San Jose State University
Department of Computer Science
CS 266, Topics in Information Security, Fall 2016

- **Course and Contact information**
  - **Instructor:** Mark Stamp
  - **Office Location:** MH 216
  - **Telephone:** 408-924-5094
  - **Email:** mark.stamp@sjsu.edu
  - **Office hours:** Tuesday noon - 2:00pm
  - **Class Days/Times:** Tuesday and Thursday, 10:30-11:45pm
  - **Classroom:** MH 422
  - **Prerequisites:** CS 166 or instructor consent

- **Course Description**
  - Advanced topics in the area of information security. Content differs with each offering. Possible topics include, but are not restricted to: Network Security, Software Reverse Engineering and Cryptanalysis. Prerequisite: CS 166 or instructor consent.

- **Learning Outcomes**
  - The focus of this course will be machine learning, with applications drawn primarily from information security. After completing this course students should have a working knowledge of a wide variety of machine learning topics, and have a deep understanding of the application of such techniques to selected challenges in information security.

- **Required Texts/Readings**
  - The primary text will be a manuscript written by your instructor. This manuscript, titled ***Machine Learning with Applications in Information Security***, covers many machine learning concepts in detail, with a large number of illustrative applications. Most of the applications are from information security, including a variety of topics related to malware, intrusion detection, spam, and cryptanalysis, among others. The manuscript will soon be published as a textbook by CRC Press.

  - Other useful resources:
    - *Computer Viruses and Malware*, John Aycock, Springer 2006. Many of the applications we will discuss are related to malware. Aycock's book is easy to read and provides a solid foundation for malware research.
    - *Information Security: Principles and Practice*, Mark Stamp, Wiley 2011. If you have not taken CS 265, then you will likely need to refer to this book at various points during this course.
    - Open Malware (at http://www.offensivecomputing.net/) includes a large collection of samples of live malware.
    - VX Heavens (at http://vx.netlux.org/) is a source for "hacker" information on viruses and malware samples.
    - Journal of Computer Virology and Hacking Techniques (at http://www.springer.com /computer/journal/11416) is a journal for malware-specific research papers. There are

also several good conferences that focus on malware and/or machine learning applications in information security.
    - ***Practical Malware Analysis: The Hands-On Guide to Dissecting Malicious Software***, Michael Sikorski and Andrew Honig, No Starch Press, 2012. This is an excellent book for details on reverse engineering (whether for malware analysis or other purposes). The book includes many hands-on exercises.
    - Software Reverse Engineering (SRE) website (at http://reversingproject.info/). This website was created by a former masters student of mine and it includes lots of good information and detailed exercises with solutions.
    - ***Security in Computing***, third edition, Charles P. Pfleeger and Shari Lawrence Pfleeger, Prentice Hall, 2003, ISBN: 0-13-035548-8. The strength of this book is its coverage of the security issues related to software. In particular, operating systems and some aspects of secure software engineering are covered well. The book also has some good basic information on viruses.
    - Recent masters project reports (at http://www.cs.sjsu.edu/~stamp /cv/mss.html#masters). The majority of these projects involve applications of machine learning to malware or other topics in information security.

  - Additional relevant material:
    - PowerPoint slides at http://www.cs.sjsu.edu/~stamp/ML/powerpoint
    - Current semester lecture videos are available at http://www.cs.sjsu.edu/~stamp /ML/lectures/CS266_Fall16/. If you are asked to login to access the videos, both the username and password are "infosec". **Note**: The instructor hereby gives students permission to record his lectures (audio and/or video). At least with respect to this class, your instructor has nothing to hide.
    - Class-related discussion will be posted on Piazza at https://piazza.com/class /isa7q4rn49f4al. You are strongly encouraged to participate by asking questions, as well as by responding to questions that other students ask. At the start of the semester, you should receive an email asking you to join this discussion group—if not, contact your instructor via email.

- **Course Requirements and Assignments**
  - SJSU classes are designed such that in order to be successful, it is expected that students will spend a minimum of forty-five hours for each unit of credit (normally three hours per unit per week), including preparing for class, participating in course activities, completing assignments, and so on. More details about student workload can be found in University Policy S12-3 at http://www.sjsu.edu/senate/docs/S12-3.pdf.

  - Schedule
    - Week 1 --- Introduction and overview
    - Week 2 --- Hidden Markov Models
    - Week 3 --- Data Analysis
    - Week 4 --- Applications of Hidden Markov Models
    - Week 5 --- Profile Hidden Markov Models
    - Week 6 --- Applications of Profile Hidden Markov Models
    - Week 7 --- Principal Component Analysis
    - Week 8 --- Applications of Principal Component Analysis

- Week 9 --- Support Vector Machines
- Week 10 --- Applications of Support Vector Machines
- Week 11 --- Clustering
- Week 12 --- Clustering Applications
- Week 13 --- k-Nearest Neighbor, Neural Networks, Boosting/AdaBoost, Random Forests
- Week 14 --- Linear Discriminant Analysis, Naive Bayes, Regression Analysis, Conditional Random Fields
- Week 15 --- Project presentations

- Homework is due *typewritten* (include source code, but not executable files) by class starting time on the due date. Each assigned problem requires a solution and an explanation and work detailing how you arrived at your solution. Cite any outside sources used to solve a problem. When grading an assignment, I may ask for additional information. Note that a *subset* of the assigned problems will typically be graded.

  Zip your homework into a file named hmk.zip. Email your work to cs266.fall16@gmail.com. The subject line of your email ***must*** be of the form:

  ```
  CS266HMK assignmentnumber yourlastname last4digitofyourstudentnumber
  ```

  The subject line must consist of the four identifiers listed. There is no space within an identifier and each identifier is separated by a space.

  - Assignment 1: Due Thursday, September 8
    Chapter 2, problems 1, 2, 3, 6, 10, 11, 14. The homework problems for Chapter 2 can be found here (at https://www.cs.sjsu.edu/~stamp/CS266/other/HMM_problems.pdf).
  - Assignment 2: Due TBD
  - Assignment 3: Due TBD
  - Assignment 4: Due TBD
  - Assignment 5: Due TBD
  - Assignment 6: Due TBD
  - Assignment 7: Due TBD
  - Assignment 8: Due TBD

        - Assignment 9: Due TBD
        - Assignment 10: Due TBD
    - NOTE that University policy F69-24 at http://www.sjsu.edu/senate/docs/F69-24.pdf states that "Students should attend all meetings of their classes, not only because they are responsible for material discussed therein, but because active participation is frequently essential to insure maximum benefit for all members of the class. Attendance per se shall not be used as a criterion for grading."

- **Grading Policy**
    - Test 1, 100 points. Date: TBD.
    - Homework, quizzes, class participation and other work as assigned, 100 points. A subset of the assigned problems will be graded.
    - Machine Learning Project, 100 points. You must obtain approval for your project (via email) by September 30. A written project report is due December 1. Note that a written report is required, and oral presentations will begin on (or shortly after) the report due date.
    - Final, 100 points. Date: Wednesday, December 14 from 9:45am-noon. The official finals schedule is here: http://info.sjsu.edu/static/catalog/final-exam-schedule-fall.html
    - Semester grade will be computed as a weighted average of the major scores listed above.
    - ***No*** make-up tests or quizzes will be given and ***no*** late homework or project (or other work) will be accepted.
    - Grading Scale:

Percentage	**Grade**
92 and above	A
90 - 91	A-
88 - 89	B+
82 - 87	B
80 - 81	B-
78 - 79	C+
72 - 77	C
70 - 71	C-
68 - 69	D+
62 - 67	D
60 - 61	D-
59 and below	F

    - Note that "All students have the right, within a reasonable time, to know their academic scores, to review their grade-dependent work, and to be provided with explanations for the determination of their course grades." See University Policy F13-1 at http://www.sjsu.edu/senate/docs/F13-1.pdf for more details.

- **Classroom Protocol**
  - Keys to success: Do the homework, complete a good project, and attend class
  - **Wireless laptop is *required***. Your laptop must remain closed (preferably in your backpack and, in any case, not on your desk) until I inform you that it is needed for a particular activity
  - **Cheating** will not be tolerated, but working together is encouraged
  - Student must be respectful of the instructor and other students. For example,
    - No disruptive or annoying talking
    - Turn off cell phones
    - Class begins on time
    - Class is not over until I say it's over
  - Valid picture ID required at all times
  - The last day to drop is Tuesday, September 6, and the last day to add is Wednesday, September 13
- **University Policies**
  - Office of Graduate and Undergraduate Programs maintains university-wide policy information relevant to all courses, such as academic integrity, accommodations, etc. You may find all syllabus related University Policies and resources information listed on GUP's Syllabus Information web page at http://www.sjsu.edu/gup/syllabusinfo/