NN Transformer BERT, LLM, CLIP Contrastive Learning	VAE GAN Diffusion	DQN Policy gradient Actor-critic
Machine Learning	Overfitting Bias, variance Regularization Validation	Linear Regression Logistic Regression Generative models SVM, Decision tree, K-NN, Emsemble methods		Finite Mixtures PCA, Bayesian Nets HMMs	Q-learning REINFORCE
Foundation Principles Algorithms	**Probability Theory** Likelihood, Bayes theorem		**Information Theory** Cross entropy Divergence		**Optimization Theory** Gradient Descent Newton Primal-dual

- In Part 5, we will cover reinforcement learning and Deep RL:
  - Introduction to RL
  - Value-Based Deep RL
  - Policy-Based Deep RL

Course Coverage

Deployment	**Adversarial Attack (Security)**	**XAI (Trust/Fairness)**	Federated Learning(Privacy), Meta-learning (Learn to Learn), Domain Adaption/Generalization, Lifelong learning, Ethics, ...		
	General Issues	**Supervised**	**Self-Supervised**	**Unsupervised**	**Reinforcement**
Deep Learning	Dropout Normalization Optimizers	Feedforward NN Convolutional NN ViT	Recurrent NN Transformer BERT, LLM, CLIP Contrastive Learning	VAE GAN Diffusion	DQN Policy gradient Actor-critic
Machine Learning	Overfitting Bias, variance Regularization Validation	Linear Regression Logistic Regression Generative models SVM, Decision tree, K-NN, Emsemble methods		Finite Mixtures PCA, Bayesian Nets HMMs	Q-learning REINFORCE
Foundation Principles Algorithms	**Probability Theory** Likelihood, Bayes theorem		**Information Theory** Cross entropy Divergence	**Optimization Theory** Gradient Descent Newton Primal-dual	

- In Part 6, we will cover deployment issues:
  - Adversarial robustness
  - Explainability

Tutorials and Hands-on Assignments

Deployment	**Adversarial Attack (Security)**	**XAI (Trust/Fairness)**	Federated Learning(Privacy), Meta-learning (Learn to Learn), Domain Adaption/Generalization, Lifelong learning, Ethics, ...		
	General Issues	**Supervised**	**Self-Supervised**	**Unsupervised**	**Reinforcement**
Deep Learning	Dropout Normalization Optimizers	Feedforward NN Convolutional NN ViT	Recurrent NN Transformer BERT, LLM, CLIP Contrastive Learning	VAE GAN Diffusion	DQN Policy gradient Actor-critic
Machine Learning	Overfitting Bias, variance Regularization Validation	Linear Regression Logistic Regression Generative models SVM, Decision tree, K-NN, Emsemble methods		Finite Mixtures PCA, Bayesian Nets HMMs	Q-learning REINFORCE
Foundation Principles Algorithms	**Probability Theory** Likelihood, Bayes theorem		**Information Theory** Cross entropy Divergence		**Optimization Theory** Gradient Descent Newton Primal-dual

Hands-on experience is very important to this course.

- There will 11 tutorials (with code and vidoes):
  PyTorch Basics; Feedforward Neural Networks in PyTorch; Convolutional Neural Networks in PyTorch; Recurrent Neural Networks in PyTorch; BERT; CLIP; GAN and VAE; Stable Diffusion; Deep Q-Network; Adversarial Attack; Explainable AI
- Students will be given hands-on assignments.

Comparison with Specialized Courses

- Pros:
  - Comprehensive coverage of ML fundamentals
  - Exposure to topics from multiple specializations
- Cons:
  - Less depth compared to specialized courses
  - Students can take specialized courses concurrently or afterward

Catering for Students with Different Backgrounds

- Students who have taken an ML course before
  - Better understanding of the foundation of machine learning
  - Better understanding of some of the new developments
- Students who have not taken an ML course before
  - Overall understanding of the machine learning field
  - Need strong foundation in Math and good programming skills, and need to make extra efforts

Feedback from Past Students

- It is better than any courses about ml that I have in the past. I can clearly know the core of each part of the machine learning and have a quick glance of the advanced technology.
- The course is challenging, especially for whom have no basic machine learning knowledge.
- The course is too tough!!
- There are too many written and hands-on assignments for most of the students.
- The Hand-on assignments are really good for me to learn the materials more.
- I have learned a lot from the course, and the hands-on assignments give me a deeper understanding of related topics. It is unparalleled by previous classes I took.
- In my opinion, a mandatary project should be added to the grading metric so that student can learn how to apply knowledge to practise.