

Syllabus

BUAN 5510: Capstone Project in Business Analytics

Summer, 2020

I. Instructor

Ben Kim, Ph.D.

Professor of Information Systems

Mailing Address:

Department of Management, Albers School of Business and Economics, Seattle University,
901 12th Ave., Seattle, WA, 98122, USA

Phone: 206-296-2806, E-mail: bkim@seattleu.edu

Office: Pigott 417

Office Hours: By appointment

II. Class

Time: 06/22/20-08/15/20 TF, 05:30PM-09:00PM and online classes

Room: online

III. Course Description

The Capstone is an application of data analytics in the planning and execution of a one-quarter-long development project. Students work in teams to define and carry out an analytics project from initial requirements analysis to final analysis/recommendation. Primary tasks include an identification of problems, datasets, data preprocessing, building data mining/statistical models, and validation. This activity will culminate in a formal presentation of results at the end of the quarter.

This course is designed to offer students the opportunity to apply the principles, concepts, and skills learned in prior classes to actual business situations. Students will be expected to use their knowledge and judgment to make decisions and recommendations concerning the business requirements and objectives that they will investigate. Instructors primarily function a coach rather than a lecturer.

During the quarter, students will need to utilize a variety of communication techniques: brainstorming, encoding, decoding, role-playing, decision-making, problem solving, library research, report writing, and more.

The overall objectives of the course are as follows:

- Provide a foundation for understanding the processes and practices of business analytics, which will assist students in their future career positions.
- Understand and develop the ability to deal with the array of critical decisions facing the management of companies engaged in business analytics.
- Develop a working familiarity with the availability and reliability of data, both qualitative and quantitative, relating to business requirements.

IV. Readings

- Data Mining: Concepts and Techniques, 3rd Edition by Han, Kamber, and Pei, Morgan Kaufmann, ISBN: 9780123814791
- Lecture Notes – Available on Canvas.

V. Software Tools

For your analysis, we will use Python with relevant libraries.

VI. Learning Objectives

- Planning and execution of data analytics
- Understanding how data mining projects, including data extraction, transformation, and loading (ETL), can be applied to solve real world problems
- Learning how to apply machine learning algorithms to large datasets for data mining
- Understanding how data mining or big data projects assist client organizations to achieve their goals
- Learning how to identify, measure, interpret, and incorporate relevant information in analyzing problems and making effective business decisions
- Demonstration of an ability to write a data analytics consulting report, make recommendations, develop conclusions from the research, and make substantial recommendations to the client organizations.
- Demonstration of an ability to orally present a consulting report to an executive audience

VII. Class Procedures and Activities

This course will be a self-directed class wherein student responsibilities will be of a different scope and nature than in the normal classroom environment. The teamwork that is required will also differ from previous team project experience, in that the projects are actual situations requiring decisions, recommendations, and action. Considerable interaction will be required between the student teams and participating companies or an instructor. Insofar as possible, students will be given a choice as to the datasets or the company which they will work with. The class will be organized into teams of four students each. Student teams will work together with a participating local company in order to accomplish a project or analyze the publicly available datasets.

Materials covered in class during the quarter will be directly and systematically linked to the projects that are being performed. Explanation and rationale for the techniques and skills that will be required to complete the projects successfully will be explained sequentially. Lectures, if any, will focus on subjects directly relevant to the objectives of student projects, and will use these projects as illustrative "live" case studies for class discussion.

Schedules permitting, executives of participating firms will be invited to class to describe their companies' products/services and industry and to discuss, informally, their business plans, objectives, activities and experiences, as well as the issues which affect the choices and decisions which they face. Consultants, government sources, and others may periodically discuss with you their perspectives and resources.

VIII. Major Tasks

- Identification of Datasets
- Problem Statement
- Literature Review
- Data Pre-Processing
 - Data Cleaning
 - Data Integration
 - Data Reduction
 - Data Transformation and Discretization
 - PCA (Primary Component Analysis) for Dimensionality Reduction
- Building Models
 - Bayesian Classifications
 - Decision Trees/Random Forests
 - Market Basket Analysis
 - Neural Networks
 - Support Vector Machines (SVM)
 - Gradient Boosting
 - Clustering
- Evaluation of Models
 - Numerical predictions
 - Classifications
- Interpretation of Models and Recommendations

Research Paper

One of our final goals is to write an article and have it published in a research journal or conference proceedings. This class is a joint research project with students and the instructor.

IX. Grading

1. Weight of Each Requirement

Individual/Group Assignments/Presentations:	30 %
Final Presentation/Paper:	60 %
Professionalism	10 %

- Oral and Written reports can be adjusted by your individual scores on peer evaluations. In other words, the team could score well, but if you received poor evaluations, your grade could be severely impacted. This is designed to make sure the team works well together and to minimize the free-rider syndrome, among other issues.

X. Academic Honesty

Seattle University is committed to the principle that academic honesty and integrity are important values in the educational process. Academic dishonesty in any form is a serious offense against the academic community. Acts of academic dishonesty will be addressed according to the Seattle University Academic Honesty Policy. The policy can be found at the address below:

http://www.seattleu.edu/regis/Policies/Policy_2004-01.htm

If you are not sure whether a particular action is acceptable according to the Academic Honesty Policy, you should check with your instructor before engaging in it.

XI. Disability

If you have, or think you may have, a disability (including an ‘invisible disability’ such as a learning disability, a chronic health problem, or a mental health condition) that interferes with your performance as a student in this class, you are encouraged to arrange support services and/or accommodations through Disabilities Services staff in the Learning Center, Loyola 100, (206) 296-5740. Disability-based adjustments to course expectations can be arranged only through this process.

Tentative Class Schedule¹

Week 1: (6/23/2020)	Orientation – Review of class objectives <ul style="list-style-type: none">• Project descriptions, guidelines and expectations.• Introductions of class members, team formation, project selection.• Identification of datasets and problem statement• Introduction: Chapter 1 (Han et al.)
Week 2: (6/30/2020)	Data Types and Data Preprocessing: Chapter 2 and 3 (Han et al.) Team Meetings – Literature Review and Problem Identification
Week 3: (7/7/2020)	Data Preprocessing - data cleaning, integration, correlation, attributes reduction, transformation, discretization, primary component analysis (Chapter 3)
Week 4: (7/14/2020)	Building data mining models
Week 5: (7/21/2020)	Building data mining models
Week 6: (7/28/2020)	Building and Evaluation of data mining models - Classification and/or Numerical Predictions, Presentations
Week 7: (8/4/2020)	Interpretations of the Data Mining Models and Discussion
Week 8: (8/11/2020)	Final Presentations/Final Paper Due

¹ Depending on the progress of the projects, the schedule can be changed.