

МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ РОССИЙСКОЙ ФЕДЕРАЦИИ
Федеральное государственное автономное образовательное учреждение высшего образования
«НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ
ТОМСКИЙ ПОЛИТЕХНИЧЕСКИЙ УНИВЕРСИТЕТ»



РАБОЧАЯ ПРОГРАММА ДИСЦИПЛИНЫ
ПРИЕМ 2019 г.
ФОРМА ОБУЧЕНИЯ очная

Машинное обучение			
Направление подготовки/специальность	09.04.04. Программная инженерия		
Образовательная программа (направленность (профиль))	Big Data Solutions / Технологии больших данных		
Специализация	Big Data Solutions / Технологии больших данных		
Уровень образования	высшее образование - магистратура		
Курс	2	семестры	3
Трудоемкость в кредитах (зачетных единицах)	4		
Виды учебной деятельности	Временной ресурс		
Лекции	24		
Лабораторные занятия	24		
ВСЕГО	48		
Самостоятельная работа, ч	96		
ИТОГО, ч	144		

Вид промежуточной аттестации

Зачёт	Обеспечивающее подразделение	ОИТ ИШИТР
-------	------------------------------	-----------

Заведующий кафедрой-
руководитель ОИТ на правах
кафедры
Руководитель ООП

Преподаватель

	Шерстнев В.С.
	Савельев А.О.
	Аксёнов С.В.

1. Цели освоения дисциплины

Целями освоения дисциплины является формирование у обучающихся определенного ООП (п. 6. Общей характеристики ООП) состава компетенций для подготовки к профессиональной деятельности.

Код компетенции	Наименование компетенции	Индикаторы достижения компетенций		Составляющие результатов освоения (дескрипторы компетенций)	
		Код индикатора	Наименование индикатора достижения	Код	Наименование
ПК(У)-5	Способен организовать разработку системного программного обеспечения	И.ПК(У)-5.1	Планирует разработку системного ПО	ПК(У)-5.1B1	Владеет навыками оценки сроков, ресурсоемкости, себестоимости проекта по разработке системного ПО
				ПК(У)-5.131	Знает основы архитектуры, устройства и функционирования вычислительных систем
				ПК(У)-5.132	Знает устройство и принципы функционирования информационных систем
ПК(У)-1	Способен управлять программно-техническими, технологическими и человеческими ресурсами	И.ПК(У)-1.1	Демонстрирует способность управлять программно-техническими, технологическими и человеческими ресурсами	ПК(У)-1.1B1	Владеет инструментальными средствами разработки
				ПК(У)-1.1У1	Умеет применять методологии управления проектами разработки программного обеспечения
				ПК(У)-1.131	Знает методологии разработки программного обеспечения

2. Место дисциплины в структуре ООП

Дисциплина относится к вариативной части Блока 1 учебного плана образовательной программы (дисциплина модуля дополнительной специализации «Технологии больших данных»).

3. Планируемые результаты обучения по дисциплине

После успешного освоения дисциплины будут сформированы результаты обучения:

Планируемые результаты обучения по дисциплине		Компетенция
Код	Наименование	
РД-1	Проводить разведочный анализ данных. Выдвигать и проверять гипотезы о взаимозависимости данных.	И.ПК(У)-5.1
РД-2	Применять методики проектирования ансамблей систем машинного обучения для конкретных предметных областей.	И.ПК (У)-1.1
РД-3	Разрабатывать программное обеспечение на языке Python для создания систем анализа данных с помощью машинного обучения.	И.ПК(У)-5.1 И.ПК (У)-1.1

Оценочные мероприятия текущего контроля и промежуточной аттестации представлены в календарном рейтинг-плане дисциплины.

4. Структура и содержание дисциплины

Основные виды учебной деятельности

Разделы дисциплины	Формируемый результат обучения по дисциплине	Виды учебной деятельности	Объем времени, ч.
Раздел 1. Основы машинного обучения	РД-1	Лекции	2
		Практические занятия	0
		Лабораторные занятия	0
		Самостоятельная работа	4

Раздел 2. Классификаторы	РД-1, РД-2	Лекции	4
		Практические занятия	0
		Лабораторные занятия	6
		Самостоятельная работа	10
Раздел 3. Регрессионные модели	РД-3	Лекции	4
		Практические занятия	0
		Лабораторные занятия	10
		Самостоятельная работа	16
Раздел 4. Обучение без учителя. Кластеризация	РД-3	Лекции	2
		Практические занятия	0
		Лабораторные занятия	8
		Самостоятельная работа	10
Раздел 5. Глубокое обучение	РД-3	Лекции	4
		Практические занятия	0
		Лабораторные занятия	8
		Самостоятельная работа	20

Содержание разделов дисциплины:

Раздел 1. Основы машинного обучения

Основные понятия теории машинного обучения: проблемы, решаемые методами машинного обучения, модели машинного обучения (геометрические, вероятностные, логические), признаки. Концептуальное обучение: пространство гипотез, поиск в пространстве гипотез, обучаемость, оценка качества решения задачи. Предварительная обработка данных. Обработка пропущенных значений. Работа с категориальными признаками.

Темы лекций:

ЛК1. Основные концепции машинного обучения

Названия лабораторных работ:

ЛР1. Манипуляции с источниками данных при построении моделей машинного обучения

Раздел 2. Классификаторы

Бинарная и многоклассовая классификация. Алгоритмы машинного обучения для построения классификационных моделей. Оценка качества работы моделей классификации. Построение ансамблей моделей машинного обучения для решения задач классификации. Бэггинг и бустинг. Оценка информативности признаков.

Темы лекций:

ЛК2. Алгоритмы построения классификаторов

Названия лабораторных работ:

ЛР2. Построение классификационных моделей и оценка их эффективности

Раздел 3. Регрессионные модели

Выдвижение и проверка статистических гипотез. Оценка значимости признаков. Алгоритмы машинного обучения для решения задач регрессии. Регуляризация в регрессионных моделях. Оценка эффективности регрессионных моделей.

Темы лекций:

ЛК3. Алгоритмы построения регрессоров

Названия лабораторных работ:

ЛР3. Построение регрессионных моделей и оценка их эффективности

Раздел 4. Обучение без учителя. Кластеризация

Оценка значимости признаков. Отбор признаков и снижение размерности признакового пространства. Конструирование признаков. Кластеризация данных (иерархические, плотностные методы, алгоритмы, использующие центроиды, нечеткие решения). Оценка эффективности решения задачи кластеризации.

Темы лекций:

ЛК4. Обучение без учителя. Выделение кластеров в данных

Названия лабораторных работ:

ЛР4. Решение задачи кластеризации данных и оценка его эффективности

Раздел 5. Глубокое обучение

Нейронные сети. Архитектуры нейронных сетей для решения задач табличных данных, сигналов, текстов и изображений. Библиотеки разработки нейросетевых моделей. Применение нейронных сетей для задач регрессии и классификации. Глубокое обучение для текста и последовательностей. Генеративное глубокое обучение.

Темы лекций:

ЛК5. Нейронные сети для анализа данных

Названия лабораторных работ:

ЛР5. Решение классификационных и регрессионных задач нейросетевыми методами

5. Организация самостоятельной работы студентов

Самостоятельная работа студентов при изучении дисциплины предусмотрена в следующих видах и формах:

- Работа с лекционным материалом, поиск и обзор литературы и электронных источников информации по индивидуально заданной проблеме курса;
- Изучение тем, вынесенных на самостоятельную проработку;
- Поиск, анализ, и структурирование информации;
- Перевод текстов с иностранных языков;
- Подготовка к лабораторным работам.

6. Учебно-методическое и информационное обеспечение дисциплины**6.1. Учебно-методическое обеспечение**

Основная литература:

1. Чيو, К. Машинное обучение и безопасность: руководство / К. Чيو, Д. Фримэн ; перевод с английского А. В. Снастина. — Москва : ДМК Пресс, 2017. — 388 с. — ISBN 978-5-97060-713-8. — Текст : электронный // Лань : электронно-библиотечная система. — URL: <https://e.lanbook.com/book/131707> (дата обращения: 01.02.2017).
2. Кук, Д. Машинное обучение с использованием библиотеки H2O / Д. Кук ; перевод с английского А. Б. Огурцова. — Москва : ДМК Пресс, 2017. — 250 с. — ISBN 978-5-97060-508-0. — Текст : электронный // Лань : электронно-библиотечная система. — URL: <https://e.lanbook.com/book/97353> (дата обращения: 01.02.2017).

3. Шалев-Шварц, Ш. Идеи машинного обучения : учебное пособие / Ш. Шалев-Шварц, Бен-Давид Ш. ; перевод с английского А. А. Слинкина. — Москва : ДМК Пресс, 2017. — 436 с. — ISBN 978-5-97060-673-5. — Текст : электронный // Лань : электронно-библиотечная система. — URL: <https://e.lanbook.com/book/131686> (дата обращения: 01.02.2017).

Дополнительная литература:

1. Вьюгин, В. В. Математические основы машинного обучения и прогнозирования : учебное пособие / В. В. Вьюгин. — Москва : МЦНМО, 2017. — 304 с. — ISBN 978-5-4439-2014-6. — Текст : электронный // Лань : электронно-библиотечная система. — URL: <https://e.lanbook.com/book/56397> (дата обращения: 01.02.2017).
2. Шарден, Б. Крупномасштабное машинное обучение вместе с Python : учебное пособие / Б. Шарден, Л. Массарон, А. Боскетти ; перевод с английского А. В. Логунова. — Москва : ДМК Пресс, 2017. — 358 с. — ISBN 978-5-97060-506-6. — Текст : электронный // Лань : электронно-библиотечная система. — URL: <https://e.lanbook.com/book/105836> (дата обращения: 01.02.2017).
3. Гаврилова, Т. А. Инженерия знаний. Модели и методы : учебник для вузов / Т. А. Гаврилова, Д. В. Кудрявцев, Д. И. Муромцев. — 3-е изд., стер. — Санкт-Петербург : Лань, 2017. — 324 с. — ISBN 978-5-8114-6473-9. — Текст : электронный // Лань : электронно-библиотечная система. — URL: <https://e.lanbook.com/book/147337> (дата обращения: 01.02.2017).

6.2. Информационное и программное обеспечение

Internet-ресурсы (в т.ч. в среде LMS MOODLE и др. образовательные и библиотечные ресурсы):

1. <http://machinelearning.ru/> – ресурс MachineLearning.ru, дата обращения 01.02.2017 г.
2. <https://azure.microsoft.com/ru-ru/services/machine-learning/> - ресурс Microsoft, дата обращения 01.02.2017г.
3. <https://aws.amazon.com/ru/machine-learning/> - ресурс Amazon Machine Learning, дата обращения 01.02.2017г.
4. http://www.sas.com/en_us/insights/analytics/machine-learning.html - ресурс SAS Machine learning, дата обращения 01.02.2017 г.
5. <https://www.ml.cmu.edu/> - ресурс Carnegie Mellon University Machine Learning, дата обращения 01.02.2017 г.

Лицензионное программное обеспечение (в соответствии с **Перечнем лицензионного программного обеспечения ТПУ**):

1. Инструменты и библиотеки разработки для языка Python – Anaconda, Jupyter Notebook, PyCharm, Keras, Scikit-learn.
2. Windows 10 Pro/ Linux Ubuntu – операционная система.

7. Особые требования к материально-техническому обеспечению дисциплины

В учебном процессе используется следующее лабораторное оборудование для практических и лабораторных занятий:

№	Наименование специальных помещений	Наименование оборудования
1.	Аудитория для проведения учебных занятий всех типов, курсового проектирования, консультаций, текущего контроля и промежуточной аттестации Доска аудиторная настенная - 1 шт. Комплект учебной мебели на 36 посадочных мест; Проектор - 1 шт. 634034 г. Томская область, Томск, ул. Советская, 84/3, КЦ, аудитория 313	Доска аудиторная настенная - 1 шт. Комплект учебной мебели на 36 посадочных мест; Компьютер - 1 шт.; Проектор - 1 шт.
2.	Аудитория для проведения учебных занятий всех типов, курсового проектирования, консультаций, текущего контроля и промежуточной аттестации (компьютерный класс) Доска аудиторная настенная - 1 шт. Компьютер - 12 шт. 634034 г. Томская область, Томск, улица Советская, 84/3, учебный корпус Кибернетический центр, аудитория 407	Доска аудиторная настенная - 1 шт. Комплект учебной мебели на 10 посадочных мест; Компьютер - 12 шт.

Рабочая программа составлена на основе Общей характеристики образовательной программы по направлению 09.04.04 Программная инженерия/ профиль Big Data Solutions / Технологии больших данных /специализации Big Data Solutions/Технологии больших данных (приема 2019 г., очная форма обучения).

Разработчик(и):

Должность	Подпись	ФИО
Доцент ОИТ ИШИТР		Аксёнов С.В.

Программа одобрена на заседании Отделения информационных технологий ИШИТР (протокол №13 от 19.06.2019 г.).

Заведующий кафедрой –
руководитель ОИТ на правах кафедры

 / В.С. Шерстнев

Лист изменений рабочей программы дисциплины¹:

Учебный год	Содержание /изменение	Обсуждено на заседании ОИТ ИШИТР (протокол)
2020/2021	1. Обновлено программное обеспечение 2. Обновлен список литературы, в том числе ссылок ЭБС	от 01.09.2020г. № 19

¹ Ежегодное обновление программы с учетом развития науки, культуры, экономики, техники и технологий, социальной сферы.