Tennessee Tech University
Department of Electrical and Computer Engineering

ECE 7970: Big Data and Artificial Intelligence---A Data-Centric Approach
Fall 2017

(Lec. 3, 3 credits)

Dates: Tue and Thur 8:00-9:20 AM

Classroom: Brown Hall 208

Instructor: Robert Qiu, BN 210 (located right across the ECE office)
Phone Ext. 3847 Email: rqiu@tntech.edu

Office hours: Tue and Thur 8:30-12:00 AM

Prerequisites (if applicable)**:** Prerequisite: ECE 3010 Co-requisite: ECE 3910

Prerequisites by Topic:
1. Fourier transforms
2. Random Variables

Texts and References:

Required:

1. Avrim Blum, John Hopcroft, and Ravindran Kannan, Foundations of Data Science, June 2017.
2. R. C. Qiu, Cognitive Radio Networking and Sensing, John Wiley, 2012.
3. R. C. Qiu and M. Wicks, Cognitive Networked Sensing and Big Data, Springer, 2014.
4. R. C. Qiu and P. Antonik, Smart Grid using Big Data Analytics, John Wiley, 2017.
5. Alex Graves, Supervised Sequence Labelling with Recurrent Neural Networks, Springer, 2012.

3. Additional reading will be provided by the instructor.

References (if applicable): MATLAB User's Manual.

Assignments:
All assignments are due *at the start of class one week after the day they are assigned* unless otherwise noted. Late assignments will be only accepted if you know you will be absent for a legitimate reason and let your instructor know ahead of time.

Attendance Policy: Class attendance is required. A failing grade (or withdrawn in the case of auditors) may be assigned for excessive absences. If a class is missed, it is your responsibility to obtain class notes and any assignments, handouts, or other materials, and to make up the necessary work. Quizzes will usually be announced in advance, and tests will be announced at least one week before the date of the test. Quizzes and tests will not be made-up unless the absence is due to unavoidable, verifiable, and exceptional circumstances.

Course Description: Big Data and Artificial Intelligence is a course for the graduate program. This course is critical to the curriculum, due to the following. (1) Big data systems are found ubiquitous in engineering and science. The analysis of big data signals measured in these linear systems is the butter and bread of daily life, e.g., in wireless industry, digital health, Internet, etc. It is very fortunate and surprising that this simple, ideal model (high-dimensional space) is sufficient for almost every problem. (2) Deep learning takes advantage of non-linear information from big data. It has become a new standard for most big data problems. (3) Familiarity with the latest research literature is critical to the graduate program.

Learning Objectives:

Big Data and Artificial Intelligence was designed with a number of learning objectives in mind. Some focus on tangible knowledge and skills while others are more abstract and long-term. All learning activities in the course rely and/or build upon one or more these objectives. The overriding goal of this course is that you better *understand* how to think like an engineer or a scientist. As a result, at the end of the semester you will be able to do the following:

- Analyze the big data systems using statistical machine learning algorithms including deep learning. MATLAB commands are developed based on this kind of analysis.
- Employ the analytical skills to solve real-life engineering and scientific problems that are be needed in your future's employers.
- Cultivate an in-depth understanding of the foundational principles of advanced problem-solving in engineering/STEM.
- Develop a toolkit of analytical practices related to major issues in STEM research.
- Conduct knowledge discovery processes in relation to large STEM data sets in multiple professional settings.

Assessment of Learning

Learning is hard! Meaningful learning—the kind of learning that last well beyond the test—is really hard. You will have to struggle through abstract concepts, grasp mathematical tricks, and continually practice the skills you learn. At times this will be frustrating, but the more you engage, the more you will learn. The essential part of learning Fourier transform is like your

learning trigonometry in high schools. A lot of exercises will sharpen your intuition. Fundamentally, this is a like calculus.

Throughout the course, you will multiple opportunities to explore a variety of signal processing problems.

Major Teaching Methods: (e.g. lectures, MATLAB simulations, demonstrations, discussion, reading, or written assignments, etc.) This is a problem-based course. By solving real-life problems, we close the gap between the textbook and the real life.

Special Instructional Platform/Materials: (**e.g.** iLearn, laptop, etc.) none

Topics to Be Covered:

- High-dimensional space
- Best-Fit Subspaces and Singular Value Decomposition (SVD)
- Machine Learning
- Algorithms for Massive Data Problems: Streaming, Sketching, and
- Sampling
- Deep Learning

Grading:	Homework Presentations:	75%
	Projects:	25%

University Plagiarism Policy (Tennessee Tech University Student Handbook – Plagiarism (Academic Regulations))**:** When you use (for example, quote or even summarize or paraphrase) someone else's media, words, data, ideas, or other works, you must cite your source. You should be especially careful to avoid plagiarizing Internet sources (for example, e-mail, chat rooms, Web sites, or discussion groups). It does not matter whether you borrow material from print sources, from the Internet, from on-line data bases, or from interviews. Failure to cite your source is plagiarism. Students who plagiarize may receive an "F" or a "0" for the assignment, or an "F" for the course. http://www.tntech.edu/ttustudenthandbook/academic-regulations/

Academic Integrity and Ethics

It is unethical to steal somebody else's work. It is also unethical to represent anybody else's work as your own. Plagiarism, cheating, and other forms of academic dishonesty are prohibited. Plagiarism includes *copying homework* and *providing homework for copying*. Violation of this policy will result in a grade of F being assigned for the course and a notation on your permanent academic record.

Disability Accommodation: Students with a disability requiring accommodations should contact the Office of Disability Services (ODS). An Accommodation Request (AR) should be completed as soon as possible, preferably by the end of the first week of the course. The ODS is located in the Roaden University Center, Room 112; phone 372-6119. **(**Disability Accommodation Policy and Procedures - Tennessee Tech University Faculty Handbook and Student Handbook http://www.tntech.edu/facultyhandbook/diabilityaccom/)

Marking Scheme
To be counted as correct, all answers must be supported by appropriate computation or explanation. Incorrect methods leading to a correct result will not receive credit. The grading scale for the course is tentatively as follows:

A	90% - 100%
B	80% - 89%
C	70% - 79%
D	60% - 69%
F	< 60%

Note that marks may be adjusted depending upon overall class performance. If you believe that there may have been an error made in grading your work, please submit the work for re-grading with an appropriate short note detailing your concern.

Keep in mind, your final grade is a reflection of how well you meet the learning objectives, not a comment on effort or self-worth.

Special Circumstances
Please inform the instructor of any special conditions or circumstances which may affect or occur during the course time frame (e.g., medical emergencies, family circumstances, etc)

Courtesy/Conduct
The classroom is a place for learning so please treat everyone with respect. Various electronic/communication devices (e.g., cell-phones, pagers, PDA's, radios, MP3 players, laptops, etc) should be turned off during class. No eating is allowed in the classroom.

Submissions
Please use standard size (8.5" x 11") paper with cut edges. Highlight all answers via underlines and/or by using boxes. Staple the pages together, number all pages, and write your name on the top of each sheet.

Calculators
The use of electronic calculators is generally permitted.

Hands-on Projects

It is critical to get hands dirty in this course. Big data analytics are about extracting relevant "useful" information from experimental datasets. The essential significance of this course is to learn how to perform this function from a real-life job environment. You are given the datasets in some domains (e.g., medical, financial, engineering); your manager asks you to do something with them. Very often, you will find that your first thing is to take a look at the spectral properties of the datasets. This practice goes back to the age of Isaac Newton.

The high-dimensional space is essential in the whole course. Algorithms are built upon this. Our job is how to apply these algorithms in real-life applications. These hands-on experiments essentially motivate us to understand the mathematical steps behind the algorithms.

I will present some worked-out examples to demonstrate the essential procedures. Different teams will be asked to work on some other datasets and report back to the class. This is the common in a job environment. Problem-solving is the best way for learning. You will understand the theory if you find it by your own doing.