

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

СОГЛАСОВАНО

Заведующий кафедрой

Кафедра прикладной
математики и компьютерной
безопасности (ПМКБ_ИКИТ)

наименование кафедры

подпись, инициалы, фамилия

«___» _____ 20__ г.

институт, реализующий ОП ВО

УТВЕРЖДАЮ

Заведующий кафедрой

Кафедра прикладной математики
и компьютерной безопасности
(ПМКБ_ИКИТ)

наименование кафедры

А.А. Кытманов

подпись, инициалы, фамилия

«___» _____ 20__ г.

институт, реализующий дисциплину

**РАБОЧАЯ ПРОГРАММА ДИСЦИПЛИНЫ
ПРИКЛАДНЫЕ ЗАДАЧИ АНАЛИЗА
ДАНЫХ (APPLIED DATA
ANALYSIS)**

Дисциплина Б1.В.04 Прикладные задачи анализа данных (Applied Data
Analysis)

Направление подготовки / 01.04.02 Прикладная математика и
специальность информатика, программа 01.04.02.09 Data
Science and Mathematical Modeling 2020г

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

010000 «МАТЕМАТИКА И МЕХАНИКА»

Направление подготовки /специальность (профиль/специализация)

Направление 01.04.02 Прикладная математика и информатика,
программа 01.04.02.09 Data Science and Mathematical Modeling 2020г.

Программу
составили

1 Цели и задачи изучения дисциплины

1.1 Цель преподавания дисциплины

Целью изучения дисциплины является ознакомление студентов с основными методами и моделями анализа данных.

1.2 Задачи изучения дисциплины

Задачами изучения дисциплины являются:

- Ознакомление с задачами обучения на размеченных данных, основными линейными моделями, методами оценки качества прогнозирования, возможными причинами низкого качества моделей;
- Выработка умений постановки задач анализа данных, обоснованного выбора и применения изученных методов в прикладных исследованиях
- Получение практических навыков программирования алгоритмов машинного обучения средствами языка R.

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

УК-3:Способен организовывать и руководить работой команды, вырабатывая командную стратегию для достижения поставленной цели.
УК-3.1:Вырабатывает стратегию сотрудничества и на ее основе организует отбор членов команды для достижения поставленной цели.
УК-3.2:Планирует и корректирует работу команды с учетом интересов, особенностей поведения и мнений ее членов.
УК-3.4:Организует дискуссии по заданной теме и обсуждение результатов работы команды с привлечением оппонентов разработанным идеям.
УК-3.5:Планирует командную работу, распределяет поручения и делегирует полномочия членам команды.
УК-2:Способен управлять проектом на всех этапах его жизненного цикла.
УК-2.1:Формулирует на основе поставленной проблемы проектную задачу и способ ее решения через реализацию проектного управления.
УК-2.2:Разрабатывает концепцию проекта в рамках обозначенной проблемы: формулирует цель, задачи, обосновывает актуальность, значимость, ожидаемые результаты и возможные сферы их применения.
УК-2.3:Планирует необходимые ресурсы, в том числе с учетом их заменимости.
УК-2.4:Разрабатывает план реализации проекта.
УК-2.5:Осуществляет мониторинг хода реализации проекта, корректирует отклонения, вносит дополнительные изменения в план реализации проекта, уточняет зоны ответственности участников проекта.
УК-1:Способен осуществлять критический анализ проблемных ситуаций на основе системного подхода, вырабатывать стратегию действий.

УК-1.1:Анализирует проблемную ситуацию как систему, выявляя ее составляющие и связи между ними.
УК-1.2:Определяет пробелы в информации, необходимой для решения проблемной ситуации, и проектирует процессы по их устранению.
УК-1.3:Критически оценивает надежность источников информации, работает с противоречивой информацией из разных источников.
УК-1.4:Разрабатывает и содержательно аргументирует стратегию решения проблемной ситуации на основе системного и междисциплинарных подходов.
ПК-1:Способен преподавать по программам бакалавриата, специалитета, магистратуры и дополнительных профессиональных программ, ориентированных на соответствующий уровень квалификации.
ПК-1.1:Знает: особенности организации образовательного процесса по программам бакалавриата, специалитета и магистратуры; преподаваемую область научного знания и профессиональной деятельности; современные образовательные технологии профессионального образования; требования, предъявляемые профессией к человеку, набор медицинских и иных противопоказаний при выборе профессии, содержание и условия труда, образ жизни работников данной профессии, возможности и перспективы карьерного роста по профессии; требования охраны труда при проведении учебных занятий в организации, осуществляющей образовательную деятельность, и вне организации.
ПК-1.2:Умеет: выполнять деятельность и (или) демонстрировать элементы деятельности, осваиваемой обучающимися, и (или) выполнять задания, предусмотренные программой учебного курса, дисциплины (модуля).
ПК-1.3:Способен проводить учебные занятия по программам бакалавриата.
ПК-3:Способен управлять разработкой продуктов, услуг и решений на основе данных.
ПК-3.1:Знает: состояние и перспективы развития информационных технологий, технологий данных в России и в мире; современные и перспективные методы сбора, хранения и передачи данных; источники данных, интенсивность генерации данных источниками; технические средства и среды сбора, хранения и обработки данных; современные и перспективные средства визуализации и интерпретации данных; исследование операций; машинное обучение; математическое моделирование; методы сравнительного анализа.
ПК-3.2:Способен проводить аналитические и поисковые исследования по тематике информационных технологий, технологий данных.
ПК-4:Способен разрабатывать и внедрять новые методы и технологии исследования данных.
ПК-4.1:Знает: состояние и перспективы развития информационных технологий, технологий данных в России и в мире; современные и перспективные методы сбора, хранения и передачи данных; источники данных, интенсивность генерации данных источниками; технические средства и среды сбора, хранения и обработки данных; современные и перспективные средства визуализации и интерпретации данных; исследование операций; машинное обучение; математическое моделирование; методы сравнительного анализа.
ПК-4.2:Способен проводить аналитические и поисковые исследования по тематике информационных технологий, технологий данных.

1.4 Место дисциплины (модуля) в структуре образовательной

программы

Продвинутые методы анализа данных

Базовые модели и методы машинного обучения

Системы хранения, обработки и управления данными

Продвинутые методы анализа данных (Advanced Methods of Data Analysis)

Базовые модели и методы машинного обучения (The Basics of Machine Learning)

Системы хранения, обработки и управления данными (Data Storage and Management Systems)

Анализ больших данных

Машинное обучение и криптография

Статистическое моделирование

выполнение и защита выпускной квалификационной работы

Анализ больших данных (Big Data)

Машинное обучение и криптография (Applications of Machine Learning in Cryptography)

Статистическое моделирование (Statistical Modeling)

Выполнение и защита выпускной квалификационной работы (Final certification)

1.5 Особенности реализации дисциплины

Язык реализации дисциплины Английский.

Дисциплина (модуль) реализуется без применения ЭО и ДОТ.

2. Объем дисциплины (модуля)

Вид учебной работы	Всего, зачетных единиц (акад.час)	Семестр
		3
Общая трудоемкость дисциплины	4 (144)	4 (144)
Контактная работа с преподавателем:	1,5 (54)	1,5 (54)
занятия лекционного типа	0,5 (18)	0,5 (18)
занятия семинарского типа		
в том числе: семинары		
практические занятия	1 (36)	1 (36)
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:	1,5 (54)	1,5 (54)
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	Data analysis for Engineering equipment monitoring	4	6	0	10	
2	Data analysis for business	4	8	0	16	
3	Learning analytics	6	12	0	14	
4	Data analysis in medicine	4	10	0	14	
Всего		18	36	0	54	

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	Anomaly detection	2	0	0
2	1	One-class support vector machines. Local Outlier Factor. Изолирующий лес	2	0	0
3	2	Customer behavior analysis	2	0	0
4	2	Sales network segmentation	2	0	0
5	3	Prediction of learning success	2	0	0
6	3	Student simulation	2	0	0

7	3	Social nets analysis	2	0	0
8	4	Pattern recognition	4	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	Data analysis for Engineering equipment monitoring	6	0	0
2	2	Data analysis for business	8	0	0
3	3	Learning analytics	12	0	0
4	4	Pattern recognition	10	0	0
Всего			36	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	Neter J., Wasserman W., Kutner M. H.	Applied Linear Statistical Models: Regression, Analysis of Variance and Experimental Designs	Boston: Irwin, 1990
6.2. Дополнительная литература			
	Авторы, составители	Заглавие	Издательство, год
Л2.1	Friedrich F.	Stochastic Simulation and Bayesian Inference for Gibbs Fields: [приложение к книге "Г. Винклер. Анализ изображений..."]	Berlin: Springer, 2003

Л2.2	Galushkin A. I.	Neural Networks Theory: with 176 figures	New York: Springer-Verlag, 2007
Л2.3	Dunn P. F.	Measurement and Data Analysis for Engineering and Science: научное издание	Boca Raton: CRC Press, 2010

8 Методические указания для обучающихся по освоению дисциплины (модуля)

Для успешного усвоения дисциплины и допуска к экзамену необходимо:

- Прослушать все лекции или, в случае пропуска занятий, изучить из материал самостоятельно по рекомендуемым источникам литературы;
- Выполнить групповой проект и защитить его в форме докладов на семинаре.

Экзамент сдается в устной форме. Каждый экзаменационный билет состоит из 2-х теоретических вопросов из списка вопросов к экзамену.

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

9.1 Перечень необходимого программного обеспечения

9.1.1	Для выполнения групповых проектов необходимы
9.1.2	- программная среда вычислений R версии не младше, чем 3.4.3, а также редактор кода RStudio версии 1.1 или старше.
9.1.3	- программная среда Python версии 3.0 и выше и редактор кода Jupiter

9.2 Перечень необходимых информационных справочных систем

9.2.1	1. Система электронного обучения Сибирского федерального университета (https://e.sfu-kras.ru)
9.2.2	2. электронные информационно-справочные ресурсы научной библиотеки СФУ (http://bik.sfu-kras.ru)
9.2.3	3. Интернет-репозиторий статистических данных Data-Planet (https://www.data-planet.com/)
9.2.4	4. Интернет-репозиторий статистических данных на платформе kaggle.com. Режим доступа: https://www.kaggle.com/datasets

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

Учебные лаборатории и классы, оснащенные современными компьютерами, объединенными в локальные вычислительные сети с выходом в Интернет, а также периферийным и проекционным оборудованием.