Items of Course Outlines

1. Instructor(s) – Name and Contact Details
   Jiguang Wang
   Email: jgwang@ust.hk
   Office: Room 4352 (Lift 13/15)
   Division of Life Science and Department of Chemical and Biological Engineering

2. Teaching Assistant(s) - Name and Contact Details
   To be determined.

3. Meeting Time and Venue – Lectures, Tutorials/ Laboratory
   Lectures: LIFS 6000A - Data Science for Biology and Medicine
   Date/Time: Tuesday (15:00 – 17:50)
   Venue: LSK1033

4. Course Description - Credit Points, Pre-requisite, Exclusion, Brief Information/synopsis
   • Credit points: 3
   • Prerequisites: NIL
   • Exclusions: NIL
   • Brief description: This will be an introductory course for the application of data science in biomedical
     research. The course will introduce the fundamental principles of data science, the technologies,
     and implementations of data analytics, as well as the mathematical modeling of several practical
     questions in biomedical research. The topics include an introduction to high-throughput biomedical
     data, statistical hypothesis testing, regression/classification methods, gene ontology analysis,
     biological networks, data visualization and the application of machine learning methods.

5. Intended Learning Outcomes
   (a) Obtain a basic understanding of data science
   (b) Understand high-throughput biomedical data
   (c) Become familiar with the most popular computational methods
   (d) Apply computational software to solve practical problems
   (e) Enhance the ability of modelling scientific questions

6. Assessment Scheme
<table>
<thead>
<tr>
<th>Assessment</th>
<th>Assessing Course ILOs</th>
</tr>
</thead>
<tbody>
<tr>
<td>10% Attendance</td>
<td>NA</td>
</tr>
<tr>
<td>20% Presentation</td>
<td>(a) (b)</td>
</tr>
<tr>
<td>30% Team Project 1</td>
<td>(a) (b) (c) (d)</td>
</tr>
<tr>
<td>40% Team Project 2</td>
<td>(a) (b) (c) (d) (e)</td>
</tr>
</tbody>
</table>
7. Student Learning Resources - Lecture Notes, Readings
   Lecture notes and supplementary reading materials will be made available on canvas.

8. Teaching and Learning Activities -
   Scheduled activities: 3 hrs (lecture)

9. Course Schedule
   February 6: Introduction
   February 13: High-throughput data in biology and medicine
   February 20: Basic math and statistical hypothesis test
   February 27: Correlation analysis and its applications
   March 6: Principle component analysis and clustering
   March 13: Linear regression and other regression methods
   March 20: Mid-term Presentation
   March 27: Gene ontology analysis
   April 3: Introduction to complex networks
   April 10: Network biology and practical applications
   April 17: Reconstruction of transcriptional regulatory networks
   April 24: Data visualization
   May 8: Brief introduction to machine learning methods