

МИНИСТЕРСТВО ОБРАЗОВАНИЯ НОВОСИБИРСКОЙ ОБЛАСТИ
ДЕПАРТАМЕНТ ОБРАЗОВАНИЯ МЭРИИ ГОРОДА НОВОСИБИРСКА

МУНИЦИПАЛЬНОЕ АВТОНОМНОЕ ОБЩЕОБРАЗОВАТЕЛЬНОЕ УЧРЕЖДЕНИЕ
ГОРОДА НОВОСИБИРСКА
«ЛИЦЕЙ № 22 «НАДЕЖДА СИБИРИ»
(МАОУ «ЛИЦЕЙ № 22 «НАДЕЖДА СИБИРИ»)
СТРУКТУРНОЕ ПОДРАЗДЕЛЕНИЕ «IT КУБ»

РАССМОТРЕНО

на педагогическом совете
МАОУ «Лицей №22 «Надежда Сибири»
протокол № 1
« 22 » августа 2022 г.

УТВЕРЖДАЮ
Директор МАОУ «Лицей №22
«Надежда Сибири»
Л.В. Потеряева
« 22 » августа 2022 г.



ДОПОЛНИТЕЛЬНАЯ ОБЩЕОБРАЗОВАТЕЛЬНАЯ ПРОГРАММА
«БОЛЬШИЕ ДАННЫЕ И МАШИННОЕ ОБУЧЕНИЕ»
2022-2023 учебный год

Направленность: техническая

Возраст обучающихся: 15-17 лет

Срок освоения программы: 9 месяцев

Автор: Кириленко Юрий
Александрович,
педагог дополнительного образования

Новосибирск - 2022

ПОЯСНИТЕЛЬНАЯ ЗАПИСКА

Программа «Большие данные и машинное обучение» разработана в соответствии с требованиями нормативных документов:

ФЗ РФ от 29.12.2012 г. № 273-ФЗ «Об образовании в Российской Федерации»;

Указ Президента РФ от 7.05.2012 г. № 599 «О мерах по реализации государственной политики в области образования и науки»;

Концепция развития дополнительного образования детей до 2030 года, утверждена распоряжением Правительства РФ от 31 марта 2022 г. N 678-р;

Приказ Минобрнауки РФ от 29.08.2013 г. № 1008 «Об утверждении Порядка организации и осуществления образовательной деятельности по дополнительным общеобразовательным программам»;

Примерные требования к содержанию и оформлению образовательных программ дополнительного образования детей (письмо Минобрнауки РФ от 11.12.2006 № 06-1844);

Постановление Главного государственного санитарного врача РФ от 28.09.2020 г. № 28 «Об утверждении санитарных правил СП 2.4.3648-20 "Санитарно-эпидемиологические требования к организациям воспитания и обучения, отдыха и оздоровления детей и молодежи"»

Методические рекомендации по проектированию дополнительных общеразвивающих программ, Письмо Минобрнауки РФ от 18.11.2015.

Содержание программы способствует развитию технических способностей, логического и алгоритмического мышления.

Уровень общеобразовательной программы: начальный.

Направленность программы: техническая.

Актуальность программы. обусловлена необходимостью обучения подростков правилам и методам безопасного использования информационного пространства, работы с информацией и защите от цифровых угроз, так как в связи с глобальной цифровизацией общества и развитием сети «Интернет» опасность таких угроз и вредоносной информации с каждым днем лишь возрастает. Также наука о данных и производные от нее профессии в настоящее время находятся на пике популярности. Благодаря обучению на данной программе подростки могут познакомиться с азами обучения искусственного интеллекта, изучить базовые аспекты программирования, лучше подготовиться к сдаче ЕГЭ и поступлению в ВУЗ, определиться с выбором профессионального развития. А принимая во внимание стремительное развитие технологий искусственного интеллекта, актуальность знаний и умений, приобретаемых обучающимися во время обучения, будет только возрастать.

Новизна программ

Большие данные и машинное обучение - стремительно развивающееся направление IT технологий. Современных детей привлекает сфера применения технологии и мотивирует к изучению курса и разработке программных продуктов.

Отличительная особенность программы состоит в изучении отдельного модуля проектной деятельности, в рамках которого обучающиеся развивают softкомпетенции, приобретают умения командного взаимодействия, работы над проектами, поскольку данные навыки играют все большее значение в современном обществе, культуре и профессиональной среде, также обучающиеся учатся применению методик Scrum и Agile, востребованные при обучении в СПО и необходимые любому современному разработчику. Кроме того, в процессе обучения, подростки углубленно изучают раздел кибергигиены. В процессе освоения модуля, обучающиеся не только научатся

анализировать входящий информационный поток и выявлять негативную информацию, но и настраивать защиту компьютеров и компьютерных сетей, работать с зараженными и вредоносными объектами, узнают причины и методики взломов и проникновений в личные данные пользователей, научатся выявлять и ликвидировать данные нарушения.

Педагогическая целесообразность данной программы заключается в том, что она является целостной и непрерывной в течение всего процесса обучения, позволяет учащимся шаг за шагом раскрывать в себе творческие возможности и самореализоваться в современном мире. В ходе практических занятий по программе дети познакомятся с виртуальной реальностью, поймут ее особенности и возможности, выявят возможные способы применения; а также определяют наиболее интересные направления для дальнейшего углубления.

Адресат программы

Дополнительная общеобразовательная программа «Большие данные и машинное обучение» ориентирована на детей 15-17 лет.

Форма обучения: очная, с применением дистанционных технологий (по необходимости).

Численность обучающихся в группе: 10-12 человек.

Объем программы: 72 часов.

Срок освоения программы: 9 месяцев (учебный год).

Режим занятий, периодичность и продолжительность занятий

Программа реализуется по 2 академических часа 1 раз в неделю в форме практических и лабораторных занятий, объединенных в тематические кейсы. 1 академический час – 40 минут с перерывом между часами – 10 минут.

Состав группы обучающихся - постоянный.

Уровень начальной подготовки: необходимы базовые знания языка программирования Python: синтаксис, алгоритмы, структуры, паттерны.

Форма работы: предусмотрено проведение комбинированных занятий.

При проведении занятий используются следующие основные формы работы:

- 1) демонстрационная – обучающиеся слушают объяснения педагога и наблюдают за демонстрационным экраном или экранами компьютеров на ученических рабочих местах;
- 2) фронтальная – обучающиеся синхронно работают под управлением педагога;
- 3) групповая и командная работа – обучающиеся совместно выполняют практические задачи, проект или кейс;
- 4) самостоятельная – обучающиеся выполняют индивидуальные задания в течение части занятия или нескольких занятий.
- 5) дистанционное обучение;

Особенности организации образовательного процесса

Большинство заданий выполняется с помощью компьютера и необходимых программных продуктов. Базовым форматом образовательного процесса является проектная деятельность. Процесс обучения включает решение лабораторных работ. Каждое занятие курса посвящено решению конкретной задачи. С этой целью педагог

дополнительного образования представляет теоретический материал и демонстрирует решение подобной задачи с помощью программных средств. Изучение нового материала носит сопровождающий характер. Обучающиеся изучают его с целью создания запланированного продукта. Каждое занятие предполагает выполнение практического задания. Завершающий этап изучения – решение итоговой лабораторной работы.

Программа является практико-ориентированной.

ЦЕЛЬ И ЗАДАЧИ ПРОГРАММЫ

Цель программы: формирование у обучающихся навыков безопасного использования ПК и сети «Интернет» и применения инструментов науки о данных для решения практических задач машинного обучения и настройки нейронных сетей. Для успешной реализации поставленной цели необходимо решить ряд обучающих, развивающих и воспитательных задач.

Задачи программы:

Обучающие:

- Обучить методам анализа данных
- Обучить основам статистики
- Обучить строить простые модели предсказаний

Развивающие:

- формировать конструктивное отношение к инженерной работе и развивать умение командной работы, координацию действий;
- расширять кругозор и культуру, межкультурную коммуникацию;
- активизировать умение анализировать возможности программного обеспечения.

Воспитательные:

- сформировать интерес к техническим наукам и, в частности, к технологиям искусственного интеллекта;
- развивать у учащихся память, внимание, логическое, пространственное и аналитическое мышление; креативность и лидерство;
- выявлять и развивать способности к исследовательской деятельности
- воспитывать уважение к интеллектуальному и физическому труду.

УЧЕБНЫЙ ПЛАН

№	Тема занятия/модуля	Кол-во часов			Формы аттестации, контроля
		Всего	теория	практика	
1	Вводный модуль	4	2	2	
1.1	Инструктаж по ТБ, знакомство с оборудованием, знакомство с системой оценивания	2	2	0	Беседа, контрольные вопросы
1.2	Введение в data science	2	2	0	Беседа, контрольные вопросы
2	ВВЕДЕНИЕ В АНАЛИЗ ДАННЫХ	8	4	4	
2.1	Основные объекты Pandas, Ввод вывод данных,	2	1	1	

	условия				
2.2	Группировка данных	2	1	1	
2.3	Агрегирующие функции, Сводные таблицы	4	2	2	Контрольное задание
3	РАБОТА С ТАБЛИЦАМИ И ПОДГОТОВКА ДАННЫХ	4	2	2	
3.1	Объединение таблиц	2	1	1	
3.2	Предварительная обработка данных	2	1	1	Контрольное задание
4	ВИЗУАЛИЗАЦИЯ ДАННЫХ	8	4	4	
4.1	Графические возможности Pандас	2	1	1	
4.2	Графические возможности SEABORN	2	1	1	
4.3	Практика на реальных данных	4	2	2	Контрольное задание
5	ЛИНЕЙНАЯ АЛГЕБРА С NUMPY	4	2	2	
5.1	Массивы NUMPY	2	1	1	
5.2	Знакомство с линейной алгеброй	2	1	1	

6	СТАТИСТИКА И ТЕОРИЯ ВЕРОЯТНОСТИ	12	6	6	
6.1	Генеральная совокупность. Типы переменных.	2	1	1	
6.2	Меры центральной тенденции. Меры изменчивости.	2	1	1	
6.3	Квартили, box-plot	2	1	1	
6.4	Нормальное распределение, центральная предельная теорема.	2	1	1	
6.5	Доверительные интервалы. Р уровень значимости.	4	2	2	Контрольное задание
7	ВВЕДЕНИЕ В МАШИННОЕ ОБУЧЕНИЕ	4	4	0	
7.1	Задачи машинного обучения	2	2	0	Беседа, контрольные вопросы
7.2	Этапы разработки, переобучение, знакомство с Sklearn	2	2	0	Беседа, контрольные вопросы
8	ЛИНЕЙНАЯ РЕГРЕССИЯ	8	3	4	

8.1	Описание линейной регрессии, метрики качества	2	1	1	
8.2	Разбор задачи линейной регрессии	2	1	1	
8.3	Практика на реальных данных	4	1	3	Контрольное задание
9	ЛОГИСТИЧЕСКАЯ РЕГРЕССИЯ	8	3	4	
9.1	Описание логистической регрессии, метрики качества	2	1	1	
9.2	Разбор задачи логистической регрессии	2	1	1	
9.3	Практика на реальных данных	4	1	3	Контрольное задание
10	ДЕРЕВО РЕШЕНИЙ И СЛУЧАЙНЫЙ ЛЕС	8	4	4	
10.1	Введения в решающие деревья	2	1	1	
10.2	Обучение, переобучение, кроссвалидация	2	1	1	
10.3	Подбор параметров	2	1	1	
10.4	Практика на реальных данных	2	1	1	Контрольное задание
11	ИТОГОВЫЙ ПРОЕКТ	4	0	4	
11.1	Разработка проекта	2	0	2	
11.2	Защита проекта	2	0	2	
Итого		72	34	38	

СОДЕРЖАНИЕ УЧЕБНОГО ПЛАНА

Модуль 1. Вводный модуль (4 часа)

- 1.1 Вводный инструктаж по технике безопасности, знакомство с оборудованием IT куба, знакомство с системой оценивания и матрицей компетенций. оценкой и технологией перевода детей на Базовый модуль. Знакомство с детскими проектами квантума
- 1.2 Знакомство с data science. Примеры проектов.

Модуль 2. ВВЕДЕНИЕ В АНАЛИЗ ДАННЫХ (8 часов)

- 2.1 Познакомимся с библиотекой для обработки и анализа данных Pandas. Научимся получать данные из внешнего файла и выводить эту информацию на экран.
- 2.2 Познакомимся с группировкой данных и поймем, зачем она нужна.
- 2.3 Попробуем различные агрегации, создадим первые сводные таблицы. Выполнение контрольного задания.

Модуль 3. РАБОТА С ТАБЛИЦАМИ И ПОДГОТОВКА ДАННЫХ (4 часов)

- 3.1 Как объединить таблицы. Трудности объединения датафреймов. Удаление дубликатов строк. Практика
- 3.2 Метод read_csv. Простая фильтрация, Функция query, Функции str.match и str.contains, пропуски и их типы, поиск пропусков, удаление и их замена. Выполнение

контрольного задания.

Модуль 4. ВИЗУАЛИЗАЦИЯ ДАННЫХ (8 часов)

4.1 Графические возможности Pandas. Как строить через Pandas или другие библиотеки. Метод plot. Отображение нескольких графиков.

4.2 Знакомство с библиотекой Seaborn. Построение различных графиков в Seaborn. Сравнение с графиками Pandas. Обсуждение. Работа с категориальными признаками.

4.3 Практика на реальных данных. Выполнение контрольного задания. Обсуждение, рефлексия.

Модуль 5. ЛИНЕЙНАЯ АЛГЕБРА С NUMPY (4 часа)

5.1 Знакомство с библиотекой numpy. Создание массива в numpy. Индексирование массивов. Особенности массивов из numpy.

5.2 Знакомство с линейной алгеброй. Векторы. Основные понятия. Базовые операции над векторами. Матрицы. Основные понятия. Сложение и вычитание матриц.

Умножение матрицы на число. Транспонирование матриц. Умножение матриц. Тренажер.

Модуль 6. СТАТИСТИКА И ТЕОРИЯ ВЕРОЯТНОСТИ (12 часов)

6.1 Понятие генеральной совокупности и выборки, репрезентативность выборки Простая случайная выборка Стратифицированная выборка Групповая выборка Типы переменных Количественные переменные Номинативные переменные Ранговые переменные

6.2 Понятие описательной статистики Мода Медиана Среднее значение Выбор меры центральной тенденции Свойства среднего Понятие меры изменчивости данных Размах Дисперсия, стандартное отклонение Свойства дисперсии и стандартного отклонения

6.3 Квантили распределения Квартили Box plot. Практическое задание

6.4 Понятие нормального распределения Стандартизация Правила двух и трех сигм, использование стандартизации Центральная предельная теорема. Выполнение практического задания.

6.5 Доверительные интервалы. Построение доверительных интервалов. Статистическая проверка гипотез Идея статистического вывода р-уровень значимости и его интерпретация. Практические задачи.

Модуль 7. ВВЕДЕНИЕ В МАШИННОЕ ОБУЧЕНИЕ (4 часа)

7.1 Задачи машинного обучения. Двоичная классификация. Входные и выходные данные при обучении. Регрессия. Кластеризация. Классификация.

7.2 Определение целей Сбор и подготовка данных. Разработка модели. Тестирование модели. Проблема переобучения. Как бороться с переобучением. Знакомство с библиотекой sklearn

Модуль 8. ЛИНЕЙНАЯ РЕГРЕССИЯ (8 часов)

8.1 Практическая задача линейной регрессии. Линейная регрессия. Применение линейной регрессии. Задача машинного обучения. Метрики качества линейной регрессии.

8.2 Для разбора возьмем один из самых популярных датасетов международной платформы соревнований по машинному обучению Kaggle. Этот набор данных был специально собран для тех, кто делает первые шаги в Data Science. Внутри датасета признаки описывают почти каждый аспект жилых домов в городе Эймс, штата Айова (США).

8.3 На примере реальных данных необходимо обучить модель регрессии. Оформить результат в виде блокнота в jupyter notebook, описать каждый этап работы. Описать выводы. Рефлексия.

Модуль 9. ЛОГИСТИЧЕСКАЯ РЕГРЕССИЯ (8 часов)

9.1 Введение в логистическую регрессию. Описание логистической регрессии. Метод максимального правдоподобия. Отступ объекта.

9.2 Для работы возьмем один популярный датасет Kaggle — bank_full.csv. Этот набор данных специально собран для тех, кто делает первые шаги в Data Science. Внутри датасета признаки описывают клиента банка и взаимодействие между банком и клиентом.

9.3 На примере реальных данных необходимо обучить модель регрессии. Оформить результат в виде блокнота в jupyter notebook, описать каждый этап работы. Описать выводы. Рефлексия.

Модуль 10. ДЕРЕВО РЕШЕНИЙ И СЛУЧАЙНЫЙ ЛЕС (8 часов)

10.1 Знакомство с деревом решений. Как строится дерево решений. Алгоритм построения дерева. Основные параметры дерева.

10.2 Реализация модели дерева решений для задач классификации. Обучение деревьев. Переобучение деревьев. Визуализация деревьев. Разбор кроссвалидации

10.3 Подбор параметров решающих деревьев. Разбор работы GridSearchCV и RandomizedSearchCV. Подбор параметров модели.

10.4 На примере реальных данных необходимо обучить модель классификации. Оформить результат в виде блокнота в jupyter notebook, описать каждый этап работы. Описать выводы. Рефлексия.

Модуль 11. ИТОГОВЫЙ ПРОЕКТ (4 часа)

5.1 Доработка проекта, подготовка к защите.

Тестирование модели, внесение правок, разработка презентации проекта. Работа в команде.

5.2 Защита проекта.

Защита проекта в формате презентаций либо стендовых защит. Ответы на вопросы. Рефлексия.

ПЛАНИРУЕМЫЕ РЕЗУЛЬТАТЫ

Обучающиеся научатся знать основы программирования, включая стандартные алгоритмы, и их реализацию в Python; уметь строить модели, оценивать их качество и сравнивать различные модели средствами Python; иметь навыки анализа реальных социологических данных в Python. Обучающиеся получают возможность научиться: уметь работать с большими объёмами данных и понимать их значимость в информационном пространстве; Прохождение данной образовательной программы должно сформировать у обучающихся следующие метакомпетенции (*soft skills*) ([Приложение 1](#)):

- критическое мышление (способность анализировать, оценивать идеи и решения, задавать правильные вопросы, аргументировать);
- креативность (способность разработать и представить принципиально новые подходы к решению ситуации или проблемы);
- коммуникация (способность выразить и понимать мысли, чувства других людей в устной и письменной форме);
- кооперация (эффективное взаимодействие с другими людьми, результативная работа в команде).

Прохождение образовательного модуля должно сформировать у обучающихся следующие профильные компетенции (*hard skills*) ([Приложение 2](#)), которые могут быть применены в ходе реализации проектов в последующих образовательных модулях:

- умение анализировать данные с помощью Pandas,
- умение работать с таблицами и подготавливать данные,

- умение работы с массивами в NumPy,
- умение создавать модели линейной и логистической регрессии
- умение решать задачи регрессии и классификации

Раздел 2. Комплекс организационно-педагогических условий, включающий формы аттестации

Календарный учебный график на 2022-2023 учебный год

Год обучения (уровень обучения)	Дата начала обучения по программе	Дата окончания обучения по программе	Сроки проведения аттестации обучающихся	Количество учебных недель	Количество учебных часов	Режим занятий
Вводный	05.09.2022	31.05.2023	32 неделя	34	72	1 раза в неделю по 2 часа

ФОРМЫ АТТЕСТАЦИИ

Реализуется текущий контроль и промежуточная аттестация обучающихся.

Формы текущего контроля включают индивидуальную оценку выполненных заданий, участие в конкурсах, а также участие в индивидуальных, командных и межгрупповых соревнованиях. Формы промежуточной аттестации учитывают данные текущего контроля, а также освоение и защиту некоторых этапов проектов командами обучающихся. На каждого обучающегося формируется индивидуальная карта, отражающая количество баллов, набранных за период обучения ([Приложение 3](#)).

ОЦЕНОЧНЫЕ МАТЕРИАЛЫ

Текущий контроль, как проверка учебных достижений, теоретических знаний и практических навыков, производится в ходе осуществления образовательной деятельности согласно учебному плану.

Промежуточная аттестация, как оценка уровня достижения результатов освоения программы обучающимися, производится в соответствии с критериями и показателями «Матрицы soft и hard компетенций» (Приложение 1, 2) в конце каждого полугодия в соответствии с учебным планом. Количество баллов, набранных обучающимися согласно «Матрице soft и hard компетенций», определяет уровень успешности освоения содержания настоящей программы, и является критерием перевода на следующий уровень программы по данному направлению при наличии вакантных мест.

Минимальное количество баллов, которое должен получить обучающийся составляет 24 балла, а максимальное количество баллов – 96. Дополнительно учащийся может получить еще 4 балла за результативность (например, получение высоких баллов в олимпиадах, конкурсах) и волонтерство.

МЕТОДИЧЕСКИЕ МАТЕРИАЛЫ

Методическое обеспечение программы включает кейсы (Приложение 4), раздаточный материал, необходимый для проведения лабораторных и практических работ.

УСЛОВИЯ РЕАЛИЗАЦИИ ПРОГРАММЫ

Материально-техническое обеспечение

В ходе реализации программы используется следующее оборудование:

1. Стационарный программно-аппаратный комплекс для изучения программирования на базе BasicRAY B161: процессор 4 ядра частотой 2,7 ГГц, ОП 8 Гб, видеокарта - 12 шт.
2. МФУ - 1 шт.
3. Проектор - 1 шт.
4. Демонстрационная доска - 1 шт.
5. Программное обеспечение: программы Anaconda Python

Информационное обеспечение

Информационные и учебно-методические ресурсы представлены презентациями и видеороликами.

Для более глубокого изучения осваиваемой темы предлагаются ссылки на электронные ресурсы и печатные издания (см. Список литературы).

Кадровое обеспечение

Освоение программы обеспечивает педагог дополнительного образования, имеющий высшее образование в профильной области или педагогике, прошедший обучение на курсах повышения квалификации педагогов-наставников.

СПИСОК ЛИТЕРАТУРЫ

Список литературы для педагога

1. Белоусов А. Д. Угрозы сети. Интернет для несовершеннолетних пользователей: психологический анализ и профилактика. – М.: «Проспект», 2021.
2. Джастин Зейтц. Black Hat Python, Программирование для хакеров и пентестеров. – СПб.: Питер, 2022 – 224 с.
3. Ли Воган. «Непрактичный» Python занимательные проекты для тех, кто хочет поумнеть. – СПб.: БХВ-Петербург, 2021. – 457 с.;
4. Сет Вейдман. Глубокое обучение: легкая разработка проектов на Python – СПб.: БХВ-Петербург, 2021. – 272 с.;
5. Солдатова Г. У., Чигарькова С. В., Дренёва А. А., Илюхина С. Н. Мы в ответе за цифровой мир: Профилактика деструктивного поведения подростков и молодежи в Интернете: Учебно-методическое пособие. – М.: Когито-Центр, 2019. – 176 с.

Список литературы для обучающихся и родителей

1. Джейми Чан. Python Быстрый старт., 352 стр. 2021 г. – СПб.: Питер, 2022 – 224 с;
2. Дэниел Г. Грэм Этичный Хакинг, практическое руководство по взлому. – СПб.: Питер, 2022 – 384 с.; 9
3. Сафронов Е.В. Азы кибергигиены. Методологические и правовые аспекты. Издательство «Проспект», 2020 г.; – 44 с.;
4. Солдатова Г. У., Чигарькова С. В., Пермякова И. Д. Тренажёр по курсу Кибербезопасность для 8 класса общеобразовательных организаций. Издательство Русское слово — учебник, Москва, 2020. — 80 с.

Ссылки на Интернет ресурсы.

1. Code Basics: обучение базовым аспектам языков программирования от образовательной платформы Hexlet. // [Электронный ресурс] URL:

<https://ru.codebasics.com/> (дата обращения: 20.04.2021);
2. Kaggle: Your Machine Learning and Data Science Community. [Электронный ресурс] URL: <https://www.kaggle.com/> (дата обращения: 30.03.2022);

Матрица метакомпетенций (soft skills) обучающихся детских технопарков

Наименование компетенции	Показатели проявления компетенций по уровням				Шкала оценивания
	Уровень 1 Пассивный участник	Уровень 2 Ведомый	Уровень 3 Инициатор	Уровень 4 Стратег	Максимальный балл за показатель
1	2	3	4	5	6
Креативность (способность разработать и представить принципиально новые подходы к решению ситуации или проблемы)					12 баллов
<i>1. Включение в деятельность</i>	1 балл	2 балла	3 балла	4 балла	4
Схватывает, исследует, создает и предлагает разные идеи и подходы	Участствует в обсуждении задания	Задает вопросы на понимание задания. Развивает предложенные кем-то идеи.	Предлагает идеи, развивающие понимание заданий, Предлагает свои идеи	Меняет, развивает предложения с учетом ситуации	
<i>2. Участие в решении</i>	1 балл	2 балла	3 балла	4 балла	4
Воспринимает, наблюдает за деятельностью группы. Исследует, учитывает подходящий опыт и информацию. Находит оригинальное решение. Устанавливает связи, интегрирует знания из разных областей для решения проблемы.	Выполняет порученную часть работы	Ищет способ приложить базовые умения к нестандартной ситуации, выделяет известное и неизвестное	Находит аналогичную ситуацию, привлекает свой опыт. Отмечает значимые факторы и условия. Формулирует вопросы по ходу решения. Находит оригинальное решение	Объясняет, обосновывает ход решения, очерчивает границы задания. Реагирует на разные идеи и решения. Корректирует неправильное или неэффективное решение.	
<i>3. Презентация результатов</i>	1 балл	2 балла	3 балла	4 балла	4
Видит новые интересные	Доволен своей	Сравнивает	Оценивает результат	Признает свои	

решения проблемы и понимает их возможные последствия.	работой	результаты своей работы с другими.	на основе критериев. Отмечает наиболее интересные другие идеи	ограничения. Видит возможности для улучшения.	
---	---------	------------------------------------	---	---	--

Критическое мышление (способность анализировать, оценивать идеи и решения, задавать правильные вопросы, аргументировать)					12 баллов
---	--	--	--	--	-----------

<i>1. Включение в деятельность</i>	1 балл	2 балла	3 балла	4 балла	
Анализирует, определяет сильные и слабые стороны ситуации/решения. Аргументирует свои идеи и решения.	Слушает, разделяет мнение большинства участников в команде.	Задаёт вопросы на понимание задания, уточняет. Развивает предложенные кем-то идеи.	Предлагает свои идеи. Контролирует ход дискуссии, обсуждения проблемы.	Анализирует идеи других, предлагает свои решения. Аргументирует и опирается на факты, критерии	4
<i>2. Участие в решении</i>	1 балл	2 балла	3 балла	4 балла	
Понимает суть и границы проблемного поля. Анализирует и сравнивает идеи/решения, аргументирует свои идеи. Контролирует точность выполнения задания.	Выполняет порученную часть работы по принципу "Копировать-вставить", не подвергая информацию анализу	Четко выделяет известное и необходимое для решения задачи/проблемы, старается анализировать информацию	Выслушивает чужое мнение, соглашается с аргументами. Формулирует решение по ходу выполнения задачи. Находит оригинальное решение. Инициатор выполнения и контроля задания	Объясняет и обосновывает ход решения проблемы. Ограничивает круг проблем, вычленяет эффективные решения.	4
<i>3. Презентация результатов</i>	1 балл	2 балла	3 балла	4 балла	
Оценивает, подтверждает мнение на основе критериев, фактов. Оценивает мнение/решение.	Доволен своей работой без критического отношения к ее результатам	Сравнивает результаты своей работы с другими. Оценивает личный вклад в общее дело.	Оценивает результат на основе критериев. Отмечает интересные идеи/решения.	Способен обобщать и выделять суть проблемы. Видит перспективы проекта/решения. Легко владеет материалом проекта.	4

				Отвечает на вопросы грамотно, аргументированно. Способен критически оценивать результаты работы команды.	
--	--	--	--	--	--

Коммуникация (способность выражать и понимать мысли, чувства других людей в устной и письменной форме)					12 баллов
<i>1. Включение в деятельность</i>	1 балл	2 балла	3 балла	4 балла	
Задаёт и отвечает на вопросы. Понятно доносит свои идеи и решения. Спрашивает, уточняет, понимает суть проблемы. Эрудирован. Готов общаться со всеми членами команды.	Внимательно слушает, пассивно участвует в обсуждении.	Имеет свое мнение и пытается отстаивать свою позицию.	Формулирует вопросы, уточняет и проясняет ситуацию. Способен к быстрому восприятию и переработке информации. Большой словарный запас.	Способен урегулировать конфликт в команде. Уважает мнение других. Владеет разными способами работы с информацией: получение, анализ, хранение, доступное изложение своих идей и мыслей. Быстро обучается.	4
<i>2. Участие в решении</i>	1 балл	2 балла	3 балла	4 балла	4
Способен регулировать конфликты и споры в команде. Уважительно относится к мнению других. Доступно доносит свою идею, позицию. Отлично владеет способами получения, анализа и переработки, хранения информации и предоставления аргументированного решения.	Малоинициативен, но исполнитель.	Эффективно работает при четком понимании задания. Инициативен. Имеет свою позицию. Не всегда удается понятно и доступно донести свои мысли.	Легко ориентируется в поиске необходимой и недостающей информации. Хорошо работает в четко очерченном пространстве идеи или задачи.	Успешно взаимодействует в команде. Готов общаться с каждым членом команды. Уважает мнение других. Способен аргументировать свою четкую позицию. Быстро	

				адаптируется к собеседнику/ситуации.	
--	--	--	--	--------------------------------------	--

<i>3. Презентация результатов</i>	1 балл	2 балла	3 балла	4 балла	4
Умеет составить текст сообщения, доклада. Умеет хорошо и правильно говорить/писать. Способен выделить главное и второстепенное в работе команды.	Помогает команде, в поиске нужной информации, выполняет поручения добросовестно.	Не всегда может выделить главное и второстепенное при презентации результатов команды. Старателен. Может представить результаты команды по за-ранее подготовлен-ному тексту другим человеком.	Умеет хорошо и правильно говорить с опорой на записи. Не всегда свободно владеет материалом проекта. Позитивная самооценка от результата совместной работы.	Самостоятельно и грамотно составляет доклад, презентацию. Речь правильная, грамотная. Высокий уровень эрудиции. Способен вести диалог, отвечать на вопросы и прояснять ситуацию.	

Кооперация (эффективное взаимодействие с другими людьми, результативная работа в команде)					12 баллов
--	--	--	--	--	-----------

<i>1. Включение в деятельность</i>	1 балл	2 балла	3 балла	4 балла	4
Социальное взаимодействие	Участвует в обсуждении задач вынуждено, только если к нему/ней обращаются	Задает вопросы на понимание, спрашивает непонятное, поддерживает идеи других членов группы без критического к ним отношения	Предлагает свои идеи по сути задачи. Контролирует выполнение задачи	Учитывает предложения членов команды, развивает предложения других, отвечает на вопросы по сути задания	
<i>2. Участие в решении</i>	1 балл	2 балла	3 балла	4 балла	4
Выполнение обязательств	Выполняет порученную часть работы	Ищет способы приложить свои умения к решению задачи, может помочь другим по их	Предлагает решение по аналогии с другими задачами. Находит оригинальное	Объясняет и обосновывает ход решения, удерживает границы задачи, реагирует на разные	

		просьбе	решение задачи. Иницирует контроль и проверку решения. Активно включается в помощь членам команды	решения, старается выбрать оптимальное решение	
--	--	---------	---	--	--

<i>3. Презентация результатов</i>	1 балл	2 балла	3 балла	4 балла	4
Самостоятельность и инициативность при оценке и презентации результатов работы команды	Помогает готовить презентацию работы команды. Принимает оценку своей работы, высказанную другими	Участвует в распределении работ при подготовке презентации. Стараются оценить свою работу и вклад в общий результат	Берет на себя ответственность по подготовке и презентации работы команды. Вносит предложения по оценке результатов работы команды и каждого его члена по определенным критериям	Отвечает на вопросы о ходе и результатах работы команды, приглашает к ответу других участников. Определяет вклад в достижение результат каждого члена команды	
Итого					48 баллов

Матрица профильных компетенций (Hard skills) обучающихся по программе «Большие Данные и Машинное Обучение»

Наименование компетенции	Показатели проявления компетенций по уровням				Шкала оценивания
	<i>Уровень 1 Начинающий</i>	<i>Уровень 2 Ученик</i>	<i>Уровень 3 Студент</i>	<i>Уровень 4 Мастер</i>	Максимальные баллы за каждый показатель
Анализ данных и машинное обучение	Умение анализировать данные и создавать собственные модели				16 баллов
<i>1. Работа с данными</i>	2 балла	4 балла	6 баллов	8 баллов	
Умение анализировать данные с помощью Pandas	Разбирается в библиотеке, знает основные функции	Разбирается в библиотеке, знает основные функции, умеет применять к данным	Разбирается в библиотеке, знает основные функции, умеет применять к данным, умеет объединять таблицы	Разбирается в библиотеке, знает основные функции, умеет применять к данным, умеет объединять таблицы, умеет строить сводные таблицы, делать выводы из полученных данных.	8
<i>2. Подготовка данных</i>	2 балла	4 балла	6 баллов	8 баллов	
Умение работать с таблицами и подготавливать данные	Умеет фильтровать таблицы	Умеет фильтровать таблицы, группировать данные	Умеет фильтровать таблицы, группировать данные, знает множество агрегаций, применяет осмысленно	Умеет фильтровать таблицы, группировать данные, знает множество агрегаций, применяет осмысленно, получает результаты из данных, умеет убирать пропуски из данных	8
<i>3.Создание линейных моделей</i>	2 балла	4 балла	6 баллов	8 баллов	

умение создавать модели линейной и логистической регрессии	Умеет создавать простые модели	Умеет создавать простые модели, понимает когда нужно применять ту или иную модель	Умеет создавать простые модели, понимает когда нужно применять ту или иную модель, умеет подбирать параметры модели	Умеет создавать простые модели, понимает когда нужно применять ту или иную модель, умеет подбирать параметры модели, умеет не переобучать модель, знает аналогичные классификаторы и умеет их использовать	8
<i>4. Работа с numpy</i>	2 балла	4 балла	6 баллов	8 баллов	
Умение работы с массивами в Numpy	Умеет создавать, обрабатывать массивы	Умеет создавать, обрабатывать массивы, применять арифметические операции	Умеет создавать, обрабатывать массивы, применять арифметические операции, функции	Умеет создавать, обрабатывать массивы, применять арифметические операции, функции, переходить от одного типа к другому, работать с N-мерными массивами	8
<i>5. Работа с моделями</i>	2 балла	4 балла	6 баллов	8 баллов	
умение решать задачи регрессии и классификации	Умеет решать задачи регрессии	Умеет решать задачи регрессии, классификации	Умеет решать задачи регрессии, классификации, score >0.6	Умеет решать задачи регрессии, классификации, score >0.8	8
				ИТОГО	40 баллов