Information Sciences and Technology

Department

AIT 614:
Big Data Essentials

Course Syllabus

Spring 2021

Detailed information on all AIT 614 sections offered in the Spring 2021 semester including the day, time, location, instructors’ names and their contact information is available through the Schedule of Classes posted on PatriotWeb.

For each section, a customized syllabus with information specific to that section will be made available to registered students via the Blackboard Learning System.

Course Description

AIT 614: Big Data Essentials

Pre-requisites:
Graduate standing
AIT 524 (Recommended)

Catalog Description:
Students participate actively through class dialogues and the crafting of IT solutions to specific problem areas. Course cannot be used to satisfy course requirements for PhD IT students.

Special Topics Description:
This is a hands-on course that discusses emerging technologies for big data analytics and their applications in the real-world environment. Students apply learned concepts and best practices using a large variety of emerging technology tools to simulate the development, implementation, and usage of a functional big data analytical system. Topics include, but are not limited to, NoSQL, MapReduce Programming paradigm, Hadoop, HDFS, HIVE, Spark, PIG, Big Data in Cloud, Internet of Things (IoT) and other technologies in the Big Data ecosystem for unstructured data analytics.

From http://catalog.gmu.edu/
Rationale
This course is designed to provide an understanding of the impetus, value proposition, and trends of emerging technologies in the big data analytics field. The course will help students fully understand the usage, opportunities, and challenges of emerging big data technologies and their applications in decision making. Emphasis for the course will be placed on a real-world orientation through the analysis of case studies and hands-on activities that emphasize the importance of a comprehensive approach in big data analytics. The course will be in demand by students working in or interested in working with big data, which is one of the fastest growing specializations within IT. As part of this specialization, students need to learn how emerging big data analytics technologies are continuing to transform the IT industry, including their use within government programs.

Objectives
On successful completion of this course, students will be able to:

- Understand the importance of Data Science and Big Data Analytics in decision-making
- Perform analytics on unstructured data using Big Data technologies, including Hadoop, HDFS, Hive, Spark, PIG and others
- Use effectively advanced SQL functions for in-database analytics
- Use R for statistical analysis
- Understand Big Data security, and how Big Data is used in combination with Cloud and Internet of Things (IoT)

References

Textbooks and Reference Materials
There are no required textbooks for this course. Students are recommended to use the following books for reference:

- EMC Data Science and Big Data Analytics. GMU Library has electronic copy of the book [here](http://example.com). The book is also available on Amazon.


- Michael Manoochehri (2013). Data Just Right: Introduction to Large-Scale Data & Analytics ©2013 • Addison-Wesley Data & Analytics Series

Faculty and Staff
AIT 614, Course Coordinator:

Ioulia Rytikova, Ph.D.
Email:  

Email:  

Virtual Office hours:  

Administrative Support

Fairfax campus

Julija Shimkute
Engineering Building, 5400
Phone: 703-993-3565

Grading

Grades will be awarded in accordance with the Mason Grading System for graduate students. See the university catalog for policies: http://catalog.gmu.edu for more information.

The grading scale for this course is:

- 97 – above  A+  Passing
- 93 – 96%  A  Passing
- 90 – 92%  A-  Passing
- 87 – 89%  B+  Passing
- 83 – 86%  B  Passing
- 77 – 82%  B-  Passing
- 70 – 76%  C  Passing
- 0 – 69%  F  Failing

Raw scores may be adjusted by the Instructor to calculate final grades.

Students are responsible for checking the currency of their grade books. Grade discrepancies must be brought to instructor's attention within one week of assignment submission and 48 hours of exam submission.

Final grades will be determined based on the following components:

<table>
<thead>
<tr>
<th>Component</th>
<th>Weight</th>
</tr>
</thead>
<tbody>
<tr>
<td>Homework Assignments</td>
<td>25%</td>
</tr>
<tr>
<td>Discussion Board</td>
<td>5%</td>
</tr>
<tr>
<td>Labs</td>
<td>20%</td>
</tr>
<tr>
<td>Midterm</td>
<td>25%</td>
</tr>
<tr>
<td>Final Exam</td>
<td>25%</td>
</tr>
<tr>
<td>Optional Project</td>
<td>20%</td>
</tr>
</tbody>
</table>
These components are outlined in the following sections.

**Homework Assignments and Discussion Board**

Homework and Discussion Board will be assigned every class during the semester. Each homework assignment is to be prepared and submitted as specified by the Instructor. Late homework may not be accepted – if accepted, a penalty may be applied. Acceptance of late homework and/or application of penalties will be at the sole discretion of the Instructor.

**Labs**

Labs will be conducted in selected class sessions throughout the semester.

**Mid-term exam**

Midterm exam will be conducted during the scheduled class session 8 and will be based on topics addressed in Classes 1-7. Midterm exam will be “closed book, closed notes” with 1-page cheat sheet allowed.

**Final exam**

The final exam will be held during the scheduled final exam session (see [http://registrar.gmu.edu](http://registrar.gmu.edu)) and will be based on topics addressed throughout the entire course. The final exam will be “closed book, closed notes” with 1-page cheat sheet allowed. Final exams will be retained by the Department of Information Sciences and Technology and will not be returned to students.

**Optional Project**

There will be an optional project. It is designed for interesting students to earn extra credit (up to 20%) by applying one or many big data and data analytics technologies to a self-defined project. The technologies are NOT limited to those studied during the class. Students are encouraged to apply other technologies beyond the course content. The project will be graded based upon scope, complexity, quality and deliverable. Up to two (2) students can team up for the project, and both will earn equal credit. Students must submit a project proposal for review by Week 9, one week after the mid-term exam. The project deliverable must be submitted by the end of Week 15 one day after the final exam.

**NET Session: The midterm exam and the final exam will be online. Exams will be held during the scheduled exam sessions. If the student cannot attend an exam due to schedule conflicts, he or she must contact the instructor in advance to schedule an alternative date and time for the exam.**

No make-up for any activity including exams, unless arranged in advance. Only in special cases, such as medical problems and family emergency, make-ups and late assignments may
be allowed with verifiable proof. Arrive promptly to exams. Late students may not be admitted.

Final grades will be posted to PatriotWeb, which is the only vehicle for students to obtain those grades. A student with a "hold" on his/her PatriotWeb account will be unable to access final grades until the hold has been removed by the Registrar.

**Schedule**

A detailed schedule will be published on Blackboard. As many factors may affect the development and progress of a class, the instructor reserves the right to alter the schedule as may be required to assure attainment of course objectives. The schedule is subject to revision before and throughout the course.

Registered students should see the Blackboard Learning System for the latest class schedule.

**Important Dates**

Dates for dropping, adding the course etc. are available via http://registrar.gmu.edu/calendars/

**Religious Holidays**

A list of religious holidays is available on the University Life Calendar page. Any student whose religious observance conflicts with a scheduled course activity must contact the Instructor at least 2 weeks in advance of the conflict date in order to make alternative arrangements.

**Attendance Policy**

Students in in-class sections are expected to attend each class, to complete any required preparatory work and to participate actively in lectures, discussions and exercises. As members of the academic community, all students are expected to contribute regardless of their proficiency with the subject matter.

Students are expected to make prior arrangements with Instructor if they know in advance that they will miss any class and to consult with the Instructor as soon as possible if they miss any class without prior notice. Any student who expects to miss more than one class session is strongly advised to drop the course and take it in a later semester when he/she can attend every class.
Departmental policy requires students to take exams at the scheduled time and place, unless there are truly compelling circumstances supported by appropriate documentation. Except in such circumstances, failure to attend a scheduled exam will result in a score of zero (0) for that exam, in accordance with Mason policy on final exams. Students should not make travel plans or other discretionary arrangements that conflict with scheduled classes and/or exams. If the University is closed due to weather or other unforeseen conditions, final exams may be rescheduled – students are strongly advised not to make plans that would prevent them from attending exams that may be rescheduled during the entire exam period.

NET Version attendance: During each week the students must perform all the requirements published for that week. A detailed week-by-week schedule of classes will be published on the net version of the course.

**Communications**

Registered students will be given access to a section of the Blackboard Learning System for this course. Blackboard will be used as the primary mechanism (outside of lectures) to disseminate course information, including announcements, lecture slides, homework and other assignments, and scores for homework and exams.

Communication with the Instructor on issues relating to the individual student should be conducted using Mason email, via telephone, or in person - not in the public forums on Blackboard. Mason Mail is the preferred method – for urgent messages, you should also attempt to contact the Instructor via telephone. Federal privacy law and Mason policy require that any communication with a student related in any way to a student's status be conducted using secure Mason systems – if you use email to communicate with the Instructor you MUST send messages from your Mason email account.

**When sending an e-mail to the instructor, please include the following:**

- Course number
- Section number
- Your full name

**Privacy**

Instructors respect and protect the privacy of information related to individual students.

As described above, issues relating to an individual student will discussed via email, telephone or in person. Instructors will not discuss issues relating to an individual student with other students (or anyone without a need to know) without prior permission of the student.

Assessable work other than final exams will be returned to individual students directly by the Instructor (or by a faculty or staff member or a Teaching Assistant designated by the Instructor, or via another secure method). Under no circumstances will a student's graded work be returned to another student.

Faculty and staff will take care to protect the privacy of each student's scores and grades.
Disability Accommodations

The Office of Disability Services (ODS) works with disabled students to arrange for appropriate accommodations to ensure equal access to university services. Any student with a disability of any kind is strongly encouraged to register with ODS as soon as possible and take advantage of the services offered.

Accommodations for disabled students must be made in advance – ODS cannot assist students retroactively, and at least one week's notice is required for special accommodations related to exams. Any student who needs accommodation should contact the Instructor during the first week of the semester so the sufficient time is allowed to make arrangements.

Honor Code

All members of the Mason community are expected to uphold the principles of scholarly ethics. Similarly, graduating students are bound by the ethical requirements of the professional communities they join. The ethics requirements for some of the communities relevant to IST graduates are available via the following links:

- ACM Code of Ethics and Professional Conduct
- IEEE Code of Ethics
- EC-Council Code of Ethics

On admission to George Mason University, students agree to comply with the requirements of the Mason Honor System and Code. The Honor Code will be strictly enforced in this course. Honor Code cases are heard by a panel consisting of students – students who meet the requirements are encouraged to nominate themselves to serve on the Honor Committee.

Any use of the words or ideas of another person(s), without explicit attribution that clearly identifies the material used and its source in an appropriate manner, is plagiarism and will not be tolerated. The Instructor reserves the right to use manual and/or automated means (including such services as SafeAssign.com) to detect plagiarism in any work submitted by students for this course, and to direct Teaching Assistants and/or other faculty and/or staff members to do likewise in support of this course.

For this course, the following requirements are specified:

- All assessable work is to be prepared by the individual student, unless the Instructor explicitly directs otherwise.
- All work must be newly created by the individual student for this course for this semester. Any usage of work developed for another course, or for this course in a prior semester, is strictly prohibited without prior approval from the instructor.

Students may seek assistance with assigned work (and are encouraged to do so if they feel the need), provided:
- The directions for the assigned work do not prohibit such assistance.
- Such assistance is acknowledged in the submitted work, clearly identifying the person(s) giving assistance and the nature of the assistance given.
- Any work to be submitted is prepared entirely an exclusively by the student submitting it. Students are expressly prohibited from sharing any assessable work for this course in any manner with other students (except students assigned as Teaching Assistants to this course and the student's section), unless all students involved have had their work graded and returned by the Instructor, or the Instructor has explicitly approved such sharing.

NET Sections only:

For the net section you will need to have the following additional resources:

- Computer with fast internet connection
- Microphone and web camera