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.

Deployment	Adversarial Attack (Security)	k XAI (Trust/Fairne		Federated Learning(Privacy), Meta- Learn), Domain Adaption/Generaliz learning, Ethics,			•
	General Issues	Supervise	Supervised Self-Supervised		Unsupervised		Reinforcement
Deep	Dropout	Feedforward NN		Recurrent NN		E	DQN
Learning	Normalization	Convolutional NN		Transformer		N	Policy gradient
	Optimizers	ViT		BERT, LLM, CLIP Contrastive Learning		fusion	Actor-critic
Machine	Overfitting	Linear Regress	ion			e Mixtures	Q-learning
Learning	Bias, variance	Logistic Regres	sion		HMN	Bayesian Nets As	REINFORCE
	Regularization	Generative mo	dels				
		SVM, Decision tree, I Emsemble methods	C-NN,				
Foundation	Probability Theory		Informatio	n Theory		Optimization	on Theory
Principles	Likelihood, Bayes the	orem	Cross entre	Cross entropy		Gradient Descent	
Algorithms			Divergence	ergence		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.

Deployment	Adversarial Attack (Security)		Al Fairness)	Federated Learning(Privacy), Meta-learning (L Learn), Domain Adaption/Generalization, Lifel learning, Ethics,			• • • • • • • • • • • • • • • • • • • •
	General Issues	Supervised	Supervised Self-Supervised		Unsupervised		Reinforcement
Deep	Dropout	Feedforward N	N Recu	Recurrent NN			DQN
Learning	Normalization	Convolutional I	NN Trans	Transformer		N	Policy gradient
	Optimizers	ViT		BERT, LLM, CLIP		usion	Actor-critic
Machine Learning	Bias, variance Regularization Validation	Linear Regressi Logistic Regress Generative mo SVM, Decision tree, K Emsemble methods	on sion dels	Stive Economic		e Mixtures Bayesian Nets Is	Q-learning REINFORCE
Foundation Principles Algorithms	Probability Theory Likelihood, Bayes the	orem	Information Cross entrop Divergence			Optimization Gradient Do 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.

Deployment	Adversarial Attack (Security)	k XAI (Trust/Fairne		Federated Learning(Privacy), Meta-learn Learn), Domain Adaption/Generalization learning, Ethics,			• • • • • • • • • • • • • • • • • • • •	
	General Issues	Supervise	Supervised Self-Supervised		Unsupervised		Reinforcement	
Deep	Dropout	Feedforward NN		Recurrent NN		VAE		DQN
Learning	Normalization	Convolutional NN		Transformer		GAN		Policy gradient
	Optimizers	ViT		BERT, LLM, CLIP		Diffusion		Actor-critic
Machine Learning		Linear Regres		contrastive ecurring		Finite Mixtures PCA, Bayesian Nets		Q-learning REINFORCE
Learning	Regularization Validation	Generative mesons tree, Emsemble methods	odels K-NN,			HMN	1s	REINFORCE
Foundation	Probability Theory		Inforn	nation	Theory		Optimization	on Theory
Principles	Likelihood, Bayes the	eorem	Cross	Cross entropy			Gradient Descent	
Algorithms			Diverg	rgence			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

Deployment	Adversarial Attack (Security)	k XAI (Trust/Fair		1 D		rning(Privacy), Meta-learning (Learn to n Adaption/Generalization, Lifelong s,		
	General Issues	Supervise	Supervised Self-Supervised		Unsupervised		Reinforcement	
Deep	Dropout	Feedforward N	NN I	Recui	rent NN	VAE		DQN
Learning	Normalization	Convolutional NN Trai		Trans	sformer GAI		N	Policy gradient
	Optimizers	ViT		BERT, LLM, CLIP		Diffusion		Actor-critic
				Contrastive Learning				
Machine	Overfitting	Linear Regress	ion				Mixtures Bayesian Nets	Q-learning
Learning	Bias, variance	Logistic Regres	ssion			HMN		REINFORCE
	Regularization	Generative mo	odels					
		SVM, Decision tree, I Emsemble methods						
Foundation	Probability Theory		Informa	ation	Theory		Optimization	on Theory
Principles	Likelihood, Bayes the	orem	Cross e	ntrop	У		Gradient Descent	
Algorithms			Diverge	ergence			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

Deployment	Adversarial Attack (Security)	k XAI (Trust/Fair		1 D		rning(Privacy), Meta-learning (Learn to n Adaption/Generalization, Lifelong s,		
	General Issues	Supervise	Supervised Self-Supervised		Unsupervised		Reinforcement	
Deep	Dropout	Feedforward N	NN I	Recui	rent NN	VAE		DQN
Learning	Normalization	Convolutional NN Trai		Trans	sformer GAI		N	Policy gradient
	Optimizers	ViT		BERT, LLM, CLIP		Diffusion		Actor-critic
				Contrastive Learning				
Machine	Overfitting	Linear Regress	ion				Mixtures Bayesian Nets	Q-learning
Learning	Bias, variance	Logistic Regres	ssion			HMN		REINFORCE
	Regularization	Generative mo	odels					
		SVM, Decision tree, I Emsemble methods						
Foundation	Probability Theory		Informa	ation	Theory		Optimization	on Theory
Principles	Likelihood, Bayes the	orem	Cross e	ntrop	У		Gradient Descent	
Algorithms			Diverge	ergence			Newton Primal-dual	

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

Deployment	Adversarial Attack (Security)	XAI (Trust/Fairne		Federated Learning(Privacy), Meta-learning (Le Learn), Domain Adaption/Generalization, Lifeld learning, Ethics,			• • • • • • • • • • • • • • • • • • • •	
	General Issues	Supervise	Supervised Self-Supervised		Unsupervised		Reinforcement	
Deep	Dropout	Feedforward NN		Recurrent NN		VAE		DQN
Learning	Normalization	Convolutional NN		Transformer		GAN		Policy gradient
	Optimizers	ViT		BERT, LLM, CLIP Contrastive Learning		Diffusion		Actor-critic
Machine Learning	Bias, variance Regularization Validation	Linear Regress Logistic Regre Generative me SVM, Decision tree, Emsemble methods	ssion odels K-NN,				e Mixtures Bayesian Nets Is	Q-learning REINFORCE
Foundation Principles Algorithms	Probability Theory Likelihood, Bayes the	orem Cross		nformation 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

Deployment	Adversarial Attack (Security)	XAI (Trust/Fairne		Federated Learning(Privacy), Meta-learnin Learn), Domain Adaption/Generalization, I learning, Ethics,			• • • • • • • • • • • • • • • • • • • •	
	General Issues	Supervise	rvised Self-Supervised		Unsupervised		Reinforcement	
Deep	Dropout	Feedforward NN		Recurrent NN		VAE		DQN
Learning	Normalization	Convolutional NN		Transformer		GAN		Policy gradient
	Optimizers	ViT		BERT, LLM, CLIP Contrastive Learning		Diffusion		Actor-critic
Machine Learning	Bias, variance Regularization Validation	Linear Regress Logistic Regre Generative me SVM, Decision tree, Emsemble methods	ssion odels K-NN,				e Mixtures Bayesian Nets Is	Q-learning REINFORCE
Foundation Principles Algorithms	Probability Theory Likelihood, Bayes the	Inform corem Cross		rmation Theory ss entropy ergence			Optimization Gradient Di Newton	

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

Deployment			NAI .		rning(Privacy), Meta-learning (Learn to n Adaption/Generalization, Lifelong s,			
	General Issues	Supervise	Supervised Self-Supervised		Unsupervised		Reinforcement	
Deep	Dropout	Feedforward	NN	Recui	rrent NN	VAE		DQN
Learning	Normalization	Convolutional	olutional NN Transformer		GAN		Policy gradient	
	Optimizers	ViT	BERT, LLM, CLIP Contrastive Learning		Diffusion		Actor-critic	
Machine Learning	Overfitting Bias, variance Regularization Validation	Logistic Regre Generative m	inear Regression ogistic Regression enerative models /M, Decision tree, K-NN,				e Mixtures Bayesian Nets As	Q-learning REINFORCE
Foundation Principles Algorithms	Probability Theory Likelihood, Bayes th	eorem		rmation Theory s entropy rgence			Optimization Gradient De Newton Primal-dual	

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

Tutorials and Hands-on Assignments

Deployment	Adversarial Attack (Security)	k XAI (Trust/Fa				rning(Privacy), Meta-learning (Learn to n Adaption/Generalization, Lifelong s,		
	General Issues	Supervised		Self-Supervised		Unsupervised		Reinforcement
Deep	Dropout	Feedforward NN		Recurrent NN		VAI	E	DQN
Learning	Normalization	Convolutional NN		Transformer		GAN		Policy gradient
	Optimizers	ViT		BERT, LLM, CLIP Contrastive Learning		Diffusion		Actor-critic
Machine	Overfitting	Linear Regres	sion				e Mixtures	Q-learning
Learning	Bias, variance	Logistic Regre	ssion			HMN	Bayesian Nets As	REINFORCE
	Regularization	Generative m	odels					
		SVM, Decision tree, Emsemble methods						
Foundation	Probability Theory		Inform	mation	Theory		Optimization Theory	
Principles	Likelihood, Bayes the	eorem Cross		entrop	у		Gradient Descent	
Algorithms			Diverg	ergence			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 Al
- 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.