

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

Ю.Ю. Якунин

подпись, инициалы, фамилия

«___» _____ 20__ г.

институт, реализующий дисциплину

**РАБОЧАЯ ПРОГРАММА ДИСЦИПЛИНЫ
АНАЛИЗ
НЕСТРУКТУРИРОВАННЫХ
ДАННЫХ**

Дисциплина Б1.В.ДВ.04.01 Анализ неструктурированных данных

Направление подготовки /
специальность 27.03.03 Системный анализ и управление
2018г.

Направленность
(профиль)

Форма обучения

очная

Год набора

2018

Красноярск 2021

РАБОЧАЯ ПРОГРАММА ДИСЦИПЛИНЫ

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

270000 «УПРАВЛЕНИЕ В ТЕХНИЧЕСКИХ СИСТЕМАХ»

Направление подготовки /специальность (профиль/специализация)

Направление 27.03.03 Системный анализ и управление 2018г.

Программу
составили

канд. техн. наук, доцент, Даничев Алексей
Александрович

1 Цели и задачи изучения дисциплины

1.1 Цель преподавания дисциплины

Целью преподавания дисциплины "Анализ 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

ОПК-2: способностью применять аналитические, вычислительные и системно-аналитические методы для решения прикладных задач в области управления объектами техники, технологии, организационными системами, работать с традиционными носителями информации, базами знаний

ОПК-7: способностью к освоению новой техники, новых методов и новых технологий

ПК-1: способностью принимать научно-обоснованные решения на основе математики, физики, химии, информатики, экологии, методов системного анализа и теории управления, теории знаний, осуществлять постановку и выполнять эксперименты по проверке их корректности и эффективности

1.4 Место дисциплины (модуля) в структуре образовательной программы

Методы искусственного интеллекта

Методы оптимизации

Теория эволюционных вычислений

Нейронные сети и их приложения

Теория и технология программирования

Математическая статистика

Объектно-ориентированное программирование

Теория вероятностей

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=22468>

2. Объем дисциплины (модуля)

Вид учебной работы	Всего, зачетных единиц (акад.час)	Семестр
		7
Общая трудоемкость дисциплины	7 (252)	7 (252)
Контактная работа с преподавателем:	3 (108)	3 (108)
занятия лекционного типа	1 (36)	1 (36)
занятия семинарского типа		
в том числе: семинары		
практические занятия	2 (72)	2 (72)
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:	3 (108)	3 (108)
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	Анализ текста	6	12	0	36	ОПК-2 ОПК-7 ПК-1
2	Структурирование информации	12	24	0	36	ОПК-2 ОПК-7 ПК-1
3	Элементы теории вероятностей и математической статистики	18	36	0	36	ОПК-2 ОПК-7 ПК-1
Всего		36	72	0	108	

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	Способы анализа текстовой информации	2	0	0
2	1	Обработка текста для машинного обучения	2	0	0
3	1	Определение тональности текста	2	0	0
4	2	Кластеризация и классификация	2	0	0
5	2	Численная оценка качества работы алгоритма	2	0	0
6	2	Ансамбль методов	2	0	0

7	2	Факторизация	2	0	0
8	2	Нейронные сети	2	0	0
9	2	Нечеткая логика	2	0	0
10	3	Библиотеки для статистических расчётов	2	0	0
11	3	Случайные величины. Характеристики выборок	2	0	0
12	3	Устойчивые распределения	2	0	0
13	3	Проверка стат. гипотез	2	0	0
14	3	Непараметрические критерии	2	0	0
15	3	Проверка закона распределения СВ	2	0	0
16	3	Связные выборки	2	0	0
17	3	Анализ мощности	2	0	0
18	3	Планирование эксперимента	2	0	0
Итого			26	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	Стеммер, лематизатор	4	0	0
2	1	Формирование n-грамм	4	0	0
3	1	Определение тональности текста	4	0	0
4	2	Определение тематики текста	4	0	0
5	2	Численная оценка качества работы алгоритма	4	0	0
6	2	Ансамбль методов	4	0	0
7	2	Факторизация	4	0	0
8	2	Нейронные сети	4	0	0
9	2	Нечеткая логика	4	0	0
10	3	Нормализация данных	4	0	0
11	3	Построение валидных гистограмм	4	0	0
12	3	Закон Бенфорда	4	0	0
13	3	Проверка гипотез	4	0	0

14	3	Непараметрические критерии	4	0	0
15	3	Проверка закона распределения СВ	4	0	0
16	3	Связные выборки	4	0	0
17	3	Предсказание ошибки эксперимента	4	0	0
18	3	Интерпретация результатов эксперимента	4	0	0
Итого			72	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
Итого					

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	Хайкин С.	Нейронные сети: полный курс	Санкт-Петербург: Вильямс, 2006
Л1.2	Шевелева И. В., Васильева А. В.	Теория вероятностей и математическая статистика: методические указания по выполнению контрол. работ для студентов заоч. формы обучения	Красноярск: ИПК СФУ, 2007
Л1.3	Сидняев Н. И., Вилисова Н. Т.	Введение в теорию планирования эксперимента: учебное пособие для вузов по направлению "Машиностроительные технологии и оборудование" специальности "Реновация средств и объектов материального производства в машиностроении"	Москва: МГТУ имени Н.Э.Баумана, 2011
6.3. Методические разработки			

	Авторы, составители	Заглавие	Издательство, год
ЛЗ.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
ЛЗ.2	Помазан В. А., Ковалевич И. А., Шестаков В. Н., Даничев А. А.	Информационные технологии в управлении человеческими ресурсами: электрон. учеб.-метод. комплекс дисциплины	Красноярск: СФУ, 2008

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

Э1	закон Бенфорда	https://habr.com/ru/post/240853/
Э2	Анализ мощности	http://statsoft.ru/home/textbook/module/s/stpowan_rus.html
Э3	Обработка текста	https://nlpub.ru/Обработка_текста

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 седьмой 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.