**Syllabus for MSBA 5510**

*Instructor:* Daqing Zhao, Yung-Jae Lee

*Office:*

*Email:* [daqing.zhao@dominican.edu](mailto:daqing.zhao@dominican.edu) and [yung-jae.lee@dominican.edu](mailto:yung-jae.lee@dominican.edu)

*Cell:* (650) 255-7360 (Zhao)

*Office :* (415) 458-3786 (Lee)

*Office hour:* By Zoom appointment

### Course Objectives

(i) Work in a team on a self-selected project of data analysis and model building; participate in all phases of the project from formulating objectives of the data analysis, exploratory data analysis, model building, to results presentation.

(ii) Use knowledge and techniques learned during the past year, apply to a realistic data analysis project.

(iii) Learn teamwork skills.

### Required Course Materials

*Textbooks: Free online version, hard copy optional*

*An Introduction to Statistical Learning*

*with Applications in R (ISLR)*

by Gareth James, Daniela Witten, Trevor Hastie and Robert Tibshirani

<http://www-bcf.usc.edu/~gareth/ISL/>

Mining of Massive Datasets (MMDS)

by [Jure Leskovec](http://cs.stanford.edu/~jure/), [Anand Rajaraman](https://twitter.com/anand_raj), [Jeff Ullman](http://infolab.stanford.edu/~ullman/)

<http://i.stanford.edu/~ullman/mmds/book0n.pdf>

Required software, R, MySQL, Anaconda (Python), TensorFlow (keras), H2O.ai

**Course weighs and assessments**

|  |  |
| --- | --- |
| Class Participation: | 10% |
| Weekly Activity Report: | 30% |
| Peer Evaluations: | 20% |
| Project Team Presentations: | 40% |
| Total | 100% |

*Group Project*

The group capstone project will be selected by group members with consultation and approval of a professor. Each group selects one project. Projects can be from work, partner companies, Kaggle competitions, or open data sources.

*Class Participation*

You are expected to actively and constructively participate in class. Reading assignments are provided in the course outline (at the end of the syllabus), and you are expected to read these materials before class to facilitate the understanding of the course materials and in-class discussions. Class participation points will be awarded for insightful concisely stated and clearly justified comments during case and project presentations.

*Weekly Activity Report*

You are expected to write activity reports and share the Google document with the professor as well as your team members. The reports should be simple descriptions of your activity, such as,

*“7/10/2023, I did 3 hours of literature search about predictive modeling for fair lending”;*

*“7/11/2023, I analyzed the data set for 4 hours and found an error in the data in column X”;*

…

You can summarize what you find or have done, if you want. This should take only a few minutes

of your time each day. Keeping a record of what you have done is a good habit to have and we

will use the record to assess your contributions to the project. You will be graded on it by your

team peers and the professor.

*Project Presentation*

Each group is required to make a 20-minute presentation. Professors and classmates and audience will ask questions and make comments.

*Grading*

Each team has a single presentation score. Each member has own class participation score and activity reporting score. Each member is evaluated by other team members on contribution and teamwork.

**Grading Scheme**

|  |  |
| --- | --- |
| **A ≥ 92 B- [77, 80)**  **A- [89, 92) C+ [74, 77)**  **B+ [85, 89) C [70, 74)**  **B [80, 85) F < 70** |  |

**Academic Honesty Expectations**

You may use any source you find useful in doing assignments, but you must cite your sources in writing.   Plagiarism is the copying or paraphrasing of any work from another source without proper written acknowledgement.  All members of a group doing a project are expected to assure that all sources are cited.

All students are bound by the Academic Honor Code, set out in the Dominican University Graduate and Professional Student Handbook that is cited in the Student Guide that you received during admission. The code, available online, sets out the requirements of academic honesty and the means by which alleged violations are administered. Please familiarize yourself with its provisions.

**Student Disability Services**

Student Disability Services extends reasonable and appropriate accommodations that take into account the context of the course and its essential elements, for individuals with qualifying disabilities. Students with disabilities are encouraged to contact the Student Disability Services Coordinator to set up a confidential appointment to discuss accommodation guidelines and available services. Additional information regarding the services available may be found at the Dominican University following address on the website: xxxx

**Class Attendance**

Attendance at all meetings of this course is expected. If you are not going to be in attendance it is your obligation to let the instructor know beforehand (except in cases where that is not possible). Two or more absences mean that you may not receive credit for the course. Hence, you **must** discuss with the instructor your lapses in attendance.

**Reminder**

Weekly lecture notes will be posted on the course website before each class. Please print a hard copy of the lecture notes and bring along to class for your own reference. No copies will be provided in class.

### Course Outline (Subject to change)

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Week** | **Date** | **Topic** | **Textbook Chapters** | **HW Due** |
| 1 | Wed 06/28 | Discuss and select project. Project must have a statistics or machine learning model.  Elect project leader, presentation leader, and technical leader.  Decide on project objectives. | Handout | Project selection and leaders |
| 2 | Sat 07/08  Wed 07/12 | Initial analysis plan and model design.  Data processing. |  | Project Plan |
| 3 | Wed 07/19  Sat 07/22 | Exploratory data analysis.  Refinement of model design. |  | Project Update |
| 4 | Wed 07/26  Sat 07/29 | Data preparation/feature engineering for modeling.  Initial model building |  | Project Update |
| 5 | Wed 08/02 | Model building and refinement |  | Model Results |
| 6 | Wed 08/9 | Insight synthesis and preparation of presentation |  | Outline of Presentation |
| 7 | Sat 08/12 | Final Project présentations. |  | Team Presentation; Code for Project |