Bioinformatics 2 for IT and Health

Course Introduction

Li Jiang

jiang@cbs.dtu.dk

Center for Biological Sequence Analysis (CBS)
Department of Systems Biology
Technical University of Denmark
Rapid growth of biological and medical data
Bioinformatics resources
Bioinformatics analysis

Methods

Data

IT tools
Learning objectives of this course

• Put together bioinformatics tools in a pipeline using Unix shell scripts and small programs (scripts).
• Write R scripts to manipulate and visualize data.
• Implement programmatic data retrieval and Bioinformatics in Python.
• Understand the principles behind the commonly used classification, modeling and prediction methods.
• Evaluate the quality and usability of bioinformatics tools using common performance measures such as sensitivity, specificity, correlation coefficients and ROC curves.
• Use web services to access to remote resources.
An example of application diagram

Data exploration

- Quantitative data
  - Classify disease stages

- Sequence data
  - Identify vaccine target

Data analysis

- Unix
- R
- Biopython
- Web Services
- Machine learning
- Alignment
- Clustering

Data retrieve
Course format

• Lecture (tutorial)
• Exercise (hands-on)
• Mini-project
• Exam
1. **Tutorial**
   - Q:
   - A:

2. **Tutorial**
   - Q:
   - A: ✓
   - Q:
   - A: ✗

3. **Quiz**

---

**Read**

**Think**

**Learn**

- Discuss
- Wrap-up
- Submit for feedbacks

- Off-site
Mini-project

• Choices
  – Category I: based on lectures
  – Category II: based on pre-course questionnaires

• Dates
  – December 15: introduction & assignment
  – December 22: a full working day
  – January 12: presentation

• “Evaluation”
  – Poster presentation
Exam

• Written exam in 4 hours
• Monday 19\textsuperscript{th} January 2015, 9:00 - 13:00
• Bring computer, lecture notes and exercise tutorials
Course information

• Course homepage

• DTU CampusNet
  – Upload your answers to the exercise
  – Submit your answers to the exam (TBD)
<table>
<thead>
<tr>
<th>Date</th>
<th>Topics</th>
<th>Teachers</th>
</tr>
</thead>
<tbody>
<tr>
<td>2014-11-17</td>
<td>Unix</td>
<td>Peter</td>
</tr>
<tr>
<td>2014-11-24</td>
<td>R</td>
<td>Li</td>
</tr>
<tr>
<td>2014-12-01</td>
<td>Biopython</td>
<td>Christian</td>
</tr>
<tr>
<td>2014-12-08</td>
<td>Machine learning</td>
<td>Paolo</td>
</tr>
<tr>
<td>2014-12-15</td>
<td>Machine learning + Introduction to project</td>
<td>Paolo + Li</td>
</tr>
<tr>
<td>2014-12-22</td>
<td>Work on project</td>
<td></td>
</tr>
<tr>
<td>2015-01-05</td>
<td>Web services</td>
<td>Kristoffer</td>
</tr>
<tr>
<td>2015-01-12</td>
<td>Project presentation + preparation for exam</td>
<td></td>
</tr>
<tr>
<td>2015-01-19</td>
<td>Exam</td>
<td></td>
</tr>
</tbody>
</table>
Schedule

Mondays

09:00  Lecture I
10:00  Exercise I
11:00  Exercise I
12:00  Lunch break
13:00  Lecture II
14:00  Exercise II
15:00  Exercise II
16:00  Wrap-up & discussion
17:00  

Lunch break
Wireless

• Eduroam
• DTU network
Student account

- For exercises and exam
- Work on CBS server: padawan
- It will be assigned by Peter (Unix) today