# MGS 662: Machine Learning for IT Managers

Section: XXX
Class Hours: M-W: 12:30-1:50 pm, YYY
Office Hours: Tues: 11:30 am -1:00 pm

Instructor: Dr. Haimonti Dutta
Office: 325P Jacobs
Email: haimonti@buffalo.edu

## Course Description

Investment in government and business infrastructure has lead to the accumulation of vast amounts of data in recent years. This course will discuss how techniques from convex optimization can be used to extract useful knowledge and business value from the data collected. It introduces students to the theory of convex optimization of relevance to managerial decision making and machine learning. Topics include convex sets and functions, formulation of convex optimization problems, and convex optimization algorithms including gradient, sub-gradient, proximal and interior point methods. Numerous examples will be chosen from machine learning problems including classification, regression and clustering. Students will have hands on experience with the R programming language and optimization packages including MOSEK. Real world examples and case studies from text mining, medical applications, fraud detection, finance, and social networks will be examined.

## Text Books and Recommended Reading

- **Text:** Stephen Boyd and Lieven Vandenberghe, "Convex Optimization". Available Online from `http://web.stanford.edu/ boyd/cvxbook/`
- **Text:** Dimitri Bertsekas, "Convex Optimization Algorithms", Athena Scientific, 2015.
- **Recommended:** Foster Provost and Tom Fawcett, "Data Science for Business: What you need to know about data mining and data analytic thinking.", O'Reily, First Edition, 2013.
- **Recommended:** Trevor Hastie, Robert Tibshirani, and Jerome Freidman, "The Elements of Statistical Learning: Data Mining, Inference and Prediction", Springer.
  Available online: `http://web.stanford.edu/hastie/local.ftp/Springer/OLD/ESLII_print4.pdf`
- **Recommended:** Tom Mitchell, "Machine Learning", McGraw Hill, 1997.

## Pre-requisities

Students are expected to have a good background in linear algebra and applied probability. Ability to program (prior experience) is expected. Familiarity with R will be useful.

## Websites

Course will be available on UBlearns. Please check regularly for updates to schedule, class presentations, reading materials and other discussions. Occasionally email messages to the whole class will be sent through UBlearns. In the event a class is missed (for example due to bad weather conditions, etc.) it will be made up based on the instructor's discretion.

## Learning Outcomes

| By the end of the course, you should be able to $\cdots$ | Method of assessment |
|---|---|
| Basic Programming in R | Assignment 1 |
| Learn about Convex Sets, Functions | Quiz 1 |
| Understand Basic Convex Optimization<br>Understand Convex Opt. Algorithms | Quiz 1<br>Quiz 2 and 3 |
| Learn use of optimization packages in R (MOSEK) | Assignment 2 |

## Course Outline

[The topics in bold reflect the new material for the 3 credit course.]

- Introduction to Machine Learning
- Programming in R
  - Convex Optimization Package MOSEK in R
- Convex optimization: An Introduction
- Convex sets and functions.
- Convex optimization problems
  - Formulating optimization problems for Machine Learning
  - Optimizing an objective function
  - Estimating loss and different kinds of loss functions
- **Convex optimization algorithms: An Overview**
  - Iterative Descent Algorithms
  - Approximation Methods
- **Sub-gradient Methods**
- **Polyhedral Approximation Methods**
  - Cutting Plane Methods

  - Simple Decomposition Techniques
- **Proximal Methods**
- Real World Use Cases from Machine Learning
  - Regression: Linear, Shrinkage Methods and Splines
  - Clustering: K-means, Hierarchical and EM
- Hands-on programming with R and R-MOSEK.

## Grading Policy

| Component | Percentage |
|---|---|
| Assignment | $(20\% + 15\%) = 35\%$ |
| Quiz | $(2 \times 30\%) = 60\%$ |
| Class Participation | 5% |

The total score from all the components of the course will be estimated. This will be subjected to a bell curve grading to obtain the letter grades. The quality points for each letter grade are illustrated below.

## Graduate Grade Options

| Grade | Quality Points |
|---|---|
| A | 4.0 |
| A- | 3.67 |
| B+ | 3.33 |
| B | 3.0 |
| B- | 2.67 |
| C+ | 2.33 |
| C | 2.00 |
| C- | 1.67 |
| D+ | 1.33 |
| D | 1.00 |
| F | 0.00 |

## Academic Integrity

Academic integrity is a fundamental university value. Through the honest completion of academic work, students sustain the integrity of the university and of themselves while facilitating the university's imperative for the transmission of knowledge and culture based upon the generation of new and innovative ideas.

As clearly stated in the student handbooks (MS, MBA, and Ph.D.) and emphasized in the university's integrity policy (`http://grad.buffalo.edu/study/progress/policylibrary.html#preamble`), faculty, students, and staff all have an obligation to each other to maintain and expect high standards of integrity. The reputation of the School of Management is derived from the performance of all its members, and faculty, students, and staff all have an obligation to be aware of their own rights and responsibilities and make every effort to maintain high standards of academic integrity and honesty.

## Accessibility Resources

If you have any disability which requires reasonable accommodations to enable you to participate in this course, please contact the Office of Accessibility Resources, 25 Capen Hall, 645-2608, and also the instructor of this course during the first week of class. The office will provide you with information and review appropriate arrangements for reasonable accommodations, which can be found on the web at: `http://www.ub-disability.buffalo.edu`.

## Tentative Course Schedule

| Week No | Readings | Assignment | Quiz |
|---|---|---|---|
| 1 | Intro to R Programming | Assgn 1 | |
| 2 | Machine Learning: An Introduction | | Quiz 1 |
| 3 | Intro to Convex Opt. | | |
| 4 | Convex sets | | |
| 5 | Convex functions | | |
| 6, 7 | Convex optimization | Assgn 2 | Quiz 2 |
| 8, 9 | Convex optimization algorithms: An introduction | | |
| 10 | Sub-gradient Methods | | |
| 11 | Polyhedral Approximation Methods | | |
| 12 | Proximal Methods | | Quiz 3 |
| 13 - 15 | Machine Learning Use Cases | | |

## Workload

1. There will be two assignments which will comprise 35% of the overall grade. These are intended to provide "hands-on" experience when dealing with real world problems.They will involve coding in R and use of R packages including MOSEK. The students(s) are expected to analyze data and present results either by using techniques learnt in class or custom designing them. All materials should be submitted through UBlearns. Assignment should be done individually by each student enrolled in the class.

2. There will be three quizzes during the course. Each quiz will be in-class, and the dates will be announced earlier. They will have multiple choice questions and/or short answer type questions or problems. Materials for quizzes will be non-inclusive – Quiz 2 will not include materials from Quiz 1 and so on. No make-up homework(s) or assignments can be given in case a quiz is not taken in class. In addition, no arrangements will be made for taking quizzes

separately in case of a failure to attend the class on the designated day of the exam, except in case of untoward circumstances (such as bad weather conditions, etc.). Failure to report for a quiz on time (a delay of more than 5 minutes) will result in a missing grade and no make-up exam will be provided for the same.

3. Of the three quizzes, the best two will be used for the purpose of grading. Students, however, must take all three quizzes to ensure they have an understanding of the entire syllabus that will be covered in class.

4. The students are required to actively participate in class. In addition, they are responsible for participating in the course evaluation process. Both of these components will contribute to the 5% of class participation grade.

5. Plagiarism policy: We strictly abide by the policy as described in the UB handbook for undergraduate, MBA, and PhD students. No exceptions! The UB plagiarism software on UBlearns will be used to verify your work – it has a high reliability 99% and students are expected to abide by the rules and write their own reports and solutions.

6. Dates for submission of assignments and quizzes can be found from the schedule on UBlearns. All materials are due at midnight. Late submissions are subject to 10% deductions in score.

7. To request for an extension on a deadline, please contact the TA by email. However, the decision to grant such an extension is left to the discretion of the TA and instructor.

## MS in MIS Program Goals

| | Program Goals | Measurable Tasks and Criteria |
|---|---|---|
| Application of MIS knowledge to to produce effective technical solutions/plan | Students will apply knowledge of MIS score $> 25$ designs and solutions for specific problems | Assignment |
| Awareness and Use of cutting-edge technology | Students will use new and emerging concepts and applications in proposing and creating IT solutions | Any one of the following: Qz 1, 2 or 3 score$>$ 20 Assignment score$>$ 20 |
| Statistical/Business Analytics | Applications of statistical methods for business decision making | Qz 1, 2 or 3 score$>$ 20 |