

МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ РОССИЙСКОЙ ФЕДЕРАЦИИ

**Федеральное государственное автономное
образовательное учреждение высшего образования
«Национальный исследовательский Нижегородский государственный университет им.
Н.И. Лобачевского»**

Институт информационных технологий, математики и механики

УТВЕРЖДЕНО
решением Ученого совета ННГУ
от 30.11.2022 г. протокол № 13

Рабочая программа дисциплины

Введение в биоинформатику

Уровень высшего образования
магистратура

Направление подготовки
010402 Прикладная математика и информатика

Направленность образовательной программы
Вычислительные методы и суперкомпьютерные технологии

Форма обучения
очная

Нижний Новгород
2022

1. Место и цели дисциплины в структуре ООП

Место дисциплины в структуре ООП

Данная дисциплина (Б1.В.ДВ.01.01) читается в первом семестре магистратуры, относится к части, формируемой участниками образовательных отношений.

Цель освоения дисциплины

Предметом рассмотрения настоящего курса являются основные задачи биоинформатики, которые могут быть сформулированы как задачи поиска, выравнивания, прочие задачи на строках и графах. Цель курса состоит в изучении соответствующих методов и алгоритмов биоинформатики. Особое внимание уделяется формированию у студентов навыков реализации и применения рассматриваемых методов к решению конкретных биоинформатических задач. Задачей дисциплины является теоретическое и практическое освоение вышеупомянутых методов и алгоритмов.

2. Планируемые результаты обучения по дисциплине, соотнесенные с планируемыми результатами освоения образовательной программы (компетенциями и индикаторами достижения компетенций)

Формируемые компетенции (код, содержание компетенции)	Планируемые результаты обучения по дисциплине (модулю), в соответствии с индикатором достижения компетенции		Наименование оценочного средства
	Индикатор достижения компетенции (код, содержание индикатора)	Результаты обучения по дисциплине	
<i>ПК-4 Способность разрабатывать и анализировать концептуальные и теоретические модели решаемых научных проблем и задач</i>	<i>ПК-4.1. Знает методы разработки и анализа концептуальных и теоретических моделей решаемых научных проблем и задач</i>	ЗНАТЬ <i>Современное состояние исследований в области биоинформатики, базовые алгоритмы биоинформатики, условия их применимости к решению научных проблем и задач.</i>	Собеседование (зачет)
	<i>ПК-4.2. Умеет применять методы разработки и анализа концептуальных и теоретических моделей решаемых научных проблем и задач</i>	УМЕТЬ <i>Определять и профессионально реализовывать необходимые вычислительные алгоритмы биоинформатики для решения научных проблем и задач, анализировать полученные результаты.</i>	Лабораторная работа (текущий контроль)
	<i>ПК-4.3 Имеет навыки применения методов разработки и анализа концептуальных и теоретических моделей решаемых</i>	ВЛАДЕТЬ <i>Навыками программной реализации вычислительных методов и алгоритмов биоинформатики для решения научных проблем и задач.</i>	Лабораторная работа (текущий контроль)

	<i>научных проблем и задач</i>		
<i>ПК-11 Способность разрабатывать и анализировать концептуальные и теоретические модели решаемых задач производственно-технологической деятельности</i>	<i>ПК-11.1. Знает методы разработки и анализа концептуальных и теоретических моделей решаемых производственно-технологических задач</i>	ЗНАТЬ <i>Современное состояние исследований в области биоинформатики, базовые алгоритмы биоинформатики, условия их применимости к решению задач производственно-технологической деятельности.</i>	Собеседование (зачет)
	<i>ПК-11.2. Умеет применять методы разработки и анализа концептуальных и теоретических моделей решаемых производственно-технологических задач</i>	УМЕТЬ <i>Определять и профессионально реализовывать необходимые вычислительные алгоритмы биоинформатики для решения задач производственно-технологической деятельности, анализировать полученные результаты.</i>	Лабораторная работа (текущий контроль)
	<i>ПК-11.3 Имеет навыки применения методов разработки и анализа концептуальных и теоретических моделей решаемых производственно-технологических задач</i>	ВЛАДЕТЬ <i>Навыками программной реализации вычислительных методов и алгоритмов биоинформатики для решения задач производственно-технологической деятельности.</i>	Лабораторная работа (текущий контроль)

3. Структура и содержание дисциплины

3.1. Трудоемкость дисциплины

	Очная форма обучения
Общая трудоемкость	2 ЗЕТ
Часов по учебному плану	72
в том числе	
аудиторные занятия (контактная работа):	33
– занятия лекционного типа	16
– занятия семинарского типа	16
– занятия лабораторного типа	0
– текущий контроль (КСР)	1
самостоятельная работа	39
Промежуточная аттестация: зачет	0

3.2. Содержание дисциплины

Наименование и краткое содержание	Всего	в том числе
--	--------------	--------------------

		Контактная работа (работа во взаимодействии с преподавателем), часы из них				Самостоятельная работа обучающегося, часы
		Занятия лекционного типа	Занятия семинарского типа	Занятия лабораторного типа	Всего контактных часов	
Введение. Основные макромолекулы живых организмов.	9	2	2		4	5
Алгоритмы полного перебора.	10	2	2		4	6
Поисковые деревья. Метод ветвей и границ.	13	3	3		6	7
Динамическое программирование. Введение.	13	3	3		6	7
Динамическое программирование. Локальное и глобальное выравнивание.	13	3	3		6	7
Графовые алгоритмы.	13	3	3		6	7
Текущий контроль (КСР)	1				1	
Промежуточная аттестация – зачет	0					
Итого	72	16	16		33	39

Практические занятия (семинарские занятия /лабораторные работы) организуются, в том числе в форме практической подготовки, которая предусматривает участие обучающихся в выполнении отдельных элементов работ, связанных с будущей профессиональной деятельностью.

Текущий контроль успеваемости реализуется в форме опросов на занятиях семинарского типа.

Промежуточная аттестация проходит в форме зачета.

4. Учебно-методическое обеспечение самостоятельной работы обучающихся

Самостоятельная работа студентов подразделяется на следующие категории:

- Изучение учебной литературы (см. перечень образовательных материалов).
- Выполнение лабораторных работ (см. пп. 5.2.2, 5.2.3).

Задания для проведения текущего контроля и промежуточной аттестации по итогам освоения дисциплины приведены в п. 5.2.

5. Фонд оценочных средств для промежуточной аттестации по дисциплине, включающий:

5.1. Описание шкал оценивания результатов обучения по дисциплине

Уровень сформированности	Шкала оценивания сформированности компетенций						
	Плохо	Неудовлетворительно	Удовлетворительно	Хорошо	Очень хорошо	Отлично	Превосходно

компетенций (индикатора достижения компетенций)	Не зачтено		Зачтено				
	<u>Знания</u>	Отсутствие знаний теоретического материала. Невозможно оценить полноту знаний вследствие отказа обучающегося от ответа	Уровень знаний ниже минимальных требований. Имели место грубые ошибки.	Минимально допустимый уровень знаний. Допущено много негрубых ошибок.	Уровень знаний в объеме, соответствующем программе подготовки. Допущено несколько негрубых ошибок	Уровень знаний в объеме, соответствующем программе подготовки. Допущено несколько несущественных ошибок	Уровень знаний в объеме, соответствующем программе подготовки, без ошибок.
<u>Умения</u>	Отсутствие минимальных умений. Невозможность оценить наличие умений вследствие отказа обучающегося от ответа	При решении стандартных задач не продемонстрированы основные умения. Имели место грубые ошибки.	Продемонстрированы основные умения. Решены типовые задачи с негрубыми ошибками. Выполнены все задания, но не в полном объеме.	Продемонстрированы все основные умения. Решены все основные задачи с негрубыми ошибками. Выполнены все задания, в полном объеме, но некоторые с недочетами.	Продемонстрированы все основные умения. Решены все основные задачи. Выполнены все задания, в полном объеме, но некоторые с недочетами.	Продемонстрированы все основные умения, решены все основные задачи с отдельными несущественными недочетами, выполнены все задания в полном объеме.	Продемонстрированы все основные умения, решены все основные задачи. Выполнены все задания, в полном объеме без недочетов
<u>Навыки</u>	Отсутствие владения материалом. Невозможность оценить наличие навыков вследствие отказа обучающегося от ответа	При решении стандартных задач не продемонстрированы базовые навыки. Имели место грубые ошибки.	Имеется минимальный набор навыков для решения стандартных задач с некоторыми недочетами.	Продемонстрированы базовые навыки при решении стандартных задач с некоторыми недочетами	Продемонстрированы базовые навыки при решении стандартных задач без ошибок и недочетов.	Продемонстрированы навыки при решении нестандартных задач без ошибок и недочетов.	Продемонстрирован творческий подход к решению нестандартных задач.

Шкала оценки при промежуточной аттестации

	Оценка	Уровень подготовки
Зачтено	Превосходно	Все компетенции (части компетенций), на формирование которых направлена дисциплина, сформированы на уровне не ниже «превосходно»
	Отлично	Все компетенции (части компетенций), на формирование которых направлена дисциплина, сформированы на уровне не ниже «отлично», при этом хотя бы одна компетенция сформирована на уровне «отлично»
	Очень хорошо	Все компетенции (части компетенций), на формирование которых направлена дисциплина, сформированы на уровне не ниже «очень хорошо», при этом хотя бы одна компетенция сформирована на уровне «очень хорошо»

	Хорошо	Все компетенции (части компетенций), на формирование которых направлена дисциплина, сформированы на уровне не ниже «хорошо», при этом хотя бы одна компетенция сформирована на уровне «хорошо»
	Удовлетворительно	Все компетенции (части компетенций), на формирование которых направлена дисциплина, сформированы на уровне не ниже «удовлетворительно», при этом хотя бы одна компетенция сформирована на уровне «удовлетворительно»
Не зачтено	Неудовлетворительно	Хотя бы одна компетенция сформирована на уровне «неудовлетворительно», ни одна из компетенций не сформирована на уровне «плохо»
	Плохо	Хотя бы одна компетенция сформирована на уровне «плохо»

Критерий оценивания лабораторной работы

Результаты работы	Оценка
Работа выполнена в полном объеме и в срок, результаты работы алгоритма корректные на тестовых примерах, проведены вычислительные эксперименты на трудоемких тестовых данных, результаты работы представлены преподавателю.	Зачтено
Работа не выполнена или выполнена не в полном объеме (программа работает некорректно на тестовых примерах, не проведены вычислительные эксперименты на трудоемких тестовых данных, результаты работы не представлены преподавателю).	Не зачтено

5.2. Типовые контрольные задания или иные материалы, необходимые для оценки результатов обучения

5.2.1. Контрольные вопросы

Вопрос	Код компетенции
1. ДНК в клетке. Расположение, структурные блоки.	ПК-4
2. Процесс репликации ДНК. Определение лидирующей и запаздывающей нити.	ПК-4
3. Механизм транскрипции.	ПК-4
4. Механизм трансляции. Уровни структуры белков.	ПК-4
5. Определение алгоритма. Основные свойства.	ПК-4
6. Вычислительная сложность алгоритмов.	ПК-4
7. Парадигмы разработки алгоритмов.	ПК-4
8. Классы сложности алгоритмов.	ПК-4
9. Основная идея алгоритмов полного перебора.	ПК-4
10. Назначение регуляторных мотивов.	ПК-4
11. Алгоритм поиска регуляторных мотивов без мутаций в регуляторных сегментах гена.	ПК-4
12. Алгоритм поиска регуляторных мотивов с возможными мутациями с учетом известных стартовых позиций мотивов в регуляторных	ПК-4

сегментах гена.	
13. Основные этапы алгоритма полного перебора для поиска мотивов с мутациями.	ПК-4
14. Расстояние Хэмминга между двумя паттернами.	ПК-4
15. Проблема поиска срединной строки.	ПК-4
16. Эквивалентность задач поиска регуляторного мотива и поиска срединной строки.	ПК-4
17. Основные этапы алгоритма полного перебора для поиска срединной строки.	ПК-4
18. Сравнение алгоритмов полного перебора для поиска мотива и срединной строки с точки зрения временной сложности.	ПК-4
19. Отличия парадигмы ветвей и границ от полного перебора.	ПК-4
20. Структура поискового дерева, имеющая применение в задаче поиска регуляторного мотива или срединной строки.	ПК-4
21. Типы перемещений в поисковом дереве.	ПК-4
22. Основные этапы алгоритмов ветвей и границ для поиска мотива и срединной строки, использующих структуру поискового дерева.	ПК-4
23. Основная идея динамического программирования.	ПК-4
24. Пример применения динамического программирования для задачи о размене монет.	ПК-4
25. Недостатки рекурсивной реализации динамического программирования для задачи о размене монет.	ПК-4
26. Пример применения динамического программирования для задачи о манхэттенском туристе.	ПК-11
27. Выравнивание последовательностей ДНК.	ПК-11
28. Определение длиннейшей общей подстроки.	ПК-11
29. Графическая интерпретация длиннейшей общей подстроки.	ПК-11
30. Основная идея оценки глобального выравнивания.	ПК-11
31. Пример оценочной матрицы для короткой последовательности нуклеотидов.	ПК-11
32. Оценочные матрицы для аминокислот. Принцип их построения.	ПК-11
33. Рекуррентное соотношение в алгоритме глобального выравнивания.	ПК-11
34. Отличие локального выравнивания от глобального выравнивания.	ПК-11
35. Отличительные особенности рекуррентной схемы для локального выравнивания.	ПК-11
36. Задача подсчета операций при трансформации строк.	ПК-11
37. Отличие аппроксимирующего выравнивания от задачи глобального и локального выравнивания.	ПК-11
38. Перекрывающееся выравнивание.	ПК-11
39. Глобальное выравнивание с аффинными штрафами за непрерывные пропуски.	ПК-11

40. Основные понятия графовых алгоритмов.	ПК-11
41. Алгоритмы поиска кратчайших путей в невзвешенных графах.	ПК-11
42. Алгоритмы поиска кратчайших путей во взвешенных графах.	ПК-11
43. Иллюстрация на примере алгоритма Дейкстры.	ПК-11
44. Эйлеров путь и Эйлеров цикл.	ПК-11
45. Гамильтонов цикл.	ПК-11
46. Восстановление генома и проблемы секвенирования.	ПК-11
47. Графовый подход в задаче восстановления генома.	ПК-11
48. Построение графа перекрытий паттернов.	ПК-11
49. Графы де Брейна.	ПК-11
50. Поиск Эйлерова цикла в графе де Брейна и восстановление генома.	ПК-11

5.2.2. Типовые темы лабораторных работ для оценки компетенции ПК-4

Задание

Разработать приложение, выполняющее поиск обратно-комплементарной цепи ДНК.

Входные параметры приложения: цепь ДНК.

Выходные параметры приложения: цепь ДНК, обратно-комплементарная к данной.

Задание

Разработать приложение, выполняющее поиск стартовых позиций паттерна в геноме.

Входные параметры приложения: цепь ДНК, паттерн.

Выходные параметры приложения: массив стартовых позиций паттерна в цепи ДНК.

Задание

Разработать приложение, выполняющее подсчет числа вхождений паттерна в геном.

Входные параметры приложения: цепь ДНК, паттерн.

Выходные параметры приложения: число вхождений паттерна в геном.

5.2.3. Типовые темы лабораторных работ для оценки компетенции ПК-11

Задание

Разработать приложение, решающее поиск длиннейшей общей подстроки.

Входные параметры приложения: две цепи ДНК.

Выходные параметры приложения: наибольшая общая подстрока.

Задание

Разработать приложение, реализующее алгоритм глобального выравнивания с использованием оценочной матрицы «BLOSUM62».

Входные параметры приложения: две строки аминокислот, матрица «BLOSUM62».

Выходные параметры приложения: максимальный Score, выравнивание, достигающее максимальный Score.

Задание

Разработать приложение, реализующее алгоритм локального выравнивания с использованием оценочной матрицы «PAM250».

Входные параметры приложения: две строки аминокислот, матрица «PAM250».

Выходные параметры приложения: максимальный Score, выравнивание, достигающее максимальный Score.

6. Учебно-методическое и информационное обеспечение дисциплины

а) основная литература:

- Курс «Алгоритмы биоинформатики» Н.И. Вяххи:
– <https://www.lektorium.tv/course/22933>.

б) дополнительная литература:

7. Материально-техническое обеспечение дисциплины

Помещения представляют собой учебные аудитории для проведения учебных занятий, предусмотренных программой (лекционного и лабораторного типа), оснащенные оборудованием и техническими средствами обучения.

Помещения для самостоятельной работы обучающихся оснащены компьютерной техникой с возможностью подключения к сети "Интернет" и обеспечены доступом в электронную информационно-образовательную среду.

Программа составлена в соответствии с требованиями ОС ННГУ по направлению 01.04.02 Прикладная математика и информатика.

Авторы: М.В. Иванченко
 И.И. Юсипов

Заведующий кафедрой М.В. Иванченко

Программа одобрена на заседании методической комиссии института информационных технологий, математики и механики от 30 ноября 2022 года, протокол № 3.