UMBC UGC New Course Request: IS 471: Data Analytics for Cybersecurity

Date Submitted: 12/15/15 Proposed Effective Date: Fall 2016

	Name	Email	Phone	Dept
Dept Chair or UPD	Carolyn Seaman	cseaman@umbc.edu	53937	IS
Other Contact	Vandana Janeja	vjaneja@umbc.edu	56238	IS

COURSE INFORMATION:

Course Number(s)	IS 471
Formal Title	Data Analytics for Cybersecurity
Transcript Title (≤30c)	Data Analytics for Cybersecurity
Recommended Course Preparation	
Prerequisite **NOTE:** Unless otherwise indicated, a prerequisite is assumed to be passed with a "D" or better.	IS 410 with C or better
Credits	3
Repeatable?	☐ Yes ☒ No
Max. Total Credits	3 This should be equal to the number of credits for courses that cannot be repeated for credit. For courses that may be repeated for credit, enter the maximum total number of credits a student can receive from this course. E.g., enter 6 credits for a 3 credit course that may be taken a second time for credit, but not for a third time. Please note that this does NOT refer to how many times a class may be retaken for a higher grade.
Grading Method(s)	☒ Reg (A-F) ☐ Audit ☐ Pass-Fail

PROPOSED CATALOG DESCRIPTION (no longer than 75 words**):**

Cyber security is pervasive in the areas of not only computer networks but also sensor networks, industrial control systems and user devices. One common thread in these types of systems and end users is data. This course provides an introduction to data analytics for multiple aspects of cyber security and focuses on using data analytics methods for discovering anomalies pertaining to cyber threats through exercises in programming and hands on data analytics tools.

RATIONALE FOR NEW COURSE:

a) Why is there a need for this course at this time?
Cyber security is a pervasive problem affecting individuals, organizations and governments. This is due to the acceptance and adoption of technology in the form of multiple types of non-traditional devices. Thus, cyber security has to address challenges emerging in the areas of not only computer networks but also sensor networks, industrial control systems and user devices. One common thread in all these types of devices and end users is data. Increasingly, the focus of cyber security is shifting to analyzing the data in not only a retrospective manner but also a prospective manner across different segments of cyber security domains. Thus, data analytics has to go beyond the traditional themes of security and seamlessly weave across several domains including networks, industrial control systems, and user roles to name a few. This course will provide a foundation towards addressing this need.

b) How often is the course likely to be taught?
Every other semester

c) How does this course fit into your department's curriculum?
The department currently does not have an undergraduate course in data analytics for cyber security. This course will fulfill that need. Moreover, a new undergraduate certificate in Cybersecurity Informatics is being proposed by the IS department. This course will serve as a core course requirement for that certificate. Also, this course can be used by IS majors as the required upper-level IS elective (if it is not being used for the certificate) or as the required third programming course.

d) What primary student population will the course serve?
Undergraduate IS majors who need a third programming course or an upper-level IS elective, or undergraduate IS majors who declare the Cybersecurity Informatics certificate.

e) Why is the course offered at the level (ie. 100, 200, 300, or 400 level) chosen?
This is an advanced upper level elective course, with a 400-level prerequisite, therefore has a 400 level designation.

f) Explain the appropriateness of the recommended course preparation(s) and prerequisite(s).
The students are expected to know the basics of programming and databases, therefore the prerequisite is IS 410. The prerequisite for IS 410 is successful completion of the IS gateway, which includes IS 147, the first Java programming course. This prerequisite ensures that students have the required preparation of basic programming (from IS 147 or equivalent) and basics of databases through IS 410.

g) Explain the reasoning behind the P/F or regular grading method.
This course will test the students' skills in programming and data analytics. There will be graded assignments and exams which will help students accumulate the points towards a final grade. The level of skill proficiency by the end of the semester will determine the level of the grade. Therefore it is a regular, letter graded course.

h) Provide a justification for the repeatability of the course.
This course is not repeatable.

ATTACH COURSE OUTLINE (mandatory):

Please see syllabus Attached.

DRAFT Syllabus

Information Systems Department
University of Maryland Baltimore County
Baltimore Maryland 21250
IS 471 "Data Analytics for Cybersecurity"

Instructor: Dr. Vandana Janeja
Office: ITE 429
e-mail vjaneja@umbc.edu : please put "cyber" in subject line
Course Delivery Site http://blackboard.umbc.edu
Office Hours: Mon 2:15-3:30 (other times by appointment)

Meeting Times: Mon-Wed 1-2:15 ITE 469

Textbook:
There is no single textbook for the course. Material will be delivered through powerpoint lecture slides and papers from conferences and journals. Reference Books and materials:

- RProgramming materials:
  - R Programming for Data Science, Robert Peng: https://www.cs.upc.edu/~robert/teaching/estadistica/rprogramming.pdf
  - R and Data Mining: Examples and Case Studies, Y Zhao: https://cran.r-project.org/doc/contrib/Zhao_R_and_data_mining.pdf
  - R for Beginners, E. Paradis: https://cran.r-project.org/doc/contrib/Paradis-rdebuts_en.pdf
- Data Mining concepts and techniques – Han and Kamber (3rd edition)
- Real Digital Forensics: Computer Security and Incident Response Keith J. Jones
- The Tao of Network Security Monitoring: Beyond Intrusion Detection, Richard Bejtlich

- Introduction to Computer Security, Matt Bishop
- Key conference and journal papers on Cybersecurity analytics

Course Description: Cyber attacks pose an increasing threat to the nation's critical infrastructure including computer networks, cyber physical systems such as industrial control systems, Sensor networks to name a few. This course is an introduction to data analytics for cybersecurity. The course will provide an introduction of cybersecurity and different aspects of it, study types of cyber attacks, anomalies and their relationship to cyber threats, introduction to data mining and big data analytics, methods for discovering anomalies, tools for data analytics and anomaly detection, hands on exercises for data analysis.

Instructional Methods: Discussion, Lectures and Demonstrations

Attendance and Participation: Regular and punctual attendance is expected of all students. In the case of absence due to emergency (illness, death in the family, accident), religious holiday, or participation in official College functions, it is the student's responsibility to confer with the instructor about the absence and missed course work. I expect that we all show mutual respect for each other during lectures and discussions. Mutual respect entails beginning the class on time, turning off cell phone ringers, pagers, and beepers, and allowing other members of the class to participate in dialogue without interruption or distraction. Adopting these practices is intended to minimize disruption to class discussion and dialogue and maximize the value of the class for all participants.

Class Preparation: All of the reading and homework assignments should be completed before the class in which the material is to be discussed.

Course Requirements:

Regular Punctual Attendance Class Assignments & Homework Tests Projects

Grading:

IS instructors are expected to have exams and evaluations consisting of a mix of class work, test, homework and programming projects, which result in a reasonable distribution of grades. Please be aware that the students will be tested **on both theory and practical aspects** of Analyitics for CyberSecurity for this course.The break up for this class is as follows:

Assignment 1 (individual): 10 Points
Assignment 2 (individual): 10 Points
Exam (In class – closed notes, individual): 20 points
Types of cyber attack: (group of two): 15 points
Attack Case Studies (group of two): 15 points
Class Project/ Term Paper (group of two): 25 points
Class Participation: 5 Points

Each task is explained in detail as follows:

Type of cyber attack: 15 points

A document will be uploaded on blackboard with an outline of the types of attacks. Each student group will select a specific type of cyber attack and explain the process of the particular cyber attack. This will be in the form of one or two slides illustrating the steps of the cyber attack describing the process by

which this attack takes place. This will be evaluated based on the level of details captured in the illustration. For example details may include the systems affected by the attack (server, desktop or mobile devices), which protocols are used, common attack mechanism etc. Examples will be posted on blackboard to show how a type of attack can be illustrated.

Assignments 1 and 2: 10 Points each

Each assignment will be an analytics task for which a sample dataset will be provided by the instructor and a specific task will have to be performed individually on the dataset. The evaluation will be based on the results of the analytics task.

Attack Case Study: 15 points

Students are expected to study a specific cyber attack case and illustrate the mechanism by which the attack was carried out. The outcome is a block diagram illustrating the process of the attack in a PowerPoint slide. The evaluation will be based on including specifics such as analysis of datasets attacked, level of breach, and financial assessments available from public sources (if available). Examples will be posted on blackboard to show how a real world attack was carried out in a block diagram.

Class Project / Term paper

The term paper will be on a topic you have selected for your final implementation project. The term paper is divided into multiple deliverables for ease of understanding of each tasks and also for better project management. The term paper can be a survey type paper or an analytics type paper. Some examples of term paper topic are: For survey type papers: Survey of types of attacks, Survey of Government databases, Mining govt databases to find trends in types of vulnerabilities, Clustering the vulnerability databases.For Cyber Data Analytics Paper: Snort log mining, Network traffic mining, Traffic data exploration, Network topography analysis, Redundancy discovery, Any type of digital forensics, Sensor network analysis, Alert generation, Communication Patterns etc.

Starting from mid semester I will have meeting times (Optional) with individuals to help keep the project on track. In the past the courses projects have been a highlight of the course. Some students have used real life examples from their jobs for project ideas. Many have used the topics in their work environment as well.

Project deliverables **(total 25 points): To be uploaded via blackboard**

- Deliverable 1 : One page write-up and potential datasets list in case of analytics or list of key sources to be surveyed (must include each members role)**2pts**
- Deliverable 2 : Analytics-preprocessed Datasets uploaded to blackboard –along with details of the data and source information. Survey – document with Key sources and detailed field information to be reviewed **4 pts**
- Deliverable 3 : Introduction and motivation **5pts**
- Deliverable 4 : Remainder of the term paper and project Presentation: Methodology and Experimental results **10 pts**, abstract, conclusions, related work and lessons learned **4 pts**

The presentation will also contribute to the class participation points – how you answer questions, how you critically evaluate other projects

Exam (20 points)

There will be only one exam for this course. The exam format will include problem solving questions and some basic theory questions. For example: patterns and how they can be interpreted to detect cyber threats, interpreting programming code outputs etc.

Class Participation (5 points)

There will be two meetings with IS 300 where groups will be formed between IS 300 and IS 498. This will be for a peer mentoring exercise based on discussions about Cybersecurity between the two classes and will be facilitated by questions provided to the students. The grade is based on participating in this exercise.

Grading criteria: With respect to final letter grades, the University's Undergraduate Catalogue states that, "A, indicates superior achievement; B, good performance; C, adequate performance; D, minimal performance; F, failure" There is specifically no mention of any numerical scores associated with these letter grades. Final letter grades in this course conform to the University's officially published definitions of the respective letter grades. In accordance with the published University grading policy, it is important to understand that final letter grades reflect academic achievement and not effort. While mistakes in the arithmetic computation of grades and grade recording errors will always be corrected, it is important to understand that in all other situations final letter grades are not negotiable and challenges to final letter grades are not entertained.

Academic Integrity: By enrolling in this course, each student assumes the responsibilities of an active participant in UMBC's scholarly community in which everyone's academic work and behavior are held to the highest standards of honesty. Cheating, fabricating, plagiarism, and helping others to commit these acts are all forms of academic dishonesty and they are wrong. Academic misconduct could result in disciplinary action that may include, but is not limited to, suspension or dismissal. Full policies on academic integrity should be available in the UMBC Student Handbook, Faculty Handbook, or the UMBC Directory.

Cheating in any form will not be tolerated in this class.*You may not copy other students' work or copy programs from the Internet. You will receive an F for any assignment found to be copied, from any source (this includes any collaboration on individual assignments), for the first time and any subsequent violations will result in immediate failure of the course. You may not reuse your own work from another class for the deliverables in this course.* **Any form of cheating will be reported and will stay on student's record for the rest of their term at UMBC with possible note on their transcripts.**

Inclement Weather: Any paper (hardcopy) assignment or test due on a class date that has been canceled due to inclement weather will be due the next class meeting. If it is an email/online submission the work should be submitted on the day it is due regardless of the class cancellation or inclement weather.

Accessibility in the classroom: If you have any special needs for technology or classroom accessibility please contact the student support Services and me as well if you like so that we can best accommodate your needs in a confidential and timely manner.

UMBC is committed to eliminating discriminatory obstacles that disadvantage students based on disability. Student Support Services (SSS) is the UMBC department designated to receive and maintain confidential files of disability-related documentation, certify eligibility for services, determine reasonable accommodations, develop with each student plans for the provision of such accommodations, and serve as a liaison between faculty members and students regarding disability-related issues. If you have a disability and want to request accommodations, contact SSS in the Math/Psych Bldg., room 213 or at 410-455-2459. SSS will require you to provide appropriate documentation of disability. If you require accommodations for this class, make an appointment to meet with me to discuss your SSS-approved accommodations.

TENTATIVE COURSE SCHEDULE
(Schedule subject to change)

There may be **guest lectures on Cybersecurity**. The schedule will be adjusted accordingly.

Date	Topic	Announcements	Deadlines	LAB
Week1	Class Introduction/ Introduction - Cyber security, Data analytics			
Week1	Cyber security, Data analytics, Vulnerability Databases			
Week2	Cybersecurity data/Methods of Data Collection			
Week2	BiG Data and Data Mining			
Week3	Data mining in context of cybersecurity			WEKA/ Excel
Week3	Clustering			WEKA
Week4	Association rules			R
Week4	*Intro to R Programming*	HW 1 announced		R
Week5	*R for Association rule mining and Clustering*	Optional One to one meetings for Project Development		WEKA
Week5	Classification	Type of attack Links posted	HW 1 Due	R
Week6	R for Classification	Optional One to one meetings for Project Development		
Week6	*R Case Study in cyber security, Project examples*		One page outline Due	WEKA
Week7	*Extending Java libraries in WEKA for DM*			
Week7	*WEKA/ Java based Programming, Case Study in cyber security, Past Project examples*		Type of attack Slide Due	
Week8	Anomaly Detection for Cyber Security	Attack Case study links posted		
Week8	**Types of Attacks - Presentations**	Optional One to one meetings for Project Development		
Week9	Anomaly Detection for Cyber Security		Preprocessed Datasets / Key source doc Due	
Week9	Outlier detection methods (IQR/ StdDev)			
Week10	Introduction to Packet sniffing	HW 2 announced		WireShark Intro
Week10	Tools for anomaly detection		Attack Case study Slide Due	Node XL
Week1	Link analysis			

1				
Week11	**Attack case studies - Presentations**		HW 2 Due	
Week12	Cybersecurity in Computer networks			
Week12	**Exam**	Optional One to one meetings for Project Development		
Week13	Cybersecurity in unstructured Web data			
Week13	Cybersecurity in unstructured Web data		Intro and Motivation Due	
Week14	Cybersecurity in Cyber physical systems	Optional One to one meetings for Project Development		
Week14	**Term Project Presentation**		ALL Final Project Presentations Due on May 4	
Week15	**Term Project Presentation**			
Exam week	Project Final Writeup Due			

Warning/Disclaimer: You will explore several Cybersecurity threats, software and potentially mal-intended websites. Please use utmost caution while learning and exploring these including visiting websites which may infect your computers. You may not use the packet sniffers to collect data without explicit permission from authorized personnel.

Safegaurds to follow to secure your own computer: http://www.fbi.gov/scams-safety/computer_protect