**生物信息学算法原理课程教学大纲**

|  |
| --- |
| 课程基本信息（Course Information） |
| 课程代码（Course Code） | BI371 | 学时（Credit Hours） | 48 | 学分（Credits） | 3 |
| 课程名称(Course Name） | （中文）生物信息学算法原理 |
| （英文）Algorithms in computational biology |
| 课程性质(Course Type) | 培养计划内课程 |
| 授课对象（Target Audience） | 生物信息学专业本科生 |
| 授课语言(Language of Instruction) | 双语/全英文（Chinese or English) |
| 开课院系（School） | 生命科学技术学院（School of Life Sciences and Biotechnology） |
| 先修课程（Prerequisite） | 概率统计（Probability and Statistics） |
| 授课教师（Teacher） | 韦朝春 | 课程网址(Course Webpage) |  |
| \*课程简介（中文） |  本课程主要介绍生物学数据处理分析中的各种算法原理，包括动态规划、隐马尔可夫模型等各方面的内容。通过学习上述较为成熟的模式识别和统计模型，培养在生命科学研究中的理性思维的习惯和定量分析的能力，开阔视野，在一个较高的层面上提升科学素养和解决实际问题的能力。 |
| \*课程简介（英文） | This course will give a systematic introduction about algorithms in computational biology, including dynamic programming, Hidden markov models and other algorithms. It is our objective that students who take this course will be able to analyze and model biological problems using mathematical knowledge and various machine learning tools.  |
| 课程教学大纲（course syllabus）（以下内容根据所选语言，显示需必填表格不同，如为中文授课，对应英文框为非必填项；如为外文授课，需必填中文、英文相对应的两部分内容，小语种课程可选填对应语言；系统开发时，会有中英文对应的两部分内容） |
| \*学习目标(Learning Outcomes) | 学习目标是掌握：1. 生物信息学算法的基本知识（A5.1，A5.2，A5.4，B10）； 2. 团队合作及沟通能力（B4，B11）； 3. 跨学科的思考方式（B3，B10）； 以及4. 生命科学研究中的计算机运用能力和技巧（A5.1，B9）。**The main outcomes of this course:**1.Basic knowledge about algorithms in Computational Biology (A5.1, A5.2, A5.4, B10); 2. Teamwork and communication skills （B4， B11）; 3. Interdisciplinary thinking （B3，B10）;4. Computer abilities and skills in life sciences research （A5.1, B9）.  |
| \*教学内容、进度安排及要求(Class Schedule& Requirements) |

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| 教学内容 | 学时 | 教学方式 | 作业及要求 | 基本要求 | 考查方式 |
| 生物信息学简介及分子生物学背景知识 | 2学时 | 课堂讲课 | 每次课堂教学和上机后有课外作业，要求独立完成，有集中问题进行讲解 | 运用一种编程语言，编写相关的算法解决实际的生物学问题，并对结果进行分析和讨论 | 作业及大作业 |
| 概论统计及信息论简介 | 6学时 | 课堂讲课 |  |
| 双序列联配，包括点阵图、动态规划及启发式搜索 | 8学时 | 6学时课堂讲课（其他学时上机，下同） |  |
| 基因结构预测方法，包括隐马尔可夫模型模型 | 8学时 | 6学时 |  |
| 多序列比对要介绍期望最大化算法等 | 4时 | 2学时 |  |
| 模体发现算法，包括MCMC,MEME等 | 2时 | 2学时 |  |
| 元基因组序列分类 | 2学时 | 2学时 |  |
| 基因芯片数据分析方法 | 6学时 | 4学时 |  |
| 蛋白质组学 | 4学时 | 2学时 |  |
| 调控网络构建及比较 | 6学时 | 4学时 |  |

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Content | Hours | Lecture/Practice | Homework and requirement | Requirement | Homework |
| Introduction to bioinformatics and molecular biology | 2  | Lecture | Homework after lecture and practice. Finish the homework independently. | Use a programming language to solve a biological problem, and discuss the results. | Homework and Projects |
| Brief Introduction to Probability and Statistics, information theory | 6 | Lecture |
| Pair-wise alignment including: Dynamic program, heuristic search | 8 | 6 hour lecture (the rest is practice) |
| Gene structure prediction, including Hidden Markov Model | 8 | 6 hour lecture |
| Multiple sequence alignment | 4 | 2 |
| Motif finding algorithm, including MCMC,MEME etc.  | 2 | 2 |
| Metagenomic sequence binning | 2 | 2 |
| Microarray data analysis | 6 | 4 |
| Proteomics data anaysls | 4 | 2 |
| Regulatory network construction and comparison | 6 | 4 |

(教师可根据课程情况添加行数，每个内容均可点击、弹出注释、提示框，对需要填写内容进行详尽解释，考查方式对应具体教学内容） |
| \*考核方式(Grading) | （说明成绩构成，如含平时、期中、期末等；需要具体表述各考核方式对教学要求达成的评测度，弹出提示框，进行范例注释）**最终成绩由平时作业、课堂表现、大作业、期终考试成绩组合而成。各部分所占比例如下：****平时作业20%， 课堂表现10%。主要考核对知识点的掌握程度、口头及文字表达能力。** **大作业及报告讨论：20%。主要考核分析解决问题、创造性工作、计算机编程及结果分析、总结等方面的能力。** 考试：50%。主要考核对生物信息学基本算法的掌握程度。**Grading**1. Homework assignments: 20% 2. Class participants: 10%2. Projects: 20% 3. Final examination: 50% |
| \*教材或参考资料(Textbooks & Other Materials) | Bioinformatics:Sequence and Genome Analysis :序列与基因组分析:[英文版]/[美]David W.Mount[著]. 北京:科学出版社,2002 生物序列分析：蛋白质和核酸的概率论模型：英文版, 生物信息学丛书, (美)R.Durbin等著, 影印版, 北京, 清华大学出版社, 2002（填写语言对英） |
| 其它（More） |  |
| 备注（Notes） |  |

备注说明：

1．带\*内容为必填项。

2．课程简介字数为300-500字；课程大纲以表述清楚教学安排为宜，字数不限。