Department of Technology Management and Innovation  
MG-GY 9753 - Machine Learning  
Fall 2019

**Professor:**  Hassane Kone, Ph.D., Adjunct Instructor

**Contact Details:**  hkk259@nyu.edu

**Office/Hours:**  By Appointment  
Rogers Hall, 303 D

**Class Schedule:**  Thursday: 6:00pm-8:30pm in JABS, Room 773

**Course Pre-requisites:**  Calculus, linear algebra, probability, statistics

**Course Description:**

Machine learning is about extracting or discovering knowledge from data. This course will cover fundamental machine learning algorithms used to understand business situations and improve business decisions. In machine learning, there are three types of commonly used algorithms: supervised (predictive), unsupervised (descriptive) and reinforcement learning algorithms. In the first part of the course, we will focus on supervised learning algorithms including K-Nearest Neighbors, Linear Regression, Logistic Regression, Decision Tree, Support Vector Machine (SVM), Naive Bayes, bagging and boosting algorithms. The second part of the course will cover unsupervised algorithms including K-means clustering and dimensionality reduction. The last part of this course will cover Reinforcement learning algorithms, especially Markov Decision Process. We will use python as our main programming language.

**Readings:**

- **Required Text(s):** n/a

- **Optional Text(s):**


Software:

This course will require the use of Python, mainly Numpy, pandas, matplotlib, seaborn and scikit-learn.

Course Assignments and Grading:

We will have three sets of homework covering basic statistics, data mining and machine learning.

**Homework (weekly codes): 20% Exams 1 & Exam 2: 40% Term project: 20% Final: 20%

Each weekly homework will be about writing model seen in class. For the purposes of computing GPAs, the following schedule is used:

<table>
<thead>
<tr>
<th>Letter Grade</th>
<th>Score Range</th>
<th>Grade Point</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>[95, 100]</td>
<td>4.0</td>
</tr>
<tr>
<td>A-</td>
<td>[90,95)</td>
<td>3.7</td>
</tr>
<tr>
<td>B+</td>
<td>[85, 90)</td>
<td>3.3</td>
</tr>
<tr>
<td>B</td>
<td>[80,85)</td>
<td>3.0</td>
</tr>
<tr>
<td>B-</td>
<td>[75,80)</td>
<td>2.7</td>
</tr>
<tr>
<td>C+</td>
<td>[70,75)</td>
<td>2.3</td>
</tr>
<tr>
<td>C</td>
<td>[65,70)</td>
<td>2.0</td>
</tr>
<tr>
<td>F</td>
<td>[0, 65)</td>
<td>0</td>
</tr>
</tbody>
</table>

Course Webpage:

All announcements and class-related documents (supplemental and suggested readings, discussion questions, etc.) will be posted on NYU Classes site (http://classes.nyu.edu/), where you must have access.

Note that some class announcements will be distributed via NYU e-mail. Thus, it is important that you actively use your NYU e-mail account, or have appropriate forwarding setup on NYU Home (https://home.nyu.edu/).
Course Topic Outline

<table>
<thead>
<tr>
<th>Lectures day</th>
<th>Lecture Topic</th>
<th>Math topics to review</th>
<th>Assignment</th>
</tr>
</thead>
<tbody>
<tr>
<td>Thursday, Sept. 6</td>
<td>Intro to Machine Learning, Installation of Jupyter Notebook, Numpy, scipy, matplotlib, pandas, scikit-learn</td>
<td>Matrix Algebra, Optimization of functions with several variables</td>
<td>Python bootcamp practice questions</td>
</tr>
<tr>
<td>Thursday, Sept. 13</td>
<td>Supervised Learning, Classification &amp; Regression</td>
<td>Inferential Statistics</td>
<td>See NYU-classes</td>
</tr>
<tr>
<td>Thursday, Sept. 20</td>
<td>Linear models: Linear Regression, Logistic Regression, Linear SVC</td>
<td>Correlation, Odds, Logit function</td>
<td>See NYU-classes</td>
</tr>
<tr>
<td>Thursday, Sept. 27</td>
<td>K-Nearest Neighbors</td>
<td>Distances and Metrics</td>
<td>See NYU-classes</td>
</tr>
<tr>
<td>Thursday, Oct. 4</td>
<td>Naïve Bayes classifier, Decision Trees, Ensemble Trees, Random forest, Gradient boosted regression trees</td>
<td>Bayes' Theorem and Conditional Probability</td>
<td>See NYU-classes</td>
</tr>
<tr>
<td>Thursday, Oct. 11</td>
<td>Kernelized Support Vector Machine, Neural Networks</td>
<td>Polynomial functions, Matrix Algebra</td>
<td>See NYU-classes</td>
</tr>
<tr>
<td>Thursday, Oct. 18</td>
<td>Exam 1 (covers supervised learning)</td>
<td>Descriptive statistics, Outliers, z-scores</td>
<td>See NYU-classes</td>
</tr>
<tr>
<td>Thursday, Oct. 25</td>
<td>Unsupervised learning, Preprocessing and scaling</td>
<td>Linear Algebra, Center of mass, Centroids, Conditional probability</td>
<td>See NYU-classes</td>
</tr>
<tr>
<td>Thursday, Nov. 1</td>
<td>Dimensionality reduction, Clustering (K-means), Associations</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Thursday, Nov. 8</td>
<td>Reinforcement learning, Markov Decision Process</td>
<td>Markov Chains, Probability</td>
<td>*Project Proposal due</td>
</tr>
<tr>
<td>Thursday, Nov. 15</td>
<td>Model evaluation and improvement, cross validation, evaluation metrics and scoring</td>
<td>Statistics</td>
<td>See NYU-classes</td>
</tr>
<tr>
<td>Thursday, Nov. 22</td>
<td>Thanksgiving Recess</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Thursday, Nov. 29</td>
<td>Exam 2 (covers unsupervised and reinforcement learning)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Thursday, Dec. 6</td>
<td>Review</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Thursday, Dec. 13</td>
<td>Project Presentation</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Thursday, Dec. 20</td>
<td>Final Exam</td>
<td></td>
<td>Project paper due</td>
</tr>
</tbody>
</table>

*Project Proposal

- This is the first in-class working session where teams will aim to scope the problems they are interested in solving. Each team should start doing research and thinking about business problems that they would like to tackle in the term project a few weeks before week 4.

- By the end of this class, each team will present a project brief outlining the context, need, vision, and outcomes of the term project. In addition, teams will submit a project plan with deliverables, milestones, and key dates.
• Project proposal should contain:

(1) Project title,

(2) Project description of the project,

(3) Why your project is important or different,

(4) The dataset you’re using including its source

(5) Description of the problem you’re solving and how you’re solving it.

(6) List of references

The team Project:

This is the capstone experience of the course where students will form groups consisting of 3 or 4 people. Teams will build a project using some publicly accessible datasets. They will motivate the business problem, conduct exploratory analysis, develop a machine learning model to answer the business problem posted, and, most importantly, generate data-driven, actionable strategic insight. Each team will give a brief class midterm presentation on the project, followed by a final presentation at the end of the course.

Grading:

We will have three sets of homework covering basic statistics, data mining and machine learning.

**Homework (weekly codes): 20% Exams 1 & Exam 2: 40% Term project: 20% Final: 20%**

**Each weekly homework will be about writing model seen in class.

Academic Integrity:

All students are responsible for understanding and complying with the NYU Statement on [Academic Integrity](#).

Academic Integrity for Students at NYU

This policy sets forth core principles and standards with respect to academic integrity for students at New York University. Each school at New York University may establish its own detailed supplemental guidelines for academic integrity, consistent with its own culture, and consistent with the University-wide general guidelines described in this document.

At NYU, a commitment to excellence, fairness, honesty, and respect within and outside the classroom is essential to maintaining the integrity of our community. By accepting membership in this community, students take responsibility for demonstrating these values in their own conduct and for recognizing and supporting these values in others. In turn, these values will create a campus
climate that encourages the free exchange of ideas, promotes scholarly excellence through active and creative thought, and allows community members to achieve and be recognized for achieving their highest potential.

In pursuing these goals, NYU expects and requires its students to adhere to the highest standards of scholarship, research and academic conduct. Essential to the process of teaching and learning is the periodic assessment of students' academic progress through measures such as papers, examinations, presentations, and other projects. Academic dishonesty compromises the validity of these assessments as well as the relationship of trust within the community. Students who engage in such behavior will be subject to review and the possible imposition of penalties in accordance with the standards, practices, and procedures of NYU and its colleges and schools. Violations may result in failure on a particular assignment, failure in a course, suspension or expulsion from the University, or other penalties.

Faculty are expected to guide students in understanding other people's ideas, in developing and clarifying their own thinking, and in using and conscientiously acknowledging resources - an increasingly complex endeavor given the current environment of widely available and continually emerging electronic resources. In addition, students come to NYU from diverse educational contexts and may have understandings regarding academic expectations that differ from those at NYU. NYU values and respects all academic traditions; however, while at NYU, students are expected to adhere to the norms and standards of academic integrity espoused by the NYU community and will be assessed in accordance with these standards. Students should ask their professors for guidance regarding these standards as well as style guide preferences for citation of sources for assignments in their courses.

Following are examples of behaviors that compromise the academic and intellectual community of NYU. The list is not exhaustive. Students should consult the websites and guidelines of their individual schools for an extended list of examples and for further clarification.

1. Plagiarism: presenting others' work without adequate acknowledgement of its source, as though it were one's own. Plagiarism is a form of fraud. We all stand on the shoulders of others, and we must give credit to the creators of the works that we incorporate into products that we call our own. Some examples of plagiarism:
   - a sequence of words incorporated without quotation marks
   - an unacknowledged passage paraphrased from another's work
   - the use of ideas, sound recordings, computer data or images created by others as though it were one's own

2. Cheating: deceiving a faculty member or other individual who assess student performance into believing that one’s mastery of a subject or discipline is greater than it is by a range of dishonest methods, including but not limited to:
   - bringing or accessing unauthorized materials during an examination (e.g., notes, books, or other information accessed via cell phones, computers, other technology or any other means)
   - providing assistance to acts of academic misconduct/dishonesty (e.g., sharing copies of exams via cell phones, computers, other technology or any other means, allowing others to copy answers on an exam)
   - submitting the same or substantially similar work in multiple courses, either in the same semester or in a different semester, without the express approval of all instructors
● submitting work (papers, homework assignments, computer programs, experimental results, artwork, etc.) that was created by another, substantially or in whole, as one's own
● submitting answers on an exam that were obtained from the work of another person or providing answers or assistance to others during an exam when not explicitly permitted by the instructor
● submitting evaluations of group members’ work for an assigned group project which misrepresent the work that was performed by another group member
● altering or forging academic documents, including but not limited to admissions materials, academic records, grade reports, add/drop forms, course registration forms, etc.

3. Any behavior that violates the academic policies set forth by the student’s NYU School, department, or division.

**Moses Center Statement of Disability**
If you are student with a disability who is requesting accommodations, please contact New York University’s Moses Center for Students with Disabilities at 212-998-4980 or mosescsd@nyu.edu. You must be registered with CSD to receive accommodations. Information about the Moses Center can be found at [www.nyu.edu/csd](http://www.nyu.edu/csd). The Moses Center is located at 726 Broadway on the 2nd floor.