

Lahore University of Management Sciences CS 535/EE 514 Machine Learning

Fall 2023-24

Course description

Machine learning (ML) techniques allow computers to adapt to data and solve new problems related to previously encountered problems more efficiently. Such methods enable machines to perform practical exploratory and predictive tasks without being explicitly programmed. ML finds applications in speech recognition and synthesis, machine translation, object recognition, chatbots, question-answering, natural language understanding, anomaly detection, medical diagnosis and prognosis, autonomous vehicles and robots, time series forecasting, and much more. This introductory course covers the theoretical foundations and practical applications of ML and the design, implementation, and analysis of various ML algorithms. Students will learn to compare across and choose the most appropriate algorithms for multiple problem types and be able to design and implement their solutions. Students will be prepared for industry and academia and for pursuing advanced courses.

Course distribution		
Elective	This is an elective course.	
Open for Student Category	Juniors, seniors, and graduates.	
Close for Student Category	Please see the prerequisites below.	

Course prerequisites

- Undergrads (Seniors/Juniors) must have passed:
 - An Ugrad/Grad course in Probability (MATH230 (Probability) OR DISC203 (Probability & Statistics) OR CS501 (Applied Probability)) OR ECON230 (Statistics and Data Analysis))
 - O And a programming course (CS200/EE201 (Intro. to Programming)
 - o And a course on Linear Algebra (MATH120 (LA with Diff. Equations))
- Grads are strongly advised to brush up their programming skills and take CS501 (Applied Probability), may be in parallel with ML
- All students must possess strong programming skills and proficiency in algorithm implementation in JAVA/C/Python/MATLAB

Course Offering Details							
Credit Hours	3 hours						
Lecture(s)	Nbr of lec(s) per week	2	Duration	75 minutes			
Recitation/Lab (per week)	Nbr of lec(s) per week		Duration				
Tutorial (per week)	Tutorial (per week) Nbr of lec(s) per week 1 (optional) Duration 50 minutes						

Instructor	Agha Ali Raza
Room No.	SBASSE 9-G49A
Office Hours	TBA
Email	agha.ali.raza@lums.edu.pk
Telephone	3336
Secretary/TA	Zohaib Khan (24100074@lums.edu.pk), Nida Tanveer (24100091@lums.edu.pk), Saad Hasan Iqbal (24020320@lums.edu.pk), Syeda Mah Noor Asad (24100243@lums.edu.pk), Alina Faisal (24100314@lums.edu.pk), Muawiz Feroze Khan (24100052@lums.edu.pk), Fatima Ali (24100048@lums.edu.pk), Rafey Rana (24100103@lums.edu.pk), Zain Ali Khokhar (24100130@lums.edu.pk), Muhammad Haad Zahid (24100134@lums.edu.pk), Mughees Ur Rehman (24100086@lums.edu.pk), Syed Kabir Ahmad (24100249@lums.edu.pk)
TA Office Hours	TBA
Course URL (if any)	From last offering: https://www.c-salt.org/courses/machine-learning-f2021

Course Teaching Methodology (Please mention the following details in plain text)

- Lectures: In-person.
- TA Sessions: TAs will conduct asynchronous and synchronous sessions (in-person and online) to cover tutorials related to assignments.
- **Exams:** Exams will be conducted in person in pre scheduled sessions.
- Quizzes: Quizzes will be conducted during announced class timings.
- Class discussions: There will be a slack channel for all discussions (general, assignments, quizzes, etc.)

PROGRAM EDUCATIONAL OBJECTIVES (PEOs)			
PEO-01	Demonstrate excellence in the profession through in-depth knowledge and skills in the field of Computing.		
PEO-02	Engage in continuous professional development and exhibit a quest for learning.		
PEO-03	Show professional integrity and commitment to societal responsibilities.		

Course Objectives

The goal of this course is to get the students excited about Machine Learning and to enable them to:

- · Develop a firm grip on the theory behind statistical learning
- · Understand and rigorously go through the phases of the design, implementation, and evaluation of fundamental ML algorithms
- · Choose the appropriate algorithm for each problem type and be able to compare the strengths and weaknesses of the algorithms
- Appreciate the end-to-end organic integration of ML in its application areas, from data sources, annotation pipelines, and choice of algorithms to societal biases, explainability of models, and potential to impact and even disrupt existing processes

COURSE LEARNING OUTCOMES (CLOs)				
	By the end of the course, students should be able to:			
CLO1:	Develop an appreciation for what is involved in learning models from data, and integrating ML in existing real-world processes			
CLO2:	Thoroughly understand the ML pipeline from design and data gathering to meaningful and relevant evaluation			
CLO3:	Learn a wide variety of learning algorithms, and formulate and implement solutions to machine learning problems			
CLO4:	Apply algorithms to real-world problems, optimize the trained models and report on the expected performance			

CLO	CLO Statement	Bloom's	PLOs/Graduate
		Cognitive Level	Attributes (Seoul Accord)
CLO1	Develop an appreciation for what is involved in learning models from data, and	C2, C3	PLO2
	integrating ML in existing real-world processes		
CLO2	Thoroughly understand the ML pipeline from design and data gathering to	C3, C4, C5	PLO2, PLO3, PLO4
	meaningful and relevant evaluation		
CLO3	Learn a wide variety of learning algorithms, and formulate and implement	C2, C5, C6	PLO4, PLO5
	solutions to machine learning problems		
CLO4	Apply algorithms to real-world problems, optimize the trained models and report	C5, C6	PLO5
	on the expected performance		

Grading Breakup and Policy

Assessment	Weight (%)	Related CLOs	ACM Recommended Disposition
Programming assignment(s)	25%	CLO2, CLO3, CLO4	D3, D4, D7, D9, D10
Quizzes	25%	CLO1, CLO2	D4, D7, D9, D10
Project	20%	CLO1 – CLO4	D1, D3, D4, D5, D6, D7, D8, D9, D11
Reading assignment(s)/homework(s)/Implementation of Research Paper(s)/viva	15%	CLO1 – CLO4	D3, D4, D7, D9, D10
Final examination + viva	15%	CLO1 – CLO4	D4, D7, D9, D10

Examination detail	l	
Midterm Exam	Yes/No: Duration: Exam Specifications:	No

Yes/No: Yes

Final Exam
Duration: 2.5 – 3 hours
Exam Specifications: In-person exam

SSE Council on Equity and Belonging

In addition to LUMS resources, SSE's **Council on Belonging and Equity** is committed to devising ways to provide a safe, inclusive and respectful learning, living, and working environment for students, faculty and staff. To seek counsel related to any issues, please feel free to approach either a member of the council or email at cbe.sse@lums.edu.pk.

Mental Health Support at LUMS

For matters relating to counselling, kindly email student.counselling@lums.edu.pk, or visit https://osa.lums.edu.pk/content/student-counselling@lums.edu.pk for more information. You are welcome to write to me or speak to me if you find that your mental health is impacting your ability to participate in the course. However, should you choose not to do so, please contact the Counseling Unit and speak to a counsellor or speak to the OSA team and ask them to write to me so that any necessary accommodations can be made.

Harassment Policy

SSE, LUMS and particularly this class, is a harassment free zone. Harassment of any kind is unacceptable, whether it be sexual harassment, online harassment, bullying, coercion, stalking, verbal or physical abuse of any kind. Harassment is a very broad term; it includes both direct and indirect behavior, it may be physical or psychological in nature, it may be perpetrated online or offline, on campus and off campus. It may be one offense, or it may comprise of several incidents which together amount to sexual harassment. It may include overt requests for sexual favors but can also constitute verbal or written communication of a loaded nature. Further details of what may constitute harassment may be found in the LUMS Sexual Harassment Policy, which is available as part of the university code of conduct.

LUMS has a Sexual Harassment Policy and a Sexual Harassment Inquiry Committee (SHIC). Any member of the LUMS community can file a formal or informal complaint with the SHIC. If you are unsure about the process of filing a complaint, wish to discuss your options or have any questions, concerns, or complaints, please write to the Office of Accessibility and Inclusion (OAI, oai@lums.edu.pk) and SHIC (shic@lums.edu.pk) —both of them exist to help and support you and they will do their best to assist you in whatever way they can. You can find more details regarding the LUMS sexual harassment policy here.

To file a complaint, please write to harassment@lums.edu.pk.

Rights and Code of Conduct for Online Teaching

A misuse of online modes of communication is unacceptable. TAs and faculty will seek consent before the recording of live online lectures or tutorials. Please ensure if you do not wish to be recorded during a session to inform the faculty member in a timely manner. Please also ensure that you prioritize formal means of communication (email, LMS) over informal means to communicate with course staff.

Cou	Course overview						
w	Topics	Recommended Readings	Relate d CLOs	ACM Comp Knowledge Landscape			
1.	Course overview What is ML? Traditional CS vs. ML, history of ML, AI vs. ML Classification and Regression with examples. Training and Testing. Rules vs. Patterns, Deterministic vs. Probabilistic, Certainty vs. Uncertainty. Learning: Supervised, unsupervised, semi-supervised Labeled data sources: Expert annotators, crowd Example ML application areas: Speech and Language Technologies Challenges and Opportunities of ML: Explainability Fairness and Societal Biases ML for Social Good, ML for Development (ML4D), Speech and Language Technologies for Development (SLT4D)	 Murphy chapter 1 Alpaydin, chapter 1 	CLO1, CLO2				
2.	 Supervised Learning Features, Labels, Training, Testing, Classification, Regression. Formalizing the supervised learning setup Feature spaces and feature vectors O Sparse and dense feature vectors, one-hot vectors O Bag-of-word features 	 Murphy: 1.1, 1.2, 1.4.2, 1.4.3, 1.4.9 Recommended topics: Goals of Cross Validation: 	CLO1, CLO2, CLO3				

Label spaces	Model selection,
 Label spaces for classification (binary 	and multiclass) and regression training, and
 Hypothesis spaces 	performance
 The No Free Lunch theorem 	estimation
 Choosing the hypothesis class H and 	hypothesis $h \in H$ • Types of Cross
 Various Algorithms for traversing hyp 	
■ Pick h randomly	Pros and Cons
	• Exhaustive
Just output the label of the	▼ Leave-D-
 Evaluating hypotheses: Loss functions and goal 	s of optimization out
o Zero-One	Leave-one-
o Squared	out
o Absolute	
 Loss reduction and Generalization in Learning 	Non-Exhaustive
o Memorizers	• k-fold
	Holdout
o Smoothing and Priors	Repeated
 Tradeoff between Bias and Variance 	random
• Sampling from the distribution $P(X, Y)$	subsamplin
 Representative datasets 	
 Training, validation, and testing 	g g
 How to split the dataset D? 	● Nested
O Time series data	ullet $k*l$ fold
o Independent and Identically Distribut	ed (IID) • k-fold with
• The weak law of large numbers $(\epsilon_{TE} \rightarrow \epsilon \ as \ D)$	$q_{TE}(1 \to +\infty)$
How to prevent overfitting to test data? Do's are the second of the	nd Don'ts sets
 Validation sets (dev sets) and Cross Validation 	
	Bootstrappi
	ng
	Stratified
	cross
	validation
	Time series
	cross validation
	(forward
	(forward chaining –
The K-Nearest Neighbor Classifier (An instance-based, la	chaining — Rolling origin)
The K-Nearest Neighbor Classifier (An instance-based, la non-parametric classifier)	zy, discriminative, non-linear, chaining – Rolling origin) zy, discriminative, non-linear, • Murphy: 1.1, CLO3,
	chaining – Rolling origin) zy, discriminative, non-linear, Murphy: 1.1, CLO3, 1.2, 1.4.2, 1.4.3, CLO4
non-parametric classifier) • KNN – The Basics	chaining – Rolling origin) zy, discriminative, non-linear, • Murphy: 1.1, 1.2, 1.4.2, 1.4.3, 1.4.9
non-parametric classifier) • KNN – The Basics • Nearest neighbor classification rule	chaining – Rolling origin) zy, discriminative, non-linear, Murphy: 1.1, 1.2, 1.4.2, 1.4.3, 1.4.9 Videos of
non-parametric classifier) • KNN – The Basics • Nearest neighbor classification rule • KNN formal definition	chaining – Rolling origin) zy, discriminative, non-linear, • Murphy: 1.1, 1.2, 1.4.2, 1.4.3, 1.4.9 • Videos of different values
non-parametric classifier) • KNN – The Basics • Nearest neighbor classification rule • KNN formal definition • KNN decision boundaries and Vorono	chaining – Rolling origin) zy, discriminative, non-linear, • Murphy: 1.1, 1.2, 1.4.2, 1.4.3, 1.4.9 • Videos of different values of k
non-parametric classifier) • KNN – The Basics • Nearest neighbor classification rule • KNN formal definition • KNN decision boundaries and Vorono • Properties of KNN: Non-parametric, u	chaining – Rolling origin) zy, discriminative, non-linear, Murphy: 1.1, 1.2, 1.4.2, 1.4.3, 1.4.9 Videos of different values of k
non-parametric classifier) • KNN – The Basics • Nearest neighbor classification rule • KNN formal definition • KNN decision boundaries and Vorono	chaining – Rolling origin) zy, discriminative, non-linear, • Murphy: 1.1, 1.2, 1.4.2, 1.4.3, 1.4.9 • Videos of different values of k
non-parametric classifier) • KNN – The Basics • Nearest neighbor classification rule • KNN formal definition • KNN decision boundaries and Vorono • Properties of KNN: Non-parametric, u	chaining — Rolling origin) zy, discriminative, non-linear, Murphy: 1.1, 1.2, 1.4.2, 1.4.3, 1.4.9 Videos of different values of k video describing nearest
non-parametric classifier) • KNN – The Basics • Nearest neighbor classification rule • KNN formal definition • KNN decision boundaries and Vorono • Properties of KNN: Non-parametric, uregression, instance-based, lazy • KNN Similarity/Distance measures and constrain	chaining – Rolling origin) zy, discriminative, non-linear, Murphy: 1.1, 1.2, 1.4.2, 1.4.3, 1.4.9 Videos of different values of k Video describing nearest neighbors
non-parametric classifier) • KNN – The Basics • Nearest neighbor classification rule • KNN formal definition • KNN decision boundaries and Vorono • Properties of KNN: Non-parametric, uregression, instance-based, lazy • KNN Similarity/Distance measures and constrait • Minkowski Distances (Manhattan, Euc	chaining — Rolling origin) zy, discriminative, non-linear, Murphy: 1.1, 1.2, 1.4.2, 1.4.3, 1.4.9 Videos of different values of k Video describing nearest neighbors lints clidean and Chebyshev) A nice
non-parametric classifier) • KNN – The Basics • Nearest neighbor classification rule • KNN formal definition • KNN decision boundaries and Vorono • Properties of KNN: Non-parametric, uregression, instance-based, lazy • KNN Similarity/Distance measures and constraition • Minkowski Distances (Manhattan, Euclidean) • The KNN algorithm and implementation	chaining — Rolling origin) zy, discriminative, non-linear, Murphy: 1.1, 1.2, 1.4.2, 1.4.3, 1.4.9 Videos of different values of k Video describing nearest neighbors h examples chaining — Rolling origin) LLO3, CLO4 Video sof different values of k Video describing nearest neighbors A nice explanation of
non-parametric classifier) • KNN – The Basics • Nearest neighbor classification rule • KNN formal definition • KNN decision boundaries and Vorono • Properties of KNN: Non-parametric, uregression, instance-based, lazy • KNN Similarity/Distance measures and constrai • Minkowski Distances (Manhattan, Euclidean Constrai) • The KNN algorithm and implementation • KNN regression and classification with	chaining — Rolling origin) izy, discriminative, non-linear, Murphy: 1.1, 1.2, 1.4.2, 1.4.3, 1.4.9 Videos of different values of k Video describing nearest neighbors ints clidean and Chebyshev) h examples CLO3, CLO4 Video describing nearest neighbors A nice explanation of nearest
non-parametric classifier) • KNN – The Basics • Nearest neighbor classification rule • KNN formal definition • KNN decision boundaries and Vorono • Properties of KNN: Non-parametric, uregression, instance-based, lazy • KNN Similarity/Distance measures and constrai • Minkowski Distances (Manhattan, Euclidean Constrai) • The KNN algorithm and implementation • KNN regression and classification with • Space and Time complexity	chaining — Rolling origin) zy, discriminative, non-linear, Murphy: 1.1, 1.2, 1.4.2, 1.4.3, 1.4.9 Videos of different values of k Video describing nearest neighbors h examples chaining — Rolling origin) LLO3, CLO4 Video sof different values of k Video describing nearest neighbors A nice explanation of
non-parametric classifier) • KNN – The Basics • Nearest neighbor classification rule • KNN formal definition • KNN decision boundaries and Vorono • Properties of KNN: Non-parametric, uregression, instance-based, lazy • KNN Similarity/Distance measures and constrai • Minkowski Distances (Manhattan, Euro • The KNN algorithm and implementation • KNN regression and classification with • Space and Time complexity • Bias/variance tradeoff as	chaining — Rolling origin) izy, discriminative, non-linear, Murphy: 1.1, 1.2, 1.4.2, 1.4.3, 1.4.9 Videos of different values of k Video describing nearest neighbors ints clidean and Chebyshev) h examples CLO3, CLO4 Video describing nearest neighbors A nice explanation of nearest
non-parametric classifier) • KNN – The Basics • Nearest neighbor classification rule • KNN formal definition • KNN decision boundaries and Vorono • Properties of KNN: Non-parametric, uregression, instance-based, lazy • KNN Similarity/Distance measures and constrai • Minkowski Distances (Manhattan, Euclidean Constrai) • The KNN algorithm and implementation • KNN regression and classification with • Space and Time complexity	chaining — Rolling origin) izy, discriminative, non-linear, Murphy: 1.1, 1.2, 1.4.2, 1.4.3, 1.4.9 Videos of different values of k Video describing nearest neighbors ints clidean and Chebyshev) h examples CLO3, CLO4 Video describing nearest neighbors A nice explanation of nearest
non-parametric classifier) • KNN – The Basics • Nearest neighbor classification rule • KNN formal definition • KNN decision boundaries and Vorono • Properties of KNN: Non-parametric, uregression, instance-based, lazy • KNN Similarity/Distance measures and constrai • Minkowski Distances (Manhattan, Euro • The KNN algorithm and implementation • KNN regression and classification with • Space and Time complexity • Bias/variance tradeoff as	chaining — Rolling origin) izy, discriminative, non-linear, Murphy: 1.1, 1.2, 1.4.2, 1.4.3, 1.4.9 Videos of different values of k Video describing nearest neighbors ints clidean and Chebyshev) h examples CLO3, CLO4 Video describing nearest neighbors A nice explanation of nearest
non-parametric classifier) • KNN – The Basics ○ Nearest neighbor classification rule ○ KNN formal definition ○ KNN decision boundaries and Vorono ○ Properties of KNN: Non-parametric, uregression, instance-based, lazy • KNN Similarity/Distance measures and constrai ○ Minkowski Distances (Manhattan, Euro • The KNN algorithm and implementation ○ KNN regression and classification with ○ Space and Time complexity ○ Bias/variance tradeoff as ○ K → 1 and K → n	chaining — Rolling origin) izy, discriminative, non-linear, Murphy: 1.1, 1.2, 1.4.2, 1.4.3, 1.4.9 Videos of different values of k Video describing nearest neighbors ints clidean and Chebyshev) h examples CLO3, CLO4 Video describing nearest neighbors A nice explanation of nearest
non-parametric classifier) • KNN – The Basics ○ Nearest neighbor classification rule ○ KNN formal definition ○ KNN decision boundaries and Vorono ○ Properties of KNN: Non-parametric, uregression, instance-based, lazy • KNN Similarity/Distance measures and constrai ○ Minkowski Distances (Manhattan, Euc • The KNN algorithm and implementation ○ KNN regression and classification with ○ Space and Time complexity ○ Bias/variance tradeoff as ○ K → 1 and K → n ○ Tuning the hyperparameter: K ○ KNN: The good, the bad and the ugly	chaining — Rolling origin) izy, discriminative, non-linear, Murphy: 1.1, 1.2, 1.4.2, 1.4.3, 1.4.9 Videos of different values of k Video describing nearest neighbors ints clidean and Chebyshev) h examples CLO3, CLO4 Video describing nearest neighbors A nice explanation of nearest
 Nearest neighbor classification rule Nearest neighbor classification rule KNN – The Basics Nearest neighbor classification rule KNN formal definition KNN decision boundaries and Vorono Properties of KNN: Non-parametric, urgression, instance-based, lazy KNN Similarity/Distance measures and constrai Minkowski Distances (Manhattan, Euro The KNN algorithm and implementation KNN regression and classification with Space and Time complexity Bias/variance tradeoff as K → 1 and K → n Tuning the hyperparameter: K KNN: The good, the bad and the ugly KNN error bounds as n → ∞ 	chaining — Rolling origin) izy, discriminative, non-linear, Murphy: 1.1, 1.2, 1.4.2, 1.4.3, 1.4.9 Videos of different values of k Video describing nearest neighbors ints clidean and Chebyshev) h examples CLO3, CLO4 Video describing nearest neighbors A nice explanation of nearest
 Non-parametric classifier) KNN – The Basics Nearest neighbor classification rule KNN formal definition KNN decision boundaries and Vorono Properties of KNN: Non-parametric, uregression, instance-based, lazy KNN Similarity/Distance measures and constrai Minkowski Distances (Manhattan, Euc The KNN algorithm and implementation KNN regression and classification with Space and Time complexity Bias/variance tradeoff as K → 1 and K → n Tuning the hyperparameter: K KNN: The good, the bad and the ugly KNN error bounds as n → ∞ Bayes Error 	chaining — Rolling origin) izy, discriminative, non-linear, Murphy: 1.1, 1.2, 1.4.2, 1.4.3, 1.4.9 Videos of different values of k Video describing nearest neighbors ints clidean and Chebyshev) h examples CLO3, CLO4 Video describing nearest neighbors A nice explanation of nearest
 Non-parametric classifier) KNN – The Basics Nearest neighbor classification rule KNN formal definition KNN decision boundaries and Vorono Properties of KNN: Non-parametric, uregression, instance-based, lazy KNN Similarity/Distance measures and constrai Minkowski Distances (Manhattan, Euclidean Minkowski Distances (Manhattan, Euclidean Minkowski Distances) KNN regression and classification with Space and Time complexity Bias/variance tradeoff as K → 1 and K → n Tuning the hyperparameter: K KNN: The good, the bad and the ugly KNN error bounds as n → ∞ Bayes Error 1-NN error as n → ∞ 	chaining — Rolling origin) izy, discriminative, non-linear, Murphy: 1.1, 1.2, 1.4.2, 1.4.3, 1.4.9 Videos of different values of k Video describing nearest neighbors ints clidean and Chebyshev) h examples CLO3, CLO4 Video describing nearest neighbors A nice explanation of nearest
 Non-parametric classifier) KNN – The Basics Nearest neighbor classification rule KNN formal definition KNN decision boundaries and Vorono Properties of KNN: Non-parametric, uregression, instance-based, lazy KNN Similarity/Distance measures and constrai Minkowski Distances (Manhattan, Euclidean Minkowski Distances (Manhattan, Euclidean Minkowsk	chaining — Rolling origin) izy, discriminative, non-linear, Murphy: 1.1, 1.2, 1.4.2, 1.4.3, 1.4.9 Videos of different values of k Video describing nearest neighbors ints clidean and Chebyshev) h examples CLO3, CLO4 Video describing nearest neighbors A nice explanation of nearest
 Non-parametric classifier) KNN – The Basics Nearest neighbor classification rule KNN formal definition KNN decision boundaries and Vorono Properties of KNN: Non-parametric, uregression, instance-based, lazy KNN Similarity/Distance measures and constrai Minkowski Distances (Manhattan, Euclidean Minkowski Distances (Manhattan, Euclidean Minkowski Distances) KNN regression and classification with Space and Time complexity Bias/variance tradeoff as K → 1 and K → n Tuning the hyperparameter: K KNN: The good, the bad and the ugly KNN error bounds as n → ∞ Bayes Error 1-NN error as n → ∞ 	chaining — Rolling origin) izy, discriminative, non-linear, Murphy: 1.1, 1.2, 1.4.2, 1.4.3, 1.4.9 Videos of different values of k Video describing nearest neighbors ints clidean and Chebyshev) h examples CLO3, CLO4 Video describing nearest neighbors A nice explanation of nearest
 Non-parametric classifier) KNN – The Basics Nearest neighbor classification rule KNN formal definition KNN decision boundaries and Vorono Properties of KNN: Non-parametric, uregression, instance-based, lazy KNN Similarity/Distance measures and constrai Minkowski Distances (Manhattan, Euclidean Minkowski Distances (Manhattan, Euclidean Minkowsk	chaining — Rolling origin) izy, discriminative, non-linear, Murphy: 1.1, 1.2, 1.4.2, 1.4.3, 1.4.9 Videos of different values of k Video describing nearest neighbors ints clidean and Chebyshev) h examples CLO3, CLO4 Video describing nearest neighbors A nice explanation of nearest
 Nearest neighbor classification rule KNN – The Basics Nearest neighbor classification rule KNN formal definition KNN decision boundaries and Vorono Properties of KNN: Non-parametric, uregression, instance-based, lazy KNN Similarity/Distance measures and constraition Minkowski Distances (Manhattan, Euclidean Minkowski Distances (Manhattan, Eucl	chaining — Rolling origin) izy, discriminative, non-linear, Murphy: 1.1, 1.2, 1.4.2, 1.4.3, 1.4.9 Videos of different values of k Video describing nearest neighbors ints clidean and Chebyshev) h examples CLO3, CLO4 Video describing nearest neighbors A nice explanation of nearest
 Nearest neighbor classification rule KNN – The Basics Nearest neighbor classification rule KNN formal definition KNN decision boundaries and Vorono Properties of KNN: Non-parametric, uregression, instance-based, lazy KNN Similarity/Distance measures and constrai Minkowski Distances (Manhattan, Euro The KNN algorithm and implementation KNN regression and classification with Space and Time complexity Bias/variance tradeoff as K → 1 and K → n Tuning the hyperparameter: K KNN: The good, the bad and the ugly KNN error bounds as n → ∞ Bayes Error 1-NN error as n → ∞ KNN Enhancements Parzen Windows and Kernels K-D trees Inverted lists 	chaining — Rolling origin) izy, discriminative, non-linear, Murphy: 1.1, 1.2, 1.4.2, 1.4.3, 1.4.9 Videos of different values of k Video describing nearest neighbors ints clidean and Chebyshev) h examples CLO3, CLO4 Video describing nearest neighbors A nice explanation of nearest
 Nearest neighbor classification rule KNN – The Basics Nearest neighbor classification rule KNN formal definition KNN decision boundaries and Vorono Properties of KNN: Non-parametric, uregression, instance-based, lazy KNN Similarity/Distance measures and constrai Minkowski Distances (Manhattan, Euro The KNN algorithm and implementation KNN regression and classification with Space and Time complexity Bias/variance tradeoff as K → 1 and K → n Tuning the hyperparameter: K KNN: The good, the bad and the ugly KNN error bounds as n → ∞ Bayes Error 1-NN error as n → ∞ KNN Enhancements Parzen Windows and Kernels K-D trees Inverted lists Locality sensitive hashing 	chaining — Rolling origin) izy, discriminative, non-linear, Murphy: 1.1, 1.2, 1.4.2, 1.4.3, 1.4.9 Videos of different values of k Video describing nearest neighbors ints clidean and Chebyshev) h examples CLO3, CLO4 Video describing nearest neighbors A nice explanation of nearest
 Nearest neighbor classification rule KNN – The Basics Nearest neighbor classification rule KNN formal definition KNN decision boundaries and Vorono Properties of KNN: Non-parametric, uregression, instance-based, lazy KNN Similarity/Distance measures and constrai Minkowski Distances (Manhattan, Eucleton KNN regression and classification with Space and Time complexity Bias/variance tradeoff as KNY = 1 and K → n Tuning the hyperparameter: K KNN: The good, the bad and the ugly KNN error bounds as n → ∞ Bayes Error 1-NN error as n → ∞ KNN Enhancements Parzen Windows and Kernels K-D trees Inverted lists Locality sensitive hashing The Curse of Dimensionality 	chaining — Rolling origin) izy, discriminative, non-linear, Murphy: 1.1, 1.2, 1.4.2, 1.4.3, 1.4.9 Videos of different values of k Video describing nearest neighbors ints clidean and Chebyshev) h examples CLO3, CLO4 Video describing nearest neighbors A nice explanation of nearest
 Nearest neighbor classification rule KNN – The Basics Nearest neighbor classification rule KNN formal definition KNN decision boundaries and Vorono Properties of KNN: Non-parametric, uregression, instance-based, lazy KNN Similarity/Distance measures and constrai Minkowski Distances (Manhattan, Euro The KNN algorithm and implementation KNN regression and classification with Space and Time complexity Bias/variance tradeoff as K → 1 and K → n Tuning the hyperparameter: K KNN: The good, the bad and the ugly KNN error bounds as n → ∞ Bayes Error 1-NN error as n → ∞ KNN Enhancements Parzen Windows and Kernels K-D trees Inverted lists Locality sensitive hashing 	chaining — Rolling origin) izy, discriminative, non-linear, Murphy: 1.1, 1.2, 1.4.2, 1.4.3, 1.4.9 Videos of different values of k Video describing nearest neighbors ints clidean and Chebyshev) h examples CLO3, CLO4 Video describing nearest neighbors A nice explanation of nearest

	 Lower dimensional subspaces and manifolds in higher dimensional ambient space 		
Tutorial 1	Topics:		
•	Review of KNNs		
	Review of Voronoi Tessellations		
•	Review (and Expansion) of Error Bounds		
	Review (and Expansion) of the Curse of Dimensionality		
	on metrics	• SLP3: 4.7-4.9	CLO2,
•	The Confusion matrix (contingency tables) – binary and multi-label		CLO3
	 True and False – Positives and Negatives 		
	Type I and type II errors		
•	Performance Metrics		
	o Accuracy, Sensitivity (recall, TPR), Specificity (TNR), Precision (Positive		
	Predictive Value), Negative predictive value		
	False acceptance rate, False rejection rate		
	O Examples – pros and cons		
•	The need for a combined measure		
	 Types of averages: AM, GM, HM Fβ-measure, F-1-measure 		
_	Multiclass Classification		
	O Any-of (multi-label) classification		
	One-of (multinomial) classification-		
	Micro and Macro averaging		
•	Gold labels and annotation of data		
	 Inter-annotator agreements 		
	 Cohen's Kappa and Krippendorff's alpha 		
•	Evaluation of Classifiers, thresholds, comparing classifiers, imbalanced classes		
	 Receiver operating Characteristic (ROC) and Precision-Recall (P-R) 		
	Curves		
	 ROC Area Under the Curve (AUC) Equal Error Rate (EER) and Biometric Systems 		
	 ROC Area Under the Curve (AUC) Equal Error Rate (EER) and Biometric Systems Topics: Review of Cohen's Kappa		
•	O ROC Area Under the Curve (AUC) O Equal Error Rate (EER) and Biometric Systems Topics: Review of Cohen's Kappa Review of ROC and Precision-Recall Curves		
•	O ROC Area Under the Curve (AUC) O Equal Error Rate (EER) and Biometric Systems Topics: Review of Cohen's Kappa Review of ROC and Precision-Recall Curves Assignment 1	• FSUILCh3	CLO2,
• • • Linear Re	O ROC Area Under the Curve (AUC) O Equal Error Rate (EER) and Biometric Systems Topics: Review of Cohen's Kappa Review of ROC and Precision-Recall Curves Assignment 1	• ESLII Ch3	CLO2, CLO3,
• • • Linear Re	O ROC Area Under the Curve (AUC) O Equal Error Rate (EER) and Biometric Systems Topics: Review of Cohen's Kappa Review of ROC and Precision-Recall Curves Assignment 1 egression	• Murphy 7-7.5.1,	
Linear Re	O ROC Area Under the Curve (AUC) O Equal Error Rate (EER) and Biometric Systems Topics: Review of Cohen's Kappa Review of ROC and Precision-Recall Curves Assignment 1 egression Motivation for linearity		CLO3,
Linear Re	O ROC Area Under the Curve (AUC) O Equal Error Rate (EER) and Biometric Systems Topics: Review of Cohen's Kappa Review of ROC and Precision-Recall Curves Assignment 1 Egression Motivation for linearity Revision O Lines, planes, hyperplanes, and vectors Lines and planes: Normal form and slope-intercept form	• Murphy 7-7.5.1,	CLO3,
Linear Re	O ROC Area Under the Curve (AUC) O Equal Error Rate (EER) and Biometric Systems Topics: Review of Cohen's Kappa Review of ROC and Precision-Recall Curves Assignment 1 Peression Motivation for linearity Revision O Lines, planes, hyperplanes, and vectors ■ Lines and planes: Normal form and slope-intercept form Decision boundaries with perpendicular weight vectors	• Murphy 7-7.5.1,	CLO3,
Linear Re	O ROC Area Under the Curve (AUC) O Equal Error Rate (EER) and Biometric Systems Topics: Review of Cohen's Kappa Review of ROC and Precision-Recall Curves Assignment 1 Egression Motivation for linearity Revision O Lines, planes, hyperplanes, and vectors ■ Lines and planes: Normal form and slope-intercept form Decision boundaries with perpendicular weight vectors Distance between a hyperplane and a point	• Murphy 7-7.5.1,	CLO3,
Linear Re	O ROC Area Under the Curve (AUC) O Equal Error Rate (EER) and Biometric Systems Topics: Review of Cohen's Kappa Review of ROC and Precision-Recall Curves Assignment 1 Egression Motivation for linearity Revision O Lines, planes, hyperplanes, and vectors ■ Lines and planes: Normal form and slope-intercept form ■ Decision boundaries with perpendicular weight vectors ■ Distance between a hyperplane and a point O The Dot Product	• Murphy 7-7.5.1,	CLO3,
Linear Re	O ROC Area Under the Curve (AUC) O Equal Error Rate (EER) and Biometric Systems **Topics:** Review of Cohen's Kappa Review of ROC and Precision-Recall Curves Assignment 1 **Egression** Motivation for linearity Revision O Lines, planes, hyperplanes, and vectors ■ Lines and planes: Normal form and slope-intercept form ■ Decision boundaries with perpendicular weight vectors ■ Distance between a hyperplane and a point O The Dot Product O The geometric interpretation of absorbing the bias term	• Murphy 7-7.5.1,	CLO3,
Linear Re	O ROC Area Under the Curve (AUC) O Equal Error Rate (EER) and Biometric Systems **Topics:** Review of Cohen's Kappa Review of ROC and Precision-Recall Curves **Assignment 1** **Egression** Motivation for linearity Revision O Lines, planes, hyperplanes, and vectors ■ Lines and planes: Normal form and slope-intercept form ■ Decision boundaries with perpendicular weight vectors ■ Distance between a hyperplane and a point O The Dot Product O The geometric interpretation of absorbing the bias term O Visualizing n dimensions	• Murphy 7-7.5.1,	CLO3,
Linear Re	O ROC Area Under the Curve (AUC) O Equal Error Rate (EER) and Biometric Systems Topics: Review of Cohen's Kappa Review of ROC and Precision-Recall Curves Assignment 1 Egression Motivation for linearity Revision O Lines, planes, hyperplanes, and vectors ■ Lines and planes: Normal form and slope-intercept form ■ Decision boundaries with perpendicular weight vectors ■ Distance between a hyperplane and a point O The Dot Product O The geometric interpretation of absorbing the bias term O Visualizing n dimensions Linear Regression	• Murphy 7-7.5.1,	CLO3,
Linear Re	O ROC Area Under the Curve (AUC) O Equal Error Rate (EER) and Biometric Systems Topics: Review of Cohen's Kappa Review of ROC and Precision-Recall Curves Assignment 1 Egression Motivation for linearity Revision O Lines, planes, hyperplanes, and vectors ■ Lines and planes: Normal form and slope-intercept form ■ Decision boundaries with perpendicular weight vectors ■ Distance between a hyperplane and a point O The Dot Product O The geometric interpretation of absorbing the bias term O Visualizing n dimensions Linear Regression O Intuition	• Murphy 7-7.5.1,	CLO3,
Linear Re	O ROC Area Under the Curve (AUC) O Equal Error Rate (EER) and Biometric Systems Topics: Review of Cohen's Kappa Review of ROC and Precision-Recall Curves Assignment 1 Egression Motivation for linearity Revision O Lines, planes, hyperplanes, and vectors Lines and planes: Normal form and slope-intercept form Decision boundaries with perpendicular weight vectors Distance between a hyperplane and a point O The Dot Product O The geometric interpretation of absorbing the bias term O Visualizing n dimensions Linear Regression O Intuition O Derivation and implementation	• Murphy 7-7.5.1,	CLO3,
Linear Re	O ROC Area Under the Curve (AUC) O Equal Error Rate (EER) and Biometric Systems Topics: Review of Cohen's Kappa Review of ROC and Precision-Recall Curves Assignment 1 Egression Motivation for linearity Revision O Lines, planes, hyperplanes, and vectors Lines and planes: Normal form and slope-intercept form Decision boundaries with perpendicular weight vectors Distance between a hyperplane and a point O The Dot Product O The geometric interpretation of absorbing the bias term O Visualizing n dimensions Linear Regression O Intuition O Derivation and implementation Linear regression with one variable	• Murphy 7-7.5.1,	CLO3,
Linear Re	O ROC Area Under the Curve (AUC) O Equal Error Rate (EER) and Biometric Systems **Topics:** Review of Cohen's Kappa Review of ROC and Precision-Recall Curves Assignment 1 **Peression** Motivation for linearity Revision O Lines, planes, hyperplanes, and vectors ■ Lines and planes: Normal form and slope-intercept form ■ Decision boundaries with perpendicular weight vectors ■ Distance between a hyperplane and a point O The Dot Product O The geometric interpretation of absorbing the bias term O Visualizing n dimensions Linear Regression O Intuition O Derivation and implementation ■ Linear regression with one variable ■ Cost function: Square Loss, Mean Square Loss (MSE),	• Murphy 7-7.5.1,	CLO3,
Linear Re	O ROC Area Under the Curve (AUC) O Equal Error Rate (EER) and Biometric Systems **Topics:** Review of Cohen's Kappa Review of ROC and Precision-Recall Curves **Assignment 1** **Egression** Motivation for linearity Revision O Lines, planes, hyperplanes, and vectors ■ Lines and planes: Normal form and slope-intercept form ■ Decision boundaries with perpendicular weight vectors ■ Distance between a hyperplane and a point O The Dot Product O The geometric interpretation of absorbing the bias term O Visualizing n dimensions Linear Regression O Intuition O Derivation and implementation ■ Linear regression with one variable ■ Cost function: Square Loss, Mean Square Loss (MSE), ■ **1/2 Mean Square Loss	• Murphy 7-7.5.1,	CLO3,
Linear Re	O ROC Area Under the Curve (AUC) O Equal Error Rate (EER) and Biometric Systems **Topics:** Review of Cohen's Kappa Review of ROC and Precision-Recall Curves **Assignment 1** **Pagression** Motivation for linearity Revision O Lines, planes, hyperplanes, and vectors ■ Lines and planes: Normal form and slope-intercept form ■ Decision boundaries with perpendicular weight vectors ■ Distance between a hyperplane and a point O The Dot Product O The geometric interpretation of absorbing the bias term O Visualizing n dimensions Linear Regression O Intuition O Derivation and implementation ■ Linear regression with one variable ■ Cost function: Square Loss, Mean Square Loss (MSE), ■ \frac{1}{2} Mean Square Loss ■ Motivating gradient descent	• Murphy 7-7.5.1,	CLO3,
Linear Re	O ROC Area Under the Curve (AUC) O Equal Error Rate (EER) and Biometric Systems **Topics:** Review of Cohen's Kappa Review of ROC and Precision-Recall Curves Assignment 1 **Pegression** Motivation for linearity Revision O Lines, planes, hyperplanes, and vectors ■ Lines and planes: Normal form and slope-intercept form Decision boundaries with perpendicular weight vectors Distance between a hyperplane and a point O The Dot Product O The geometric interpretation of absorbing the bias term O Visualizing n dimensions Linear Regression O Intuition O Derivation and implementation ■ Linear regression with one variable ■ Cost function: Square Loss, Mean Square Loss (MSE), ■ \frac{1}{2} Mean Square Loss ■ Motivating gradient descent The Gradient Descent Algorithm	• Murphy 7-7.5.1,	CLO3,
Linear Re	O ROC Area Under the Curve (AUC) O Equal Error Rate (EER) and Biometric Systems **Topics:** Review of Cohen's Kappa Review of ROC and Precision-Recall Curves Assignment 1 **Egression** Motivation for linearity Revision O Lines, planes, hyperplanes, and vectors ■ Lines and planes: Normal form and slope-intercept form ■ Decision boundaries with perpendicular weight vectors ■ Distance between a hyperplane and a point O The Dot Product O The geometric interpretation of absorbing the bias term O Visualizing n dimensions Linear Regression O Intuition O Derivation and implementation ■ Linear regression with one variable ■ Cost function: Square Loss, Mean Square Loss (MSE), ■ ½ Mean Square Loss ■ Motivating gradient descent The Gradient Descent Algorithm O Gradients	• Murphy 7-7.5.1,	CLO3,
Linear Re	O ROC Area Under the Curve (AUC) O Equal Error Rate (EER) and Biometric Systems Fopics: Review of Cohen's Kappa Review of ROC and Precision-Recall Curves Assignment 1 Egression Motivation for linearity Revision O Lines, planes, hyperplanes, and vectors ■ Lines and planes: Normal form and slope-intercept form ■ Decision boundaries with perpendicular weight vectors ■ Distance between a hyperplane and a point O The Dot Product O The geometric interpretation of absorbing the bias term O Visualizing n dimensions Linear Regression O Intuition O Derivation and implementation ■ Linear regression with one variable ■ Cost function: Square Loss, Mean Square Loss (MSE), ■ ½ Mean Square Loss ■ Motivating gradient descent The Gradient Descent Algorithm O Gradients O The step size,	• Murphy 7-7.5.1,	CLO3,
Linear Re	O ROC Area Under the Curve (AUC) O Equal Error Rate (EER) and Biometric Systems **Topics:** Review of Cohen's Kappa Review of ROC and Precision-Recall Curves Assignment 1 **Peression** Motivation for linearity Revision O Lines, planes, hyperplanes, and vectors ■ Lines and planes: Normal form and slope-intercept form ■ Decision boundaries with perpendicular weight vectors ■ Distance between a hyperplane and a point O The Dot Product O The geometric interpretation of absorbing the bias term O Visualizing n dimensions Linear Regression O Intuition O Derivation and implementation ■ Linear regression with one variable ■ Cost function: Square Loss, Mean Square Loss (MSE), ■ ½ Mean Square Loss ■ Motivating gradient descent The Gradient Descent Algorithm O Gradients O The step size, O α	• Murphy 7-7.5.1,	CLO3,
Linear Re	O ROC Area Under the Curve (AUC) O Equal Error Rate (EER) and Biometric Systems **Topics:** Review of Cohen's Kappa Review of ROC and Precision-Recall Curves Assignment 1 **Peression** Motivation for linearity Revision O Lines, planes, hyperplanes, and vectors ■ Lines and planes: Normal form and slope-intercept form ■ Decision boundaries with perpendicular weight vectors ■ Distance between a hyperplane and a point O The Dot Product O The geometric interpretation of absorbing the bias term O Visualizing n dimensions Linear Regression O Intuition O Derivation and implementation ■ Linear regression with one variable ■ Cost function: Square Loss, Mean Square Loss (MSE), ■ ½ Mean Square Loss ■ Motivating gradient descent The Gradient Descent Algorithm O Gradients O The step size, O α O Convex and non-convex cost functions — Global and local optima	• Murphy 7-7.5.1,	CLO3,
Linear Re	O ROC Area Under the Curve (AUC) O Equal Error Rate (EER) and Biometric Systems Review of Cohen's Kappa Review of ROC and Precision-Recall Curves	• Murphy 7-7.5.1,	CLO3,
Linear Re	O ROC Area Under the Curve (AUC) O Equal Error Rate (EER) and Biometric Systems **Topics:** Review of Cohen's Kappa Review of ROC and Precision-Recall Curves **Assignment 1** **Segression** Motivation for linearity Revision O Lines, planes, hyperplanes, and vectors ■ Lines and planes: Normal form and slope-intercept form ■ Decision boundaries with perpendicular weight vectors ■ Distance between a hyperplane and a point O The Dot Product O The geometric interpretation of absorbing the bias term O Visualizing n dimensions Linear Regression O Intuition O Derivation and implementation ■ Linear regression with one variable ■ Cost function: Square Loss, Mean Square Loss (MSE), ■ ½ Mean Square Loss ■ Motivating gradient descent The Gradient Descent Algorithm O Gradients O The step size, O α O Convex and non-convex cost functions — Global and local optima O The batch gradient Descent: Batch, Minibatch, and Stochastic	• Murphy 7-7.5.1,	CLO3,
Linear Re	O ROC Area Under the Curve (AUC) O Equal Error Rate (EER) and Biometric Systems Review of Cohen's Kappa Review of ROC and Precision-Recall Curves	• Murphy 7-7.5.1,	CLO3,

	 Practical Issues of linear regression 		
	 Feature Scaling, local minima, ravines, saddle points, tracking progress 		
	in GD		
	 Hyperparameters: Learning rate 		
	Polynomial regression		
	Tutorial Topics:		
	Review of Hyperplanes		
	Review of Gradient Descent		
	Demo on Polynomial Regression		
	Review (and Expansion) of Standardization for Linear Regression		
6.	Linear Regression: Bias and Variance	• Dan Taalaan/a	CLO2,
0.	How to recognize high variance/high bias scenarios?	Ben Taskar's	CLO3,
	Underfitting and Overfitting	under- and	CLO3,
	O How to reduce bias and variance?	overfitting	CLO4
	■ Cross validation	MLaPP: 1.4.7	
		 Andrew Ng's 	
	Feature selection	<u>lecture</u> – ML	
	Manual Feature Selection Seatter Disgrams and Blate	debugging	
	Scatter Diagrams and Plots Scatter Diagrams and Plots	• Ben	
	O Eyeballing those Correlations!	Taskar's Notes	
	Regularization	on Bias Variance	
	Motivation – The fitting problem	Notes by Scott	
	o L2 Regularization or Ridge Regression	Foreman-Roe	
	o L1 Regularization or Lasso Regression		
	 Automatic Feature selection 	ELSII Chapter	
	 Comparison of Ridge and Lasso regression 	2.9	
	Elastic Net Regression – Intuition	Murphy: 6.2.2	
	Logistic Regression (A linear, discriminative, parametric classifier)	SLP3 Ch5, ESLII	
	Intuition and derivation	Ch4	
	 Regression for classification 		
	o "Squishing" between 0 and 1 using a non-linear activation function:	 Murphy 8, 13.3, 13.5.3 	
	The sigmoid		
	A simple sentiment classifier	 Ben Taskar's 	
	Visualizing the logistic regression decision boundary	<u>notes</u>	
	Hyperplanes, linear and non-linear decision boundaries	TM chapter:	
	Cost function: Derivation of the cross-entropy loss function (log loss)	Naive Bayes and	
	Learning algorithm: Batch, Stochastic and Mini-batch Gradient Descent	Logistic	
	Multiclass (multinomial) classification: One-vs-all (one-vs-rest), One-vs-one	Regression	
	The SoftMax activation function and multivariate log loss	• <u>Nice</u>	
	The solution failure in a failu	blogpost on	
	Tutorial Topics:	Gradient	
	Review of Bias and Variance: detection and techniques to deal with them	Descent.	
	Review (and Expansion) of Feature Selection	Adagrad,	
	Review (and Expansion) of Fedure Selection Review (and Expansion) of Sigmoid and Softmax: derivatives, visualizations,	Newton's	
	dealing with overflow	method	
	Review of using Binary Classifiers to setup Multiclass Classification		
	Assignment 2		
7.	The Perceptron (A discriminative, linear, parametric classifier)		CLO2
/.		The Perceptron	CLO2,
		Wiki page	CLO3,
	The Perceptron and its limitations The Heaviside step function	 Murphy 8.5.4 	CLO4
	O The Heaviside step function		
	O Boolean functions: AND, OR and XOR!	• Murphy: 14.5 -	
	O One perceptron, two perceptrons,	14.5.2.2	
	Linear separability in low and high dimensional spaces	Ben Taskar's	
	From the step function to other activation functions	Notes on SVMs	
	The perceptron learning algorithm and its geometric interpretation	·	
	Proof of convergence	• <u>Kernel</u>	
	 Relation between margin and rate of convergence 	Cookbook by	
		David Duvenaud	
	Maximum Margin Classifiers: Support Vector Machines (SVMs) (A discriminative,	• <u>Laurent El</u>	
	linear/non-linear classifier)	Ghaoui's lecture	
	Intuition and motivation	s on duality	
	 The perceptron and the optimal separating hyperplane 	 "Idiot's guide to 	
	Hard Margin Linear Support Vector Machines: Derivation	SVM"	
	I.		

Octorstained optimization and Lagrange Multipliers Soft Margin Linear Support Vector Mulchiner: Derivation Hinge-loss Kernels and Kernel SVMs Tutorial Topics: Review of Margin, Hinge Loss, and optimizing SVMs Review India Crypansion) of Soft Margin Classifiers and their Hyperparameters Review of Margin, Hinge Loss, and optimizing SVMs Review Charles (Linear, Polynomial, RBTs Supplementary: Ne Kernel Trick and Quadratic Programming Neural Networks The Neuron and linear decision boundaries. Can we do better? Review Logistic repression and gradient descent Changing the representation of the data Narials Non-linear activation Multi-layer Perception Luly Hollian How Nis work? Non-linear activation Luly Hollian How Nis work? Non-linear and complex decision boundaries Universal approximation and regression Layers: Depth vs. width Gradient descent for Nis Overfitting and the importance of Stochastic Gradient Descent (SGD) Formal Notation for ton Nis Overfitting and the importance of Stochastic Gradient Descent (SGD) Formal Notation for Logistic Regression and Nis Vectorizing, It and Nish Forward propagation (IR, Ni) Simple, vectored with a single instance Vectorizing, It and Nish Review of the rorward Pass and Backward Pass (in equations) Using Decision Boundaries to infer the architecture of a Ni Review of the rorward Pass and Backward Pass (in equations) Using Decision Boundaries to deal with Overfitting and Underfitting Narian Properties of the Nish O			T	
Mingeloss Kernels and Kernel SVMs		 Constrained optimization and Lagrange Multipliers 		
Mingeloss Kernels and Kernel SVMs		Soft Margin Linear Support Vector Machines: Derivation		
Recrieds and Kernels SVMs Decision Boundaries and Hyperplanes of SVMs Review (and Expansion) of Soft Margin Classification of Leaving Classification of Lea				
Tutorial Tapics: Decision Boundaries and Hyperplanes of SVMs Review (and Expansion) of Soft Margin Classifiers and their Hyperparameters Review of Margins, Hinge Loss, and optimizing SVMs Review of Kernesis Linear, Polynomial, RBFs Supplementary: the Kernel Trick and Quadratic Programming Neural Networks The Neuron and linear decision boundaries. Can we do better? Review of Margins, Hinge Loss, and a politic regression and gradient descent Review of Sernesis Review of Kernesis Review Re				
Decision Boundaries and Hyperplanes of SVMs Review (and Expansion) of Soft Margin Classifiers and their Hyperparameters Review of Kernetis Linear, Polynomial, RBTs Supplementary: the Kernel Trick and Quadratic Programming Neural Networks The Neuron and linear decision boundaries. Can we do better? On Review Logistic regression and gradient descent On Review Logistic regression and gradient descent On Neural Networks On Neural networks On Neural Networks On Multi-layer Perceptron Neural Networks Understand Approximation and regression Neural Networks Understand Approximation and regression Understand Approximation and regression Layer Department of Complex decision boundaries On-Invest Networks One-firting and the importance of Stochastic Gradient Descent (SGD) Formal Notation for Logistic Regression and NNs Overfirting and the importance of Stochastic Gradient Descent (SGD) Formal Notation for Logistic Regression and NNs Overfirting and the importance of Stochastic Gradient Descent (SGD) Formal Notation for Logistic Regression and NNs Overfirting and the importance of Stochastic Gradient Descent (SGD) Formal Notation for Logistic Regression and NNs Overfirting and the importance of Stochastic Gradient Descent (SGD) Formal Notation for Logistic Regression and NNs Overfirting and the importance of Stochastic Gradient Descent (SGD) Formal Notation for Logistic Regression and NNs Overfirting and the importance of Stochastic Gradient Descent (SGD) Formal Notation for Logistic Regression and NNs Overfirting and the International NNs Overfirting and the International Nns Overfirting and Nns Overfirting and Underfirting Assignment S Notation of Sequence Modeling tasks Or Feet and Speech (One-Hot) Overfirting and Underfirting Assignment S Notation of Sequence Modeling tasks Or Feet and Speech (One-Hot) Overfirting Assignment Classification Overfirting Assignment S Notation of Sequence Modeling tasks Or Feet and Speech (One-Hot) Overfirting Assignment S Notation of Franciscular Overfirting Assignment S Notation of Fran		Remeis and Remei Sams		
Decision Boundaries and Hyperplanes of SVMs Review (and Expansion) of Soft Margin Classifiers and their Hyperparameters Review of Kernetis Linear, Polynomial, RBTs Supplementary: the Kernel Trick and Quadratic Programming Neural Networks The Neuron and linear decision boundaries. Can we do better? On Review Logistic regression and gradient descent On Review Logistic regression and gradient descent On Neural Networks On Neural networks On Neural Networks On Multi-layer Perceptron Neural Networks Understand Approximation and regression Neural Networks Understand Approximation and regression Understand Approximation and regression Layer Department of Complex decision boundaries On-Invest Networks One-firting and the importance of Stochastic Gradient Descent (SGD) Formal Notation for Logistic Regression and NNs Overfirting and the importance of Stochastic Gradient Descent (SGD) Formal Notation for Logistic Regression and NNs Overfirting and the importance of Stochastic Gradient Descent (SGD) Formal Notation for Logistic Regression and NNs Overfirting and the importance of Stochastic Gradient Descent (SGD) Formal Notation for Logistic Regression and NNs Overfirting and the importance of Stochastic Gradient Descent (SGD) Formal Notation for Logistic Regression and NNs Overfirting and the importance of Stochastic Gradient Descent (SGD) Formal Notation for Logistic Regression and NNs Overfirting and the importance of Stochastic Gradient Descent (SGD) Formal Notation for Logistic Regression and NNs Overfirting and the International NNs Overfirting and the International Nns Overfirting and Nns Overfirting and Underfirting Assignment S Notation of Sequence Modeling tasks Or Feet and Speech (One-Hot) Overfirting and Underfirting Assignment S Notation of Sequence Modeling tasks Or Feet and Speech (One-Hot) Overfirting Assignment Classification Overfirting Assignment S Notation of Sequence Modeling tasks Or Feet and Speech (One-Hot) Overfirting Assignment S Notation of Franciscular Overfirting Assignment S Notation of Fran		Tutovial Tonica		
Review of Margins, Hinge Loss, and optimizing SVMs Review of Kernels: Linear, Polynomial, RBFs Supplementary: the Kernel Trick and Quadratic Programming Neural Networks In Neural networks Neural Networks Neural Networks Neural Networks Neural Networks Neural Networks Non-linear activation Non-linear activ		•		
Review of Margins, Hinge Loss, and optimizing SVMs Review of Kernels: Linear, Polymonial, RBrs Supplementary: the Kernel Trick and Quadratic Programming The Neuron and linear decision boundaries. Can we do better? OR Review Lugistic Regression and gradient descent Changing the representation of the data Kernels ON-Incera rativation Multi-layer Perceptron Neural networks Deep Learning Intuition: How Nhs work? ON-Inlear and complex decision boundaries O Universal approximation and regression Layers: Depth vs. width Gradient descent for Nhs O Verfitting and the importance of Stochastic Gradient Descent (SGD) Formal Notation for Logistic Regression and Nhs O Verctoriging IR and Nhs O Forward propagation (IR, Nh) Simple, vectored with a single instance Vectored with m instances Activation functions, gradients, and pros and cons O Sigmoid O Tanh OR REU and Leaky ReU Backward propagation (IR, Nh) Simple, vectored with a single instance Vectored with m instances Training the Nn Tutorial Topics Review of the Forward Pass and Backward Pass (in equations) Using Decision Boundaries to Infer the architecture of a Nh Review of strategies to deal with Overfitting and Underfitting Assignment 3 Supplementary: programming demo on using MLPs for character-level Language Modeling Notion of Sequence Modeling tasks Times Series Videos as sequences of images Types of Sequence Modeling tasks Notion of Sequence Modeling tasks Notion of Sequence Modeling morblems and specifications Videos as sequences of images Types of Sequence Modeling problems and specifications On Many-to-1- Video Classification, Sentiment Classification On Many-to-1- Video Classification, Sentiment Classi				
Review of Kernels: Unear, Polynomial, RBFs Supplementary: the Kernel Trick and Quadratic Programming Neural Networks The Neuron and linear decision boundaries. Can we do better? Review Logistic regression and gradient descent Changing the representation of the data Kernels Neural networks Non-linear activation Null-layer Perceptron Neural Networks Neural Networks Neural Networks Neural Networks Non-linear activation Null-layer Perceptron Neural Networks Non-linear activation Null-layer Perceptron Null-layer Perceptron Null-layer Perceptron Non-linear activation Null-layer Perceptron Null-layer		 Review (and Expansion) of Soft Margin Classifiers and their Hyperparameters 		
Neural Networks The Neural Prick and Quadratic Programming		 Review of Margins, Hinge Loss, and optimizing SVMs 		
Neural Networks The Neural Prick and Quadratic Programming		 Review of Kernels: Linear, Polynomial, RBFs 		
Neural Networks The Neuron and linear decision boundaries. Can we do better?				
The Neuron and linear decision boundaries. Can we do better? Review Logistic regression and gradient descent Changing the representation of the data Nemels Nemels Nemels Neural networks Non-linear activation Multi-layer Perceptron Noural networks Deep Learning Universial approximation and regression Layers: Depth vs. width Gradient descent for NNs On-linear and complex decision boundaries Universial approximation and regression Layers: Depth vs. width Gradient descent for NNs One-linear and complex decision boundaries Universial approximation and regression Layers: Depth vs. width Gradient descent for NNs One-linear and complex decision boundaries Universial approximation and regression Layers: Depth vs. width Gradient descent for NNs One-linear and complex decision boundaries Universial approximation and regression Layers: Depth vs. width Gradient descent for NNs One-linear and complex decision boundaries Forward propagation (IK, NN) Simple, vectored with a single instance Vectored with m instances Activation functions, gradients, and pros and cons Osigmoid Tanh Ret U and Leaky Ret U Backward propagation (IK, NN) Simple, vectored with a single instance and m instances Training the NN Tutorial Topics: Review of the Forward Pass and Backward Pass (in equations) Using Decision Boundaries to infer the architecture of a NN Review of strategies to deal with Overfitting and Underfitting Assignment 3 Supplementary: programming demo on using MLPs for character-level Language Andre (Karpathy's The Unreasonable Effectiveness of RNNs Supplementary: programming demo on using MLPs for character-level Language Andre (Karpathy's The Unreasonable Effectiveness of RNNs Supplementary: programming demo on using MLPs for character-level Language Andre (Karpathy's The Unreasonable Effectiveness of RNNs Supplementary: programming demo on using MLPs for character-level Language The University of the Forward Pass (In equations) Training the NN Andre (Karpathy's The Unreasonable Effectiveness of RNNs O Vector of the Holden State N	Q		• CLD CL7 FCLU	(1.02
CLO4 Changing the representation of the data CLO4 Changing the representation of the data Nor-linear activation Nor-linear activation Nor-linear activation Nor-linear activation Nor-linear and complex decision boundaries Ouniversal approximation and regression Layers: Depth vs. width Gradient descent for NNs Overfitting and the importance of Stochastic Gradient Descent (SGD) Formal Notation for Logistic Regression and NNs Vectorizing IR and NNs Vectorizing IR and NNs Vectorizing IR and NNs Vectorizing IR, and NNs Vectored with a single instance Vectored with minstances Activation functions, gradients, and pros and cons Signold Tanh Relul and Lesky ReLU Backward propagation (IR, NN) Simple, vectored with a single instance and m instances Training the NN Tutorial Topics: Review of the Forward Pass and Backward Pass (lin equations) Using Decision Boundaries to infer the architecture of a NN Review of strategies to deal with Overfitting and Underfitting Assignment 3 Supplementary: programming demo on using MLPs for character-level Language Modeling Nodeling Text and Speech (One-Hot) Time Series Videos as sequence Modeling tasks The Unreasonable Effectiveness of RNNs Notion of Sequence Modeling tasks The Unreasonable Effectiveness of RNNs Notion of Sequence Modeling topolems and specifications The Videos as sequence of images Videos as sequence Modeling topolems and specifications On Many-to-1-Video Classification, Sentiment Classification Many-to-1-Video Classification, Sentiment Classification Many-to-1-Video Classification, Sentiment Classification Many-to-1-Video Classification, Sentiment Classification On Many-to-1-Video Classification, Sentiment Classification On Many-to-Many (Seq2Seq) - Machine Translation Univolving in unit Introducing RNNs and how to deal with them Challenges of training RNNs, and how to deal with them	0.			
Changing the representation of the data O Kernels O Neural networks Non-linear activation Multi-layer Perceptron Perceptron Neural Networks Deep Learning Intuition: How Nhs work? O Non-linear and complex decision boundaries Universal approximation and regression Layers: Deph vs. width Gradient descent for Nhs O Overfitting and the importance of Stochastic Gradient Descent (SGD) Formal Notation for Logistic Regression and Nhs O Vectoriang Rt and Nhs O Forward propagation (IX, NN) Simple, vectored with a single instance Vectored with m instances Activation functions, gradients, and pros and cons O Sigmoid Tanh RetU and Leaky RetU Backward propagation (IX, NN) Simple, vectored with a single instance and m instances Training the NN Tutorial Topics: Review of the Forward Pass and Backward Pass (in equations) Using Decision Boundaries to infer the architecture of a NN Review of strategies to deal with Overfitting and Underfitting Assignment 3 Supplementary: programming demo on using MLPs for character-level Language Modeling Assignment 3 Supplementary: programming demo on using MLPs for character-level Language Modeling Text and Speech (One-Hot) Te			Ch11	/
O Kernels O Neural networks Non-inear activation Multi-layer Perceptron Neural Networks Deep Learning Intuition- How Niks work? O Non-linear and complex decision boundaries O Universal approximation and regression Layers: Depth vs. width Gradient descent for Nis O Overfitting and the importance of Stochastic Gradient Descent (SGD) Formal Notation for Logistic Regression and Nis O Vectorizing IR and Nis O Vectorizing IR, and Nis O Simple, vectored with a single instance Vectored with minstances Activation functions, gradients, and pros and cons O Sigmoid O Tanh O Relu and Leaky Rel U Backward propagation (IR, Ni) O Simple, vectored with a single instance and m instances O Trailing the Nin Tutorial Topics: Review of the Forward Pass and Backward Pass (in equations) Using Decision Boundaries to infer the architecture of a Ni Review of strategies to deal with Overfitting and Underfitting Assignment 3 Supplementary; programming demo on using MLPs for character-level Language Modeling Assignment 3 Supplementary; programming demo on using MLPs for character-level Language Modeling O Simple, vectored with a single instance and m instances O Trailing the Nin Tutorial Topics: Review of the Forward Pass and Backward Pass (in equations) Using Decision Boundaries to infer the architecture of a Ni Review of strategies to deal with Overfitting and Underfitting Andreign Andreign State of the University of the August Andreign State of the University of the Univer				CLO4
O Neural networks Non-linear activation Multi-layer Perceptron Neural Networks Deep Learning Intuition: How Nins work? O Non-linear and complex decision boundaries O Universal approximation and regression Layers: Depth vs. width Gradlent descent for NNs O Vertifuting and the importance of Stochastic Gradient Descent (SGD) Formal Notation for Logistic Regression and NNs Neurorizing IR and NNs Forward propagation (IR, NN) Simple, wectored with a single instance Vectored with m instances Activation functions, gradients, and pros and cons Sigmoid Tanh Ret U and Leaky Ret U Backward propagation (IR, NN) Simple, vectored with a single instance and m instances Training the NN Tutorial Topics: Review of the Forward Pass and Backward Pass (in equations) Using becision Boundaries to infer the architecture of a NN Review of strategies to deal with Overfitting and Underfitting Assignment 3 Supplementary: programming demo on using MLPs for character-level Language Modelling Assignment 3 Supplementary: programming demo on using MLPs for character-level Language Modelling The Unireasonable Effectiveness of RNNs Notion of Sequence Modeling tasks Text and Speech (One-Hot) Time-Series Videos as sequences of images Types of Sequence Modeling problems and specifications Text and Speech (One-Hot) Many-to-Many (Seq2Seq) - Machoine Translation, Summarization etc. With 9th Federovard NNs perform poorly at Sequence Modeling tasks? Introducing RNNs Machine Translation Model SLP3 Andrej Karpathy's Studieng A Neural Machine Translation, Summarization etc. With 9th Federovard NNs perform poorly at Sequence Modeling tasks? Introducing RNNs Machine Translation Model SLP3 CLO2, The Unreasonable Effectiveness of RNNs Notion of Turnicated BPTT Challenges of training RNNs, and how to deal with them		Changing the representation of the data		
Non-linear activation Multi-layer Perceptron Neural Networks Deep Learning Intuition: How NNs work? O Non-linear and complex decision boundaries O Universal approximation and regression Layers: Depth vs. width Gradient descent for NNs O Overfitting and the importance of Stochastic Gradient Descent (SGD) Formal Notation for Logistic Regression and NNs O Vectorizing IR and NNs Forward propagation (IR, NN) Simple, vectored with a single instance Vectorizing IR, and pros and cons Sigmoid Tanh RetU and Leaky RetU Backward propagation (IR, NN) Simple, vectored with a single instance Vectorizing IR, NN) Simple, vectored with a single instance Vectorizing IR, NN) Simple, vectored with a single instance and minstances Training the NN Tutorial Topics: Review of the Forward Pass and Backward Pass (in equations) Using Decision Boundaries to infer the architecture of a NN Review of Strategies to deal with Overfitting and Underfitting Assignment 3 Supplementancy: programming demo on using MLPs for character-level Language Modeling Notion of Sequence Modeling tasks International Complex of Sequence Modeling tasks Videos as sequences of images Videos as sequences of images Types of Sequence Modeling trablems and specifications O 1-to-Many - Image Captioning O Many-to-Many (Seq2Seq) - Machine Translation, Summarization etc. Why do Feedforward NNs perform poorly at Sequence Modeling tasks? Introducing RNNs Oldea of the Hidden State The overall architecture Vectorizing RNNs Challenges of training RNNs, and how to deal with them		o Kernels		
Multi-layer Perceptron Neural Networks Deep Learning Intuition: How NNs work? O Non-linear and complex decision boundaries O Universal approximation and regression Layers: Depth vs. width Gradient descent for NNs O Vertifting and the importance of Stochastic Gradient Descent (SGD) Formal Notation for Logistic Regression and NNs O Vertorizing IR and NNs O Vertorizing IR and NNs O Forward propagation (IR, NN) I Simple, vectored with a single instance Vectored with m instances Activation functions, gradients, and pros and cons O Sigmoid O Tanh O Ret U and Leaky Ret U Backward propagation (IR, NN) O Simple, vectored with a single instance and m instances Training the NN Tutorial Topics: Review of the Forward Pass and Backward Pass (in equations) Using Decision Boundaries to infer the architecture of a NN Review of strategies to deal with Overfitting and Underfitting Assignment 3 Supplementary: programming demo on using MLPs for character-level Language Modeling Notion of Sequence Modeling tasks Treat and Speech (One-Hot) Time-Series Videos as sequences of images Types of Sequence Modeling tasks Types of Sequence Modeling problems and specifications O Time-Series Types of Sequence Modeling problems and specifications O Many-to-Many (Seq2Seq) - Machine Translation, Summarization etc. With Q5 Decdforward NNs perform poorly at Sequence Modeling tasks? Introducing RNNs Introducing RNNs O Idea of the Hidden State The overall architecture O "Unrolling" a unit United Translation of (truncated) BPTT Challenges of training RNNs, and how to deal with them		 Neural networks 		
Neural Networks Deep Learning Intuition: How NNs work? Non-linear and complex decision boundaries Universal approximation and regression Layers: Depth vs. width Gradient descent for NNs O Vertorizing IR and Nns O Vertorizing IR and Nns No Porward propagation (IR, NN) Forward propagation (IR, NN) Simple, vectored with a single instance Training the NN Tutorial Topics: Review of the Forward Pass and Backward Pass (in equations) Using Decision Boundaries to infer the architecture of a NN Review of Strategies to deal with Overfitting and Underfitting Assignment 3 Supplementary: programming demo on using MLPs for character-level Language Modeling Sequence Models Notion of Sequence Modeling tasks Types of Sequence Modeling problems and specifications Training An Neural Notion of Sequence Modeling problems and specifications The Universaciable of the Midden State Types of Sequence Modeling problems and specifications Namy-to-1-Video Classification, Sentiment Classification Many-to-1-Video Classification, Sentiment Classification Many-to-Many (Seq25eq) - Machine Translation, Summarization etc. Why do Feedforward NNs perform poorly at Sequence Modeling tasks? Introducing RNNs Under Machine Translation Model Challenges of training RNNs, and how to deal with them		Non-linear activation		
Neural Networks Deep Learning Intuition: How NNs work? Non-linear and complex decision boundaries Universal approximation and regression Layers: Depth vs. width Gradient descent for NNs O Vertorizing IR and Nns O Vertorizing IR and Nns No Porward propagation (IR, NN) Forward propagation (IR, NN) Simple, vectored with a single instance Training the NN Tutorial Topics: Review of the Forward Pass and Backward Pass (in equations) Using Decision Boundaries to infer the architecture of a NN Review of Strategies to deal with Overfitting and Underfitting Assignment 3 Supplementary: programming demo on using MLPs for character-level Language Modeling Sequence Models Notion of Sequence Modeling tasks Types of Sequence Modeling problems and specifications Training An Neural Notion of Sequence Modeling problems and specifications The Universaciable of the Midden State Types of Sequence Modeling problems and specifications Namy-to-1-Video Classification, Sentiment Classification Many-to-1-Video Classification, Sentiment Classification Many-to-Many (Seq25eq) - Machine Translation, Summarization etc. Why do Feedforward NNs perform poorly at Sequence Modeling tasks? Introducing RNNs Under Machine Translation Model Challenges of training RNNs, and how to deal with them				
Deep Learning Intuition: How NNs work? O Non-linear and complex decision boundaries Universal approximation and regression Layers: Depth vs. width Gradient descent for NNs O Vectriting and the importance of Stochastic Gradient Descent (SGD) Formal Notation for Logistic Regression and NNs Vectoral with a single instance Vectored with a single instance Vectored with m instances Activation functions, gradients, and pros and cons Simple, vectored with a single instance Activation functions, gradients, and pros and cons Sigmoid Training the NN Simple, vectored with a single instance and m instances Review of the Forward Pass and Backward Pass (in equations) Simple, vectored with a single instance and m instances Training the NN Tutorial Topics: Review of the Forward Pass and Backward Pass (in equations) Supplementary: programming demo on using MLPs for character-level Language Modeling Assignment 3 Supplementary: programming demo on using MLPs for character-level Language Modeling Assignment 3 Totat and Speech (One-Hot) The Universal approximation and Underfitting The Universal Apple of Control of Sequence Modeling tasks Types of Sequence Modeling problems and specification Many-to-1- video Classification, Sentiment Classification Many-to-1- Windeo Classification, Sentiment Classification Many-to-1- Video Classification, Sentiment Classification Many-to-1- Video Classification, Sentiment Classification The owerall architecture The Oweral architecture The Ow				
Intuition: How NNs work? O Non-linear and complex decision boundaries Universal approximation and regression Layers: Depth vs. width Gradient descent for NNS O Verefitting and the importance of Stochastic Gradient Descent (SGD) Formal Notation for Logistic Regression and NNS O Vectorizing IX and NNS Forward propagation (IX, NN) Simple, vectored with a single instance Vectored with mistances Activation functions, gradients, and pros and cons Sigmoid Tanh Review of Layers (IX, NN) Simple, vectored with a single instance Vectored with a s				
O Non-linear and complex decision boundaries O Universal approximation and regression Layers: Depth vs. width Gradient descent for NNs O Vectoriting and the importance of Stochastic Gradient Descent (SGD) Formal Notation for Logistic Regression and NNs O Vectorizing LR and NNs O Vectored with a single instance Vectored with m instances Activation functions, gradients, and pros and cons O Sigmoid O Tanh RetU and Leaky RetU Backward propagation (LR, NN) Sigmoid O Tanh RetU and Leaky RetU Backward propagation (LR, NN) Sigmoid O Simple, vectored with a single instance and m instances Training the NN Vising Decision Boundaries to infer the architecture of a NN Review of strategies to deal with Overfitting and Underfitting Assignment 3 Supplementary: programming demo on using MLPs for character-level Language Modeling Assignment 3 Supplementary: programming demo on using MLPs for character-level Language Modeling Assignment 3 Supplementary: programming demo on using MLPs for character-level Language Noteling Andrej Karpathy's The Unreasonable Effectiveness of Effectiveness of RNNs Types of Sequence Modeling tasks Types of Sequence Modeling tasks Types of Sequence Modeling tasks Types of Sequence Modeling problems and specifications University of Many-to-1- Video Classification, Sentiment Classification Many-to-Many - Image Captioning Many-to-Many - Image Captioning Many-to-Many - Image Captioning Many-to-Many - Image Captioning Introducing RNNs Introducing RNNs Olde of the Hidden State The overall architecture University of the Format Nns perform poorly at Sequence Modeling tasks? Introducing RNNs University of the Format Nns perform poorly at Sequence Modeling tasks? CLO2, The Overall Andrews of the Model Model Machine Translation Model Supplementary: Victorizing A Neural Machine Translation Model CLO3, The Overal Many - Introducing RNNs University of the Model Many - Introducing RNNs University of the Model Many - Introducing RNNs Univers				
Universal approximation and regression Layers: Depth vs. width Gradient descent for NNs Overfitting and the importance of Stochastic Gradient Descent (SGD) Formal Notation for Logistic Regression and NNs Overtorizing LR and NNs Forward propagation (LR, NN) Simple, vectored with a single instance Sigmoid Oranh RetU and Leaky RetU Backward propagation (LR, NN) Simple, vectored with a single instance Sigmoid Oranh RetU and Leaky RetU Backward propagation (LR, NN) Simple, vectored with a single instance and m instances Training the NN Tutorial Topics: Review of the Forward Pass and Backward Pass (in equations) Using Decision Boundaries to infer the architecture of a NN Review of strategies to deal with Overfitting and Underfitting Assignment 3 Supplementary: programming demo on using MLPs for character-level Language Modeling Notion of Sequence Modeling tasks Notion of Sequence Modeling tasks Text and Speech (One-Hot) O Time-Series Videos as sequences of images Types of Sequence Modeling problems and specifications O It-o-Many: Image Captioning O Many-to-1-Video Classification, Sentiment Classification O Many-to-Many (Seq2Seq) - Machine Translation, Summarization etc. Why do Feedforward NNs perform poorly at Sequence Modeling tasks? Introducing RNNs Olde of the Hidden State The overall architecture "Unrolling" aunit Inturding of (truncated) BPTT Challenges of training RNNs, and how to deal with them				
Layers: Depth vs. width Gradient descent for NNs O Voerfitting and the importance of Stochastic Gradient Descent (SGD) Formal Notation for Logistic Regression and NNs O Vectorizing LR and NNs O Simple, vectored with a single instance Vectored with m instances O Activation functions, gradients, and pros and cons O Sigmoid O Tanh O ReLU and Leaky ReLU Backward propagation (LR, NN) O Simple, vectored with a single instance and m instances O Training the NN Tutorial Topics: Review of the Forward Pass and Backward Pass (in equations) Using Decision Boundaries to infer the architecture of a NN Review of strategies to deal with Overfitting and Underfitting Assignment 3 Supplementary: programming demo on using MLPs for character-level Language Modeling Notion of Sequence Modeling tasks O Text and Speech (One-Hot) O Time-Series O Videos as sequences of images Types of Sequence Modeling tasks O Text and Open Cyber of Sequence Modeling tasks O The Namy - Image Captioning O Many-to-14 Video Classification, Semiment Classification O Many-to-14 Video Classification, S		·		
Gradient descent for NNS O Overfitting and the importance of Stochastic Gradient Descent (SGD) Formal Notation for Logistic Regression and NNS Vectorizing LR and NNS Forward propagation (LR, NN) Simple, vectored with a single instance Vectored with m instances Activation functions, gradients, and pros and cons Sigmoid Tanh No RELU and Leaky ReLU Backward propagation (LR, NN) Simple, vectored with a single instance and m instances Training the NN Tutorial Topics: Review of the Forward Pass and Backward Pass (in equations) Sigment 3 Review of strategies to deal with Overfitting and Underfitting Assignment 3 Supplementary: programming demo on using MLPs for character-level Language Modeling Assignment 3 Supplementary: programming demo on using MLPs for character-level Language Modeling Trext and Speech (One-Hot)		.,,		
O Overfitting and the importance of Stochastic Gradient Descent (SGD) Formal Notation for Logistic Regression and NNs O Vectorizing IR and NNs O Forward propagation (IR, NN) I Simple, vectored with a single instance O Sigmoid O Tanh O ReLU and Leaky ReLU Backward propagation (IR, NN) O Simple, vectored with a single instances Training the NN Tutorial Topics: Review of the Forward Pass and Backward Pass (in equations) Using Decision Boundaries to infer the architecture of a NN Review of strategies to deal with Overfitting and Underfitting Assignment 3 Supplementary: programming demo on using MLPs for character-level Language Modeling Sequence Models Notion of Sequence Modeling tasks O Text and Speech (One-Hot) O Time-Series O Videos as sequences of images Types of Sequence Modeling problems and specifications O 1-to-Many - Image Captioning O Many-to-Jany (Seq2Seq) - Machine Translation, Summarization etc. Why do Feedforward NNs perform poorly at Sequence Modeling tasks? Introducing RNNs O Idea of the Hidden State O The overall architecture O "Unrolling" aunit O Intuition of (truncated) BPTT Challenges of training RNNs, and how to deal with them		Layers: Depth vs. width		
Formal Notation for Logistic Regression and NNs O Vectorizing LR and NNs Forward propagation (LR, NN) Simple, vectored with a single instance Vectored with m instances Activation functions, gradients, and pros and cons Sigmoid Tanh ReLU and Leaky ReLU Backward propagation (LR, NN) Simple, vectored with a single instance and m instances Training the NN Tutorial Topics: Review of the Forward Pass and Backward Pass (in equations) Using Decision Boundaries to infer the architecture of a NN Review of strategies to deal with Overfitting and Underfitting Assignment 3 Supplementary: programming demo on using MLPs for character-level Language Modeling Notion of Sequence Modeling tasks Time-Series Notion of Sequence Modeling tasks Time-Series Videos as sequences of images Types of Sequence Modeling problems and specifications Time-Series Types of Sequence Modeling problems and specifications Time-Series Types of Sequence Modeling problems and specifications Time-Series Types of Sequence Modeling problems and specifications Time-Series Types of Sequence Modeling problems and specifications Time-Series Types of Sequence Modeling problems and specifications Time-Series Types of Sequence Modeling problems and specifications Time-Series Types of Sequence Modeling problems and specifications Time-Series Types of Sequence Modeling problems and specifications Time-Series Types of Sequence Modeling problems and specifications Time-Series Types of Sequence Modeling problems and specifications Time-Series Types of Sequence Modeling problems and specifications Time-Series Types of Sequence Modeling tasks? Introducing RNNs Time-Series Time-Ser		Gradient descent for NNs		
Formal Notation for Logistic Regression and NNs O Vectorizing LR and NNs Forward propagation (LR, NN) Simple, vectored with a single instance Vectored with m instances Activation functions, gradients, and pros and cons Sigmoid Tanh ReLU and Leaky ReLU Backward propagation (LR, NN) Simple, vectored with a single instance and m instances Training the NN Tutorial Topics: Review of the Forward Pass and Backward Pass (in equations) Using Decision Boundaries to infer the architecture of a NN Review of strategies to deal with Overfitting and Underfitting Assignment 3 Supplementary: programming demo on using MLPs for character-level Language Modeling Notion of Sequence Modeling tasks Time-Series Notion of Sequence Modeling tasks Time-Series Videos as sequences of images Types of Sequence Modeling problems and specifications Time-Series Types of Sequence Modeling problems and specifications Time-Series Types of Sequence Modeling problems and specifications Time-Series Types of Sequence Modeling problems and specifications Time-Series Types of Sequence Modeling problems and specifications Time-Series Types of Sequence Modeling problems and specifications Time-Series Types of Sequence Modeling problems and specifications Time-Series Types of Sequence Modeling problems and specifications Time-Series Types of Sequence Modeling problems and specifications Time-Series Types of Sequence Modeling problems and specifications Time-Series Types of Sequence Modeling problems and specifications Time-Series Types of Sequence Modeling problems and specifications Time-Series Types of Sequence Modeling tasks? Introducing RNNs Time-Series Time-Ser		O Overfitting and the importance of Stochastic Gradient Descent (SGD)		
O Vectorizing LR and NNs O Forward propagation (LR, NN) Simple, vectored with a single instance Vectored with m instances Activation found in Signoid O Tanh O RetU and Leaky ReLU Backward propagation (LR, NN) O Simple, vectored with a single instance and m instances O Training the NN Tutorial Topics: Review of the Forward Pass and Backward Pass (in equations) Using Decision Boundaries to infer the architecture of a NN Review of strategies to deal with Overfitting and Underfitting Assignment 3 Supplementary: programming demo on using MLPs for character-level Language Modeling Notion of Sequence Modeling tasks O Text and Speech (One-Hot) O Time-Series Videos as sequences of images Types of Sequence Modeling problems and specifications O 1-to-Many - Image Captioning Many-to-1 - Video Classification, Sentiment Classification Many-to-Many (Seq2Seq) - Machine Translation, Summarization etc. Why do Feedforward NNs perform poorly at Sequence Modeling tasks? Introducing RNNs Idea of the Hidden State The overall architecture My do Feedforward NNs perform poorly at Sequence Modeling tasks? LP3 Challenges of training RNNs, and how to deal with them				
Simple, vectored with a single instance Vectored with ministances Activation functions, gradients, and pros and cons Sigmoid Tanh RetU and Leaky RetU Backward propagation (IR, NN) Simple, vectored with a single instance and ministances Training the NN Tutorial Topics: Review of the Forward Pass and Backward Pass (in equations) Using Decision Boundaries to infer the architecture of a NN Review of strategies to deal with Overfitting and Underfitting Assignment 3 Supplementary: programming demo on using MLPs for character-level Language Modeling Sequence Models Notton of Sequence Modeling tasks Treat and Speech (One-Hot) Time-Series Types of Sequence Modeling problems and specifications Many-to-Many- Image Captioning Many-to-Many (Seq2Seq) - Machine Translation, Summarization etc. My do Feedforward NNs perform poorly at Sequence Modeling tasks? Introducing RNNs Idea of the Hidden State The overall architecture Univolling aunit Intuition of (truncated) BPTT Challenges of training RNNs, and how to deal with them				
Simple, vectored with a single instance Vectored with m instances Activation functions, gradients, and pros and cons Sigmoid Tanh Ret U and Leaky Ret U Backward propagation (LR, NN) Simple, vectored with a single instance and m instances Training the NN Tutorial Topics: Review of the Forward Pass and Backward Pass (in equations) Single Decision Boundaries to infer the architecture of a NN Review of strategies to deal with Overfitting and Underfitting Assignment 3 Supplementary: programming demo on using MLPs for character-level Language Modeling Andrej Karpathy's Modeling Sequence Models Notion of Sequence Modeling tasks The Unreasonable Effectiveness of CLO2, Time-Series Videos as sequences of images Types of Sequence Modeling problems and specification Time-Series Many-to-1 - Video Classification, Sentiment Classification Machine Translation Model Why do Feedforward NNS perform poorly at Sequence Modeling tasks? Introducing RNNS Idea of the Hidden State The overall architecture Multion of (truncated) BPTT Challenges of training RNNs, and how to deal with them				
Activation functions, gradients, and pros and cons Sigmoid Tanh ReLU and Leaky ReLU Backward propagation (LR, NN) Simple, vectored with a single instance and m instances Training the NN Tutorial Topics: Review of the Forward Pass and Backward Pass (in equations) Using Decision Boundaries to infer the architecture of a NN Review of strategies to deal with Overfitting and Underfitting Assignment 3 Supplementary: programming demo on using MLPs for character-level Language Modeling Notion of Sequence Modeling tasks Text and Speech (One-Hot) Time-Series Videos as sequences of images Types of Sequence Modeling problems and specifications Videos as sequences of images Types of Sequence Modeling problems and specifications Nany-to-1 - Video Classification, Sentiment Classification Many-to-1 - Video Classification, Sentiment Classification Model Machine Translation Model SLP3 CLO2, The Unreasonable				
Activation functions, gradients, and pros and cons Sigmoid Tanh ReLU and Leaky ReLU Backward propagation (LR, NN) Simple, vectored with a single instance and m instances Training the NN Tutorial Topics: Review of the Forward Pass and Backward Pass (in equations) Using Decision Boundaries to infer the architecture of a NN Review of strategies to deal with Overfitting and Underfitting Assignment 3 Supplementary: programming demo on using MLPs for character-level Language Modeling 9. Sequence Models Notion of Sequence Modeling tasks Text and Speech (One-Hot) Time-Series Videos as sequences of images Types of Sequence Modeling problems and specifications Videos as sequences of images Types of Sequence Modeling problems and specification Many-to-1-Video Classification, Sentiment Classification Many-to-1-Video Classification, Sentiment Classification Many-to-1-Video Classification, Sentiment Classification Many-to-1-Video Classification, Sentiment Classification Many-to-Many (Seq2Seq) - Machine Translation, Summarization etc. Why do Feedforward NNs perform poorly at Sequence Modeling tasks? Introducing RNNs Introducing RNNs, and how to deal with them				
o Sigmoid O Tanh O ReLU and Leaky ReLU Backward propagation (LR, NN) O Simple, vectored with a single instance and m instances Training the NN Tutorial Topics: Review of the Forward Pass and Backward Pass (in equations) Using Decision Boundaries to infer the architecture of a NN Review of strategies to deal with Overfitting and Underfitting Assignment 3 Supplementary: programming demo on using MLPs for character-level Language Modeling Notion of Sequence Modeling tasks Text and Speech (One-Hot) Time-Series Videos as sequences of images Videos as sequences of images Types of Sequence Modeling problems and specifications Videos as sequence Modeling roblems and specifications Anary-to-Many - Image Captioning Many-to-I - Video Classification, Sentiment Classification Many-to-Many (Seq2Seq) - Machine Translation, Summarization etc. Why do Feedforward NNs perform poorly at Sequence Modeling tasks? Introducing RNNs Idea of the Hidden State The overall architecture Univolling a unit Intuition of (truncated) BPTT Challenges of training RNNs, and how to deal with them				
O Tanh O ReLU and Leaky ReLU Backward propagation (LR, NN) O Simple, vectored with a single instance and m instances O Training the NN Tutorial Topics: Review of the Forward Pass and Backward Pass (in equations) Using Decision Boundaries to infer the architecture of a NN Review of strategies to deal with Overfitting and Underfitting Assignment 3 Supplementary: programming demo on using MLPs for character-level Language Modeling Sequence Models Notion of Sequence Modeling tasks The Unreasonable Effectiveness of RNNs Types of Sequence Modeling problems and specifications O Ito-Many - Image Captioning O Many-to-1 - Video Classification, Sentiment Classification O Many-to-Many (Seq2Seq) - Machine Translation, Summarization etc. Why do Feedforward NNs perform poorly at Sequence Modeling tasks? Introducing RNNs O Idea of the Hidden State O The overall architecture O "Unrolling" a unit Intuition of (truncated) BPTT Challenges of training RNNs, and how to deal with them		 Activation functions, gradients, and pros and cons 		
Backward propagation (LR, NN) Simple, vectored with a single instance and m instances Training the NN Tutorial Topics: Review of the Forward Pass and Backward Pass (in equations) Using Decision Boundaries to infer the architecture of a NN Review of strategies to deal with Overfitting and Underfitting Assignment 3 Supplementary: programming demo on using MLPs for character-level Language Modeling Notion of Sequence Modeling tasks Notion of Sequence Modeling problems and specifications Notion of Sequence Modeling problems and specifications Notion of Sequence Modeling problems and specification Notion of Sequence Modeling tasks? Introducing RNNs Notion of Sequence Modeling tasks? Introducing RNNs Idea of the Hidden State Notion of truncated) BPTT Challenges of training RNNs, and how to deal with them		o Sigmoid		
Backward propagation (LR, NN) Simple, vectored with a single instance and m instances Training the NN Tutorial Topics: Review of the Forward Pass and Backward Pass (in equations) Using Decision Boundaries to infer the architecture of a NN Review of strategies to deal with Overfitting and Underfitting Assignment 3 Supplementary: programming demo on using MLPs for character-level Language Modeling Notion of Sequence Modeling tasks Notion of Sequence Modeling problems and specifications Notion of Sequence Modeling problems and specifications Notion of Sequence Modeling problems and specification Notion of Sequence Modeling tasks? Introducing RNNs Notion of Sequence Modeling tasks? Introducing RNNs Idea of the Hidden State Notion of truncated) BPTT Challenges of training RNNs, and how to deal with them		o Tanh		
Backward propagation (LR, NN) Simple, vectored with a single instance and m instances Training the NN Tutorial Topics: Review of the Forward Pass and Backward Pass (in equations) Using Decision Boundaries to infer the architecture of a NN Review of strategies to deal with Overfitting and Underfitting Assignment 3 Supplementary: programming demo on using MLPs for character-level Language Modeling Sequence Models Notion of Sequence Modeling tasks The Unreasonable Effectiveness of CLO3, CLO3, CT The Unreasonable Effectiveness of CLO4, CLO3, CLO3, CT The Unreasonable Effectiveness of CLO4, CLO3, CT The Unreasonable Effectiveness of CLO4, CLO3, CLO3, CT The Unreasonable Effectiveness of CLO4, CLO3, CT The Unreasonable Effectiveness of CLO4, CLO3, CLO3, CT The Unreasonable Effectiveness of CLO4, CLO3, CT The Unreasonable Effectiveness of CLO4, CLO3, CLO3, CT The Unreasonable Effectiveness of CLO4, CLO3, CLO3, CT The Unreasonable Effectiveness of CLO4, CLO3, CLO3, CLO3, CLO3, CLO3, CLO3, CLO3, CLO3, CLO3, CLO4, CLO3, CLO3				
O Simple, vectored with a single instance and m instances O Training the NN Tutorial Topics: Review of the Forward Pass and Backward Pass (in equations) Using Decision Boundaries to infer the architecture of a NN Review of strategies to deal with Overfitting and Underfitting Assignment 3 Supplementary: programming demo on using MLPs for character-level Language Modeling Notion of Sequence Models Notion of Sequence Modeling tasks Text and Speech (One-Hot) Time-Series Videos as sequences of images Types of Sequence Modeling problems and specifications Types of Sequence Modeling problems and specification Many-to-1 - Video Classification, Sentiment Classification Many-to-Many (Seq2Seq) - Machine Translation, Summarization etc. My do Feedforward NNs perform poorly at Sequence Modeling tasks? Introducing RNNs Under Graph News Introducing RNNs Under Graph News Unrolling" a unit Unrolling" a unit Unrolling" a unit Unrolling RNNs, and how to deal with them		,		
O Training the NN Tutorial Topics: Review of the Forward Pass and Backward Pass (in equations) Using Decision Boundaries to infer the architecture of a NN Review of strategies to deal with Overfitting and Underfitting Assignment 3 Supplementary: programming demo on using MLPs for character-level Language Modeling Sequence Models Notion of Sequence Modeling tasks The Unreasonable Effectiveness of CLO3, Effectiveness of Time-Series Videos as sequences of images Types of Sequence Modeling problems and specifications O 1-to-Many - Image Captioning O Many-to-1 - Video Classification, Sentiment Classification O Many-to-Many (Seq2Seq) - Machine Translation, Summarization etc. Why do Feedforward NNs perform poorly at Sequence Modeling tasks? Introducing RNNs O Idea of the Hidden State The Overall architecture O "Unrolling" a unit O Intuition of (truncated) BPTT Challenges of training RNNs, and how to deal with them				
Tutorial Topics: Review of the Forward Pass and Backward Pass (in equations) Using Decision Boundaries to infer the architecture of a NN Review of strategies to deal with Overfitting and Underfitting Assignment 3 Supplementary: programming demo on using MLPs for character-level Language Modeling 9. Sequence Models Notion of Sequence Modeling tasks Text and Speech (One-Hot) Time-Series Videos as sequences of images Types of Sequence Modeling problems and specifications Types of Sequence Modeling problems and specifications Anny-to-1 - Video Classification, Sentiment Classification Many-to-1 - Video Classification, Sentiment Classification Model Mary-to-Many (Seq2Seq) - Machine Translation, Summarization etc. My do Feedforward NNs perform poorly at Sequence Modeling tasks? Introducing RNNs I dea of the Hidden State The overall architecture The overall archite		• •		
Review of the Forward Pass and Backward Pass (in equations) Using Decision Boundaries to infer the architecture of a NN Review of strategies to deal with Overfitting and Underfitting Assignment 3 Supplementary: programming demo on using MLPs for character-level Language Modeling Sequence Models Notion of Sequence Modeling tasks The Unreasonable Effectiveness of RNNs Time-Series Types of Sequence Modeling problems and specifications Types of Sequence Modeling problems and specifications Many-to-1 - Video Classification, Sentiment Classification Many-to-Many (Seq2Seq) - Machine Translation, Summarization etc. Myndo Feedforward NNs perform poorly at Sequence Modeling tasks? Introducing RNNs Idea of the Hidden State The overall architecture The overall architectu		O Training the NN		
Review of the Forward Pass and Backward Pass (in equations) Using Decision Boundaries to infer the architecture of a NN Review of strategies to deal with Overfitting and Underfitting Assignment 3 Supplementary: programming demo on using MLPs for character-level Language Modeling Sequence Models Notion of Sequence Modeling tasks The Unreasonable Effectiveness of RNNs Time-Series Types of Sequence Modeling problems and specifications Types of Sequence Modeling problems and specifications Many-to-1 - Video Classification, Sentiment Classification Many-to-Many (Seq2Seq) - Machine Translation, Summarization etc. Myndo Feedforward NNs perform poorly at Sequence Modeling tasks? Introducing RNNs Idea of the Hidden State The overall architecture The overall architectu				
Using Decision Boundaries to infer the architecture of a NN Review of strategies to deal with Overfitting and Underfitting Assignment 3 Supplementary: programming demo on using MLPs for character-level Language Modeling 9. Sequence Models Notion of Sequence Modeling tasks Text and Speech (One-Hot) Time-Series Videos as sequences of images Types of Sequence Modeling problems and specifications Time-Series Videos as sequence Modeling problems and specifications Anny-to-1 - Video Classification, Sentiment Classification Many-to-Many (Seq2Seq) - Machine Translation, Summarization etc. Why do Feedforward NNs perform poorly at Sequence Modeling tasks? Introducing RNNs Introducing RNNs Unrolling" a unit Unrolling" a unit Intuition of (truncated) BPTT Challenges of training RNNs, and how to deal with them		Tutorial Topics:		
Review of strategies to deal with Overfitting and Underfitting Assignment 3 Supplementary: programming demo on using MLPs for character-level Language Modeling Sequence Models Notion of Sequence Modeling tasks Notion of Sequence Modeling tasks Text and Speech (One-Hot) Time-Series Videos as sequences of images Types of Sequence Modeling problems and specifications Anny-to-Many - Image Captioning Many-to-1 - Video Classification, Sentiment Classification Many-to-Many (Seq2Seq) - Machine Translation, Summarization etc. Why do Feedforward NNs perform poorly at Sequence Modeling tasks? Introducing RNNs Idea of the Hidden State The overall architecture Microbian and Underfitting Text and Speech (One-Hot) The overall architecture Microbian and Specification of Sequence Modeling tasks? The Unreasonable Effectiveness of RNNs Signal Sequence Modeling Nodel Translation Model SLP3 CLO2, The Unreasonable Effectiveness of RNNs Signal Sequence Modeling Nodel Translation Machine Translation Model Translation Model SLP3		 Review of the Forward Pass and Backward Pass (in equations) 		
Review of strategies to deal with Overfitting and Underfitting Assignment 3 Supplementary: programming demo on using MLPs for character-level Language Modeling Sequence Models Notion of Sequence Modeling tasks Notion of Sequence Modeling tasks Text and Speech (One-Hot) Time-Series Videos as sequences of images Types of Sequence Modeling problems and specifications Anny-to-Many - Image Captioning Many-to-1 - Video Classification, Sentiment Classification Many-to-Many (Seq2Seq) - Machine Translation, Summarization etc. Why do Feedforward NNs perform poorly at Sequence Modeling tasks? Introducing RNNs Idea of the Hidden State The overall architecture Microbian and Underfitting Text and Speech (One-Hot) The overall architecture Microbian and Specification of Model CLO2, The Unreasonable Effectiveness of RNNs Visualizing A Neural of Machine Translation of Model Machine Translation of Model SLP3 CLO2, The Unreasonable Effectiveness of RNNs Signal Specification of Model CLO4 CLO3, Effectiveness of RNNs Visualizing A Neural of Machine Translation of Model Machine Translation of Union of Uni		 Using Decision Boundaries to infer the architecture of a NN 		
Assignment 3 Supplementary: programming demo on using MLPs for character-level Language Modeling 9. Sequence Models Notion of Sequence Modeling tasks O Text and Speech (One-Hot) O Time-Series O Videos as sequences of images Types of Sequence Modeling problems and specifications O 1-to-Many - Image Captioning O Many-to-1 - Video Classification, Sentiment Classification O Many-to-Many (Seq2Seq) - Machine Translation, Summarization etc. Why do Feedforward NNs perform poorly at Sequence Modeling tasks? Introducing RNNs O Idea of the Hidden State O The overall architecture O "Unrolling" a unit O Intuition of (truncated) BPTT Challenges of training RNNs, and how to deal with them				
Supplementary: programming demo on using MLPs for character-level Language Modeling 9. Sequence Models Notion of Sequence Modeling tasks The Unreasonable Effectiveness of CLO3, CLO3, CLO4 Time-Series Videos as sequences of images Types of Sequence Modeling problems and specifications Many-to-1 - Video Classification, Sentiment Classification Many-to-Many (Seq2Seq) - Machine Translation, Summarization etc. My do Feedforward NNs perform poorly at Sequence Modeling tasks? Introducing RNNs Introducing RNNs Idea of the Hidden State The overall architecture "Unrolling" a unit Intuition of (truncated) BPTT Challenges of training RNNs, and how to deal with them				
9. Sequence Models Notion of Sequence Modeling tasks Text and Speech (One-Hot) Time-Series Videos as sequences of images Types of Sequence Modeling problems and specifications Many-to-1 - Video Classification, Sentiment Classification Many-to-Many (Seq2Seq) - Machine Translation, Summarization etc. Why do Feedforward NNs perform poorly at Sequence Modeling tasks? Introducing RNNs Idea of the Hidden State The Overall architecture "Unrolling" a unit Intuition of (truncated) BPTT Challenges of training RNNs, and how to deal with them				
9. Sequence Models Notion of Sequence Modeling tasks Text and Speech (One-Hot) Time-Series Videos as sequences of images Types of Sequence Modeling problems and specifications Text-on-Many - Image Captioning Many-to-1 - Video Classification, Sentiment Classification Many-to-Many (Seq2Seq) - Machine Translation, Summarization etc. Why do Feedforward NNs perform poorly at Sequence Modeling tasks? Introducing RNNs Introducing RNNs Introducing RNNs Introducing RNNs CLO2, The Unreasonable Effectiveness of RNNs Jay Alammar's Visualizing A Neural Machine Translation Model SLP3 SLP3				
 Notion of Sequence Modeling tasks Text and Speech (One-Hot) Time-Series Videos as sequences of images Types of Sequence Modeling problems and specifications 1-to-Many - Image Captioning Many-to-1 - Video Classification, Sentiment Classification Many-to-Many (Seq2Seq) - Machine Translation, Summarization etc. Why do Feedforward NNs perform poorly at Sequence Modeling tasks? Introducing RNNs Idea of the Hidden State The Unreasonable Effectiveness of RNNs SLP3 	-	<u> </u>		01.00
O Text and Speech (One-Hot) O Time-Series O Videos as sequences of images Types of Sequence Modeling problems and specifications O 1-to-Many - Image Captioning O Many-to-1 - Video Classification, Sentiment Classification O Many-to-Many (Seq2Seq) - Machine Translation, Summarization etc. Why do Feedforward NNs perform poorly at Sequence Modeling tasks? Introducing RNNs O Idea of the Hidden State O The overall architecture O "Unrolling" a unit O Intuition of (truncated) BPTT CLO4 Effectiveness of RNNs Jay Alammar's Visualizing A Neural Machine Translation Model SLP3 SLP3	9.			
O Time-Series O Videos as sequences of images Types of Sequence Modeling problems and specifications O 1-to-Many - Image Captioning O Many-to-1 - Video Classification, Sentiment Classification O Many-to-Many (Seq2Seq) - Machine Translation, Summarization etc. Why do Feedforward NNs perform poorly at Sequence Modeling tasks? Introducing RNNs O Idea of the Hidden State O The overall architecture O "Unrolling" a unit O Intuition of (truncated) BPTT Challenges of training RNNs, and how to deal with them				· '
 Videos as sequences of images Types of Sequence Modeling problems and specifications 1-to-Many - Image Captioning Many-to-1 - Video Classification, Sentiment Classification Many-to-Many (Seq2Seq) - Machine Translation, Summarization etc. Why do Feedforward NNs perform poorly at Sequence Modeling tasks? Introducing RNNs Idea of the Hidden State The overall architecture "Unrolling" a unit Intuition of (truncated) BPTT Challenges of training RNNs, and how to deal with them			Effectiveness of	CLO4
 Types of Sequence Modeling problems and specifications 1-to-Many - Image Captioning Many-to-1 - Video Classification, Sentiment Classification Many-to-Many (Seq2Seq) - Machine Translation, Summarization etc. Why do Feedforward NNs perform poorly at Sequence Modeling tasks? Introducing RNNs Idea of the Hidden State The overall architecture "Unrolling" a unit Intuition of (truncated) BPTT Challenges of training RNNs, and how to deal with them 		o Time-Series	RNNs	
 Types of Sequence Modeling problems and specifications 1-to-Many - Image Captioning Many-to-1 - Video Classification, Sentiment Classification Many-to-Many (Seq2Seq) - Machine Translation, Summarization etc. Why do Feedforward NNs perform poorly at Sequence Modeling tasks? Introducing RNNs Idea of the Hidden State The overall architecture "Unrolling" a unit Intuition of (truncated) BPTT Challenges of training RNNs, and how to deal with them 		 Videos as sequences of images 		
 1-to-Many - Image Captioning Many-to-1 - Video Classification, Sentiment Classification Many-to-Many (Seq2Seq) - Machine Translation, Summarization etc. Why do Feedforward NNs perform poorly at Sequence Modeling tasks? Introducing RNNs Idea of the Hidden State The overall architecture "Unrolling" a unit Intuition of (truncated) BPTT Challenges of training RNNs, and how to deal with them 			Jay Alammar's	
 Many-to-1 - Video Classification, Sentiment Classification Many-to-Many (Seq2Seq) - Machine Translation, Summarization etc. Why do Feedforward NNs perform poorly at Sequence Modeling tasks? Introducing RNNs Idea of the Hidden State The overall architecture "Unrolling" a unit Intuition of (truncated) BPTT Challenges of training RNNs, and how to deal with them 				
 Many-to-Many (Seq2Seq) - Machine Translation, Summarization etc. Why do Feedforward NNs perform poorly at Sequence Modeling tasks? Introducing RNNs Idea of the Hidden State The overall architecture "Unrolling" a unit Intuition of (truncated) BPTT Challenges of training RNNs, and how to deal with them 				
 Why do Feedforward NNs perform poorly at Sequence Modeling tasks? Introducing RNNs Idea of the Hidden State The overall architecture "Unrolling" a unit Intuition of (truncated) BPTT Challenges of training RNNs, and how to deal with them 				
 Introducing RNNs Idea of the Hidden State The overall architecture "Unrolling" a unit Intuition of (truncated) BPTT Challenges of training RNNs, and how to deal with them 			iviouei	
 Idea of the Hidden State The overall architecture "Unrolling" a unit Intuition of (truncated) BPTT Challenges of training RNNs, and how to deal with them				
 The overall architecture "Unrolling" a unit Intuition of (truncated) BPTT Challenges of training RNNs, and how to deal with them 			SLP3	
 "Unrolling" a unit Intuition of (truncated) BPTT Challenges of training RNNs, and how to deal with them 		 Idea of the Hidden State 		
 "Unrolling" a unit Intuition of (truncated) BPTT Challenges of training RNNs, and how to deal with them 		o The overall architecture		
 Intuition of (truncated) BPTT Challenges of training RNNs, and how to deal with them 		o "Unrolling" a unit		
Challenges of training RNNs, and how to deal with them				
		interior or to directed by 11		
		Challanges of training DNNs and have to deal with the sec		
O Exploding/Vanishing Gradients				
		O Exploding/Vanishing Gradients		

	Choice of Activation functions			
	Gradient Clipping			
	 Changing the architecture (intro to LSTMs) 			
	Introducing LSTMs			
	O Changes to the architecture			
	o Idea of storing the "memory" in a cell			
	Exploring Machine Translation Section on a Section much laws			
	Setup as a Seq2Seq problem			
	o Embeddings			
	 Using the Encoder-Decoder framework 			
	 Information bottleneck: passing on only one hidden state 			
	o Improvements			
	Passing all the hidden states			
	 Intuition of the Attention Mechanism 			
	Tutorial Topics:			
	Review of the representations of text: Label Encoding, One-Hot Encoding,			
	Embeddings			
	Programming demo on Embeddings in Python			
	Review of RNNs and LSTMs Review of the Attention Machine in Machine Translation			
	Review of the Attention Mechanism in Machine Translation			
10.	Attention Mechanism and Transformers	Jay Alammar's <u>The</u>		
	The Attention Mechanism in Machine Translation	Illustrated		
	Self Attention	<u>Transformer</u>		
	 Dot Product Attention 			
	 The idea of contextualized word embeddings 	Andrej Karpathy's		
	Introducing the Transformer and its contributions	Let's Build GPT		
	O Highly Parallelizable			
	Contextualized Embeddings vs. Plain Embeddings	SLP3		
	O Long-Term Dependencies	02.0		
	Transformer Architecture in a nutshell			
	Role of a Decoder and Masked Attention			
	o Query, Keys, Values			
	The Transformer in equations			
	 Positional Embeddings 			
	 Projections to QKV 			
	 Self Attention as Dot Product Attention 			
	Role of Feedforward layers			
	o Multi Headed Self Attention: the role and its equations			
	o (ignore Layer Norms for brevity)			
	Case Studies: BERT, T5, and GPT-3/LLMs			
	Case Studies. DEIVI, 13, and GIT 3/LEIVIS			
	Tutorial Tonics			
	Tutorial Topics:			
	Review (and Expansion) of Self-Attention Pifference between the Feed decade Report Line and Control of Self-Attention Output Description:			
	Differences between the Encoder and Decoder: masking, role, architecture			
	Demo on coding out GPT in Python			
11.	Decision (Classification/Regression) Trees (Discriminative, non-linear, parametric/non-	 Decision Tree 	CLO3,	
	parametric)	wiki page	CLO4	
	Clustering using K-D Trees and nearest neighbor methods	Ben Taskar's		
	• Why do we still need nearest neighbors?	old <u>notes</u>		
	Revision of Trees			
	 Graph: Nodes, edges, directed/undirected, path, cyclic/acyclic, 	• Murphy: 16.2		
	Tree: A rooted directed acyclic graph (Rooted DAG)	● ESLII 8.7, ESLII		
	■ Parent, children, siblings, root, leaves, degree, height, arity,	Ch10, 15, 16		
	various relationships			
	·			
	■ Forests			
	From K-D trees to decision trees			
	Decision tree examples: classification and regression			
	 Categorical and real-valued attributes 			
	Bias and variance			
	o Tree height			
	Growing tree automatically – Decision tree training			
	The ID3 and CART algorithms			
		I	1	

	Splits and purity/impurity			
	■ Gini and Entropy			
	 Information Gain and Information Gain Ratio 			
	■ Multiclass			
	Regression: Variance			
	 Real-valued attributes 			
	 Strengths and weakness of CARTs 			
	 Automatic feature selection 			
	 Generalizability of splitting on attributes with a large set of values 			
	Missing data			
	 Axis-aligned splits 			
	O Over fitting			
	Side notes on Huffman coding			
	A deep dive into Entropy Open tife in a proportion of the control of the co			
	Quantifying expected surprise			
	Events in isolationProbability distributions			
	Probability distributionsInformation, surprise, and uncertainty			
	Evaluating Language Models			
	■ Comparing distributions using cross entropy			
	A discussion on Parametric/Non-parametric models			
	Ensemble methods: Bagging and Random Forests			
	Decomposition of Generalization Error			
	o Bias/Variance/Noise			
	 Detecting high bias and high variance regimes 			
	Variance reduction			
	 The weak law of large numbers 			
	 M datasets sampled from P 			
	 Bootstrapping 			
	Bootstrapped Aggregation (Bagging)			
	• Are we even drawing from the same <i>P</i>			
	Summary and advantages of bagging			
	Random Forests All a sixthers			
	O Algorithm			
	O Examples and benefits			
	Out-of-box performanceThe hyper parameters			
	\blacksquare m and k			
	■ No need of normalization and feature-scaling			
	Resilience to the curse of dimensionality			
	■ Feature selection			
	Missing data and clustering			
	o Variants			
	Tutorial Topics:			
	 Review of Gini/Entropy and Decision Trees 			
	Demo on Random Forests in Python: exploring Feature Importances for Feature			
	Selection			
	Standardization in the context of Decision Trees			
12	Decision Trees vs. Neural Networks From blomathods: Posting Gradient Posted Trees, and AdaPost		CL C1	
12.	Ensemble methods: Boosting, Gradient Boosted Trees, and AdaBoost ■ Bias Reduction	• ESLII 8.7, ESLII	CLO1, CLO2,	
	Bias Reduction Intuition	Ch10, 15, 16	CLO2, CLO3,	
	o Vectors		CLO3, CLO4	
	o Gradient Descent in function space	• SLP Ch4	5254	
	o Generic Boosting (AnyBoost)	• Murphy 2.2, 3.1-		
	Algorithm and geometric interpretation	3.4		
	Gradient Boosted Regression Trees	• ESLII 6.6.3		
	o Algorithm			
	O Detailed walk-through			
	 AdaBoost 			
	o Setting			
	O Odds Ratio and log-odds			
	Step size proportional to error reduction			

	O limited and the first transfer that the first transfer the first transfer transfer that the first transfer tr		
	 Instance weights proportional to (mis)classification and the say of the classifier 		
	O Algorithm and detailed walk-through		
	o Properties and Summary		
	, , , , , , , , , , , , , , , , , , , ,		
	Bayes Theorem		
	Review of probability, joint and conditional probability, and derivation of the		
	Bayes Theorem		
	Maximum a posteriori (MAP) and Maximum Likelihood Estimation (MLE)		
	 Posterior, likelihood, prior and evidence 		
	 Classification using MAP and MLE 		
	 Example problems and solutions using Bayes Theorem 		
	Binary and multiclass		
	Monty Hall problem, medical testing, Language Modeling		
	Generative and Discriminative classifiers		
4.2	O Solving SPAM vs. Not-SPAM		0.03
13.	The Naïve Bayes Classifier (A linear, generative, parametric classifier)		CLO2,
	Derivation and implementation Classification using the Reuse Theorem	Ben Taskar's	CLO3,
	O Classification using the Bayes Theorem	notes on Naïve	CLO4
	 Learning by example: The SPAM vs. Not-SPAM problem The "zeros" and how to get rid of them! 	<u>Bayes</u>	
	Independence, mutual exclusion, and conditional independence	TM chapter	
	The challenge of "how much of the context to use?" - Ngrams	on <u>Naive</u>	
	Naïve Assumptions: Conditional Independence and Bag-of-Words	Bayes (ch 1-3)	
	Data sparsity and Out-of-vocabulary (OOV) items	 Xiaojin Zhu's 	
	O Laplace Add-1 smoothing	notes	
	Another example: Sentiment analysis	on <u>Multinomial</u>	
	Text generation using Naïve Bayes	<u>Naïve Bayes</u>	
	 Infinite monkeys on typewriters 	• Mannings'	
	 The Shannon visualization method for Ngrams 	description	
	 Approximating Shakespeare and the Wall Street Journal 	of <u>Multinomial</u>	
	Real-valued features: Gaussian Naïve Bayes	Naive Bayes	
	Probability vs. likelihood		
	A worked example		
	 Naïve Bayes decision boundary Under assumptions and general case 		
	 Under assumptions and general case Naïve Bayes: Strengths and weaknesses 		
	Waive Dayes. Strengths and Weakhesses		
	Tutorial Topics:		
	Probability vs. Likelihood (and other terminologies)		
	Review of Naive Bayes		
	Review (and Expansion) of add-k smoothing		
	Demo on using Bigrams for Language Modeling		
14.	Unsupervised Learning	● ESLII 8.7, ESLII	CLO3,
	The unsupervised learning setup	Ch10, 15, 16	CLO4
	Use cases of unsupervised learning		
	o Clustering		
	o Anomaly detection		
	O Feature selection and dimensionality reduction		
	Types of clustering Monothetic and Polythetic		
	 Monothetic and Polythetic Hard and Soft 		
	o Flat and Hierarchical		
	Clustering		
	o K-D Trees		
	■ Monothetic, hard boundary, hierarchical, divisive (top-down)		
	■ Motivation		
	■ Algorithm		
	o Vector Quantization		
	Motivation and method		
	 Codebook and distance metric 		
	■ Euclidean and Mahalanobis distances		
	o K-means		
	■ Polythetic, hard boundary, flat		
	■ Lloyd/Forgy method		

- Expectation Maximization (EM) and K-means
- The K-means objective
- Optimal Number of Clusters
- Categorical data and K-modes
- Vector Quantization using K-means
- Evaluating Clustering
 - Extrinsic and Intrinsic evaluation
- Gaussian Mixture Models
 - Polythetic, soft boundary, flat, probabilistic
 - K-means vs. GMMs
 - EM for GMMs
 - Mixture models in 1-dimension and n-dimensions
 - Likelihoods, cluster assignments, and cluster update rules
 - The covariance matrix
 - How many Gaussians?
- Hierarchical clustering
 - Recursive K-means
 - Polythetic, hard boundary, hierarchical, top-down
 - Agglomerative Clustering
 - Polythetic, hard boundary, hierarchical, bottom-up
 - Examples
 - Distances: Single link, complete link, average link, centroids, Ward's method
- Dimensionality Reduction
 - Feature selection vs. feature reduction
 - Motivation
 - Visualization
 - Redundant and correlated features
 - Real vs. apparent dimensionality
 - The curse of dimensionality
 - Principal Component Analysis (PCA)
 - The dimension of greatest variability
 - A side-note on Matrices, Linear Transformations, the determinant, Eigenvalues and Eigenvectors
 - The Eigenvectors of the Covariance Matrix
 - How many dimensions?
 - Strengths and weaknesses
 - Linear Discriminant Analysis (LDA)
 - Supervised setup
 - Discrimination vs. spread
- Anomaly Detection
 - Anomalies
 - O How do we define anomalies?
 - O Why unsupervised?
 - o Examples and challenges
 - o Detection
 - One Class Classification
 - Density estimation
 - One feature and multiple features
 - Algorithm
 - Example
 - Evaluation
 - Unsupervised vs. supervised

Big Challenges and Opportunities in AI and ML

- The case for Explainable AI
- The case for Fair AI
- Societal biases
- Imbalanced classification
- Machine Learning for Development (ML4D)

Tutorial Topics:

- Review of Clustering
- Review of PCA

	Programming demo of PCA In Python as a preprocessing step									
	Other topics – to be covered if we have time									
15.	Bayes: Advanced topics (supplementary) Hypothesis spaces Frequentist viewpoint Olintuition, derivation, pros and cons, extreme data Bayesian viewpoint Olintuition, derivation, MAP, pros and cons OConjugate priors: The Beta and Dirichlet distributions Comparison of MAP and MLE OLaplace smoothing The Bayes Optimal Classifier	 SLP Ch4 Murphy 2.2, 3.1-3.4 ESLII 6.6.3 	CLO2, CLO3, CLO4							
16.	 Graphical Sequence Processing Models Hidden Markov Models (HMMs) Maximum Entropy Markov Models (MEMMs) Undirected Graphical Models (Markov Random Fields) Conditional Random Fields (CRFs) Directed Graphical Models (Bayes Nets) 	• SLP A, ESLII Ch17	CLO3, CLO4							

Textbook(s)/Supplementary Readings

Text Books

- Machine Learning, Tom Mitchell, McGraw Hill, 1997 TM
- The Elements of Statistical Learning: Data mining, Inference, and Prediction, Hastie, Trevor, Robert Tibshirani, and Jerome Friedman, Springer Science & Business Media, 2009 ESLII

Reference Books

- Speech and Language Processing by Jurafsky and Martin, Ed 3 (online draft) SLP
- Machine Learning: A Probabilistic Perspective, Murphy, Kevin P. MIT press, 2012 Murphy.
- Pattern Recognition and Machine Learning, Christopher M. Bishop, Springer, 2006 Bishop.
- Introduction to Machine Learning, Ethem Alpaydin, Ed 2, MIT Press, 2010 Alpaydin.
- Deep Learning, Ian Goodfellow and Yoshua Bengio and Aaron Courville, 2016 Goodfellow

Course policies

Use of electronic devices (e.g., mobile phones and laptops) in the class is strictly forbidden. A violation could result in deduction of marks and other strict penalties

Late arrival: You may not be allowed in the class 10 minutes after the start time

Plagiarism: All work MUST be done independently. In certain assignments students will be allowed to have discussions with peers, in which case they must mention the name and roll number of the student with whom the discussion took place and the nature of the discussion. Even in those assignments, all implementations need to be done independently. Any plagiarism or cheating of work from others or the internet will be immediately referred to the DC. If you are confused about what constitutes plagiarism, it is YOUR responsibility to consult with the instructor or the TA in a timely manner. No "after the fact" negotiations will be possible.

• Submitting someone else's assignment as your own "by mistake" would count as plagiarism. If this indeed happens accidentally, please let us know immediately (within minutes) along with an explanation and do not wait until we find it out on our own. In the latter case, it would be considered plagiarism.

Quizzes: Quizzes will be unannounced. We will be following an n-x (x=2) policy for the quizzes. There is no makeup for a missed quiz. If you have missed up to x quizzes, you will be covered only using the n-x policy (even if you have an approved petition with the OSA). If you have missed more than x quizzes, then you would be awarded the average marks (across all the quizzes that you attempted) for each missed quiz, provided that your case has been approved by the Office of Student Affairs.

Non-uniform weightage: All subcomponents (e.g., quizzes, assignments) may not carry the same weight. These weights may not be announced prior to the submission of the components and will be determined by the course instructor based on factors including (but not limited to) the length, difficulty level, amount of help available, etc. for each subcomponent.

Programming: Strong programming skills are expected for this course. Please keep in mind that this is a programming intensive course, and you will be spending a lot of time designing and coding up your solutions.

Assignments: There is negative marking for skipped assignments and there is no n-x policy for assignments. Assignments are a basic building block of this course, and it will be ensured that students, who pass the course, have significant hands-on experience.

- You will be awarded 0 marks or investigated for plagiarism for submitting incorrect/corrupted files and/or older assignments. We will not accept resubmissions in these cases even if the system date shows that the file was not modified after the deadline.
- You are allowed 5 grace days for the entire semester. No late submission of assignments is allowed after your grace days have expired. We do not have any deduction policy for late submissions in addition to the grace days. All grace days must be utilized before the start of the dead week and any remaining grace days will expire as soon as the dead week begins.
- Please do not wait until the last moment to submit assignments and other components. Any requests to accommodate late submissions due to last minute issues (submission of partial or incorrect files, assignment server down-time, internet and power failures, personal problems, etc.) would not be accommodated.

Appendix A Bloom's Taxonomy

BLOOM's TAXONOMY* 1 - Remember 2 - Understand 3 - Apply 4 - Analyze 5 - Evaluate 6 - Create • Recall facts and basic concepts • Explain ideas or concepts • Use information in new situations • Draw connection among ideas • Justify a stand or decision • Produce new or original work

https://cft.vanderbilt.edu/guides-sub-pages/blooms-taxonomy/

Appendix B

ACM Dispositions Table - I

ACM Dispositions										
Element	Elaboration	Element	Elaboration							
D1 Adaptable:	Flexible; agile, adjust in response to change	D7 Professional:	Professionalism, discretion, ethical, astute							
D2 Collaborative:	Team player; willing to work with others	D8 Purpose-	Goal driven, achieve goals, business acumen							
D3 Inventive:	Exploratory; Look beyond simple solutions	driven:	Use judgment, discretion, act appropriately							
D4 Meticulous:	Attentive to detail; thoroughness, accurate	D9 Responsible:	Respectful; react quickly and positively							
D5 Passionate:	Conviction, strong commitment, compelling	D10 Responsive:	Self-motivated, determination, independent							
D6 Proactive:	With initiative, self-starter, independent	D11 Self-directed:								

ACM Dispositions Table - II

Class Assessments and Proposed Dispositions												
Assessment Type	D1 Adaptable	D2 Collaborative	D3 Inventive	D4 Meticulous	D5 Passionate	D6 Proactive	D7 Professional	D8 Purpose- driven	D9 Responsible	D10 Responsive	D11 Self- directed	Included
Quiz				✓			✓		✓			Yes
Assignment- Individual			✓	✓			✓		✓			Yes
Assignment- Group		✓	✓	✓			✓		✓	√		Yes
Project- Individual	✓		✓	✓	✓	✓	✓	✓	✓		✓	Yes
Project- Group	✓	✓	✓	✓	✓	✓	✓	✓	✓			Yes
Presentation- Individual				✓			✓		✓	✓	✓	Yes
Presentation- Group		✓		✓			✓		✓	✓		Yes
Labs- Individual			√	✓			✓		✓			Yes
Labs- Group		✓	✓	✓			✓		✓	✓		Yes
Exams				✓			✓		✓			Yes
Included	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	

Appendix C ACM Computing Knowledge Landscape Table

ACM Computing Knowle	ACM Computing Knowledge Landscape (CK)									
1. Users and Organizations	CK1.1: Social Issues and Professional Practice CK1.2: Security Policy and Management CK1.3: IS Management and Leadership CK1.4: Enterprise Architecture CK1.5: Project Management CK1.6: User Experience Design	4. Software Development	CK4.1: Software Quality, Verification and Validation CK4.2: Software Process CK4.3: Software Modeling and Analysis CK4.4: Software Design CK4.5: Platform-Based Development							
2. Systems Modeling	CK2.1: Security Issues and Principles CK2.2: Systems Analysis & Design CK2.3: Requirements Analysis and Specification CK2.4: Data and Information Management	5. Software Fundamentals	CK5.1: Graphics and Visualization CK5.2: Operating Systems CK5.3: Data Structures, Algorithms and Complexity CK5.4: Programming Languages CK5.5: Programming Fundamentals CK5.6: Computing Systems Fundamentals							
3. Systems Architecture and Infrastructure	CK3.1: Virtual Systems and Services CK3.2: Intelligent Systems (AI) CK3.3: Internet of Things CK3.4: Parallel and Distributed Computing CK3.5: Computer Networks	6. Hardware	CK6.1: Architecture and Organization CK6.2: Digital Design CK6.3: Circuits and Electronics CK6.4: Signal Processing							