EN.650.672 Security Analytics – Fall 2018

Meeting Time:	6:00-8:30pm, Monday;
Instructor:	Lei Ding
Email:	lding10@jhu.edu
Class Location:	Gilman 219

Course Information:
Security analytics refers to information technology solutions that gather and analyze security events to bring situational awareness and enable IT staff to understand and analyze events that pose the greatest risk. Increasingly, detecting and preventing cyber attacks require sophisticated use of data analytics and machine learning tools. This course will cover fundamental theories and methods in data science, modern security analytical tools, and practical use cases of security analytics. Students of this course learn concepts, tasks, and methods of data science; and how to apply data science to cyber security problems. Students also learn how to use modern software in security analytics.

Use of Course Website:

- You should/will be enrolled into the BlackBoard course website for this course.
- Class notes, assignments, and announcements will be posted on the course website for each class period. You are responsible for printing out them for reference in class.

Use of Computer and Software:
In addition to Internet access, you need a personal computer or access to other computer systems for some course assignments and course projects. Knowledge of working with computers is expected including Python programming skills. You may be asked to install analytical development environment on your computer. This course will use open source software and resources including:

- Python and its data analysis modules
- Jupyter Notebook (jupyter.org)

Reading and Sample Textbook Recommendations:
Two textbooks are recommended.

- "Data Science for Business" by Foster Provost and Tom Fawcett, O'Reilly Media, 2013
- "Data Driven Security" by Bob Rudis and Jay Jacobs, Wiley, 2014

Tentative Topics:

- Introduction to security analytics
  - Commercial needs, available methodologies, state-of-the-art, use cases
  - Analytical software tools and platform to be used, course dataset, etc.
- Typical security analytics goals, tasks, and procedures
  - Think cybersecurity problems as data analytical tasks
  - Supervised and unsupervised machine learning
  - Data mining process
  - Typical analytical techniques and technologies
- Security data preparation and exploration
  - Security data sources
  - Common problems in collecting and preparing security data
  - Basic security data exploration techniques
- Building predictive models for cyber security applications
  - Typical security applications using predictive models
  - Feature selection
  - SVM, linear regression, logistic regression models
  - Practical examples of using predictive models in cybersecurity
- Clustering security data

    - Typical security applications using clustering
    - Most common clustering algorithms
    - Practical examples of using clustering in cybersecurity
- Evaluating security analytical models
    - Problems to be addressed
    - Available methods of evaluating analytical model performance
    - Practical examples
- Mining in text and its applications in cybersecurity
    - Preparing security text for mining tasks
    - Text mining techniques
    - Practical example of using text mining in cybersecurity
- Introduction to Deep Learning
    - Deep Neural Network
    - TensorFlow
- Visualizing security data
    - The goal of security visualization
    - Visual security analysis
        - Reporting
        - Historical data analysis
        - Real-time monitoring and analysis
    - Data visualization tools
    - Security analytical dashboard
- Big data for cybersecurity
    - Introduction to big data technologies and tools
    - Use big data technologies in security analytics
    - Practical examples

Student Evaluation:

- Homework (30%) – 3 homework assignments to be fulfilled individually or in groups as specified. These assignments include short-essay responses similar to those in exams or analytical and programming exercises with required computer software tools.
- Quizzes (30%) – 2 quizzes on understanding of basic concepts in security analytics and knowledge to address common requirements of security analytics. All the quizzes require written answers by individual students in class. No computational aids and Internet access are allowed in exams.
- Term Project (40%) – 1 project proposal report to include proposal, literature review and problem definition; 1 final report (including any design/implementation) to include problem/challenge addressed in cybersecurity, analytical solution adopted in the project, and project results. Students are encouraged to complete projects in teams; each team should not have more than 3 students. There will be a presentation at the end of semester. Detailed guidelines will be given in class.

Grading:

A letter grade will be assigned according to this formula: A - 85% and above; B - 70% to 85%; C - 60% to 70%; D - 50% to 60%; F - less than 50%. Appropriate upward curving will be made as necessary.

General Requirements:

- Collaboration is not allowed for individual homework assignments and quizzes. Discussion and collaboration is encouraged in team projects. In projects, every student is expected to work with his/her teammates and contribute equally. Individual contributions should be indicated in the report. Only typed or electronic reports (hand drawing okay for figures if necessary and legible) are allowed for homework and term project.
- Highly ethical behavior is expected when using computing tools and techniques especially when working at on-campus or remote computing facilities.

- Each student has a responsibility to understand, accept, and comply with the university's standards of academic conduct, as well as policies established by the schools. For example, getting help from fellow students or helping others in examination will be a violation. Copying other students' work or online materials, or quoting others' work in report without contribution declarations, will be plagiarism activity. At my discretion, any or all papers submitted in this course may be subject to a plagiarism detection service.

Ethics:

The strength of the university depends on academic and personal integrity. In this course, you must be honest and truthful, abiding by the *Computer Science Academic Integrity Policy*:

> Cheating is wrong. Cheating hurts our community by undermining academic integrity, creating mistrust, and fostering unfair competition. The university will punish cheaters with failure on an assignment, failure in a course, permanent transcript notation, suspension, and/or expulsion. Offenses may be reported to medical, law or other professional or graduate schools when a cheater applies.
>
> Violations can include cheating on exams, plagiarism, reuse of assignments without permission, improper use of the Internet and electronic devices, unauthorized collaboration, alteration of graded assignments, forgery and falsification, lying, facilitating academic dishonesty, and unfair competition. Ignorance of these rules is not an excuse.
>
> Academic honesty is required in all work you submit to be graded. Except where the instructor specifies group work, you must solve all homework and programming assignments without the help of others. For example, you must not look at anyone else's solutions (including program code) to your homework problems. However, you may discuss assignment specifications (not solutions) with others to be sure you understand what is required by the assignment.
>
> *If* your instructor permits using fragments of source code from outside sources, such as your textbook or on-line resources, you must properly cite the source. Not citing it constitutes plagiarism. Similarly, your group projects must list everyone who participated.
>
> Falsifying program output or results is prohibited.
>
> Your instructor is free to override parts of this policy for particular assignments. To protect yourself: (1) Ask the instructor if you are not sure what is permissible. (2) Seek help from the instructor, TA or CAs, as you are always encouraged to do, rather than from other students. (3) Cite any questionable sources of help you may have received.
>
> On every exam, you will sign the following pledge: "I agree to complete this exam without unauthorized assistance from any person, materials or device. [Signed and dated]". Your course instructors will let you know where to find copies of old exams, if they are available.

Report any violations you witness to the instructor.

You can find more information about university misconduct policies on the web at these sites:

- For undergraduates: http://e-catalog.jhu.edu/undergrad-students/student-life-policies/
- For graduate students: http://e-catalog.jhu.edu/grad-students/graduate-specific-policies/

Students with Disabilities

Any student with a disability who may need accommodations in this class must obtain an accommodation letter from Student Disability Services, 385 Garland, (410) 516-4720, studentdisabilityservices@jhu.edu.

Campus and Online Resources:

Writing and independent study are important in graduate study and future industrial/academic work. Resources are available on the campus at various places including libraries, the CLE program, or online. You can get help in writing programs and can access many journals related to the topics of this class.