MED7005: AI and Big Data for Biomedical Research

Joo Sang Lee | joosang.lee@skku.edu Tue 9:30-12:00 | AI classroom, 3rd Floor of Bldg. 71 (Medicine) TA Jaeyoung Ahn | <u>talbot05@naver.com</u> (as of Aug 19, 2024, still in progress)

COURSE DESCRIPTION

This class is **open for both undergraduate students and graduate students**. We will invite expert data scientists and AI researchers dedicated to biomedical and healthcare research in Sungkyunkwan University and other institutions to cover the most up-to-date research topics and methodology in the field. We believe the best way to learn how to perform research is to perform mentored research. The class aims to equip students with AI/data science-based biomedical and healthcare research by executing term projects. The students will form a team of 1-4 members with a goal to write a research paper. Each team will be guided by the instructor to follow the basic steps to establish experimentally or clinically testable hypotheses based on AI/data science approaches.

COURSE OBJECTIVES

This course aims to expose the students to the current approaches of artificial intelligence and data science in biomedical and healthcare research. This class will broadly cover the basics of bioinformatics, data science, and artificial intelligence with strong emphasis on the applications to biomedical and healthcare research. The basic mathematical and algorithmic background will be introduced, and the current research trends will be reviewed. The students are expected to have hands-on experience with artificial intelligence and data science approaches by participating in the term projects in the class with other students as a team.

COMMUNICATION

We will use iCampus and kakaotalk/email for communication. iCampus is for official communications as well as for anything that contains personal and sensitive information. Kakaotalk is for day-to-day information sharing, Q&As, team discussions, and other casual conversations.

PREREQUISITES

You are expected to read assigned reading materials prior to the class meetings if there is any. There will be extra credits based on your strong engagement in the course, in terms of participating in discussions and helping other students. Because coding is an integral part of the course, it is a plus to have a good understanding and working knowledge of programming, as well as working knowledge of using open-source libraries. Contact the instructor if you are uncertain about your background.

EXPECTATIONS AND REQUIREMENTS

The primary assessment will be through the assignments and the final course project. The topic of the final project will be of your (team's) choice, but I encourage everyone to consult with the instructors. You are required to submit a final paper that contains not only the results but also detailed explanation of the coding process to demonstrate your knowledge on the computational principles and techniques. You are expected to attend every class and engage in class discussions. You are expected to complete all assignments and are not allowed to use your phone or computer during the class unless explicitly asked to do so. You are expected to read assigned reading materials prior to the class meetings if there is any. There will be extra credits based on your strong engagement in the course, in terms of participating in discussions and helping other students.

· Attendance and Participation: 20%

· Assignments: 30%

· Final project: 50%

COURSE SCHEDULE

The schedule may change.

Key dates

· Project proposal due: 10/8

· Project presentation and final paper due: 12/10

Schedule and Readings

Week 1 (9/3): Introduction to the Course | Joo Sang Lee

• Assignment #1 and Assignment #2 will be announced before the class and the outcome of the Assignment #2 will be reviewed in the class. Each student is supposed to give a short introduction to the class.

· Course Project Team Assignment will be suggested and fixed before 9/10.

Week 2 (9/10): Precision Cancer Medicine | Joo Sang Lee

Advancing precision cancer medicine and drug discovery via synthetic lethality

• Lee JS[#] et al. Synthetic lethality-mediated precision oncology via the tumor transcriptome. **Cell.** (2021) doi: 10.1016/j.cell.2021.03.030.

 Feng X, Arang N, Rigiracciolo DC, Lee JS[#], et al. A Platform of Synthetic Lethal Gene Interaction Networks Reveals that the GNAQ Uveal Melanoma Oncogene Controls the Hippo Pathway through FAK. Cancer Cell (2019).
doi: 10.1016/j.ccell.2019.01.009.

· Further readings: https://nexgem.github.io/research/

Week 3 (9/17): Next Generation Sequencing | Eric Chow

•This is midterm week and we do not have class. Assignment #3 is to submit one-page graphical hand-written summary of the online lectures of Prof. Eric Chow, University of California, San Francisco:

Next Generation Sequencing I <u>https://www.youtube.com/watch?v=mI0Fo9kaWqo</u> Next Generation Sequencing II <u>https://www.youtube.com/watch?v=PFwSe09dJX0</u> Single Cell Sequencing <u>https://www.youtube.com/watch?v=k9VFNLLQP8c</u>

Week 4 (9/24): Brain Imaging and Network Modeling | Suk Jun Hong

How do the children with autism represent the external world in the brain?

• Pellicano E, Burr D. When the world becomes 'too real': a Bayesian explanation of autistic perception. **Trends Cogn Sci.** (2012) doi: 10.1016/j.tics.2012.08.009.

· Robertson CE, Baron-Cohen S. Sensory perception in autism. Nat Rev Neurosci. (2017). doi: 10.1038/nrn.2017.112.

Week 5 (10/1): Computational Pathology | Joo Sang Lee

 \cdot As a part of the class, the students will participate in the Future of Medicine: at the crossroads of computational pathology and spatial multi-omics conference.

· Assignment #4 is to submit one-page graphical hand-written summary of the lecture of Faisal Mahmood, Harvard University.

Week 6 (10/8): Nutritional Epidemiology | Jinhee Hur

Diet and Colorectal Cancer Prevention: The Role of Nutritional Epidemiology

• Hur J, Otegbeye E, Joh HK, et al. Sugar-sweetened beverage intake in adulthood and adolescence and risk of earlyonset colorectal cancer among women. **Gut** (2021).

• Hur J, Smith-Warner SA, Rimm EB, et al. Alcohol intake in early adulthood and risk of colorectal cancer: three large prospective cohort studies of men and women in the United States. **Eur J Epidemiol.** (2021). (https://rdcu.be/cfffF)

Week 7 (10/15): Precision Psychiatry | Yoonie Joo

Advancing Precision Psychiatry: Predicting and Improving Mental Health Outcomes through Multimodal DNAlinked Biobank Data

• Park J, Lee E, Cho G, Hwang H, Kim BG, Kim G, Joo YY*, Cha J*. Gene–environment pathways to cognitive intelligence and psychotic-like experiences in children. **Elife.** (2024).

· Joo YY, Moon SY, Wang HH, Kim H, Lee EJ, Kim JH, Posner J, Ahn WY, Choi I, Kim JW, Cha J. Association of genome-wide polygenic scores for multiple psychiatric and common traits in preadolescent youths at risk of suicide. **JAMA network open** (2022).

Week 8 (10/22): Spatial Transcriptomics | Jun Hee Lee

•This is midterm week and we do not have class. Assignment #5 is to summarize the online lecture of Prof. Jun Hee Lee from University of Michigan: <u>https://www.youtube.com/watch?v=5pfnNXY7xK0</u>

• Further reading: Cho CS, ..., Lee JH. Microscopic examination of spatial transcriptome using Seq-Scope. Cell. (2021) doi: 10.1016/j.cell.2021.05.010.

Week 9 (10/29): AI-based Medical Devices | Kyu-Hwan Jung

· Moor et al., Foundation models for generalist medical artificial intelligence Nature (2023).

· Tu et al., Toward Generalist Biomedical AI NEJM (2024).

Week 10 (11/5): Attention | Kwangsun Yoo

Brain MRI big data and predictive modeling

• Yoo K*, et al. A brain-based general measure of attention. **Nature Human Behavior.** (2022) doi: 10.1038/s41562-022-01301-1.

• Jiang R, Noble S, Sui J, Yoo K, Rosenblatt M, Horien C, Qi S, Liang Q, Sun H, Calhoun VD, Scheinost D. Associations of physical frailty with health outcomes and brain structure in 483 033 middle-aged and older adults: a population-based study from the UK Biobank. Lancet Digit Health. (2023) doi: 10.1016/S2589-7500(23)00043-2.

Week 11 (11/12): Single Cell and Spatial Proteomics | Heejoo Ryu

From single cell to spatial landscape of tumor microenvironment

· Ryu H, Bi TM, Pulliam TH, Sarkar K, Church CD, Kumar N, Mayer-Blackwell K, Jani S, Ramchurren N, Hansen UK, Hadrup SR, Fling SP, Koelle DM, Nghiem P, Newell EW. Merkel cell polyomavirus-specific and CD39+CLA+
CD8 T cells as blood-based predictive biomarkers for PD-1 blockade in Merkel cell carcinoma. Cell Rep Med. (2024). doi: 10.1016/j.xcrm.2023.101390.

· Jones DC, Elz AE, Hadadianpour A, Ryu H, et al. Cell Simulation as Cell Segmentation biorxiv (2024). doi: https://doi.org/10.1101/2024.04.25.591218

Week 12 (11/19): Tumor Virome | Noam Auslander

Machine learning for virus identification

Elbasir A, .., Auslander N. A deep learning approach reveals unexplored landscape of viral expression in cancer. **Nat Commun.** (2023) doi: 10.1038/s41467-023-36336-z.

Auslander N, Gussow AB, Benler S, Wolf YI, Koonin EV. Seeker: alignment-free identification of bacteriophage genomes by deep learning. **Nucleic Acids Res.** (2020) doi: 10.1093/nar/gkaa856.

Week 13 (11/26): Smart Healthcare | Sujee Lee

Sensor-Based AI Models for Digital Healthcare

• Lee et al., Multimodal sensor fusion models for real-time exercise repetition counting with IMU sensors and respiration data. **Information Fusion** (2024).

Week 14 (12/3): Term Projects: Interim Review | Joo Sang Lee

Week 15 (12/10): Term Projects: Presentation | Joo Sang Lee

Term Project presentations will be reviewed by the instructors and other students of the course.

Week 16 (12/17): Term Projects: Evaluation | Joo Sang Lee

Term Project written reports will be reviewed by the instructors and other students of the course.

POLICIES

1. Be honest. Don't be a cheater. Your assignments and papers should be your own work. If you find useful resources for your assignments, share them and cite them. If your friends helped you, acknowledge them. You should feel free to discuss both online and offline (except for the exam), but do not show your code directly. Any cases of academic misconduct (cheating, fabrication, plagiarism, etc) will be reported to the University, following the standard procedure. Cheating is not cool.

2. You have the responsibility of backing up all your data and code. Always back up your code and data. You should at least use Google Drive or Dropbox at the minimum. Loss of data, code, or papers (e.g. due to malfunction of your laptop) is not an acceptable excuse for delayed or missing submission.