

Документ подписан простой электронной подписью

Информация о владельце:

ФИО: Волков В.В.

Должность: Ректор

Дата подписания: 14.07.2025 14:14:42

Уникальный программный ключ:

ed68fd4b85b778e0f0b1bfea5dbc56cf4148f1229917e799a70e51517ff6d591

**Автономная некоммерческая образовательная организация высшего образования
«Европейский университет в Санкт-Петербурге»**

УТВЕРЖДАЮ

Ректор

« 25 » июля 2025 г.

Протокол Ученого Совета

№ 6 от 25 июля 2025 г.



**УЧЕБНО-ТЕМАТИЧЕСКИЙ ПЛАН
программы профессиональной переподготовки
«Прикладной анализ данных»**

№ п/ п	Наименование и содержание тем	Основные понятия (категории) и проблемы, рассматриваемые в теме	Объем дисциплины час.			Форма промежуточн ой аттестации/ итоговой аттестации	
			Всего	Аудиторная работа по видам учебных занятий	СР ⁴		
				Л	СЗ ⁵		
Язык программирования Python							
1.	Что такое Python и почему мы выбрали именно его	История языка. Области применения. Дистрибутивы. Редакторы кода. Установка. Настройка. Лицензии на использование. Простые программы.	7	2	2	3	
2.	Что такое ноутбуки и как ими пользоваться	Что такое ноутбуки и как ими пользоваться Интерактивный режим. Понятие алгоритма. Переменные. Запуск Jupyter. Архитектура. Настройка.	7	2	2	3	
3.	Типы данных	Числа, строки, списки, словари, множества. Объекты. Методы объекта. Взаимосвязь алгоритма и структур данных	9	4	2	3	
4.	Циклы, функции, генераторы, list comprehension	Структурное программирование. Утиная типизация. Императивное и декларативное программирование.	10	2	4	4	
5.	Работа файлами различных форматов	Работа с распространенными форматами файлов: лог-файлы, бинарные файлы, CSV, JSON, XML. Модули для работы с ними. API. Работа с большими файлами, событийные	13	4	4	5	

⁴ Самостоятельная работа, включает в себя часы на текущий контроль

⁵ Могут включать в себя: лабораторные работы, круглые столы, мастер-классы, мастерские, деловые игры, ролевые игры, тренинги, семинары по обмену опытом, выездные занятия, консультации

№ п/ п	Наименование и содержание тем	Основные понятия (категории) и проблемы, рассматриваемые в теме	Объем дисциплины час.			Форма промежуточн ой аттестации/ итоговой аттестации	
			Всего	Аудиторная работа по видам учебных занятий	СР ⁴		
				Л	СЗ ⁵		
		обработчики. Деревья объектов файла с данными.					
6.	Функции и их свойства	Стандартная библиотека. Создание своих функций	11	4	2	5	
7.	Pandas.	Установка библиотеки. Работа с файлами данных как с таблицами. Индексация, селекция, агрегация. NumPy, SciPy и Matplotlib. Взаимосвязь библиотек, области применения, типичные приемы использование. Примеры.	17	6	6	5	
8.	Работа со случайными величинами (ipython notebook)	ГПСЧ, получение распределения с заданными характеристиками. Оценка распределения по выборке. Доверительный интервал. Параметры распределения случайной величины. Практика.	14	4	6	4	
9.	Промежуточная аттестация		2	-	-	2	
Всего:			90	28	28	34	

Высшая математика и введение в статистику

1.	Векторы и матрицы	Определение вектора. Операции с векторами. Длина вектора и расстояние между векторами. Определение матрицы. Собственные векторы, собственные числа. Операции с матрицами - сложение, умножение, диагонализация. Обратная матрица	7	3	2	2	
2.	Сингулярное разложение матрицы	Сингулярное разложение матрицы, область применимости	7	2	3	2	
3.	Функции логарифмы	Функции, глобальные и локальные экстремумы функций. Определение логарифма. Натуральный, десятичный логарифм. Графическое представление	14	6	4	4	
4.	Производные	Определение предела. Определение производной. Геометрический смысл производной. Полная производная. Частная производная. Производная тригонометрических функций. Производная сложной функции. Градиент	14	4	6	4	
5.	Градиентный спуск	Знакомство с алгоритмом градиентного спуска, область	10	4	2	4	

№ п/ п	Наименование и содержание тем	Основные понятия (категории) и проблемы, рассматриваемые в теме	Объем дисциплины час.			Форма промежуточн ой аттестации/ итоговой аттестации	
			Всего	Аудиторная работа по видам учебных занятий	СР ⁴		
				Л	СЗ ⁵		
		применения, графическое представление					
6.	Обратное распространение	Знакомство с алгоритмом обратного распространения, область применения, графическое представление	10	4	2	4	
7.	Описательные статистики и графическая презентация данных	классификация переменных: качественные и количественные, дискретные и непрерывные; алгебраические локация и разброс: среднее, дисперсия, среднеквадратичное отклонение; структурные локация и разброс: мода, медиана, квартили, квантили, межквартильный размах; графические методы представления данных: боксплот с выбросами, отличие столбиковой диаграммы от гистограммы, плотность частоты; скосленность данных и способы её измерить.	12	5	5	2	
8.	Теория вероятностей	ключевые определения: эксперимент, исходы, вероятностное пространство, события, вероятность; экспериментальная вероятность как частота, геометрическая интуиция вероятности; презентация вероятностного пространства диаграммами Эйлера и Венна, разница между ними; формула сложения вероятностей для двух событий; формула условной вероятности; независимость событий; формула полной вероятности; формула Байеса, байесова ловушка и её интуитивное объяснение	18	7	10	1	
9.	Случайные величины	дискретные и непрерывные случайные величины; распределение дискретных случайных величин: определение, свойства, табличная запись, кумулятивная функция вероятности; распределение непрерывных случайных величин: функция плотности вероятности и кумулятивная функция распределения; математическое	16	7	8	1	

№ п/ п	Наименование и содержание тем	Основные понятия (категории) и проблемы, рассматриваемые в теме	Объем дисциплины час.			Форма промежуточн ой аттестации/ итоговой аттестации	
			Всего	Аудиторная работа по видам учебных занятий	СР ⁴		
				Л	СЗ ⁵		
		ожидание и дисперсия: расчёт для дискретных, общий концепт для непрерывных случайных величин; Биномиальное распределение: расчет коэффициентов, интерпретация отдельных вероятностей и интервалов; некоторые часто употребляемые распределения с примерами					
10.	Промежуточная аттестация		10	10	-	-	Зачет / Опрос по вопросам (устно)
	Всего		118	42	42		
Визуализация данных							
1.	Визуализация данных, где взять данные	Что такое визуализация данных, история визуализации. Гештальтовский принципы восприятия, восприятие форм, восприятие цветов (цвета и культур, цветовые ассоциации). TAGS. DataMiner. Простой скрипт на R с парсингом страниц Визуализация данных в онлайн сервисах.	6	2	2	2	
2.	Misleading diagrams	Misleading diagrams, выбор графиков, какой лучше выбрать, убрать лишнее с графиков (data ink ratio). Практика графиков в Tableau.	6	2	2	2	
3.	Сети и карты, онлайн-сервисы	Сети и карты, онлайн-сервисы	5	2	1	2	
4.	Текст в визуализации изменяемые графики	Текст в визуализации. Изменяемые графики. Plotly R.	4	1	1	2	
5.	Создание приложений для визуализации	Создание приложений для визуализации, Shiny R.	6	1	2	3	
6.	Используемые программы	Знакомство с основными программами, которые будут использоваться при обучении. Настройка окружения. Настройка git. Настройка VSC. Настройка QGIS 3	7	2	2	3	
7.	Программирование из консоли внутри QGIS 3	Работа с векторными слоями. Создание векторного слоя. Работа с различными форматами данных. Работа с таблицей атрибутов. Редактирование объектов.	7	2	2	3	

№ п/ п	Наименование и содержание тем	Основные понятия (категории) и проблемы, рассматриваемые в теме	Объем дисциплины час.			Форма промежуточн ой аттестации/ итоговой аттестации	
			Всего	Аудиторная работа по видам учебных занятий	СР ⁴		
				Л	СЗ ⁵		
		Добавление/удаление объектов. Добавление/Удаление полей. Работа с геометрией Добавление/удаление геометрии. Редактирование геометрии. Работа с выборкой. Работа с отображением векторных слоев. Работа с точечными символами. Работа с линейными символами. Работа с полигональными символами. Работа с растровыми слоями. Работа с интерфейсом программы. Добавление/Удаление пунктов меню. Управление канвасом					
8.	Написание кастомных выражений в QGIS 3	Основы логических операторов. Проверка выражений. Основные виды выражений. Выражения для фильтрации. Обработка ошибок. Структура файлов плагина. Метаданные плагина. Документация/Спарвка. Структура написания основного кода плагина. Дополнительные инструменты для создания плагина. Plugin Reloader. Публикация плагина	7	2	2	3	
9.	Промежуточная аттестация		2	-	-	2	Зачет / проект
	Всего		50	14	14	22	
Машинное обучение и нейронные сети							
1.	kNN. Общий вид метрических классификаторов, kNN, WkNN, отбор эталонов	Общий вид метрических классификаторов, kNN, WkNN, отбор эталонов. Оценка классификаторов, Precision, Recall, ROC, AUC. Валидация. Кластеризация. kMeans, Mean Shift, DBSCAN, Agglomerative Clustering. Метрики кластеризации. Кластеризация и обучение с частичным привлечением учителя (semi-supervised).	15	4	3	8	
2.	Деревья. Построение деревьев, информационный выигрыш.	Энтропия, критерий Джини, девиация. CART. Прунинг. Ансамбли деревьев, бустинг деревьев. Случайный лес, адаптивный бустинг, градиентный бустинг.	15	3	4	8	

№ п/ п	Наименование и содержание тем	Основные понятия (категории) и проблемы, рассматриваемые в теме	Объем дисциплины час.			Форма промежуточн ой аттестации/ итоговой аттестации	
			Всего	Аудиторная работа по видам учебных занятий	СР ⁴		
				Л	СЗ ⁵		
3.	Нейронные сети	Логистическая регрессия, градиентный спуск. Полносвязные сети, обратное распространение градиента. Сверточные нейронные сети и анализ изображений. Свертки, пулинг. LeNet, AlexNet, VGG, ResNet.	15	4	3	8	
4.	Векторные представления и анализ естественного языка	Векторное представление слов. Рекуррентные нейронные сети, LSTM, GRU. Механизм внимания, Трансформер, BERT.	15	3	4	8	
5.	SVM	SVM. Прямая и двойственная задача. Решение двойственной задачи. Линейно неразделимый случай. Ядерный трюк.	15	4	3	8	
6.	Байес. Байесовский классификатор	Оценка распределения признаков. EM-алгоритм. Регрессия. Линейная регрессия, гребневая регрессия, LASSO, LARS. Борьба с выбросами.	15	3	4	8	
7.	Локальный поиск	Байесовская оптимизация. Hill Climb, отжиг, генетический алгоритм.	15	4	3	8	
8.	Обучение подкреплением	Постановка задачи. Функция ценности и функция качества. Q-обучение. Глубокое обучение с подкреплением, DQN, Actor-Critic.	15	3	4	8	
9.	Промежуточная аттестация		4	-	-	4	Зачет / Опрос по вопросам (устно)
	Всего:		124	28	28	68	

Компьютерное зрение

1.	Обработка изображений	Пространственная область. Частотная область, преобразование Фурье, спектральный анализ. Выделение компонент связности. Выделение краев. Математическая морфология	4	2	1	1	
2.	Глобальные и локальные признаки изображений	Свойства признаков изображений. Текстовые и визуальные признаки. Пространства признаков. Ключевые точки. Детектор угловых точек. Детектор Моравица	4	2	1	1	
3.	Параметрические модели	Выбор модели. Оценка параметров модели. Метод наименьших квадратов, М-оценки, RANSAC,	4	2	1	1	

№ п/ п	Наименование и содержание тем	Основные понятия (категории) и проблемы, рассматриваемые в теме	Объем дисциплины час.			Форма промежуточн ой аттестации/ итоговой аттестации	
			Всего	Аудиторная работа по видам учебных занятий	СР ⁴		
				Л	СЗ ⁵		
		преобразование Хафа					
4.	Классификация объектов на изображении	Бинарная и многоклассовая классификация. Категории объектов. Извлечение фрагментов. Вычисление признаков фрагментов. Обучение словаря.	4	2	1	1	
5.	Поиск изображений по содержанию	Поиск по визуальному подобию. Поиск нечетких дубликатов. Поиск объектов на фотографии. Поиск сцен	4	1	1	2	
6.	Детектирование объектов	Метрика качества IoU. Задача многоклассовой детекции. Классификация окон. HUG.	4	1	1	2	
7.	Сегментация изображений	Сегментация объектов. Семантическая сегментация. Суперпикселизация	3	1	1	1	
8.	Генеративные сети	Реконструкция изображения. Реконструкция стилей. Матрица Грама. Реконструкция текстур. Генеративносостязательные нейронные сети.	4	1	2	1	
9.	Трекинг объектов	Видеопоток и видеопоследовательность. Разметка объектов. Определение поз и жестов. Распознавание событий	5	2	2	1	
10	Промежуточная аттестация		6	-	3	3	Зачет / Опрос по вопросам (устно)
	Всего		42	14	14	14	

Дисциплина по выбору

Практический минимум

1.	Системное и прикладное ПО	Операционная система как вид ПО. Виды ОС. Linux-подобные ОС. Вычислительные ядра.	14	4	2	8	
2.	Архитектура клиент-сервер	Архитектура клиент-сервер, взаимодействие между компьютерами через интернет. Протоколы.	12	4	4	4	
3.	Удаленный доступ.	Консоль управления на примере Яндекс.Облаца.	18	2	4	12	
4.	Работа в ОС Linux	Команды. Утилиты. Настройки базовых компонент.	16	4	4	8	
5.	Серверное ПО. Базы данных. Понятие бекенда.	Серверное ПО. Базы данных. Понятие бекенда.	12	4	4	4	
6.	Службы	Превращение своей программы в сервер. Запуск по расписанию. Логирование. Базовые приемы администрирования. Установка и настройка программ.	16	4	4	8	

№ п/ п	Наименование и содержание тем	Основные понятия (категории) и проблемы, рассматриваемые в теме	Объем дисциплины час.			Форма промежуточн ой аттестации/ итоговой аттестации	
			Всего	Аудиторная работа по видам учебных занятий	СР 4		
				Л	СЗ ⁵		
		Практика.					
7.	Объединение компонент в ИС	Объединение компонент в ИС. Введение в проектирование ИС. Рассмотрение архитектур существующих ИС.	10	4	2	4	
8.	Параллельная обработка данных	Параллельная обработка данных. Практикум по созданию вычислительных кластеров.	14	2	4	8	
9.	Промежуточная аттестация		24	-	-	24	Зачет / Проект
	Всего:		136	28	28	80	
Проектный менеджмент							
1.	Введение	Разделение обучающихся на группы по интересам (всего 4 проекта). Формулировка темы исследования, паззла, постановка задачи минимум и задачи максимум. Составление временного плана (диаграмма Ганнта), распределение задач, назначение тимлида внутри команды.	16	4	4	8	
2.	Этика	Обсуждение первых результатов и сложностей командной работы. Разработка эффективных инструментов взаимодействия внутри команды. Этические аспекты. Вклад каждого члена команды и общий продукт.	16	4	4	8	
3.	Технические аспекты работы	Обсуждение технических аспектов выполнения выбранного проекта с учетом характерных особенностей выбранной темы исследования.	16	4	4	8	
4.	Отчет о ходе исследования.	Отчет о ходе исследования. Анализ наработанного обучающимися в ходе реализации проекта материала. Как написать хороший отчет о ходе исследования?	24	8	4	12	
5.	Работа с первым черновиком.	Как сделать текст интересным?	16	4	4	8	
6.	Визуальная версия	Предпродажная подготовка визуальной	24	4	8	12	

№ п/ п	Наименование и содержание тем	Основные понятия (категории) и проблемы, рассматриваемые в теме	Объем дисциплины час.			Форма промежуточн ой аттестации/ итоговой аттестации	
			Всего	Аудиторная работа по видам учебных занятий	СР 4		
				Л	СЗ ⁵		
	продукта.	версии продукта.					
7.	Промежуточная аттестация		24	-	-	24	Зачет / Проект
	Всего:		136	28	28	80	
Базы данных и SQL							
1.	Реляционные СУБД	Основы реляционной алгебры. Табличное хранение информации. Первичные ключи. Внешние ключи. Операции над таблицами	16	4	4	8	
2.	SQL	Теория: SQL как декларативный язык запроса в данных. Основные конструкции. Версии. Отличия в разных СУБД. Запросы SELECT, INSERT, UPDATE, DELETE. Встроенные функции языка. Работа с числами, строками, датами. Агрегация	16	4	4	8	
3.	Проектирование структуры реляционной БД	Нормализация и денормализация. Нормальные формы. ПО для проектирования баз данных, диаграммы. Ускорение запросов.	16	4	4	8	
4.	Настройка и развертывание СУБД	Настройки по умолчанию. Конфигурационные файлы. Профилирование запросов. Кеширование и логирование. Ведущий и ведомые сервера. Горизонтальное масштабирование. Ускорение чтения. Встроенные БД. БД в оперативной памяти	16	4	4	8	
5.	Документноориентированные базы данных	Принципы работы. JSON-документы. API и библиотеки для взаимодействия. Использование в качестве хранилища и кеширующего сервера. Индексы. Запросы к данным. Взаимосвязь SQL и декларативных языков запросов к данным в СУБД и языках программирования	16	4	4	8	
6.	Развертывание БД в облачной инфраструктуре	На примере Yandex.Cloud создание инфраструктуры для разных СУБД, совместимость с облачными функциями обработки и сохранения данных. Использование хранилища	16	4	4	8	
7.	Библиотеки для	Библиотеки для работы с	16	4	4	8	

№ п/ п	Наименование и содержание тем	Основные понятия (категории) и проблемы, рассматриваемые в теме	Объем дисциплины час.			Форма промежуточн ой аттестации/ итоговой аттестации	
			Всего	Аудиторная работа по видам учебных занятий	СР ⁴		
				Л	СЗ ⁵		
	работы с СУБД	разными СУБД в языках программирования R и Python. ORM. Вебфреймворк Django. Написание классов для ORM реляционных СУБД					
8.	Промежуточная аттестация		24	-	-	24	Зачет / Проект
	Всего:		136	28	28	80	

Программирование в R

1.	Введение	R как язык программирования. Функциональное программирование vs. ООП. Знакомство с R-Studio. Основные типы данных в R (num, int, factor, char, bool, date&time). Основные типы структур данных (vector, data.frame, list). Основные типы библиотек в R и их структура. Объяснение принципов векторизации. Базовые функции в R. Как загружать данные разных типов из разных источников. Выполнение в R простейших операций, которые слушатели уже умеют на языке Python (ifelse, gsub, grep и проч.)	16	4	4	8	
2.	Основы R	Как работать с типом данных “data.frame”? Обращение к переменным и строкам, индексация. Фильтрация, выборка переменных. Преобразование данных, long2wide/wide2long: merge, rbind, cbind, aggregate, dcast. Практика объединений и преобразования датафреймов. Повторение пройденного на предыдущем занятии вместе с новым материалом (фильтрация и преобразование переменных с помощью функций ifelse, gsub и проч.).	16	4	4	8	
3.	Библиотека Dplyr	Объяснение общей логики работы в этой библиотеки. Повторение старого материала на dplyr. Основы визуализации данных в ggplot2	16	4	4	8	
4.	Data.table	Преимущества data.table. Основные принципы работы и	24	8	4	12	

№ п/ п	Наименование и содержание тем	Основные понятия (категории) и проблемы, рассматриваемые в теме	Объем дисциплины час.			Форма промежуточн ой аттестации/ итоговой аттестации	
			Всего	Аудиторная работа по видам учебных занятий	СР ⁴		
				Л	СЗ ⁵		
		специфические операторы data.table. Повторение предыдущего материала на базе этой библиотеки.					
5.	Циклы, специфика apply family	Повторение материала по data.table. Объяснения циклов в R (предполагается, что слушатели уже знают о них после курса по Python). Работа с разными типами apply family. Введение в работу с листами в R (иллюстрация работы типа apply family - lapply)	16	4	4	8	
6.	Методы кластеризации и кластеризация в R	Основные методы кластеризации и их применение в R: K-Means, Mean-Shift, DBSCAN.	24	4	8	12	
7.	Промежуточная аттестация		24	-	-	24	
	Всего:		136	28	28	80	

Алгоритмы и структуры данных

1.	Введение в алгоритмы	Свойства алгоритмов. Анализ алгоритмов. Псевдокод. Простейшие операции. Анализ лучших и худших показателей. Асимптотическая нотация.	16	4	4	8	
2.	Абстрактные типы данных	Векторы, списки, последовательности. Вектор. Список. Стек. Очередь. Дек.	16	4	4	8	
3.	Поиск	Наивный метод. Алгоритм Кнута-Морриса-Прата. Алгоритм Бойера-Мура. Алгоритм Робина-Карпа.	16	4	4	8	
4.	Сортировка	Сортировка подсчетом. Сортировка включением. Сортировка Шелла. Сортировка извлечением. Пирамидальная сортировка. Обменная сортировка. Быстрая сортировка. Сортировка слиянием.	24	8	4	12	
5.	Деревья	Прохождение бинарных деревьев. Поиск заданного ключа. Поиск минимума и максимума. Прелшествующий и последующий элементы. Вставка и удаление. Сбалансированные деревья. Сильноветвящиеся деревья.	16	4	4	8	
6.	Хеширование данных	Таблицы с прямой адресацией. Хеш-таблицы. Хеш-функции. Методы разрешения коллизий.	24	4	8	12	

№ п/ п	Наименование и содержание тем	Основные понятия (категории) и проблемы, рассматриваемые в теме	Объем дисциплины час.			Форма промежуточн ой аттестации/ итоговой аттестации	
			Всего	Аудиторная работа по видам учебных занятий	СР ⁴		
				Л	СЗ ⁵		
7.	Промежуточная аттестация		24	-	-	24	Зачет /проект
	Всего:		136	28	28	80	
	Итоговая аттестация		174	-	74	100	Зачет / Проектная работа
	ИТОГО		734 ⁶	64	71	115	

⁶ В случае реализации образовательной программы с применением дистанционных образовательных технологий, данные часы реализуются онлайн, в синхронном формате в соответствии с расписанием занятий