

Министерство науки и высшего образования Российской Федерации
Федеральное государственное бюджетное образовательное учреждение
высшего образования
«Амурский государственный университет»

УТВЕРЖДАЮ

Проректор по ИиНОТ

 А.А. Остапенко

«21» 08 2020 г.

РАБОЧАЯ ПРОГРАММА
**ВВЕДЕНИЕ В БИОИНФОРМАТИКУ И АНАЛИЗ БИОЛОГИЧЕСКИХ
ПОСЛЕДОВАТЕЛЬНОСТЕЙ**

Уровень образования: дополнительное

Направление: Малая академия

Форма обучения: очная

Возраст обучающихся: 17 – 18 лет

Общая трудоемкость дисциплины – 72 (час.)

Составитель – Мишаченко К.Г.

ЦРСКД «АмурТехноЦентр» (ДНК им. академика РАН М.Т. Луценко)

2020 г.

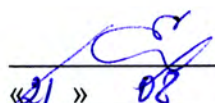
Рабочая программа одобрена на заседании педагогического совета ЦРСКД
«АмурТехноЦентр» (ДНК им. академика РАН М.Т. Луценко)

«21» 08 2020 г., протокол № 5

Председатель  В.В. Еремина
подпись И.О.Ф.

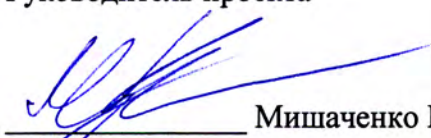
СОГЛАСОВАНО

Директор

 Еремина В.В.
«21» 08 2020 г.

СОГЛАСОВАНО

Руководитель проекта

 Мишаченко К.Г.
«21» 08 2020 г.

ПОЯСНИТЕЛЬНАЯ ЗАПИСКА

Биоинформатика – это наука о хранении, извлечении, организации, анализе, интерпретации и использовании биологической информации.

Содержание программы направлено на формирование у детей научно-технических и естественнонаучных знаний, профессионально-прикладных навыков и создание условий для социального, культурного и профессионального самоопределения, творческой самореализации личности ребенка в окружающем мире.

Актуальность программы. Биоинформатика является перспективной областью для применения образовательных методик в процессе обучения за счет объединения в себе различных инженерных и естественнонаучных дисциплин. Программа даёт возможность обучить детей профессиональным навыкам в области информатики и предоставляет условия для проведения педагогом профориентационной работы. Кроме того, обучение по данной программе способствует развитию творческой деятельности.

Новизна данной дополнительной образовательной программы заключается в том, что по форме организации образовательного процесса она является модульной.

Возраст детей, участвующих в реализации данной дополнительной образовательной программы, колеблется от 17 до 18 лет.

Сроки реализации программы: 1 год. Режим занятий – 1 раза в неделю по 2 академических часа, наполняемость в группе – 12 учащихся.

Цель: введение в компьютерно-ориентированные методы решения информационных задач в области промышленных биотехнологий.

Задачи:

дать основные знания о молекулярной биологии;
дать основные знания о современных алгоритмах обработки текста и поиска;
способствовать формированию творческого отношения к выполняемой работе;
воспитывать умение работать в коллективе, эффективно распределять обязанности;
развивать творческую инициативу и самостоятельность;
развивать психофизиологические качества обучающихся: память, внимание, способность логически мыслить, анализировать, концентрировать внимание на главном;
развивать умения излагать мысли в четкой логической последовательности, отстаивать свою точку зрения, анализировать ситуацию и самостоятельно находить ответы на вопросы путем логических рассуждений.

Для реализации данной программы используются педагогические технологии уровневой дифференциации обучения, технологии на основе личностной ориентации, а также следующие методы и формы обучения и контроля: фронтальный опрос, индивидуальная, парная, групповая работа.

Методами обучения являются: поисковый метод, самостоятельная работа, метод кейсов, дидактические игры, метод проблемного обучения, дискуссия, практическая деятельность, ИКТ.

ЛИЧНОСТНЫЕ, МЕТАПРЕДМЕТНЫЕ И ПРЕДМЕТНЫЕ РЕЗУЛЬТАТЫ ОСВОЕНИЯ УЧЕБНОГО ПРЕДМЕТА

Личностные результаты:

- ответственное отношение к выполнению заданий и стремление к получению результата;
- навык самостоятельного решения задач;
- умение работать в команде при решении задач.

Метапредметные результаты: программа направлена на развитие мышления учащихся и воспитания у них информационной культуры. На занятиях выполняются задания, развивающие творчество учащихся, умение анализировать, систематизировать, визуализировать информацию. Учащиеся учатся моделировать реально происходящие процессы, т.е. создавать информационную модель поставленной задачи.

Предметные результаты:

- представлять биологическую информацию в виде текста, таблицы, графика, диаграммы и делать выводы на основании представленных данных;
- оценивать результаты взаимодействия человека и окружающей среды, прогнозировать возможные последствия деятельности человека для существования отдельных биологических объектов и целых природных сообществ;
- сравнивать биологические объекты между собой по заданным критериям, делать выводы и умозаключения на основе сравнения;
- формулировать гипотезы на основании предложенной биологической информации и предлагать варианты проверки гипотез;
- понимать, описывать и применять на практике взаимосвязь между естественными науками: биологией, информатикой, математикой; устанавливать взаимосвязь природных явлений;
- использовать основные методы научного познания в учебных биологических исследованиях, проводить эксперименты по изучению биологических объектов и явлений, объяснять результаты экспериментов, анализировать их, формулировать выводы;

Содержание курса представлено в составе пяти модулей:

«Предмет и инфраструктура биоинформатики», «Введение в программирование на Python», «Машинное обучение в биоинформатике», «Анализ биологических последовательностей», «Разработка приложения для работы с биоинформационными данными в системе MATLAB».

УЧЕБНО-ТЕМАТИЧЕСКИЙ ПЛАН

Структура курса рассчитана на 72 часа

№ урока	Содержание учебного материала	Количество часов		
		всего	теория	практика
1	Предмет и инфраструктура биоинформатики	6	6	0
2	Введение в программирование на Python	9	4	5
3	Машинное обучение в биоинформатике	4	3	1
4	Анализ биологических последовательностей	8	4	4
5	Разработка приложения для работы с биоинформационными данными в системе MATLAB	46	15	31
	ИТОГО	72	24	48

СОДЕРЖАНИЕ ТЕМ УЧЕБНОГО КУРСА

Модуль 1 «Предмет и инфраструктура биоинформатики» (6 часов)

Биоинформатика – это наука о хранении, извлечении, организации, анализе, интерпретации и использовании биологической информации возникновение и интенсивное развитие биоинформатики происходило синхронно с возникновением и повсеместным распространением компьютерных технологий. Предметом учебной дисциплины «Введение в биоинформатику и анализ биологических последовательностей» являются компьютерно-ориентированные методы решения информационных задач в области промышленной и фармацевтической биотехнологий.

Цель модуля: Ознакомление с целями и задачами биоинформатики, ее предметом, с особенностями биологических данных, темпами их увеличения, перспективами использования биоинформатики, биоинформационными базами, сетями, данными.

Задачи модуля: знакомство с языками программирования, сценариями и разметками, нашедшими свое применение в биоинформатике; знакомство с важными

пакетами программ для вычислительной биологии рассчитаны на работу в ОС Linux; знакомство с форматом FASTA, EMBnet, SRS, NCBI, Entrez.

Учебно-тематический план Модуля 1

№ п/п	Наименование темы	Количество часов			Формы аттестации/контр оля
		всего	теория	практика	
1	Предмет, цели, задачи, перспективы биоинформатики. Основные понятия.	1	1	0	Выполнение кейса №1
2	Компьютерные компоненты и интернет – компоненты биоинформатики.	1	1	0	Дискуссия №1
3	Биоинформационные данные, сети и базы. Алгоритмы, графы, строки.	1	1	0	Выполнение кейса №2
4	Разбор алгоритмов сравнения строк. Знакомство с Java.	1	1	0	Дискуссия №1
5	Алгоритмы секвенирования.	1	1	0	Самостоятельная работа №1
6	Сборка генома.	1	1	0	Выполнение кейса №3
	Итого:	6	6	0	

Модуль 2 «Введение в программирование на Python» (9 часов)

Python обладает рядом преимуществ перед другими языками для начинающих изучать программирование, прежде всего благодаря ясности кода и скорости реализации. Вот что говорит создатель языка Python Гвидо ван Россум: «Python - интерпретируемый, объектно-ориентированный высокоуровневый язык программирования с динамической семантикой. Встроенные высокоуровневые структуры данных в сочетании с динамической типизацией и связыванием делают язык привлекательным для быстрой разработки приложений (RAD, Rapid Application Development). Кроме того, его можно использовать в качестве сценарного языка для связи программных компонентов. Синтаксис Python прост в изучении, в нем придается особое значение читаемости кода, а это сокращает затраты на сопровождение программных продуктов. Python поддерживает модули и пакеты, поощряя модульность и повторное использование кода. Интерпретатор Python и большая стандартная библиотека доступны бесплатно в виде исходных и исполняемых кодов для всех основных платформ и могут свободно распространяться.»

Цель модуля: получение представлений о базовых понятиях модульного и объектно ориентированного программирования;

Задачи модуля: знакомство со средой VS Code, основными операторами и структурами данных языка Python, функциями, ООП;

Учебно-тематический план Модуля 2

№ п/п	Наименование темы	Количество часов			Формы аттестации /контроля
		всего	теория	практика	
1	Знакомство со средой VS Code, основными операторами языка Python, циклами.	1	1	0	Самостоятельная работа №2
2	Множества, строки, списки, кортежи, словари.	1	0	1	Выполнение кейса №4
3	Функции, область видимости переменных, переменное число параметров.	1	0	1	Выполнение кейса №5
4	Рекурсия.	1	1	0	Самостоятельная работа №3
5	Решение прикладных задач на языке программирования Python.	1	1	0	Самостоятельная работа №4
6	Введение в ООП.	1	1	0	Выполнение кейса №6
7	Полиморфизм, наследование.	1	1	0	Дискуссия №2
8	Знакомство с Rosalind, Github, Git.	1	1	0	Самостоятельная работа №5
9	Решение задач на Rosalind.	1	1	0	Самостоятельная работа №6
Итого:		9	7	2	

Модуль 3 «Машинное обучение в биоинформатике» (4 часов)

Данный модуль направлен на получение начальных знаний об алгоритмах машинного обучения и их применения в биоинформатике;

Цель модуля: введение в прогрессивные методы биоинформационного анализа;

Задачи модуля: основные понятия и алгоритмы машинного обучения и их значение в биоинформатике;

Учебно-тематический план Модуля 3

	Наименование темы	Количество часов			Формы аттестации/контроля
		всего	теория	практика	

1	Машинное обучение в биоинформатике. Математические основы машинного обучения	1	0	1	Самостоятельная работа №7
2	Основные алгоритмы машинного обучения.	1	0	1	Выполнение кейса №7
3	Знакомство с Jupyter. Реализация алгоритмов ML с scikit-learn.	1	0	1	Дискуссия №3
4	Введение в глубокое обучение с Keras. Архитектуры нейронных сетей.	1	0	1	Самостоятельная работа №8
Итого:		4	0	4	

Модуль 4 «Анализ биологических последовательностей» (8 часа)

Данный модуль прикладной характер и направлен на решение конкретных задач биоинформатики с использованием различных компьютерных технологий.

Цель модуля: получение опыта разработки биоинформационных приложений.

Задачи модуля: знакомство с базами данных, биоинформационными библиотеками Python;

Учебно-тематический план Модуля 4

№ п/п	Наименование темы	Количество часов			Формы аттестации/контроля
		всего	теория	практика	
1	Решение сложных задач на Rosalind.	1	1	0	Выполнение кейса №8
2	Знакомство с БД. Необходимость в БД.	1	0	1	Выполнение кейса №9
3	Разработка БД. Запросы SQL. Разнообразие СУБД.	1	1	0	Выполнение кейса №10
4	Написание запросов SQL. Интеграция с pandas.	3	1	2	Дискуссия №4
	Знакомство с BioPython.	1	0	1	Выполнение кейса №11
5	Знакомство с Nucleus.	1	1	0	Самостоятельная работа №9
Итого:		8	4	4	

Модуль 5 «Разработка приложения для работы с биоинформационными данными в системе MATLAB» (46 часа)

Данный модуль направлен на ознакомление с системой автоматического проектирования MATLAB и ее возможностями, пакетом Bioinformatics Toolbox, визуальным проектированием GUI. Ученики должны применить свои знания о генах и генетике для работы с последовательностями генетического кода, микромассивами, спектральными данными, кластеризацией и т. д.

Цель модуля: разработка приложения в среде MATLAB;

Задачи модуля: знакомство с MATLAB, Bioinformatics Toolbox, визуальным проектированием GUI;

Учебно-тематический план Модуля 5

№ п/п	Наименование темы	Количество часов			Формы аттестации/контроля
		всего	теория	практика	
1	Введение в MATLAB. Работа с документацией MATLAB.	1	1	0	Выполнение кейса №12
2	Знакомство с пакетами расширений MATLAB.	1	1	0	Дискуссия №5
3	Дополнительные возможности MATLAB.	16	5	11	Самостоятельная работа №10
4	Разработка программы.	6	1	5	Самостоятельная работа №11
5	Защита проекта.	22	7	15	Представление и техническое описание робо-руки
	Итого:	46	15	31	

МАТЕРИАЛЬНО-ТЕХНИЧЕСКОЕ ОБЕСПЕЧЕНИЕ:

- компьютеры с установленным необходимым программным обеспечением (Python-IDLE, VS Code, Anaconda, Jupyter);
- для запуска моделей машинного обучения и программы MATLAB необходимы ноутбуки с минимальными характеристиками: тактовая частота процессора – 3,5 ГГц; оперативная память – 4 Гбайт; объем жесткого диска должен быть не менее 50 Гб;
- проектор.

ПЕРЕЧЕНЬ УЧЕБНО-МЕТОДИЧЕСКОГО ОБЕСПЕЧЕНИЯ

Список литературы для учителя:

1. Меншуткин В.В. Искусство моделирования (экология, физиология, эволюция). – Петрозаводск — Санкт-Петербург. 2010. 416 с.
2. Огурцов А.Н. Основы биоинформатики. – Харьков: НТУ «ХПИ», 2013. – 400 с.
3. Огурцов А.Н. Выравнивание белковых последовательностей. – Харьков: НТУ «ХПИ», 2015. – 80 с.
4. Сетубал Ж., Мейдаиис Ж. Введение в вычислительную молекулярную биологию. – Москва–Ижевск: НИЦ «Регулярная и хаотическая динамика», Институт компьютерных исследований, 2007. – 420 с.

5. Романовский Ю. М., Степанова Н.В., Чернавский Д. С. Математическое моделирование в биофизике. – Москва-Ижевск: Институт компьютерных исследований, 2003, 402 стр.

Список литературы для ученика:

1. Огурцов А.Н. Основы биоинформатики. – Харьков: НТУ «ХПИ», 2013. – 400 с.
2. Меншуткин В.В. Искусство моделирования (экология, физиология, эволюция). – Петрозаводск — Санкт-Петербург. 2010. 416 с.
3. Сетубал Ж., Мейдаиис Ж. Введение в вычислительную молекулярную биологию. – Москва–Ижевск: НИЦ «Регулярная и хаотическая динамика», Институт компьютерных исследований, 2007. – 420 с.
4. Романовский Ю. М., Степанова Н.В., Чернавский Д. С. Математическое моделирование в биофизике. – Москва-Ижевск: Институт компьютерных исследований, 2003, 402 стр.