

Министерство науки и высшего образования РФ
Федеральное государственное автономное образовательное учреждение
высшего образования
«СИБИРСКИЙ ФЕДЕРАЛЬНЫЙ УНИВЕРСИТЕТ»

СОГЛАСОВАНО

Заведующий кафедрой

Базовая кафедра
Интеллектуальные системы
управления (ИСУ_ИКИТ)

наименование кафедры

подпись, инициалы, фамилия

«___» _____ 20__ г.

институт, реализующий ОП ВО

УТВЕРЖДАЮ

Заведующий кафедрой

Базовая кафедра
Интеллектуальные системы
управления (ИСУ_ИКИТ)

наименование кафедры

Якунин Юрий Юрьевич

подпись, инициалы, фамилия

«___» _____ 20__ г.

институт, реализующий дисциплину

**РАБОЧАЯ ПРОГРАММА ДИСЦИПЛИНЫ
АНАЛИТИЧЕСКАЯ ОБРАБОТКА
СТРУКТУРИРОВАННЫХ И
НЕСТРУКТУРИРОВАННЫХ
ДАННЫХ**

Дисциплина Б1.В.ДВ.01.01 Аналитическая обработка
структурированных и неструктурированных данных

Направление подготовки / 09.04.04 Программная инженерия,
специальность программа 09.04.04.02 Технологии
индустриального производства

Направленность
(профиль)

Форма обучения

очная

Год набора

2020

Красноярск 2021

РАБОЧАЯ ПРОГРАММА ДИСЦИПЛИНЫ

составлена в соответствии с Федеральным государственным образовательным стандартом высшего образования с учетом профессиональных стандартов по укрупненной группе

090000 «ИНФОРМАТИКА И ВЫЧИСЛИТЕЛЬНАЯ ТЕХНИКА»

Направление подготовки /специальность (профиль/специализация)

Направление 09.04.04 Программная инженерия, программа 09.04.04.02

Технологии индустриального производства программного обеспечения интеллектуальных систем управления 2020г.

Программу
составили

канд. техн. наук, доцент, Даничев Алексей
Александрович

1 Цели и задачи изучения дисциплины

1.1 Цель преподавания дисциплины

Целью преподавания дисциплины "Аналитическая обработка структурированных и неструктурированных данных является изучение основных задач и методов обработки и анализа данных и освоение методов статистического анализа при работе с неструктурированными данными.

1.2 Задачи изучения дисциплины

Для достижения цели изучения дисциплины предполагается решение студентами следующих задач:

- изучение методов статистического анализа данных;
- изучение методов систематизации информации неструктурированных данных;
- изучение принципов организации и проведения аналитического исследования;
- решение практических задач, прогнозирование и выработка рекомендаций.

1.3 Перечень планируемых результатов обучения по дисциплине (модулю), соотнесенных с планируемыми результатами освоения образовательной программы

ПК-1:Применять методологии разработки программного обеспечения в управлении инфраструктурой коллективной среды разработки
ПК-1.1:Знать методологии разработки программного обеспечения при управлении инфраструктурой коллективной среды разработки
ПК-1.2:Уметь применять методологии разработки программного обеспечения при управлении инфраструктурой коллективной среды разработки
ПК-1.3:Иметь навыки управления инфраструктурой коллективной среды разработки

1.4 Место дисциплины (модуля) в структуре образовательной программы

Методы и алгоритмы распознавания и обработки данных

Основы научных исследований

Знание материала дисциплины «Аналитическая обработка структурированных и неструктурированных данных» может использоваться при выполнении выпускной квалификационной работы.

1.5 Особенности реализации дисциплины

Язык реализации дисциплины Русский.

Дисциплина (модуль) реализуется с применением ЭО и ДОТ

<https://e.sfu-kras.ru/course/view.php?id=16388>

2. Объем дисциплины (модуля)

Вид учебной работы	Всего, зачетных единиц (акад.час)	Семестр
		3
Общая трудоемкость дисциплины	7 (252)	7 (252)
Контактная работа с преподавателем:	2 (72)	2 (72)
занятия лекционного типа	0,5 (18)	0,5 (18)
занятия семинарского типа		
в том числе: семинары		
практические занятия	1,5 (54)	1,5 (54)
практикумы		
лабораторные работы		
другие виды контактной работы		
в том числе: групповые консультации		
индивидуальные консультации		
иная внеаудиторная контактная работа:		
групповые занятия		
индивидуальные занятия		
Самостоятельная работа обучающихся:	4 (144)	4 (144)
изучение теоретического курса (ТО)		
расчетно-графические задания, задачи (РГЗ)		
реферат, эссе (Р)		
курсовое проектирование (КП)	Нет	Нет
курсовая работа (КР)	Нет	Нет
Промежуточная аттестация (Экзамен)	1 (36)	1 (36)

3 Содержание дисциплины (модуля)

3.1 Разделы дисциплины и виды занятий (тематический план занятий)

№ п/п	Модули, темы (разделы) дисциплины	Занятия лекционного типа (акад. час)	Занятия семинарского типа		Самостоятельная работа, (акад. час)	Формируемые компетенции
			Семинары и/или Практические занятия (акад. час)	Лабораторные работы и/или Практикумы (акад. час)		
1	2	3	4	5	6	7
1	Подготовка данных	4	12	0	36	ПК-1.1 ПК-1.2 ПК-1.3
2	Планирование эксперимента	4	12	0	36	ПК-1.1 ПК-1.2 ПК-1.3
3	Структурирование информации	4	12	0	36	ПК-1.1 ПК-1.2 ПК-1.3
4	Анализ текста	6	18	0	36	ПК-1.1 ПК-1.2 ПК-1.3
Всего		18	54	0	144	

3.2 Занятия лекционного типа

№ п/п	№ раздела дисциплины	Наименование занятий	Объем в акад. часах		
			Всего	в том числе, в инновационной форме	в том числе, в электронной форме
1	1	Описательная статистика, характеристики выборок	2	0	0
2	1	Устойчивые распределения, белый шум	2	0	0
3	2	Планирование эксперимента	2	0	0
4	2	Анализ мощности	2	0	0
5	3	Кластеризация, факторизация, нечеткая логика	2	0	0

6	3	Оценка качества работы алгоритма; ансамбли методов	2	0	0
7	4	Обработка текста для машинного обучения	2	0	0
8	4	Способы анализа текстовой информации	2	0	0
9	4	Словари	2	0	0
Всего			18	0	0

3.3 Занятия семинарского типа

№ п/п	№ раздела дисциплины	Наименование занятий	Объем в акад. часах		
			Всего	в том числе, в инновационной форме	в том числе, в электронной форме
1	1	Нормализация данных	2	0	0
2	1	Построение валидных гистограмм	4	0	0
3	1	Закон Бенфорда	2	0	0
4	1	Белый шум	4	0	0
5	2	Проверка гипотез	4	0	0
6	2	Интерпретация результатов эксперимента	2	0	0
7	2	Предсказание ошибки эксперимента	4	0	0
8	2	Интервальный анализ	2	0	0
9	3	Кластеризация	2	0	0
10	3	Факторизация	2	0	0
11	3	Нечеткая логика	2	0	0
12	3	Оценка качества работы алгоритма	2	0	0
13	3	Ансамбль методов	4	0	0
14	4	Стеммер	2	0	0
15	4	Лематизатор	4	0	0
16	4	Формирование n-грамм	2	0	0
17	4	Классификатор	4	0	0
18	4	Определение тональности текста	2	0	0
19	4	Морфологический, синтаксический, семантический анализ текста	4	0	0
Всего			54	0	0

3.4 Лабораторные занятия

№ п/п	№ раздела дисциплины	Наименование занятий	Объем в акад. часах		
			Всего	в том числе, в инновационной форме	в том числе, в электронной форме
Всего					

5 Фонд оценочных средств для проведения промежуточной аттестации

Оценочные средства находятся в приложении к рабочим программам дисциплин.

6 Перечень основной и дополнительной учебной литературы, необходимой для освоения дисциплины (модуля)

6.1. Основная литература			
	Авторы, составители	Заглавие	Издательство, год
Л1.1	Кузенков М. В., Середкин В.Г.	Методы, техника измерений и математическая обработка данных: электрон. учеб.-метод. комплекс дисциплины	Красноярск: ИПК СФУ, 2007
Л1.2	Кравцова Е. Д., Никифорова Э. М.	Математическое планирование эксперимента и статистическая обработка результатов: учеб.-метод. пособие для самост. работы студентов спец. 150108 «Порошковая металлургия, композиционные материалы, покрытия», 150701 «Физико-химические методы исследования металлургических процессов»	Красноярск: СФУ, 2012
Л1.3	Чашкин Ю. Р.	Математическая статистика. Анализ и обработка данных: учеб. пособие для вузов	Ростов-на-Дону: Феникс, 2010
Л1.4	Жолен Л., Кифер М., Дидри О., Вальтер Э.	Прикладной интервальный анализ	Москва, Ижевск: Регулярная и хаотическая динамика, Ижевский институт компьютерных исследований, 2007
6.2. Дополнительная литература			
	Авторы, составители	Заглавие	Издательство, год

Л2.1	Крамер Д., Митин О. В.	Математическая обработка данных в социальных науках: современные методы: перевод с английского	Москва, 2007
Л2.2	Шокин Ю. И., Яненко Н. Н.	Интервальный анализ: монография	Новосибирск: Наука. Сибирское отделение [СО], 1981
Л2.3	Бондаренко В. Н.	Помехоустойчивость приема спектрально-эффективных шумоподобных сигналов: монография	Красноярск: СФУ, 2015
6.3. Методические разработки			
	Авторы, составители	Заглавие	Издательство, год
Л3.1	Даничев А. А.	Компьютерные технологии поддержки принятия решений: учеб.-метод. пособие для лаб. работ студентов направления 220100.62 «Системный анализ и управление».	Красноярск: СФУ, 2012

7 Перечень ресурсов информационно-телекоммуникационной сети «Интернет», необходимых для освоения дисциплины (модуля)

Э1	Градиентный бустинг	https://neurohive.io/ru/osnovy-data-science/gradientyj-busting/
Э2	Компьютерная лингвистика	https://kelijah.livejournal.com/

8 Методические указания для обучающихся по освоению дисциплины (модуля)

Источники для самостоятельного изучения теоретического материала по разделам, указанным в п.3, приведены, соответственно, в пп. 6 и 7 настоящей рабочей программы.

Изучение дисциплины рассчитано только на третий семестр и завершается экзаменом. Минимальное количество набранных в течение семестра баллов для получения допуска к экзамену должно быть объявлено студентам в начале семестра.

Баллы набираются за выполнение работ и их защиту, которая может проводиться в аудитории или, по согласованию с преподавателем, дистанционно. При оценивании используется шкала от 0 до 5 баллов (критерии оценивания зависят от конкретной работы).

Оформление отчетов выполняется согласно СТУ 7.5-07-2021, если это не оговорено особо. Преподаватель осуществляет промежуточный контроль выполнения всех практических работ.

9 Перечень информационных технологий, используемых при осуществлении образовательного процесса по дисциплине (модулю) (при необходимости)

9.1 Перечень необходимого программного обеспечения

9.1.1	Любой инструментарий для статистического анализа.
-------	---

9.2 Перечень необходимых информационных справочных систем

9.2.1	Не требуется
-------	--------------

10 Материально-техническая база, необходимая для осуществления образовательного процесса по дисциплине (модулю)

Учебная аудитория для проведения лекционных и практических занятий, групповых и индивидуальных консультаций, текущего контроля и промежуточной аттестации, а также помещение для самостоятельной работы, содержащее специализированную мебель, компьютеры с подключением к сети Интернет (неограниченный доступ) и обеспечением доступа в электронную информационно-образовательную среду Университета, доступом к системе виртуальных машин; демонстрационное оборудование: интерактивная доска обратной проекции, проектор, экран для проектора; доступ к беспроводной сети WI-FI, маркерная доска.