This course introduces the advanced machine learning techniques for extending the boundary of life sciences. Topics to be covered include but are not limited to the recent successful stories of AI for life sciences study. We are going to cover reinforcement learning-based drug design (GENTRL), graph neural network-assisted antibiotics discovery, deep learning-enhanced super-resolution microscopy, ground-breaking molecular folding algorithm (AlphaFold), deep learning-based disease diagnosis and prediction, deep learning for single-cell and spatial transcriptomics, multi-modality/omics learning, model interpretability, privacy-preserving learning for life sciences. Along the course, we will further discuss the challenges and opportunities of AI in life sciences, such as foundation models with unlabeled data for solving data scarcity issue, out-of-distribution learning for drug design. This is a research-orientated course. By the end of the course, the students are expected to finish a substantial project, aiming at publication.

Grade Descriptor:

A
EXCELLENT – exceptionally good performance and far exceeding expectation in all or most of the course learning outcomes; demonstration of superior understanding of the subject matter, the ability to analyze problems and apply extensive knowledge, and skillful use of concepts and materials to derive proper solutions.

B
GOOD – good performance in all course learning outcomes and exceeding expectation in some of them; demonstration of good understanding of the subject matter and the ability to use proper concepts and materials to solve most of the problems encountered.

C
FAIR – adequate performance and meeting expectation in all course learning outcomes; demonstration of adequate understanding of the subject matter and the ability to solve simple problems.
有關等級說明的資料，請參閱英文版本。

D

MARGINAL - performance barely meets the expectation in the essential course learning outcomes; demonstration of partial understanding of the subject matter and the ability to solve simple problems.

有關等級說明的資料，請參閱英文版本。

F

FAILURE - performance does not meet the expectation in the essential course learning outcomes; demonstration of serious deficiencies and the need to retake the course.

有關等級說明的資料，請參閱英文版本。

Equivalent Offering:
Units: 3 (Min) / 3 (Max) / 3 (Acad Progress)
Grading Basis: Graded
Repeat for Credit: N
Multiple Enroll: N
Course Attributes: MSc Computer Science, MPhil-PhD Computer Sci & Erg

Topics:

COURSE OUTCOMES

Learning Outcomes:
At the end of the course of studies, students will be able to:
1. know the different kinds of computational problems in life sciences.
2. understand how to use the AI techniques to resolve the above computational problems.
3. apply and develop AI techniques to resolve the above problems.
4. analyze the performance of the machine learning method and understand their limitations.
5. enhance problem-solving ability, communication and presentation skills by doing project.
Course Syllabus:

Week 1: The overview of ML and DL
Week 2: Molecular folding algorithm based on deep learning (AlphaFold)
Week 3: Reinforcement learning-based drug design
Week 4: Graph neural network-assisted antibiotics discovery
Week 5: Deep learning-enhanced super-resolution microscopy
Week 6: Deep learning-based disease diagnosis and prediction
Week 7: Deep learning for single-cell and spatial transcriptomics
Week 8: Multi-modality/omics learning
Week 9: Model interpretability in life sciences
Week 10: Privacy-preserving learning for life sciences
Week 11: Foundation models in life sciences
Week 12: Out-of-distribution learning for drug design
Week 13: Course project presentation

Assessment Type:

<table>
<thead>
<tr>
<th>Assessment Type</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Homework or assignment</td>
<td>20%</td>
</tr>
<tr>
<td>Project</td>
<td>60%</td>
</tr>
<tr>
<td>Test or quiz</td>
<td>20%</td>
</tr>
</tbody>
</table>

Feedback for Evaluation:

1. Results of assignments and quiz
2. Course evaluation and questionnaire
3. Reflection of teachers
4. Question-and-answer sessions during class
5. Student consultation during office hours or online

Required Readings:

To be provided by course teacher.

Recommended Readings:

"Introduction to data mining" by Michael Steinbach, Pang-Ning Tan, and Vipin Kumar (ISBN: 9780321321367)
"Deep Learning " by Ian Goodfellow, Yoshua Bengio, and Aaron Courville (ISBN: 9780262035613)
ENROLMENT REQUIREMENTS

1. CSCI5660

Enrollment Requirement Group:
For undergraduate students in Computer Science (CSCIN) or Computer Engineering (CENGN) or Artificial Intelligence – Systems & Technologies (AISTN); or
For students in MSc Computer Science; or
For students in MPhil-PhD Computer Science & Engineering

New Enrollment Requirement(s):
Other Requirement = For undergraduate students in Computer Science (CSCIN) or Computer Engineering (CENGN) or Artificial Intelligence – Systems & Technologies (AISTN); or
For students in MSc Computer Science; or
For students in MPhil-PhD Computer Science & Engineering

<END OF REPORT>