PURDUE
U N I V E R S I T Y™

Instructor Info

- Z. Berkay Celik
- Office Hrs: Th 2pm - 4pm
- TBD
- https://beerkay.github.io
- celik.berkay@gmail.com

Course Info

- Prereq: Bachelor degree in Computer Science or equivalent.
- Tuesday & Thursday
- 12pm - 1:15pm
- Lawson CS Bldg 1106

Recitations Info

- TBD
- TBD
- Lab Space (TBD)

TA Info

- TBD
- Office Hrs: Th&Th 10am - 12am
- TBD

CS 529 Security Analytics (Fall 2019)

Overview

This graduate-level course will provide students with materials to discuss the intersection of two ubiquitous concepts: Security and Machine Learning. The course is structured in two parts: (1) Machine Learning for Security and (2) Security of Machine Learning Systems. The focus of the first part will be on building a principled understanding of key learning algorithms and techniques, and their applications within the security domain, as well as general questions related to analyzing and handling datasets. The first part will provide students with the necessary background to understand the second half of the course. The second part covers recently discovered security implications of deploying machine learning algorithms in the physical realm. Students will learn about attacks against computer systems leveraging machine learning algorithms, as well as defense techniques to mitigate such attacks during learning and inference. The course aims to motivate the exploration of new problems that advance the state-of-the-art; thus, the course will include reviewing recent papers from top-tier conferences. Students successfully completing this class will be able to evaluate machine learning systems in academic and commercial security, and will have rudimentary skills in security and machine learning research.

Prerequisites

The course assumes knowledge of programming, calculus, basic probability and mathematical statistics. You must be comfortable writing code to process and analyze data, and be familiar with basic algorithmic design and analysis.

Material

There is no official textbook for the class. Slides will be provided and reading materials for each topic will be assigned from the following references

Recommended Texts

1. Pattern Recognition and Machine Learning, Christopher Bishop, Springer (2006)
2. Machine Learning A Probabilistic Perspective, Kevin P Murphy, MIT Press (2012)
3. Deep Learning, Ian Goodfellow, Yoshua Bengio, and Aaron Courville, MIT press (2016)
4. Adversarial Machine Learning, Yevgeniy Vorobeychik and Murat Kantarcioglu, Morgan & Claypool Publishers (2018)
5. Machine Learning and Security: Protecting Systems with Data and Algorithms, Clarence Chio and David Freeman, O'Reilly Media (2018)

Recitations

Recitations are complement to lectures. One- or two-hour recitations will be held. Recitations are aimed to address background material, individualized studying that is unclear and practical application of materials presented in lectures. Recitations also provides a more intimate opportunity to ask questions of and to interact with the course staff.

Announcements

Course announcements will be made through the course e-mail list. The e-mail list uses your Purdue e-mail address. You are expected to check this account regularly for information related to the class. Please sign-up on Piazza to ask/answer questions. We will send class announcements through this site.

FAQs

Why should I take this course?

You will learn key machine learning algorithms and techniques, and their security applications, as well as you will explore the attack surface of systems built upon ML, and vulnerabilities of ML and the countermeasures used to defend against them.

Why is this course important?

ML algorithms are used in a large variety of computer security applications, including the discovery of new malware families, uncovering vulnerabilities in software, and analyzing malicious code. There is a growing recognition that ML applications expose vulnerabilities such as adversarial perturbations. This course brings these two complementary views together by (a) exploring ML algorithms for security, and (b) investigating the security of ML algorithms.

(Tentative) Grading Scheme

The course will be graded on assignments, exams, paper reviews, and class participation in the following proportions:

- 40% Homework
- 15% Midterm Exam
- 15% Final Exam
- 30% Paper reviews, extra homework and class participation (random quizzes)

Students are guaranteed to receive the letter grade based on the standard scale: A = 93% - 100%; A- = 90% - 93%; B+ = 87% - 90%; B = 87% - 83%; B- = 83% - 80%; C+ = 80% - 77%; C = 77% - 70%; D = 70% - 60%; F = 0% - 60%. Curving is at the discretion of the professor.

Homework Assignments

Homework assignments (along with instructions) will be posted on the Blackboard. We aim to have five homework assignments (both implementation/empirical and problem solving/mathematical). But this is subject to change as the semester progresses. Each homework write-up must be neatly typeset as a PDF document. You can use LaTeX or any other system that produces typesetting of equal quality and legibility (especially for mathematical symbols and expressions). Please write your solutions as succinctly as possible while including all the necessary details. Ensure that the following appear at the top of the first page of the write-up: **your name**, **your Purdue ID**, and **the ID's of any students** with whom you discussed the assignment. Submit your write-up as a single PDF file and corresponding code implementations (if any) on Blackboard by 11:59 PM of due date (always going to be Sunday). It is your responsibility to ensure that the submission is successfully received by Blackboard.

Exams

There will be a midterm and a final exam. The first exam covers topics from the first two parts of the course; the second exam covers topics from the entire course, but with an emphasis on the material covered after the midterm exam. Exams are closed book.

Quizzes

There will be about 3-5 random quizzes during the semester that will mostly contribute to 10% attendance. Each quiz will be based on most recent lecture and reading assignments. Please make sure you carefully read the assigned material for each lecture.

Paper Reviews

Understanding research papers is a key task in computer science research. In this class, students will provide one-page reviews research papers assigned as readings. Roughly one review will be due per week. These reviews are due at the beginning of class. Most of the course readings will come from seminal papers in the field. Links to these papers will be provided on the course pages.

Missing or Late Work

The score for a late homework, a paper review, a missed quiz and exam is 0. Exceptions will be made in case of serious illness or bereavement. If a student has a planned absence for a class when an exam will be given, the student should make arrangement before the planned absence to take the exam early or take a makeup exam after returning to campus.

Grade Disputes

Feedback on graded material will be posted on Blackboard in as timely a manner as possible. Once feedback for a graded assignment is posted, you will have 1 week from the posting date to dispute a grade. No re-grade requests will be honored after 1 week from posting feedback.

FAQs

❓ **How this course is different from other security and data science courses?**

❗ Data science is a broad term for multiple disciplines, machine learning fits within data science. While the focus of data science is mainly on skills and tools (e.g., Spark, MapReduce, etc) required to tackle big data including feature engineering, data cleaning, preparation, exploratory data analysis, and utilizing ML algorithms, the course emphasis will be on machine learning algorithms and security applications, with a perspective in evaluating security that focuses on attacks and defenses with respect to the machine learning pipeline. Yet, several software libraries and publicly available datasets will be used to illustrate the application of these algorithms (e.g., pandas for loading and exploratory data analysis and scikit-learn for algorithms). In your homework assignments, you will also be asked to design and implement new learning algorithms from scratch.

❓ **How to succeed in class?**

❗ Review slides before each lecture to familiarize yourself with the topics to be covered, and complete reading assignments after each lecture. Attend lectures recitations and office hours, and ask questions. TAs and myself are always available to help you.

Collaboration Policy

You are encouraged to discuss course materials and reading assignments, and homework assignments with each other in small groups (two to three people). You must list all discussants in your homework write-up. Discussion about homework assignments may include brainstorming and verbally discussing possible solution approaches, but must not go as far as one person telling others how to solve a problem. In addition, you must write-up your solutions by yourself, and you may not look at another student's homework write-up/solutions (whether partial or complete).

Conduct and Courtesy

Students are expected to maintain a professional and respectful classroom environment. This includes: silencing cellular phones, arriving on time for class, speaking respectfully to others and participating in class discussion. You may use non-disruptive personal electronics for the purpose class participation (e.g., taking notes).

Correspondence with the instructor: The best way to correspond in this class is by emailing the instructor. Please prefix all course-related emails with the string **CS-529** to help filter email. The instructor will make every effort to answer promptly (within 48 hours). However, replies could be delayed due to circumstances outside the instructor's control.

Academic Integrity

Behavior consistent with cheating, copying, and academic dishonesty is not tolerated. Depending on the severity, this may result in a zero score on the assignment or exam, and could result in a failing grade for the class or even expulsion. Purdue prohibits "dishonesty in connection with any University activity. Cheating, plagiarism, or knowingly furnishing false information to the University are examples of dishonesty." (Part 5, Section III-B-2-a, University Regulations) Furthermore, the University Senate has stipulated that "the commitment of acts of cheating, lying, and deceit in any of their diverse forms (such as the use of substitutes for taking examinations, the use of illegal cribs, plagiarism, and copying during examinations) is dishonest and must not be tolerated. Moreover, knowingly to aid and abet, directly or indirectly, other parties in committing dishonest acts is in itself dishonest." (University Senate Document 7218, December 15, 1972). You are expected to read both Purdue's guide to academic integrity and Prof. Gene's Spafford's guide as well. You are responsible for understanding their contents and how it applies to this class.

Posting Class Material: Posting material associated with this class (e.g., solutions to homework sets or exams) without the written permission of the instructor is forbidden and may be a violation of copyright.

Students with Disabilities

Purdue University is required to respond to the needs of the students with disabilities as outlined in both the Rehabilitation Act of 1973 and the Americans with Disabilities Act of 1990 through the provision of auxiliary aids and services that allow a student with a disability to fully access and participate in the programs, services, and activities at Purdue University. If you have a disability that requires special academic accommodation, please make an appointment to speak with the instructor within the first three (3) weeks of the semester in order to discuss any adjustments. It is the student's responsibility to notify the Disability Resource Center of an impairment/condition that may require accommodations and/or classroom modifications. We cannot arrange special accommodations without confirmation from the Disability Resource Center.

Instructor Absence

The instructor might be away for a few classes. There will be a guest instructor for these classes. If we need to reschedule additional classes, we will do so on an as-needed basis.

FAQs

? **What if I lack critical background skills?**

! We will start at the beginning and do our best to make the background material accessible to everyone through recitations. If you cannot attend recitations, we will make the materials covered in recitations available.

? **Are e-books versions of the texts available at Purdue library?**

! Review slides before each lecture to familiarize yourself with the topics to be covered, and complete reading assignments after each lecture. Attend lectures recitations and office hours, and ask questions. TAs and myself are always available to help you.

Emergencies

In the event of a major campus emergency, course requirements, deadlines and grading percentages are subject to changes that may be necessitated by a revised semester calendar or other circumstances beyond the instructor's control. Relevant changes to this course will be posted onto the course website and/or announced via email. You are expected to read your purdue.edu email on a frequent basis. Emergency Preparedness: Emergency notification procedures are based on a simple concept: If you hear an alarm inside, proceed outside. If you hear a siren outside, proceed inside. Indoor Fire Alarms are mean to stop class or research and immediately evacuate the building. Proceed to your Emergency Assembly Area away from building doors. Remain outside until police, fire, or other emergency response personnel provide additional guidance or tell you it is safe to leave. All Hazards Outdoor Emergency Warning sirens mean to immediately seek shelter (Shelter in Place) in a safe location within the closest building. "Shelter in place" means seeking immediate shelter inside a building or University residence. This course of action may need to be taken during a tornado, a civil disturbance including a shooting or release of hazardous materials in the outside air. Once safely inside, find out more details about the emergency. Remain in place until police, fire, or other emergency response personnel provide additional guidance or tell you it is safe to leave. In both cases, you should seek additional clarifying information by all means possible: Purdue Home page, email alert, TV, radio, etc. Review the Purdue Emergency Warning Notification System multi-communication layers. Please review the Emergency Response Procedures . Please review the evacuation routes, exit points, emergency assembly area and shelter in place procedures and locations for our building. Video resources include a 20-minute active shooter awareness video that illustrates what to look for and how to prepare and react to this type of incident.

Violent Behavior Policy

Purdue University is committed to providing a safe and secure campus environment for members of the university community. Purdue strives to create an educational environment for students and a work environment for employees that promote educational and career goals. Violent Behavior impedes such goals. Therefore, Violent Behavior is prohibited in or on any University Facility or while participating in any university activity.

CAPS Information

Purdue University is committed to advancing the mental health and well-being of its students. If you or someone you know is feeling overwhelmed, depressed, and/or in need of support, services are available. For help, such individuals should contact Counseling and Psychological Services (CAPS) at (765)494-6995 during and after hours, on weekends and holidays, or through its counselors physically located in the Purdue University Student Health Center (PUSH) during business hours.

Nondiscrimination

Purdue University is committed to maintaining a community which recognizes and values the inherent worth and dignity of every person; fosters tolerance, sensitivity, understanding, and mutual respect among its members; and encourages each individual to strive to reach his or her own potential. In pursuit of its goal of academic excellence, the University seeks to develop and nurture diversity. The University believes that diversity among its many members strengthens the institution, stimulates creativity, promotes the exchange of ideas, and enriches campus life. Purdue University prohibits discrimination against any member of the University community on the basis of race, religion, color, sex, age, national origin or ancestry, marital status, parental status, sexual orientation, disability, or status as a veteran. The University will conduct its programs, services and activities consistent with applicable federal, state and local laws, regulations and orders and in conformance with the procedures and limitations as set forth in Executive Memorandum No. D-1, which provides specific contractual rights and remedies.

Upon successful completion, the students are expected to have a good understanding of fundamental concepts of ML, familiarity of main algorithms, evaluate the security of ML systems and required programming skills to be able to design data-driven solutions for challenging security problems. The tentative list of topics to be covered is:

Background: The Basics of Machine Learning

1. Introduction to Machine Learning
2. The Process of Learning and Key Concepts
3. Background I: Review of Probability and Statistics
4. Background II: Linear Algebra and Vector Calculus

PART 1: Machine Learning Algorithms for Security

These lectures will cover the practical technical understanding of machine learning algorithms applied to security. Examples include malware classification, finding insider threats, and detection of IoT sensor anomalies.

1. Regression: Linear Regression
2. Logistic Regression
3. Probabilistic classifier: Naive Bayes
4. Decision Trees
5. Instance-based Learning (Lazy Learning): k-nearest neighbors (k-NN))
6. Clustering: k-means and Gaussian Mixture Models
7. Dimensionality Reduction: Principle Component Analysis and Linear Discriminant Analysis
8. Neural Networks (Perceptron & Deep Neural Networks)
9. Convolutional Neural Networks

PART 2: Security of Machine Learning Systems

We will view systems built on ML through confidentiality, integrity, and availability (CIA) model, articulate a comprehensive threat model for machine learning systems, and categorize attacks and defenses within an adversarial framework.

1. Attacks on the Machine Learning Pipeline: Poisoning attacks, model theft attacks, adversarial examples, recovery of sensitive training data, and physical-world attacks
2. Threat Models: White Box, Black Box, and Grey Box
3. Transferability
4. Types of Defenses: Pre-processing, and robust optimization
5. Introduction to Privacy in Machine Learning: Membership inference and model inversion attacks

Recitations (The schedule of recitations will be determined later.)

1. **Background:** This recitation will provide students an overview of probability, linear algebra, and vector and matrix calculus.
2. **Computational Tools for Data Science:** This recitation will focus on developing practical skills in working with security data. This will be a coding-intensive recitation. We are using Python, since it allows for fast prototyping and is supported by a great variety of scientific (and, specifically, data related) libraries. We will cover a set of tools such as Scikit-learn (a free software machine learning library), NumPy (Python mathematical functions library) and Pandas (Python Data Analysis Library).
3. **Evaluating Machine Learning Algorithms:** This recitation will cover correct evaluation methodology of machine learning systems, including case studies of methodological errors. You will learn terms to evaluate machine learning systems, such as stratified cross validation, confusion matrix, and ROC curves.
4. **Introduction of Deep Learning Libraries in Python:** We will introduce Keras, a library for deep learning in Python, especially for beginners. We will demonstrate the steps to build a CNN using Keras with Theano or TensorFlow backend.

Changes to the syllabus: This syllabus is subject to change. Updates will be posted and dated on the course website.

(Tentative) Schedule

Background: The Basics of Machine Learning

Week 1 (Aug 19 - Aug 23)	Logistics & Course Overview	What is machine learning: A brief history and applications
	The Processes of Learning and Overview of Key Concepts	Data-driven problem solving, learning stages, feature engineering, training, model selection, generalization, optimization, prediction overfitting, underfitting, and regularization
	Background Readings (do them before class)	Bishop Chapter 1 (excluding 1.2.5, 1.2.6 and 1.6) and Mathematics for Machine Learning
	📅 Recitation #1	Date & time and location TBD

Part 1: Machine Learning Algorithms for Security

Week 2 (Aug 26 - Aug 30)	Linear Models for Regression	The Ordinary Least Squares (OLS) and regularization
	Loss Functions and Optimization	Gradient descent (GD) and stochastic gradient descent (SGD)
	Background Readings (do them before class)	TBD
	📅 Recitation #2	Date & time and location TBD
	⏲ Homework #1	(Due Sunday 9/8/2019 11:59 PM)
Week 3 (Sep 2 - Sep 6)	Logistic Regression	Introduction to supervised learning, the predicting probabilities, the logit function, Logistic regression for classification and regularization, the exponential loss and loss minimization viewpoint, Gradient Descent (GD) for LR
	Decision Trees	Entropy, Information gain, Decision Trees, Regularization
	Background Readings (do them before class)	TBD
Week 4 (Sep 9 - Sep 13)	Naive Bayes Classifier	Bayes' theorem, independence between features
	Nearest Neighbors	Similarity measures, decision boundaries, model selection for NN, and metric learning
	⚡ Paper Review	Sikder et al., 6thSense: A Context-aware Sensor-based Attack Detector for Smart Devices
	Background Readings (do them before class)	TBD
	⏲ Homework #2	(Due Sunday 9/22/2019 11:59 PM)
Week 5 (Sep 16 - Sep 20)	Clustering	Introduction to unsupervised learning, k-means, Mixture of Gaussians and EM algorithm
	Background Readings (do them before class)	TBD
	⚡ Paper Review	Yen et al., Beehive: Large-scale log analysis for detecting suspicious activity in enterprise networks
	📅 Recitation #3	Date & Time and location TBD

Week	Topic	Details
Week 6 (Sep 23 - Sep 27)	Feature Space Reduction	Filter and Wrapper methods
	Feature Projection	Principal Component Analysis (PCA), Linear Discriminant Analysis
	Background Readings (do them before class)	TBD
	⊘ Homework #3	(Due Sunday 10/6/2019 11:59 PM)
	⚡ Paper Review	Rossow et al., Prudent practices for designing malware experiments: Status quo and outlook
Week 7 (Sep 30 - Oct 4)	Artificial Neural Networks	Binary classification, Perceptron algorithm and its convergence, multiple layer perceptron, overview of Deep Neural Networks and its key challenges
	Background Readings (do them before class)	TBD
	⚡ Paper Review	Shin et al., Recognizing Functions in Binaries with Neural Networks
Week 8 (Oct 7 - Oct 12)	Convolutional Neural Networks	CNN Architectures, and Convolution/Pooling/ReLu layers
	Background Readings (do them before class)	TBD
Week 9 (Oct 14 - Oct 18)	Review Part 1 and Part 2	Wrap up the material, ask me anything
	❗ Midterm Exam (Oct 17)	Covers all topics

Part 2: Security of Machine Learning Systems

Week	Topic	Details
Week 10 (Oct 21 - Oct 25)	Overview of key concepts	Poisoning attacks, model theft attacks, adversarial examples, and recovery of sensitive training data
	Background Readings (do them before class)	Saltzer's and Schroeder's Design Principles
	⚡ Paper Review	Papernot et al., Towards the Science of Security and Privacy in Machine Learning
	📅 Recitation #4	Date & time and location TBD
Week 11 (Oct 28 - Nov 1)	Perturbation Attack Strategies	Adversarial samples, misclassification and targeted attacks, FGSM, JSMA, and CW attacks
	Background Readings (do them before class)	Adversarial Examples: Attacks and Defenses for Deep Learning
	⊘ Homework #4	(Due Sunday 11/3/2019 11:59 PM)
	⚡ Paper Review	Booler et al. Simple Physical Adversarial Examples against End-to-End Autonomous Driving Models
Week 12 (Nov 11 - Nov 15)	Adversarial Examples in Physical-world	Noisy physical environments, dynamic physical conditions including different viewpoint angles and distances
	Adversarial Examples in Constrained Domains	Semantics of features and capability of only controlling a subset of features
	Background Readings (do them before class)	Adversarial Examples for Malware Detection
	⚡ Paper Review	Kevin et al., Robust Physical-World Attacks on Deep Learning Visual Classification

	⊘ Homework #5	(Due Sunday 11/24/2019 11:59 PM)
Week 13 (Nov 4 - Nov 8)	Transferability & Black-Box Attacks	Substitute models, intra and cross-technique transferability
	Evaluating Robustness of ML systems	Defense evaluations and adversarial example defenses
	Background Readings (do them before class)	When Does Machine Learning FAIL? Generalized Transferability for Evasion and Poisoning Attacks and On Evaluating Adversarial Robustness
	⚡ Paper Review	Ling et al., DeepSec: A Uniform Platform for Security Analysis of Deep Learning Models
Week 14 (Nov 18 - Nov 22)	Defenses Against Adversarial Attacks	Pre-processing and robust optimization techniques, and empirical and theoretic approaches
	Thanksgiving break (Nov 27 - Nov 30)	No class on Thursday
	Background Readings (do them before class)	TBD
Week 15 (Nov 25 - Nov 29)	Generative Adversarial Networks (GANs)	Fake data generation and its applications
	Introduction to Privacy in ML Models	Model reconstruction, model inversion, membership inference, and privacy preserving ML
	⚡ Paper Review	Hitaj et al., Deep Models Under the GAN: Information Leakage from Collaborative Deep Learning
Week 16 (Dec 3 - Dec 7)	Future Research Directions	Introduction of related conferences and open research topics
	Review Part 2	Wrap up the material, ask me anything
❗ Final Exam	December 9-14 (TBD)	Covers all topics

⊘ Homework #5   (Due Sunday 11/24/2019 11:59 PM)