# City University of Hong Kong
## Course Syllabus

## offered by School of Data Science
## with effect from Semester A 2022/23

---

### Part I    Course Overview

| | |
|---|---|
| **Course Title:** | Stochastic Optimization for Machine Learning |
| **Course Code**: | SDSC6015 |
| **Course Duration:** | One semester |
| **Credit Units**: | 3 |
| **Level**: | P6 |
| **Medium of Instruction:** | English |
| **Medium of Assessment:** | English |
| **Prerequisites**: *(Course Code and Title)* | Nil |
| **Precursors**: *(Course Code and Title)* | Nil |
| **Equivalent Courses**: *(Course Code and Title)* | Nil |
| **Exclusive Courses**: *(Course Code and Title)* | Nil |

**Part II    Course Details**

**1.    Abstract**

Stochastic optimization plays a vital role in machine learning where the full batch of data is either unavailable or too large to process in practice. This course introduces the theoretical foundations and algorithmic development in this area. The topics will start form the basic convex optimization theories as well as numerical methods, and we then focus on the stochastic approximation for stochastic optimization and its various accelerations in many statistical and machine learning models, supplemented with the most recent progress from research literature. Basic theoretic understanding of these stochastic optimization algorithms will also be explained. After this class, the students with some preliminaries of classic optimizations and probability theories are expected to transit into the new optimization world in the machine learning, in which significant progresses have been made during the past decades.

**2.    Course Intended Learning Outcomes (CILOs)**
*(CILOs state what the student is expected to be able to do at the end of the course according to a given standard of performance.)*

| No. | CILOs | Weighting (if applicable) | Discovery-enriched curriculum related learning outcomes (please tick where appropriate) | | |
|---|---|---|---|---|---|
| | | | *A1* | *A2* | *A3* |
| 1. | Understand the methodologies and the underlying mathematical structures in optimization | 20% | ✓ | | |
| 2. | Identify the essential principles and nature of stochastic approximation for modern machine learning optimization problem | 25% | | ✓ | |
| 3. | Analyze basic forms of stochastic optimization and learning algorithms across various machine learning models. | 30% | | ✓ | |
| 4. | Develop and analyze the practical algorithms for related optimization problems. | 25% | | | ✓ |
| | | 100% | | | |

*A1:      Attitude*
*Develop an attitude of discovery/innovation/creativity, as demonstrated by students possessing a strong sense of curiosity, asking questions actively, challenging assumptions or engaging in inquiry together with teachers.*
*A2:      Ability*
*Develop the ability/skill needed to discover/innovate/create, as demonstrated by students possessing critical thinking skills to assess ideas, acquiring research skills, synthesizing knowledge across disciplines or applying academic knowledge to self-life problems.*
A3:      *Accomplishments*
*Demonstrate accomplishment of discovery/innovation/creativity through producing /constructing creative works/new artefacts, effective solutions to real-life problems or new processes.*

## 3. Teaching and Learning Activities (TLAs)

*(TLAs designed to facilitate students' achievement of the CILOs.)*

| TLA | Brief Description | CILO No. | | | | | Hours/week (if applicable) |
|-----|-------------------|---|---|---|---|---|----------------------------|
| | | 1 | 2 | 3 | 4 | | |
| Lecture | Learning through teaching is primarily based on lectures and demonstrations. | ✓ | ✓ | ✓ | ✓ | | 26 hours/sem |
| Tutorial | Assist students with tutorial questions, computer programming of implementing algorithms, and the in-depth discussion. | | ✓ | ✓ | ✓ | | 13 hours/sem |

## 4. Assessment Tasks/Activities (ATs)

*(ATs are designed to assess how well the students achieve the CILOs.)*

| Assessment Tasks/Activities | CILO No. | | | | | Weighting | Remarks |
|-----------------------------|---|---|---|---|---|-----------|---------|
| | 1 | 2 | 3 | 4 | | | |
| Continuous Assessment:  60  % | | | | | | | |
| Test/Quiz<br>Students need to participate actively in in-class quizzes and tests designed to facilitate their understanding of knowledge taught in class. | ✓ | ✓ | ✓ | | | 30% | |
| Take-home Assignment<br>Students will critically analyze, apply and develop the theoretical concepts and quantitative skills by solving the individual samples of questions and exercises. | | ✓ | ✓ | ✓ | | 30% | |
| Examination:  40  % (duration: 2 hours, if applicable) | | | | | | | |
| Examination<br>Students will be assessed via the examination their understanding of concepts learned in class, textbooks, reading materials and their ability to apply subject-related knowledge. | ✓ | ✓ | ✓ | ✓ | | 40% | |
| | | | | | | 100% | |

For a student to pass the course, at least 30% of the maximum mark for the examination should be obtained.

## 5. Assessment Rubrics

*(Grading of student achievements is based on student performance in assessment tasks/activities with the following rubrics.)*

**Applicable to students admitted in Semester A 2022/23 and thereafter**

| Assessment Task | Criterion | Excellent (A+, A, A-) | Good (B+, B) | Marginal (B-, C+, C) | Failure (F) |
|---|---|---|---|---|---|
| 1. Test/Quiz | Demonstrate the learning process, the intellectual ability and achievements of understanding the materials in lecture time. | High | Significant | Moderate | Not even reaching marginal levels |
| 2. Take-home Assignment | Show the capability to apply the knowledge and methods to practical exercises and the generalization to new context. | High | Significant | Moderate | Not even reaching marginal levels |
| 3. Examination | Present the overall academic performance in understanding fundamentals and achieving measurable progress in the taught discipline. | High | Significant | Moderate | Not even reaching marginal levels |

**Applicable to students admitted before Semester A 2022/23**

| Assessment Task | Criterion | Excellent (A+, A, A-) | Good (B+, B, B-) | Fair (C+, C, C-) | Marginal (D) | Failure (F) |
|---|---|---|---|---|---|---|
| 1. Test/Quiz | Demonstrate the learning process, the intellectual ability and achievements of understanding the materials in lecture time. | High | Significant | Moderate | Basic | Not even reaching marginal levels |
| 2. Take-home Assignment | Show the capability to apply the knowledge and methods to practical exercises and the generalization to new context. | High | Significant | Moderate | Basic | Not even reaching marginal levels |
| 3. Examination | Present the overall academic performance in understanding fundamentals and achieving measurable progress in the taught discipline. | High | Significant | Moderate | Basic | Not even reaching marginal levels |

**Part III  Other Information** (more details can be provided separately in the teaching plan)

1.  **Keyword Syllabus**
    *(An indication of the key topics of the course.)*

- Machine Learning models for optimization.
- Convex optimization theory : gradient and sub-gradient descent, mirror descent, acceleration; constrained optimization; primal-dual; ADMM.
- Stochastic convex optimization: stochastic gradient descent; Nesterov acceleration; stochastic momentum method: Adagrad, RMSprop, Adam; stochastic mirror descent; noise-reduction technique.
- Python programming for related algorithms.
- Selected Topics that arise from stochastic optimization in adversarial training, reinforcement learning, etc.

2.  **Reading List**
2.1  **Compulsory Readings**
    *(Compulsory readings can include books, book chapters, or journal/magazine articles. There are also collections of e-books, e-journals available from the CityU Library.)*

| 1. | Lecture note |
|---|---|

2.2  **Additional Readings**
    *(Additional references for students to learn to expand their knowledge about the subject.)*

| 1. | First-Order and Stochastic Optimization Methods for Machine Learning<br>Author(s): Guanghui Lan<br>Series: Springer in the Data Sciences<br>Publisher: Springer Nature, Year: 2020 |
|---|---|
| 2. | Optimization Methods for Large-Scale Machine Learning. SIAM Review Vol. 60, No. 2, pp. 223–311, 2018 |
| 3. | ADAM: a method for stochastic optimization,   Diederik P. Kingma, Jimmy Ba, ICLR 2015 |