

Учебный план дополнительной профессиональной программы переподготовки:

«Системные основы программирования и баз данных»

№ п/п	Наименование темы	Описание темы	Всего, час	В том числе			
				Лекции	Практ. раб.	Сам. работа	Аттест.
Модуль 1 «Программирование на языке Python. Уровень 1. Основы программирования»							
1.1	Язык программирования Python и его место среди других языков и систем программирования	Обзор существующих языков программирования. Основные принципы работы интерпретатора языка программирования Python. Язык программирования Python в различных операционных системах. Установка и запуск интерпретатора языка программирования Python. Установка и тестирование среды разработки языка программирования Python. Основные элементы программирования	5	3	1	1	-
1.2	Базовые конструкции языка программирования Python	Переменные и выражения. Знакомство и типами данных: числа, строки, списки, логический тип, None. Типы данных. Принцип динамической типизации. Оператор del. Структура программы. Блок Ветвления Базовая форма цикла Операторы break и continue. Оператор pass. Простой ввод и простой вывод.	4	3	1	-	-
1.3	Кортежи, списки, словари, множества	Кортеж. Основные операции с кортежем. Список. Основные операции со списком. Словарь. Основные операции со словарем. Цикл по итератору.	7	3	2	2	-

		Перебор (for)					
1.4	Функции и коллекции	Создание функции. Вызов функции. Именованные и неименованные аргументы функций. Функции с переменным числом аргументов. Перебор (for). Функция как объект первого порядка. Понятие коллекции. Генераторное выражение. Генератор-функция. Генератор списков, генератор словарей, генератор множеств. Встроенные функции: map, zip, filter, reduce. Лямбда-функция.	5	3	2	-	-
1.5	Модули и пакеты	Создание модуля. Импорт пакета. Создание собственного пакета. Важнейшие стандартные пакеты. Подсистема pip. Установка стороннего модуля.	4	3	1	-	-
1.6	Работа с файловой системой	Работа с файлами и каталогами. Основные операции с файлами. Основные операции с путями к файлам. Рекурсивный обход каталога. Чтение файла. Запись в файл.	5	2	2	1	-
1.7	Исключения и обработка ошибок	Понятие об исключении. Выброс исключения. Перехват исключения. Стандартные исключения.	4	3	1	-	-
1.8	Регулярные выражения	Понятие о регулярном выражении. Синтаксис регулярных выражений. Применение регулярных выражений.	4	2	2	-	-
1.9	Элементы функционального программирования	Каррирование. Замыкание. Понятие о декораторе	6	3	2	1	-
1.10	Промежуточная аттестация	Тестирование	1	-	-	1	-

Модуль 2 «Язык программирования SQL. Уровень 1: создание запросов и анализ данных»

2.1	Введение. Постановка задачи	Анализ бизнес-требований, определение масштаба проблемы. Метрики оценки модели. Определение конечной цели проекта интеллектуального анализа данных. Элементы программирования. Переменные в языке SQL. Объявление переменных. Присвоение значения. Логические выражения и условные вычисления. Приоритет операторов.	4	2	-	2	-
2.2	Типы данных, функции, операторы и программные модули	Логические выражения и условные вычисления. Типы данных. Математические функции. Работа с календарём. Статистические функции. Работа с текстом. Представления, процедуры и пользовательские функции. Табличные выражения. Программирование и динамические запросы. Работа с базой данных и представление данных. Использование функций различных видов.	11	1	6	4	-
2.3	Структуры данных	Нормализованные базы данных. Хранилища и витрины данных. Объектные данные. Данные в формате XML. Работа с большими таблицами. Иерархии. Сети и графы. Работа со структурой баз данных. Работа с хранилищем и витриной данных (DWH, Data Marts). Эффективная работа с программой Management Studio (SSMS).	9	1	6	2	-

2.4	Методы анализа	Вычисления и поиск. Работа с упорядоченными сериями. Распределение. Сегментирование. Регрессия. Корреляция. Анализ связанных событий. Вероятностные модели. Другие методы анализа данных. Статистический анализ данных. Инструменты анализа данных. Использование функций на практике для построения отчетов.	9	1	6	2	-
2.5	Промежуточная аттестация	Тестирование	1	-	-	1	-

Модуль 3 «Наука о данных: профессиональный уровень (Data Science PRO)»

Тема 1. Базовый модуль. Большие данные

3.1.1	Введение в большие данные. Знакомство с основными понятиями. Статистика	Введение в предмет. Какие бывают источники данных, характеристики, корреляция. Типы структурированности данных. Введение в статистику. Типы распределений данных. Проверка гипотез, матрицы ошибок Структуры данных. Техническое задание и требование к проекту больших данных, науки о данных.	11	8	-	3	-
3.1.2	Базы данных, их типы и моделирование. Взаимодействие пользователей с данными	Системы хранения данных. Базы данных. Реляционные базы данных и их моделирование. Программа PGAdmin. Язык программирования SQL, синтаксис обращения к базам данных. Простые и сложные запросы. Решение практического кейса: «Создание базы данных университета».	16	3	10	3	-

3.1.3	Языки программирования, среды исполнения и основные операционные системы, которые используются в задачах больших данных	Введение в Python, среды исполнения (IDE). Типы данных в языке программирования Python. Базовые конструкции и структуры. Циклы и условия. Функции и классы. Библиотеки Numpy, Scipy для научных вычислений. Библиотека Pandas, как стандарт исследования данных. Визуализация в Python и срезы данных. Библиотеки Matplotlib, seaborn, plotly. Работа с командной строкой. Операционные системы Linux, Windows. Решение практического кейса: «Визуализация данных с использованием основных библиотек языка программирования Python в среде «Jupyter Notebook».	24	4	11	9	-
3.1.4	Основная последовательность машинного обучения и потоков данных. Облачные платформы	Последовательность машинного обучения. Особенности архитектурных решений. Sklearn - основная библиотека машинного обучения. Ленивые вычисления. Решение практического кейса: «Создание последовательности действий для проекта по машинному обучению с использованием библиотеки Sklearn».	12	3	4	5	-
3.1.5	Автоматизированный сбор и структурирование данных из открытых источников с помощью скриптов, подготовка выборки, преобразование и балансировка данных	Ключевые задачи в подготовке датасетов и их важность. Нормализация и стандартизация. Поиск аномалий и выбросов. Методы обработки и визуализации. Разбалансированные датасеты и методы балансировки. Доверительные интервалы. Обработка категориальных признаков. Инструменты	16	5	5	6	-

		LabelEncoder, One Hot encoding. Решение практического кейса: «Преобразование данных для подготовки датасета к решению задач машинного обучения».					
3.1.6	Обучение без учителя. Работа с размерностью и структурой данных	Обучение на размеченных и неразмеченных данных. Проблема снижения размерности данных. Метод главных компонент. Алгоритмы кластеризации. Метод k-средних, c-средних, иерархическая кластеризация. Выбор алгоритма кластеризации. Решение практического кейса: «Снижение размерности в задаче определения кредитного потенциала заемщика» Решение практического кейса: «Кластеризация городов Российской Федерации по типу прироста (убыли) населения».	14	2	7	5	-
3.1.7	Теория вероятности и вероятностные классификаторы	Полная и условная вероятность, теорема Байеса Байесовский вероятностный классификатор. Решение практического кейса: «Решение задачи классификации осколков стекла по их химическому составу».	7	3	2	2	
3.1.8	Классификация. Оценка качества алгоритма	Задача классификации. Матрица ошибок (Confusion -matrix). Точность классификации. Модель анализа качества для задач бинарной классификации, показывающая способность модели разделять классы. Методы классификации. Метод ближайших соседей. Метод опорных векторов. Решение практического	12	4	5	3	

		кейса: «Оценка и классификация кредитного потенциала заемщиков».					
3.1.9	Ансамбли и повышение точности алгоритмов	Ансамбли решателей. Решающие деревья. Случайный лес. Ансамбли алгоритмов. Алгоритмы улучшения стабильности и точности алгоритмов. Основы составления сеток обучения. Решение практического кейса: «Решение задач медицинской диагностики с применением ансамблей классификаторов» Решение практического кейса: «Решение задачи анализа сетевой активности».	11	3	6	2	
3.1.10	Регрессия	Линейная регрессия. Логистическая регрессия. Метод наименьших квадратов. Средняя квадратичная ошибка, средняя абсолютная ошибка. Авторегрессия. Решение практического кейса: «Прогнозирование объемов продаж товаров торгового предприятия». Решение практического кейса: «Оценка кредитоспособности клиентов банка - физических лиц».	23	3	12	8	
3.1.11	Нейронные сети	Введение в искусственные нейронные сети. Биологическая модель нейрона. Активационные функции. Полносвязные нейронные сети. Методы обучения нейронных сетей. Борьба с переобучением. Нейронные сети Кохонена. Сверточные нейронные сети. Решение практического кейса: «Моделирование объемов продаж торгового предприятия с	16	5	7	4	

		использованием нейронных сетей». Решение практического кейса: «Анализ банковского сектора с применением нейронных сетей Кохонена».					
3.1.12	Промежуточная аттестация	Тестирование	10	-	-	10	-
Тема 2. Профильный модуль. Нейронные сети							
3.2.1	Разработка компьютерных программ: написание программного кода и решение практических задач с применением нейронных сетей	Решение практического кейса: «Разработка приложения на языке программирования Python для решения задачи компьютерного зрения с использованием нейронных сетей».	25	0	16	9	-
3.2.2	Рекомендательные системы	Введение. Коллаборативная фильтрация и фильтрация на основе содержания. Сингулярное разложение матрицы. Достоинства и недостатки. Решение практического кейса: «Создание рекомендательной системы торгового предприятия».	11	2	6	3	-
3.2.3	Создание приложений, интеграция в ИТ ландшафт	Сортирование и сохранение моделей. Работа с предобученными моделями. Фласк-приложение (Flask-приложение). Выведение моделей в производство. Контроль версии моделей с серверной архитектурой. Облачная платформа Kaggle для решения задач науки о данных. Решение практического кейса: «Создание фласк-приложения (Flask-приложения) для запуска моделей машинного обучения».	16	3	8	5	-

3.2.4	Промежуточная аттестация	Тестирование	10	-	-	10	-
4	Итоговая аттестация	Защита ВКР	28	-	-	-	28
	ИТОГО		341	78	131	104	28