

Министерство науки и высшего образования РФ
Федеральное государственное автономное образовательное учреждение
высшего образования
«СИБИРСКИЙ ФЕДЕРАЛЬНЫЙ УНИВЕРСИТЕТ»
РАБОЧАЯ ПРОГРАММА ДИСЦИПЛИНЫ (МОДУЛЯ)

Б1.В.ДВ.01.01 Аналитическая обработка
структурированных и неструктурированных данных
наименование дисциплины (модуля) в соответствии с учебным планом

Направление подготовки / специальность

09.04.04 Программная инженерия

Направленность (профиль)

09.04.04.02 Технологии индустриального производства программного
обеспечения интеллектуальных систем управления

Форма обучения

очная

Год набора

2021

РАБОЧАЯ ПРОГРАММА ДИСЦИПЛИНЫ (МОДУЛЯ)

Программу составили _____

_____ канд. техн. наук, доцент, Даничев Алексей Александрович

_____ должность, инициалы, фамилия

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

Код и наименование индикатора достижения компетенции	Запланированные результаты обучения по дисциплине
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	Знать: методы статистического анализа данных и систематизации информации неструктурированных данных.
ПК-1.2: Уметь применять методологии разработки программного обеспечения при управлении инфраструктурой коллективной среды разработки	Уметь организовывать и проводить аналитические исследования
ПК-1.3: Иметь навыки управления инфраструктурой коллективной среды разработки	Владеть навыками решения практических задач прогнозирования и выработки рекомендаций.

1.4 Особенности реализации дисциплины

Язык реализации дисциплины: Русский.

Дисциплина (модуль) реализуется с применением ЭО и ДОТ

URL-адрес и название электронного обучающего курса: <https://e.sfu->

kras.ru/course/view.php?id=16388.

2. Объем дисциплины (модуля)

Вид учебной работы	Всего, зачетных единиц (акад.час)	е
		1
Контактная работа с преподавателем:	2 (72)	
занятия лекционного типа	0,5 (18)	
практические занятия	1,5 (54)	
Самостоятельная работа обучающихся:	4 (144)	
курсовое проектирование (КП)	Нет	
курсовая работа (КР)	Нет	
Промежуточная аттестация (Экзамен)	1 (36)	

3 Содержание дисциплины (модуля)

3.1 Разделы дисциплины и виды занятий (тематический план занятий)

№ п/п	Модули, темы (разделы) дисциплины	Контактная работа, ак. час.							
		Занятия лекционного типа		Занятия семинарского типа				Самостоятельная работа, ак. час.	
				Семинары и/или Практические занятия		Лабораторные работы и/или Практикумы			
		Всего	В том числе в ЭИОС	Всего	В том числе в ЭИОС	Всего	В том числе в ЭИОС	Всего	В том числе в ЭИОС
1. Подготовка данных									
	1. Описательная статистика, характеристики выборок	2							
	2. Нормализация данных			2					
	3. Построение валидных гистограмм			4					
	4. Устойчивые распределения, белый шум	2							
	5. Закон Бенфорда			2					
	6. Белый шум			4					
	7. Платформы для обработки неструктурированных данных							36	
2. Планирование эксперимента									
	1. Планирование эксперимента	2							
	2. Проверка гипотез			4					
	3. Интерпретация результатов эксперимента			2					
	4. Анализ мощности	2							

5. Предсказание ошибки эксперимента			4					
6. Интервальный анализ			2					
7. Планирование эксперимента							36	
3. Структурирование информации								
1. Кластеризация, факторизация, нечеткая логика	2							
2. Кластеризация			2					
3. Факторизация			2					
4. Нечеткая логика			2					
5. Оценка качества работы алгоритма; ансамбли методов	2							
6. Оценка качества работы алгоритма			2					
7. Ансамбль методов			4					
8. Структурирование информации							36	
4. Анализ текста								
1. Обработка текста для машинного обучения	2							
2. Стеммер			2					
3. Лематизатор			4					
4. Способы анализа текстовой информации	2							
5. Формирование n-грамм			2					
6. Классификатор			4					
7. Словари	2							
8. Определение тональности текста			2					
9. Морфологический, синтаксический, семантический анализ текста			4					
10. Способы анализа текстовой информации							36	
Всего	18		54				144	

4 Учебно-методическое обеспечение дисциплины

4.1 Печатные и электронные издания:

1. Кузенков М. В., Середкин В.Г. Методы, техника измерений и математическая обработка данных: электрон. учеб.-метод. комплекс дисциплины(Красноярск: ИПК СФУ).
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).
3. Чашкин Ю. Р. Математическая статистика. Анализ и обработка данных: учеб. пособие для вузов(Ростов-на-Дону: Феникс).
4. Жолен Л., Кифер М., Дидри О., Вальтер Э. Прикладной интервальный анализ(Москва, Ижевск: Регулярная и хаотическая динамика, Ижевский институт компьютерных исследований).
5. Крамер Д., Митин О. В. Математическая обработка данных в социальных науках: современные методы: перевод с английского (Москва).
6. Шокин Ю. И., Яненко Н. Н. Интервальный анализ: монография (Новосибирск: Наука. Сибирское отделение [СО]).
7. Бондаренко В. Н. Помехоустойчивость приема спектрально-эффективных шумоподобных сигналов: монография(Красноярск: СФУ).
8. Даничев А. А. Компьютерные технологии поддержки принятия решений: учеб.-метод. пособие для лаб. работ студентов направления 220100.62 «Системный анализ и управление».(Красноярск: СФУ).

4.2 Лицензионное и свободно распространяемое программное обеспечение, в том числе отечественного производства (программное обеспечение, на которое университет имеет лицензию, а также свободно распространяемое программное обеспечение):

1. Любой инструментарий для статистического анализа.

4.3 Интернет-ресурсы, включая профессиональные базы данных и информационные справочные системы:

1. Не требуется

5 Фонд оценочных средств

Оценочные средства находятся в приложении к рабочим программам дисциплин.

6 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.