

ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ АВТОНОМНОЕ  
ОБРАЗОВАТЕЛЬНОЕ УЧРЕЖДЕНИЕ ВЫСШЕГО ОБРАЗОВАНИЯ  
«НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ УНИВЕРСИТЕТ  
«ВЫСШАЯ ШКОЛА ЭКОНОМИКИ»  
НИУ ВШЭ  
факультет компьютерных наук  
центр непрерывного образования

УТВЕРЖДАЮ  
Заместитель первого проректора НИУ ВШЭ  
Г.В. Можаяева

УЧЕБНО-ТЕМАТИЧЕСКИЙ ПЛАН  
дополнительная профессиональная программа – программа профессиональной переподготовки  
для получения компетенций, необходимых для выполнения нового вида профессиональной деятельности  
«Специалист по Data Science»

Учебный год: 2024/2025  
Трудоемкость программы: 21 зач. ед., 714 академических час. (в том числе 440 конт. час.).  
Форма обучения: Очная  
Формат ДОП: Смешанный

№ п/п	Названия учебных модулей, дисциплин	Трудоемкость		Объем контактной работы						Обучение с использованием ДОТ и (или) ЭО, в том числе в полном объеме (акад. ч.)	Самостоятельная работа (акад.ч.)	Форма промежуточной аттестации
		в зач. ед.	в акад. ч.	всего, конт. ч	Аудиторная работа (акад.ч.)				Обучение с использованием ДОТ и (или) ЭО, в том числе в полном объеме (акад. ч.)			
					всего, ауд. ч.	лекции	практические, семинарские, лабораторные занятия	консультации, иные виды занятий		синхр.		
1	2	3	4	5	6	7	8	9	10	11	12	13
1.1	Python для автоматизации и анализа данных	3	102	72	72	36	36				30	Зачет
1.1.1	Введение в Python, Jupyter. Ввод-вывод, основные типы данных (int, float, str). Первые функции.		4	4	4	2	2					
1.1.2	Логические переменные, операции и операторы сравнения. Условия if-elif-else. Условный цикл while. Списки, кортежи, последовательности		4	4	4	2	2					

1.1.3	Множества. Словари. Вложенные структуры данных. Методы строк, списков. Срезы		4	4	4	2	2					
1.1.4	Цикл for. Вложенный цикл for		6	4	4	2	2				2	
1.1.5	Функции. Рекурсия. О-нотация		6	4	4	2	2				2	
1.1.6	Текстовые файлы и таблицы		6	4	4	2	2				2	
1.1.7	Регулярные выражения		6	4	4	2	2				2	
1.1.8	Классы и основы ООП. Декораторы		6	4	4	2	2				2	
1.1.9	Telegram-бот на Python. Работа с SQL и базами данных с помощью Python		6	4	4	2	2				2	
1.1.10	Введение в numpy. Работа с векторами и матрицами. Введение в pandas		6	4	4	2	2				2	
1.1.11	Pandas продолжение. Разведывательный анализ данных		6	4	4	2	2				2	
1.1.12	Сбор данных: скрейпинг, requests, BeautifulSoup		6	4	4	2	2				2	
1.1.13	Сбор данных: requests, BeautifulSoup – продолжение		6	4	4	2	2				2	
1.1.14	Сбор данных: работа с сервисами через API		6	4	4	2	2				2	
1.1.15	Сбор данных: Selenium, работа с сервисами через API		6	4	4	2	2				2	
1.1.16	Визуализация для презентации данных: matplotlib, seaborn		6	4	4	2	2				2	
1.1.17	EDA особенности		6	4	4	2	2				2	
1.1.18	Предобработка текстов: лемматизация, удаление стоп-слов		6	4	4	2	2				2	
1.2	SQL	1	34	20	20	10	10				14	Зачет
1.2.1	Введение в базы данных. Обращение к базе данных. Типы данных. Фильтрация данных		8	4	4	2	2				4	
1.2.2	Агрегирующие функции. Функции работы с датами. Функции работы со строками		8	4	4	2	2				4	
1.2.3	Взаимотношения между таблицами. Операторы JOIN и UNION		6	4	4	2	2				2	
1.2.4	Подзапросы и временные таблицы		6	4	4	2	2				2	
1.2.5	Оконные функции		6	4	4	2	2				2	
1.3	Алгоритмы и структуры данных	2	68	40	40	20	20				28	Зачет
1.3.1	Асимптотика. Нотация O-большое. Измерение времени работы программы.		7	4	4	2	2				3	
1.3.2	Базовые структуры данных (список, стек, дек, очередь).		6	4	4	2	2				2	

1.3.3	Базовые структуры данных (сет, мэл, хэш-таблицы).		6	4	4	2	2				2	
1.3.4	Сортировки: итеративные, линейные, рекурсивные. Поиск.		7	4	4	2	2				3	
1.3.5	Динамическое программирование.		7	4	4	2	2				3	
1.3.6	Алгоритмы на строках. Сравнение строк. Расстояние редактирования.		7	4	4	2	2				3	
1.3.7	Теория графов. Введение. Алгоритмы обхода графов.		7	4	4	2	2				3	
1.3.8	Поиск кратчайших путей в графе.		7	4	4	2	2				3	
1.3.9	Минимальное остовное дерево.		7	4	4	2	2				3	
1.3.10	Эвристические и приближенные алгоритмы.		7	4	4	2	2				3	
1.4	Математика для анализа данных	3	102	84	84	42	42				18	Зачет
1.4.1	Дискретная математика: множества и логика.		4	4	4	2	2					
1.4.2	Дискретная математика: комбинаторика и вероятность.		5	4	4	2	2				1	
1.4.3	Дискретная математика: неориентированные графы.		5	4	4	2	2				1	
1.4.4	Дискретная математика: ориентированные графы и алгоритмы на графах.		5	4	4	2	2				1	
1.4.5	Линейная алгебра: системы линейных уравнений, матрицы, операции, блочные операции, обратимости и невырожденности.		5	4	4	2	2				1	
1.4.6	Линейная алгебра: Определители ориентированные объемы, явные формулы обратной матрицы, характеристический многочлен, полиномиальное исчисление от матриц, спектр.		5	4	4	2	2				1	
1.4.7	Линейная алгебра: Векторные пространства и подпространства, размерности, ранги матриц. Свойства рангов и неравенства на ранги.		5	4	4	2	2				1	
1.4.8	Линейная алгебра: Линейные отображения и их матричное описание, смена координат. Образ и ядро, их геометрический смысл, связи на размерности. Инварианты линейного оператора: след, определитель, характеристический многочлен. Собственные значения и векторы, связь со спектром.		5	4	4	2	2				1	

1.4.9	Линейная алгебра: Билинейные формы. Квадратичные формы и симметричные билинейные формы. Сигнатура, ее геометрический смысл, методы определения сигнатуры. Связи с LU-разложением. Скалярные произведения, углы и расстояния. Ортогонализация и QR-разложение. Линейные многообразия и линейные классификаторы, отступы.		5	4	4	2	2				1	
1.4.10	Линейная алгебра: Операторы в евклидовых пространствах. Движения и ортогональные матрицы их классификация. Самосопряженные операторы и симметрические матрицы, их диагонализуемости. Сингулярное разложение (SVD). Нахождение SVD.		5	4	4	2	2				1	
1.4.11	Математический анализ: функции одной переменной, пределы, производные и их геометрический смысл, вычисление производных, касательные, критические точки, поиск минимумов и максимумов.		9	8	8	4	4				1	
1.4.12	Математический анализ: интегралы и ориентированные площади, скорости и расстояния, введение в вычисление интегралов, свойства интегрирования.		9	8	8	4	4				1	
1.4.13	Математический анализ: функции нескольких переменных, градиент, производная по направлению, линии уровня, касательная плоскости, критические точки, поиск минимумов и максимумов.		5	4	4	2	2				1	
1.4.14	Математический анализ: оптимизационные задачи, лагранжиан и его геометрический смысл, нахождение минимума или максимума с заданными ограничениями.		5	4	4	2	2				1	

1.4.15	Теория вероятностей: Пространство элементарных исходов. События. Вероятность и её свойства. Дискретное пространство элементарных исходов. Условная вероятность. Формула полной вероятности. Формула Байеса.		5	4	4	2	2				1	
1.4.16	Теория вероятностей: Дискретные случайные величины и их распределения. Примеры дискретных распределений: бернуллиевское, биномиальное, пуассоновское. Независимость случайных величин в дискретном случае. Распределение функции от дискретной случайной величины. Числовые характеристики распределений: математическое ожидание и дисперсия.		5	4	4	2	2				1	
1.4.17	Теория вероятностей: Случайные величины, имеющие плотности. Математическое ожидание случайной величины, имеющей плотность. Примеры распределений, имеющих плотность: равномерное, экспоненциальное, нормальное.		5	4	4	2	2				1	
1.4.18	Теория вероятностей: Функция распределения. Распределение функции от случайной величины, имеющей плотность. Многомерные случайные величины. Ковариация случайных величин.		5	4	4	2	2				1	
1.4.19	Теория вероятностей: условное распределение, свойства условного распределения, теоремы умножения для с.в., формула полной вероятности с.в., индикатор события, условное математическое ожидание, формула полной вероятности для математического ожидания, свойства условного математического ожидания, неравенство Маркова, неравенство Чебышева, правило трех сигм, ЗБЧ, ковариация, коэффициент корреляции, некоррелированные с.в., вариационная матрица, ЦПТ, неравенство Чернова.		5	4	4	2	2				1	

1.5	Прикладная статистика для машинного обучения	2	68	40	40	20	20				28	Зачет
1.5.1	Основы теории вероятностей. Описательные статистики.		7	4	4	2	2				3	
1.5.2	Взаимосвязи между случайными величинами. Взаимосвязи в данных.		6	4	4	2	2				2	
1.5.3	Методы оценивания параметров распределения. Свойства статистических оценок.		7	4	4	2	2				3	
1.5.4	Предельные теоремы. Асимптотические и точные доверительные интервалы.		7	4	4	2	2				3	
1.5.5	Процедура проверки гипотез. Параметрические критерии.		6	4	4	2	2				2	
1.5.6	Непараметрические критерии. Критерии согласия. Бутстреп, множественное тестирование.		7	4	4	2	2				3	
1.5.7	Введение в А/Б тестирование: постановка задачи, особенности сбора данных, взаимосвязь размера выборки, уровня значимости, мощности теста и MDE. Особенности метрик.		7	4	4	2	2				3	
1.5.8	Основные методы А/Б тестирования: стратификация, бакетное преобразование, линеаризация, CUPED.		7	4	4	2	2				3	
1.5.9	Статистический взгляд на линейную регрессию.		7	4	4	2	2				3	
1.5.10	Основы анализа временных рядов. Модель SARIMA и её подгонка.		7	4	4	2	2				3	
1.6	Машинное обучение	3	102	56	56	28	28				46	Зачет
1.6.1	Введение в машинное обучение. Основные понятия. Типы задач. Разбиение на train и test. Обучение и оценка качества модели		7	4	4	2	2				3	
1.6.2	Линейные методы регрессии. Функционал ошибки. Метрики качества. Переобучение		7	4	4	2	2				3	
1.6.3	Градиентный спуск		8	4	4	2	2				4	
1.6.4	Регуляризация в линейной регрессии. Кросс-валидация. Кодирование категориальных признаков		7	4	4	2	2				3	
1.6.5	Линейные методы классификации (бинарная). Логистическая регрессия. Метрики качества		7	4	4	2	2				3	
1.6.6	SVM. Многоклассовая классификация		7	4	4	2	2				3	

1.6.7	Нелинейные методы классификации. Байесовский классификатор. KNN. Kernel SVM.		7	4	4	2	2				3	
1.6.8	Деревья решений. Критерии информативности		7	4	4	2	2				3	
1.6.9	Ансамблевые методы. Случайный лес, бэггинг. Разложение ошибки на bias, variance		8	4	4	2	2				4	
1.6.10	Градиентный бустинг. XGBoost, CatBoost, LightGBM.		8	4	4	2	2				4	
1.6.11	Обучение без учителя: кластеризация		8	4	4	2	2				4	
1.6.12	Понижение размерности: PCA, TSNE, feature importance		7	4	4	2	2				3	
1.6.13	Рекомендательные системы		7	4	4	2	2				3	
1.6.14	Статистический взгляд на модели МО. Анализ временных рядов		7	4	4	2	2				3	
1.7	Промышленное машинное обучение на Spark	2	68	32	32	16	16				36	Зачет
1.7.1	Введение: как работают и где находятся большие данные.		8	4	4	2	2				4	
1.7.2	Среда Spark. Spark RDD / Spark SQL.		8	4	4	2	2				4	
1.7.3	Advanced SQL.		9	4	4	2	2				5	
1.7.4	Spark ML / Spark TimeSeries.		8	4	4	2	2				4	
1.7.5	Advanced ML и проверка результатов качества моделей.		9	4	4	2	2				5	
1.7.6	Spark GraphX /Spark Streaming.		8	4	4	2	2				4	
1.7.7	Экосистема Spark (MLFlow, AirFlow,H2O AutoML).		8	4	4	2	2				4	
1.7.8	Spark в архитектуре проекта / Spark CI/CD.		10	4	4	2	2				6	
1.8	Глубинное обучение	2	68	40	40	20	20				28	Зачет
1.8.1	Введение в глубокое обучение. От линейной регрессии к нейронной сети. Смотрим на базовые возможности tensorflow/pytorch, собираем первую нейросеть.		7	4	4	2	2				3	
1.8.2	Обучение нейронных сетей. Алгоритм обратного распространения ошибки.		7	4	4	2	2				3	
1.8.3	Оптимизация. Эвристики для обучения нейронных сетей. Нормализация по батчам, инициализация и др.		7	4	4	2	2				3	
1.8.4	Свёрточные нейронные сети. Классификация изображений.		6	4	4	2	2				2	

1.8.5	Архитектуры свёрточных сетей. Что видят свёрточные сети. Transfer learning.		7	4	4	2	2				3	
1.8.6	Обзор задач компьютерного зрения. Детекция, сегментация, перенос стиля, распознавание лиц.		6	4	4	2	2				2	
1.8.7	Автокодировщики и генеративные модели (Generative Adversarial Networks).		7	4	4	2	2				3	
1.8.8	Глубокое обучение для NLP. Векторные представления текстов: word2vec, fasttext.		7	4	4	2	2				3	
1.8.9	Рекуррентные нейронные сети, работа с последовательностями. ELMO-эмбеддинги.		7	4	4	2	2				3	
1.8.10	Sequence2sequence, архитектуры с механизмом внимания. Трансформеры. BERT.		7	4	4	2	2				3	
1.9	Прикладные задачи анализа данных	2	68	40					40		28	Зачет
1.9.1	Введение в цифровую обработку сигналов и классификация звуков.		8	4					4		4	
1.9.2	Автоматическое распознавание речи.		8	4					4		4	
1.9.3	Преобразование текста в речь.		8	4					4		4	
1.9.4	Введение в обработку текстов и классификация текстов.		8	4					4		4	
1.9.5	Языковое моделирование.		6	4					4		2	
1.9.6	Машинный перевод		6	4					4		2	
1.9.7	Распознавание лиц и метрическое обучение.		6	4					4		2	
1.9.8	Семантическая сегментация изображений.		6	4					4		2	
1.9.9	Детекция объектов на изображениях.		6	4					4		2	
1.9.10	Глубинное обучение в производстве: обслуживание и ускорение.		6	4					4		2	
	<b>ВСЕГО</b>	<b>20</b>	<b>680</b>	<b>424</b>	<b>384</b>	<b>192</b>	<b>192</b>		<b>40</b>		<b>256</b>	
<b>ИА</b>	<b>Итоговая аттестация:</b>											
1	Подготовка и защита аттестационной работы	1	34	16	16			16			18	
	<b>ИТОГО:</b>	<b>21</b>	<b>714</b>	<b>440</b>	<b>400</b>	<b>192</b>	<b>192</b>	<b>16</b>	<b>40</b>		<b>274</b>	