CSC 362 – DATA ANALYTICS I

CREDIT HOURS: 3
PREREQUISITES: CSC 302, 340 and MTH 220
GRADE REMINDER: Must have a grade of C or better in each prerequisite course.

CATALOG DESCRIPTION

Introduction to the study of data analytics including programming for problem solving, ethics, data science process, statistical methods, and machine learning techniques.

PURPOSE OF COURSE

The purpose of this course is to provide a broad knowledge of the fundamental concepts of data analytics. This knowledge will enable the student to apply the data analytics process to problems involving data. Students will acquire a knowledge of data analytics and understand its use in data based decision making.

EDUCATIONAL OBJECTIVES

Upon successful completion of the course, students should be able to:

1. Demonstrate a broad knowledge of the fundamental concepts of data analytics.

2. Describe the main issues of the data science process.

3. Identify current trends in data analytics.

4. Demonstrate a knowledge of data representation and visualization.

COURSE CALENDAR

This course meets for a minimum of 37.5 lecture contact hours during the semester, including the final exam. Students have significant weekly reading assignments. Students are required to complete a major project and make at least one class presentation, weekly homework/programming assignments, and 2-3 periodic exams in addition to the final exam. Students are expected to prepare for any class assignments or quizzes over the material covered in class or in the reading material. Successful completion of these activities requires at a minimum six additional hours of outside of classroom work each week.

CONTENT   Hours

Introduction to Data Analytics..3

Programming for problem solving..12

Ethics in data analytics...3

Data Science process...6
    Data pipeline
    Understanding data quality
    Data pre-processing: data munging, wrangling, cleaning
    Introduction of feature selection
    Data analysis and interpretation

                                         TOTAL   45

REFERENCES

James et al. (2013) *An Introduction to Statistical Learning: With applications in R* . Springer.

Zumel, N. and Mount, J. (2014). *Practical Data Science with R*. Manning Publications.

Lantz, B. (2013). *Machine Learning with R*. Packt Publishing.

Leskovec, J., Rajaraman, A. and Ullman, J. (2011). *Mining of Massive Datasets* . Cambridge University Press.

Zafarani, R., Abbasi, M. A. and Liu, H. (2014) *Social Media Mining: An introduction* . Cambridge University Press.

SQL The Complete Reference, 3rd Edition / Edition 3 by James R Groff, Paul N. Weinberg, Andy Oppel