



МИНОБРНАУКИ РОССИИ
федеральное государственное бюджетное образовательное учреждение
высшего образования
«ИРКУТСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ»
ФГБОУ ВО «ИГУ»
Кафедра алгебраических и информационных систем

«УТВЕРЖДАЮ»
Директор ИМИТ ИГУ

М. В. Фалалеев
«19» мая 2021 г.



Рабочая программа дисциплины

Наименование дисциплины (модуля) Б1.В.08 Интеллектуальный анализ данных

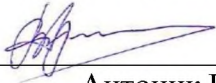
Направление подготовки 02.03.02 Фундаментальная информатика и информационные технологии

Направленность (профиль) подготовки Фундаментальная информатика и информационные технологии


Квалификация выпускника – бакалавр

Форма обучения очная

Согласовано с УМК Института математики
и информационных технологий
Протокол № 4 от «19» мая 2021 г.

Председатель  _____
Антоник В.Г.

Рекомендовано кафедрой Алгебраических и
информационных систем ИМИТ ИГУ:
Протокол № 11 от «29» апреля 2021 г.

Зав. кафедрой  _____
Пантелеев В.И.

СОДЕРЖАНИЕ

1. Цели и задачи дисциплины	4
2. Место дисциплины в структуре опоп во.....	4
3. Требования к результатам освоения дисциплины	4
4. Содержание и структура дисциплины	7
4.1. СОДЕРЖАНИЕ ДИСЦИПЛИНЫ, СТРУКТУРИРОВАННОЕ ПО ТЕМАМ, С УКАЗАНИЕМ ВИДОВ УЧЕБНЫХ ЗАНЯТИЙ И ОТВЕДЕННОГО НА НИХ КОЛИЧЕСТВА АКАДЕМИЧЕСКИХ ЧАСОВ.....	7
4.2. ПЛАН ВНЕАУДИТОРНОЙ САМОСТОЯТЕЛЬНОЙ РАБОТЫ ОБУЧАЮЩИХСЯ ПО ДИСЦИПЛИНЕ	8
4.3. СОДЕРЖАНИЕ УЧЕБНОГО МАТЕРИАЛА	8
4.4. МЕТОДИЧЕСКИЕ УКАЗАНИЯ ПО ОРГАНИЗАЦИИ САМОСТОЯТЕЛЬНОЙ РАБОТЫ СТУДЕНТОВ	11
4.5. ПРИМЕРНАЯ ТЕМАТИКА КУРСОВЫХ РАБОТ (ПРОЕКТОВ).....	14
5. Учебно-методическое и информационное обеспечение дисциплины (модуля)....	15
6. Материально-техническое обеспечение дисциплины	15
7. Образовательные технологии	15
8. Оценочные материалы для текущего контроля и промежуточной аттестации	15

1. ЦЕЛИ И ЗАДАЧИ ДИСЦИПЛИНЫ

Цель

Целью курса является изучение современных подходов, моделей, алгоритмов анализа данных и решения задач распознавания, классификации, нахождения зависимостей.

Задачи:

Задачи дисциплины: изучение сущности и роли анализа информации; ознакомление бакалавров с основными средствами и методами анализа данных при исследовании сложных систем; знакомство с основными технологиями анализа данных; формирование умений и навыков в области поддержки принятия решений в организации с применением современных методов и средств анализа данных.

2. МЕСТО ДИСЦИПЛИНЫ В СТРУКТУРЕ ОПОП ВО

2.1. Учебная дисциплина (модуль) относится к части программы, формируемой участниками образовательных отношений, и изучается на третьем курсе.

2.2. Для изучения данной учебной дисциплины (модуля) необходимы знания, умения и навыки, сформированные дисциплинами «Математический анализ», «Алгебра», «Программирование».

2.3. Знания, умения и навыки, формируемые в процессе освоения дисциплины, используются при выполнении научно-исследовательских работ, для решения прикладных задач и выполнения выпускной квалификационной работы.

3. ТРЕБОВАНИЯ К РЕЗУЛЬТАТАМ ОСВОЕНИЯ ДИСЦИПЛИНЫ

Процесс освоения дисциплины направлен на формирование компетенций (элементов следующих компетенций) в соответствии с ФГОС ВО по соответствующему направлению подготовки.

Перечень планируемых результатов обучения по дисциплине (модулю), соотнесенных с индикаторами достижения компетенций

Компетенция	Индикаторы компетенций	Результаты обучения
ПК-2 Способность проводить под научным руководством локальные исследования на основе существующих методов в конкретной области профессиональной деятельности, принимать участие в научных дискуссиях, выступать с сообщениями и докладами, устного, письменного и виртуального (размещение в информационных сетях) характера, представлять материалы собственных исследований; проводить корректуру, редактирование, ре-	ИДК ПК2.3 Способен проводить под научным руководством исследования на основе существующих методов в области профессиональной деятельности	Знает методы и подходы решения практических задач анализа данных и классификации композициями алгоритмов. Умеет пользоваться своими знаниями для решения фундаментальных, прикладных и технологических задач в различных предметных областях; делать правильные выводы из сопоставления результатов теории и эксперимента, выбирать правильно параметры методов, адекватные размерности обучающих выборок.

<p>ферирование работ.</p>		<p>Владеет навыками анализа большого объема частично противоречивых и неполных признаков описаний; навыками грамотной обработки статистических многомерных данных, оформления результатов численных расчетов и их сопоставления с теоретическими оценками; практикой исследования и решения теоретических и прикладных задач; навыками анализа реальных задач из различных предметных областей на уровне отдельных подходов и коллективами алгоритмов.</p>
<p>ПК-5 Способность собирать, обрабатывать и интерпретировать экспериментальные данные, необходимые для проектной и производственно-технологической деятельности; способность к разработке новых алгоритмических, методических и технологических решений в конкретной сфере профессиональной деятельности, в том числе с применением технологий искусственного интеллекта</p>	<p>ИДК ПК5.1 Способен собирать, обрабатывать и интерпретировать экспериментальные данные, необходимые для проектной и производственно-технологической деятельности</p>	<p>Знает современные проблемы анализа данных, теории распознавания, классификации, поиска зависимостей; методы и подходы решения практических задач анализа данных и классификации композициями алгоритмов; программные средства решения основных задач анализа данных и классификации.</p> <p>Умеет делать правильные выводы из сопоставления результатов теории и эксперимента, выбирать правильно параметры методов, адекватные размерности обучающих выборок; получать оптимальные алгоритмы классификации и правильно оценивать степень их точности и достоверности.</p> <p>Владеет навыками грамотной обработки статистических многомерных данных, оформления результатов численных расчетов и их сопоставления с теоретическими оценками; практикой исследования и решения теоретических и прикладных задач; навыками анализа реальных задач из различных предметных областей на уровне отдельных подходов и коллективами алгоритмов.</p>
	<p>ИДК ПК5.3</p>	<p>Умеет пользоваться своими</p>

	<p>Способен применять методы искусственного интеллекта при решении типовых задач в профессиональной области</p>	<p>знаниями для решения фундаментальных, прикладных и технологических задач в различных предметных областях; эффективно использовать информационные технологии и компьютерную технику для достижения необходимых теоретических и прикладных результатов.</p> <p>Владеет навыками анализа большого объема частично противоречивых и неполных признаков описаний; практикой исследования и решения теоретических и прикладных задач; навыками анализа реальных задач из различных предметных областей на уровне отдельных подходов и коллективами алгоритмов.</p>
<p>ПК-3 Способность понимать и применять в научно-исследовательской и прикладной деятельности современный математический аппарат, теоретические основы информатики</p>	<p>ИДК ПК3.1 Способен понимать современный математический аппарат и теоретические основы информатики</p>	<p>Знает современные проблемы анализа данных, теории распознавания, классификации, поиска зависимостей; методы и подходы решения практических задач анализа данных и классификации коллективами алгоритмов; программные средства решения основных задач анализа данных и классификации.</p>
	<p>ИДК ПК3.2 Способен применять в научно-исследовательской и прикладной деятельности современный математический аппарат</p>	<p>Владеет навыками грамотной обработки статистических многомерных данных, оформления результатов численных расчетов и их сопоставления с теоретическими оценками; практикой исследования и решения теоретических и прикладных задач.</p>
	<p>ИДК ПК3.3 Способен применять в научно-исследовательской и прикладной деятельности теоретические основы информатики</p>	<p>Умеет пользоваться своими знаниями для решения фундаментальных, прикладных и технологических задач в различных предметных областях; эффективно использовать информационные технологии и компьютерную технику для достижения необходимых теоретических и прикладных результатов.</p>

4. СОДЕРЖАНИЕ И СТРУКТУРА ДИСЦИПЛИНЫ

Объем дисциплины составляет 5 зачетных единиц, 180 часов, в том числе 35 часов на контроль, практическая подготовка _____.
 Форма промежуточной аттестации: 6 семестр - экзамен.

4.1. СОДЕРЖАНИЕ ДИСЦИПЛИНЫ, СТРУКТУРИРОВАННОЕ ПО ТЕМАМ, С УКАЗАНИЕМ ВИДОВ УЧЕБНЫХ ЗАНЯТИЙ И ОТВЕДЕННОГО НА НИХ КОЛИЧЕСТВА АКАДЕМИЧЕСКИХ ЧАСОВ

№ п/п	Раздел дисциплины/темы	Семестр	Виды учебной работы, включая самостоятельную работу обучающихся и трудоемкость (в часах)				Формы текущего контроля успеваемости
			Контактная работа преподавателя с обучающимися			Самостоятельная работа + контроль	
			Лекции	Семинарские (практические занятия)	Контроль обучения		
1	Основные понятия анализа данных.		2				лаб.
2	Линейная регрессия.		6	8	1	15	лаб.
3	Логистическая регрессия.		2	2	1	15	лаб.
4	Регуляризация.		2	4	1	15	лаб.
5	Метрики классификации.		2		1		лаб.
6	Метод k ближайших соседей.		2	2	1		лаб.
7	Метод опорных векторов.		2	2	1		лаб.
8	Деревья решений.		2	2	1		лаб.
9	Композиции деревьев.		4	2	1		лаб.
10	Рекомендательные системы.		4	4	1	15	лаб.
11	Нейронные сети.		4	6	1	11	лаб.
12	Подготовка к экзамену.					35	
Итого часов			32	32	10	106	

4.2. ПЛАН ВНЕАУДИТОРНОЙ САМОСТОЯТЕЛЬНОЙ РАБОТЫ ОБУЧАЮЩИХСЯ ПО ДИСЦИПЛИНЕ

Се- местр	Название раздела, темы	Самостоятельная работа обучающихся			Оценочное сред- ство	Учебно- методическое обеспечение са- мостоятельной работы
		Вид самостоя- тельной работы	Сроки выполне- ния	Затраты времени (час.)		
6	Линейная регрессия.	Выполнение практической работы	01.03	15	Проверка до- машней работы	Материалы кур- са на платформе ИОС DOMIC
6	Логистическая регрессия.	Выполнение практической работы	01.04	15	Проверка до- машней работы	Материалы кур- са на платформе ИОС DOMIC
6	Регуляризация.	Выполнение практической работы	01.04	15	Проверка до- машней работы	Материалы кур- са на платформе ИОС DOMIC
6	Рекомендательные системы.	Выполнение практической работы	01.05	15	Проверка до- машней работы	Материалы кур- са на платформе ИОС DOMIC
6	Нейронные сети.	Выполнение практической работы	01.06	11	Проверка до- машней работы	Материалы кур- са на платформе ИОС DOMIC
6	Подготовка к экзамену.		14.06	35		
Общая трудоемкость самостоятельной работы по дисциплине (час)				106		
Из них объем самостоятельной работы с использованием электронного обучения и дистанци- онных образовательных технологий (час)				106		

4.3. СОДЕРЖАНИЕ УЧЕБНОГО МАТЕРИАЛА

1. Основные понятия анализа данных.
Введение. Задачи анализа данных. Основные понятия анализа данных. Регрессия и классификация. Основы Python.
2. Линейная регрессия.
Линейная регрессия. Гипотеза и функция потерь. Градиентный спуск. Визуализация. Библиотека Matplotlib.

3. Логистическая регрессия.
Задача классификации. Логистическая регрессия. Гипотеза и функция потерь. Библиотеки Numpy, Pandas.
4. Регуляризация.
Смещение моделей машинного обучения. Разброс моделей машинного обучения. Компромисс между смещением и разбросом. Недообучение и переобучение. Регуляризация.
5. Метрики классификации.
Точность, полнота, F-мера. AUC-ROC.
6. Метод k ближайших соседей.
Метод k ближайших соседей. Основные параметры. Метод k ближайших соседей в scikit-learn.
7. Метод опорных векторов.
Метод опорных векторов. Основные параметры. Метод опорных векторов в scikit-learn.
8. Деревья решений.
Деревья решений. Основные параметры. Деревья решений в scikit-learn.
9. Композиции деревьев.
Композиции деревьев решений. Бэггинг. Случайный лес.
10. Рекомендательные системы.
Задачи и примеры рекомендательных систем. Рекомендательные системы на основе контента. Ассоциативные правила. Совместная фильтрация. Рекомендательные системы второго поколения. Контекстно-зависимые рекомендательные системы.
11. Нейронные сети.
Модель нейрона. Персептрон. Линейный нейрон. Функции активации. Метод обратного распространения ошибки. Библиотека TensorFlow. Распознавание образов. Сверточные нейронные сети. Рекуррентные нейронные сети.

4.3.1. Перечень семинарских, практических занятий и лабораторных работ

№ п/н	№ раздела и темы	Наименование семинаров, практических и лабораторных работ	Трудоемкость (час.)		Оценочные средства	Формируемые компетенции (индикаторы)*
			Всего часов	Из них практическая подготовка		
1	2	3	4	5	6	7
1	2	Линейная регрессия с одной переменной.	2	2	Проверка лабораторной работы	ПК-2, ПК-3, ПК-5

2	2	Визуализация.	2	2	Проверка лабораторной работы	ПК-2, ПК-3, ПК-5
3	2	Линейная регрессия в scikit-learn.	1	1	Проверка лабораторной работы	ПК-2, ПК-3, ПК-5
4	2	Линейная регрессия с несколькими переменными.	2	2	Проверка лабораторной работы	ПК-2, ПК-3, ПК-5
5	2	Линейная регрессия с несколькими переменными в scikit-learn.	1	1	Проверка лабораторной работы	ПК-2, ПК-3, ПК-5
6	3	Логистическая регрессия.	2	2	Проверка лабораторной работы	ПК-2, ПК-3, ПК-5
7	4	Логистическая регрессия с регуляризацией в scikit-learn.	4	4	Проверка лабораторной работы	ПК-2, ПК-3, ПК-5
8	6	Метод k ближайших соседей.	1	1	Проверка лабораторной работы	ПК-2, ПК-3, ПК-5
9	6	Подбор параметров для метода kNN.	1	1	Проверка лабораторной работы	ПК-2, ПК-3, ПК-5
10	7	Метод опорных векторов.	2	2	Проверка лабораторной работы	ПК-2, ПК-3, ПК-5
11	8	Деревья решений.	1	1	Проверка лабораторной работы	ПК-2, ПК-3, ПК-5
12	8	Визуализация деревьев решений.	1	1	Проверка лабораторной работы	ПК-2, ПК-3, ПК-5
13	9	Случайный лес.	2	2	Проверка лабораторной работы	ПК-2, ПК-3, ПК-5
14	10	Поиск ассоциативных правил.	2	2	Проверка лабораторной работы	ПК-2, ПК-3, ПК-5
15	10	Метод коллаборативной фильтрации.	2	2	Проверка лабораторной работы	ПК-2, ПК-3, ПК-5
16	11	Алгоритм обучения перцептрона.	2	2	Проверка лабораторной работы	ПК-2, ПК-3, ПК-5
17	11	TensorFlow playground.	2	2	Проверка лабораторной работы	ПК-2, ПК-3, ПК-5
18	11	Полносвязная сеть в TensorFlow.	2	2	Проверка лабораторной работы	ПК-2, ПК-3, ПК-5
		Всего	32	32		

4.3.2. Перечень тем (вопросов), выносимых на самостоятельное изучение студентами в рамках самостоятельной работы (СР)

Тема	Задание	Формируемые компетенции
Линейная регрессия.	Решить задачи для самостоятельного выполнения	ПК-2, ПК-3, ПК-5
Логистическая регрессия.	Решить задачи для самостоятельного выполнения	ПК-2, ПК-3, ПК-5
Регуляризация.	Решить задачи для самостоятельного выполнения	ПК-2, ПК-3, ПК-5
Рекомендательные системы.	Решить задачи для самостоятельного выполнения	ПК-2, ПК-3, ПК-5
Нейронные сети.	Решить задачи для самостоятельного выполнения	ПК-2, ПК-3, ПК-5

4.4. МЕТОДИЧЕСКИЕ УКАЗАНИЯ ПО ОРГАНИЗАЦИИ САМОСТОЯТЕЛЬНОЙ РАБОТЫ СТУДЕНТОВ

Во время изучения дисциплины студент посещает лекции, практические занятия, выполняет лабораторные задания, готовится к тестам, зачетам и экзаменам. Для каждого вида деятельности необходимо правильно организовать самостоятельную работу.

Лекции. В высшем учебном заведении лекция является важной формой учебного процесса. На лекции студенты получают глубокие и разносторонние знания. Лекция способствует развитию творческих способностей, формирует идейную убежденность, позволяет устанавливать связь учебного материала с производством, новейшими научными достижениями. Лекция требует три вида деятельности: подготовку к лекции, работу на лекции и работу после лекции.

После прослушивания лекции студент должен проработать и осмыслить полученный материал. На каждый пример, приведенный на лекции, желательно, (если это возможно) привести свой. Материал, изложенный в лекции, можно просмотреть в других источниках.

В процессе лекционного занятия студент должен выделять важные моменты, выводы, анализировать основные положения. Недостаточно только «слушать» лекцию. Возможности памяти человека не универсальны. Как бы внимательно студент ни слушал лекцию, большая часть информации вскоре после восприятия будет забыта. Чтобы восстановить лекционный материал, его нужно повторить, а для этого лекцию необходимо конспектировать. Конспект лекций должен быть в отдельной тетради, в которой не должно быть ничего, кроме лекции. Не надо стремиться подробно слово в слово записывать всю лекцию. Конспектируйте только самое важное в рассматриваемой теме: ключевые слова и их значения, примеры использования конструкций, что старается выделить лектор, на чем акцентирует внимание студентов.

Тетрадь для конспекта лекций также требует особого внимания. Ее нужно сделать удобной, практичной и полезной, ведь именно она является основным информативным источником при подготовке к различным отчетным занятиям, зачетам, экзаменам. Конечно, оформление лекционной тетради – это дело вкуса. Но целесообразно выделить поля, где студент мог бы изложить свои мысли, вопросы, появившиеся в

ходе лекции. Полезно одну из страниц оставлять свободной. Она потребуется потом, при самостоятельной подготовке. Сюда можно будет занести дополнительную информацию по данной теме, полученную из других источников: рисунки, схемы, примеры кода и т.д.

Лабораторное занятие. Лабораторные занятия по решению задач существенно дополняют лекции. В процессе анализа и решения задач студенты расширяют и углубляют знания, полученные из лекционного курса и учебников, приобретают умение применять общие закономерности к конкретным случаям.

Необходимо, чтобы студенты готовили теоретический материал, т.к. именно невыполнение этого требования приводит к неудаче при решении задач.

Несмотря на различие в видах задач, их решение можно проводить по следующему общему плану (некоторые пункты плана могут выпадать в некоторых конкретных случаях): а) прочесть внимательно условие задачи; б) посмотреть, все ли термины в условиях задачи известны и понятны (если что-то неясно, следует обратиться к учебнику, просмотреть решения предыдущих задач, посоветоваться с преподавателем); в) произвести анализ задачи, (нужно чётко понимать, в чем будет заключаться решение задачи); г) решить задачу; д) протестировать полученное решение на данных из примеров к задаче, а также на дополнительных данных.

Если задача не решена или «не решается», то необходимо еще раз вернуться к пунктам а) и б). Сколько раз нужно возвращаться к этим пунктам? Практика показывает, что не более десяти раз. Если и после этого задача «не решается», то можно попытаться найти решение этой или похожей задачи в различных источниках.

Домашнее задание. При выполнении домашнего задания необходимо просмотреть текст лекции, разобраться с новыми определениями, посмотреть задания, которые были выполнены на лабораторной работе и применить полученные знания для выполнения домашней работы.

Тест. В первую очередь постарайтесь узнать чего ждать от теста, какие примерно там будут задания. Если вам доступны образцы теста (как, например, при сдаче ЕГЭ), необходимо этим воспользоваться и ежедневно тренироваться.

Не оставляйте все на самый последний момент. Если будете постоянно готовиться к тесту, вы наверняка улучшите свои знания. Для этого составьте план на каждый день, чтобы правильно распределять свое время.

Делайте небольшие перерывы во время учебы. В промежутках можно дать себе небольшую физическую нагрузку. Мозг лучше всего работает, когда умственный труд сменяется физическим. Прогуляйтесь, побегайте, поиграйте в баскетбол, попинайте мяч – помимо стимуляции умственной деятельности, это снимет стресс.

Отдых и контроль над волнением — одни из главных составляющих успеха при подготовке к тесту. Часто ошибки совершаются только из-за стресса, который мешает сконцентрироваться и собраться. Чтобы быть отдохнувшим и расслабленным, соблюдайте составленный режим и старайтесь высыпаться.

Экзамен. На экзамене оцениваются: 1) понимание и степень усвоения теории; 2) методическая подготовка; 3) знание фактического материала; 4) знакомство с обязательной литературой; 5) умение приложить теорию к практике, решать практические задачи и т. д.; 6) логика, структура и стиль ответа, умение защищать выдвигаемые положения. Но значение экзаменов не ограничивается проверкой знаний. Являясь естественным завершением работы студента, они способствуют обобщению и закреплению знаний и умений, приведению их в строгую систему, а также устранению возникших в процессе занятий пробелов.

Студенты готовятся к экзаменам по-разному. Одни из них прорабатывают лишь некоторые вопросы, выбранные наугад, другие стремятся запомнить весь материал подряд, не вникая глубоко в его суть. Работа при этом концентрируется на одном стремлении – сдать экзамен. Недостатки такой системы очевидны. Очевидно также, что подготовка не должна ограничиваться чтением лекционных записей. Первоначальные необработанные конспекты студента содержат факты, определения, выводы, сделанные преподавателем, но в них, как правило, слабо просматривается связующая идея курса, так как студент, записывая каждую лекцию в отдельности, редко способен сразу и достаточно точно уловить общую направляющую мысль. Поэтому конспект требует дополнительной обработки на основе использования учебников и рекомендованной литературы.

Существенные недостатки имеет и такой способ подготовки к экзаменам, как беглый просмотр всего материала. Он эффективен только на некоторых этапах планирования и закрепляющего повторения. Более надежный и целесообразный путь – это тщательная систематизация материала при вдумчивом повторении, установлении внутриспредметных связей, увязке различных тем и разделов, закреплении путем решения задач.

Перед экзаменом назначается консультация. Цель ее – дать ответы на вопросы, возникшие в ходе самостоятельной подготовки. Хотелось бы обратить особое внимание на важность предэкзаменационных консультаций. Здесь студент имеет полную возможность получить ответ на все неясные ему вопросы. А для этого он должен проработать до консультации весь курс. Кроме того, преподаватель будет отвечать на вопросы других студентов, что будет для вас повторением и закреплением знаний. И еще очень важное обстоятельство: лектор на консультации, как правило, обращает внимание на те разделы, по которым на предыдущих экзаменах ответы были неудовлетворительными, а также фиксирует внимание на наиболее трудных разделах курса. Некоторые студенты не приходят на консультации либо потому, что считают, что у них нет вопросов к лектору, либо полагают, что у них и так мало времени и лучше самому почитать материал по конспекту или в учебнике. Это глубокое заблуждение. Никакая другая работа не сможет принести столь значительного эффекта накануне экзамена, как консультация преподавателя.

Подготовку к экзамену следует начинать с первого дня изучения дисциплины. Как правило, на лекциях подчеркиваются наиболее важные и трудные вопросы или разделы курса, требующие внимательного изучения и обдумывания. Нужно эти вопросы выделить и обязательно постараться разобраться в них, не дожидаясь экзамена, проработать их, готовясь к лабораторным занятиям, попробовать самостоятельно решить несколько типовых задач. И если, несмотря на это, часть материала осталась неувоенной, ни в коем случае нельзя успокаиваться, надеясь на то, что это не попадет на экзамене. Факты говорят об обратном: если те или другие вопросы курса не вошли в экзаменационный билет, преподаватель может их задать (и часто задает) в виде дополнительных вопросов. Точно такое же отношение должно быть выработано к вопросам и задачам, перечисленным в экзаменационной программе, выдаваемой студентам еще до экзамена. Обычно эти же вопросы и аналогичные задачи содержатся в экзаменационных билетах. Не следует оставлять без внимания ни одного раздела курса; если не удалось в чем-то разобраться самому, нужно обратиться к товарищам; если и это не помогло выяснить какой-либо вопрос до конца, нужно обязательно задать этот вопрос преподавателю на предэкзаменационной консультации. Чрезвычайно важно приучить себя к умению самостоятельно мыслить, учиться думать, понимать суть дела. Очень полезно после проработки каждого раздела восстановить в памяти содержание изученного материала, кратко записав это на листе бумаги. Если этого не сделать, то большая часть материала останется не понятой, а лишь формально заученной, и при первом же вопросе экзаменатора студент убедится в том, насколько поверхностно он усвоил материал.

4.5. ПРИМЕРНАЯ ТЕМАТИКА КУРСОВЫХ РАБОТ (ПРОЕКТОВ)

Не предусмотрено.

5. УЧЕБНО-МЕТОДИЧЕСКОЕ И ИНФОРМАЦИОННОЕ ОБЕСПЕЧЕНИЕ ДИСЦИПЛИНЫ (МОДУЛЯ)

а) перечень литературы

1. Грас, Джоэл. Data Science. Наука о данных с нуля [Текст] : учеб. пособие / Д. Грас. - СПб. : БХВ - Петербург, 2019. - 336 с. ; 23 см. - ISBN 978-5-9775-3758-2 (30 экз.).
2. Цильковский, И. А. Методы анализа знаний и данных [Электронный ресурс] : уч.-методич. пособие / И. А. Цильковский. — Новосибирск : Изд-во НГТУ, 2010. — 68 с. — Режим доступа: ЭБС «Руконт».
3. Чубукова, Ирина Александровна. Data Mining [Текст] : учеб. пособие / И. А. Чубукова. - М. : Интернет-Ун-т информ. технологий : Бином. Лаб. знаний, 2006. - 382 с. (9 экз.)
4. Сапрыкин, О. Н. Интеллектуальный анализ данных : учебное пособие / О. Н. Сапрыкин. — Самара : Самарский университет, 2020. — 80 с. — ISBN 978-5-7883-1563-8. — Текст : электронный // Лань : электронно-библиотечная система. — URL: <https://e.lanbook.com/book/188906> — Режим доступа: для авториз. пользователей.

б) периодические издания

б) периодические издания

в) список авторских методических разработок:

Материалы курса, опубликованные в ИОС «DOMIC».

г) базы данных, информационно-справочные и поисковые системы

Google Colaboratory <https://colab.research.google.com/>

6. МАТЕРИАЛЬНО-ТЕХНИЧЕСКОЕ ОБЕСПЕЧЕНИЕ ДИСЦИПЛИНЫ

6.1. УЧЕБНО-ЛАБОРАТОРНОЕ ОБОРУДОВАНИЕ:

Для проведения лекционных занятий необходима аудитория с презентационным оборудованием, для проведения лабораторных занятий необходима аудитория на 25-30 рабочих мест (в зависимости от численности учебной группы), оборудованная доской, презентационной техникой, компьютерами.

6.2. ПРОГРАММНОЕ ОБЕСПЕЧЕНИЕ:

1. Браузер.
2. Python 3.
3. Anaconda.

6.3. ТЕХНИЧЕСКИЕ И ЭЛЕКТРОННЫЕ СРЕДСТВА:

ИОС DOMIC, презентационное оборудование, персональный компьютер с возможностью демонстрации презентаций в формате pdf.

7. ОБРАЗОВАТЕЛЬНЫЕ ТЕХНОЛОГИИ

При реализации данного курса используются следующие образовательные технологии: технологии традиционного обучения, игровые технологии, технологии проблемного обучения, технологии обучения в сотрудничестве, технологии контекстного обучения, интерактивные технологии, технологии дистанционного обучения, активные педагогические технологии.

8. ОЦЕНОЧНЫЕ МАТЕРИАЛЫ ДЛЯ ТЕКУЩЕГО КОНТРОЛЯ И ПРОМЕЖУТОЧНОЙ АТТЕСТАЦИИ

8.1. ОЦЕНОЧНЫЕ МАТЕРИАЛЫ ДЛЯ ВХОДНОГО КОНТРОЛЯ

Не предусмотрено.

8.2. ОЦЕНОЧНЫЕ МАТЕРИАЛЫ ТЕКУЩЕГО КОНТРОЛЯ

Полный фонд оценочных средств находится в ИОС «Домик» по адресу domic.isu.ru.

8.3. ОЦЕНОЧНЫЕ МАТЕРИАЛЫ ПРОМЕЖУТОЧНОГО КОНТРОЛЯ

Контроль за изучением дисциплины осуществляется следующим способом: в течение семестра студенты выполняют лабораторные работы, заключающие в написании программ на языке Python с последующим анализом результатов, и отвечают на устные вопросы по некоторым темам. В течение семестра можно набрать максимум 80 баллов.

Экзамен проходит в устной форме и принимается по билетам. Каждый билет содержит один теоретический вопрос.

Критерии оценки:

1. Промежуточный (текущий) контроль призван определить насколько глубоко студент изучил материал дисциплины согласно рабочей программы дисциплины, знает и правильно понимает содержание предмета.

2. Промежуточный (текущий) контроль выявить степень умения студента правильно применять полученные знания при анализе процессов и явления согласно предметной области в своей практической деятельности.

3. Промежуточный (текущий) контроль демонстрирует способность студента к самостоятельной работе в предметной области дисциплины по приобретению представлений, знаний и навыков.

Оценка «отлично» выставляется, если ответ студента в полной мере соответствует названным выше принципам. Твердо знает предмет, рекомендованную обязательную литературу, грамотно, аргументировано излагает основные вопросы и проблемы дисциплины, полно отвечает на дополнительные вопросы, умеет применять теоретические знания при анализе явлений, ситуаций и проблем в предметной области дисциплины.

Оценка «хорошо» предполагает наличие прочных знаний в объеме рабочей программы дисциплины, знание минимума рекомендованной литературы, умение анализировать явления, ситуации и проблемы в предметной области дисциплины.

Оценка «удовлетворительно» ставится, если студент показал в целом верные знания по всем вопросам билета, обнаружил понимание основного содержания экзаменационных вопросов, умение в целом применять теоретические знания при анализе явлений, ситуаций и проблем в предметной области дисциплины.

Оценка «неудовлетворительно» ставится в случае, если ответы студента не удовлетворяют требованиям к оценкам «удовлетворительно» или студент не выполнил требования квалификации.

При ответе на билеты установлено следующее соответствие.

При текущем контроле оценке «отлично» соответствует 30 баллов, оценке «хорошо» соответствует 25 баллов, оценке «удовлетворительно» соответствует 20 баллов.

При промежуточном контроле оценке «отлично» соответствует 80 баллов, оценке «хорошо» соответствует 70 баллов, оценке «удовлетворительно» соответствует 60 баллов.

Дополнительные баллы ставятся за ответы на сложные вопросы, активную работу в семестре, участие в олимпиадах по математике и информационным технологиям (до 20 баллов).

Общая оценка за промежуточный контроль выставляется в соответствии с балльно-рейтинговой системой (от 60 до 69 баллов – удовлетворительно, от 70 до 79 баллов – хорошо, от 80 баллов – отлично).

Примерные вопросы к экзамену:

1. Понятие анализа данных, его основные задачи. Роль анализа данных в современном бизнесе.

2. Специфика работы с информацией. Свойства информации. Информация и данные. Виды и методы анализа данных.
3. Понятие интеллектуального анализа данных (Data Mining). Задачи и стадии DM.
4. Модели Data Mining. Сферы применения технологии Data Mining.
5. Методы Data Mining: классификация. Основные алгоритмы классификационного анализа.
6. Методы Data Mining: регрессия. Основные алгоритмы регрессионного анализа.
7. Методы Data Mining: поиск ассоциативных правил. Основные алгоритмы ассоциативного анализа.
8. Методы Data Mining: кластеризация. Основные алгоритмы кластерного анализа.
9. Линейная регрессия.
10. Логистическая регрессия.
11. Переобучение и регуляризация.
12. Деревья решений и алгоритмы обучения.
13. Нейронные сети. Метод обратного распространения ошибки.
14. Сверточные нейронные сети.
15. Рекуррентные нейронные сети.
16. Рекомендательные системы. Совместная фильтрация.
17. Рекомендательные системы нового поколения.
18. Анализ текстовой информации (Text Mining), его задачи.

Разработчики:


(подпись)

доцент
(занимаемая должность)

Казимиров А.С.
(Ф.И.О.)

Программа составлена в соответствии с требованиями ФГОС ВО по направлению подготовки 02.03.02 «Фундаментальная информатика и информационные технологии» (уровень бакалавриата), утвержденный приказом Министерства образования и науки Российской Федерации от 23 августа 2017 г. N 808, зарегистрированный в Минюсте России «14» сентября 2017 г. № 48185 с изменениями и дополнениями с изменениями и дополнениями от: 26 ноября 2020 г., 8 февраля 2021 г.

Программа рассмотрена на заседании кафедры алгебраических и информационных систем ИМИТ ИГУ «29» апреля 2021 г.

Протокол № 11 Зав. кафедрой _____ Пантелеев В.И.

Настоящая программа не может быть воспроизведена ни в какой форме без предварительного письменного разрешения кафедры-разработчика программы.