



Дополнительная профессиональная программа
повышения квалификации
«Большие данные»

УЧЕБНО-ТЕМАТИЧЕСКИЙ ПЛАН

Цель: формирование теоретических знаний и практических навыков в сфере работы с большими массивами данных, в части анализа и обработки данных; формирование теоретических знаний и практических навыков в области машинного обучения: построение и оценка качества моделей машинного обучения.

Программно-аппаратной основой обучения является оборудование и программное обеспечение Мастерской «Машинное обучение и большие данные» ФГБОУ ВО «СГУ им. Питирима Сорокина».

Целевой аудиторией программы являются лица, имеющие среднее профессиональное или высшее образование, или имеющие необходимость приобрести или повысить свои теоретические знания и практические навыки при работе с большими массивами данных. В силу этого освоение программы требует наличия базовых компетенций и трудовых функций, включающих знания, умения и навыки, необходимые для работы с компьютером на уровне пользователя, а также знания, умения и навыки, необходимые для установки, настройки, применения, разработки компьютерных программ, желательно наличие компетенций, включающих владение языками программирования высокого уровня, например, Python, знания основ математической статистики, математического анализа.

Минимальные требования: владение навыками работы с компьютером на уровне среднего пользователя, владение приемами работы с ресурсами и технологиями сети Интернет.

Рекомендуемые требования: владение навыками работы с компьютером на уровне опытного пользователя, владение приемами работы с ресурсами и технологиями сети Интернет, знания и навыки в области программирования, владение языками программирования высокого уровня (Python и т.д.), знание основ высшей математики (основы математической статистики, математического анализа).

Форма обучения: заочная с применением дистанционных образовательных технологий.

Трудоемкость: 18 академических часов.

Период реализации программы: 2020/2021 учебный год.

№ п/п	Наименование раздела	Всего, часов	в том числе					Форма контроля
			Лекции	Практических и лабораторных занятий	Занятия с использованием ДОТ	Самостоятельная работа	Практика (стажировка)	
1	Основы синтаксиса Python	3	-	3	3	-	-	-
1.1.	Типы данных, операции, операторы, функции, структуры данных Python	3	-	3	3	-	-	-

№ п/п	Наименование раздела	Всего, часов	в том числе					Форма контроля
			Лекции	Практических и лабораторных занятий	Занятия с использованием ДОТ	Самостоятельная работа	Практика (стажировка)	
2	Первичный анализ больших данных с Pandas	3	-	3	3	-	-	-
2.1.	Загрузка, подготовка, первичный анализ данных с Pandas	3	-	3	3	-	-	-
3	Визуальный анализ больших данных с Seaborn	3	-	3	3	-	-	-
3.1.	Визуальный анализ: функции plot, pairplot, distplot, jointplot, boxplot, heatmap, hist, corr_matrix, violin_plot, t-SNE-представление данных, библиотека plotly	3	-	3	3	-	-	-
4	Задача классификации	3	-	3	3	-	-	-
4.1	Задача классификации, план решения задачи классификации, построение дерева решений	1	-	1	1	-	-	-
4.2	Метод ближайших соседей, кросс-валидация	1	-	1	1	-	-	-
4.3	Случайный лес, измерение качества модели классификации	1	-	1	1	-	-	-
5	Задача регрессии	3	-	3	3	-	-	-
5.1	Задача регрессии, план решения задачи регрессии, линейная модель	1	-	1	1	-	-	-
5.2	Модели регрессии с регуляризацией	1	-	1	1	-	-	-
5.3	Измерение качества модели регрессии	1	-	1	1	-	-	-
	Обучение без учителя	2	-	2	2	-	-	-
6.1.	Понижение размерности. Метод главных компонент	1	-	1	1	-	-	-
6.2.	Задача кластеризация. Метод k-means	1	-	1	1	-	-	-
	Всего	17	-	17	17	-	-	-
	Итоговая аттестация	1	-	-	-	-	-	Тестирование

№ п/п	Наименование раздела	Всего, часов	в том числе					Форма контроля
			Лек- ции	Прак- тиче- ских и лабо- ратор- тор- ных заня- тия	Занятия с ис- пользо- ванием ДОТ	Само- стоя- тель- ная работа	Прак- тика (ста- жиров- ка)	
	Итого	18	-	17	17	-	-	1