

Министерство науки и высшего образования Российской Федерации
Федеральное государственное бюджетное образовательное учреждение
высшего образования
«Сибирский государственный индустриальный университет»
Кафедра прикладных информационных технологий и программирования

УТВЕРЖДАЮ
Проректор по учебной работе
_____ М.В. Темлянец
подпись
« ____ » _____ 20__ г.

РАБОЧАЯ ПРОГРАММА ДИСЦИПЛИНЫ

Математические и инструментальные методы анализа данных
09.04.03 «Прикладная информатика»
(направленность (профиль): «Системы корпоративного управления»)

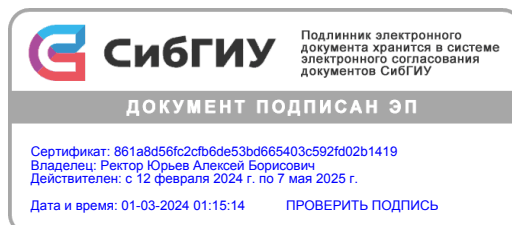
Квалификация выпускника
Магистр

Форма обучения
Очная форма

Срок обучения: 2 года

Год начала подготовки 2022

Новокузнецк
2022



1 Цели и задачи освоения учебной дисциплины

Целями учебной дисциплины являются:

- формирование целостного представления о методах и алгоритмах сбора, хранения и анализа больших объемов данных.

Задачами учебной дисциплины являются:

- знакомство с основными понятиями в области обработки «больших данных»;
- знакомство с методами и инструментами получения, хранения и анализа данных;
- получение практических навыков обработки и анализа данных с использованием специализированного программного обеспечения.

2 Место учебной дисциплины в структуре ООП по направлению подготовки (специальности)

Учебная дисциплина относится к учебным дисциплинам части, формируемой участниками образовательных отношений **Блока 1 «Дисциплины (модули)»** ООП по направлению подготовки (специальности) 09.04.03 «Прикладная информатика».

Учебная дисциплина опирается на базовые знания и компетенции, полученные в процессе получения предыдущего образования.

Учебная дисциплина дополняет знания, умения и навыки, получаемые по одновременно изучаемым и последующим дисциплинам:

- Python для анализа данных;
- Анализ и оптимизация бизнес-процессов;
- Машинное обучение для задач управления;
- Разработка и реализация проектов 1;
- Разработка и реализация проектов 2;
- Научно-исследовательская работа;
- Выполнение и защита выпускной квалификационной работы.

3 Планируемые результаты обучения по учебной дисциплине

Процесс изучения учебной дисциплины направлен на формирование следующих компетенций:

– Профессиональные компетенции

Наименование категории (группы) ПК	Код и наименование ПК	Код и наименование индикатора достижения ПК	Планируемые результаты обучения
	ПК-2: Способен разрабатывать и применять	ПК-2.2 Собирает данные из различных	– знать: методы сбора и интеграции данных из

	алгоритмы интеллектуального анализа больших объемов данных для управления технологическими системами	испечников и осуществляет их подготовку для анализа	различных источников, их подготовки, первичной обработки и визуализации . – уметь: осуществлять сбор, первичную обработку и визуализацию данных при решении профессиональных задач. – владеть: навыками сбора, подготовки, первичной обработки и визуализации данных с использованием специализированных аналитических платформ.
		ПК-2.3 Выбирает и использует методы и инструментальные средства для проведения анализа больших объемов данных, строит модели на основе данных	– знать: теоретические и прикладные основы анализа больших объемов данных и построения моделей на основе данных. – уметь: использовать типовые алгоритмы анализа и обработки больших объемов данных. – владеть: навыками построения моделей на основе данных в специализированных аналитических платформах.

4 Объем и содержание учебной дисциплины

Учебные занятия по учебной дисциплине проводятся в форме контактной работы и в форме самостоятельной работы обучающихся.

Контактная работа обучающихся с педагогическим работником включает в себя занятия лекционного типа (лекции), занятия

семинарского типа (семинары, практические занятия, практикумы), промежуточную аттестацию обучающихся и иную контактную работу, предусматривающую групповую или индивидуальную работу обучающихся с педагогическим работником. Контактная работа обучающихся с педагогическим работником может быть аудиторной, внеаудиторной, а также проводиться в электронной информационно-образовательной среде.

Объем учебной дисциплины

Семестр / курс		ИТОГО	1 семестр
Форма промежуточной аттестации			экзамен, зачет с оценкой по КР
Трудоёмкость	академ. час.	216	216
	зачетных единиц	6	6
Лекции, академ. час.		16	16
в форме практической подготовки		0	0
Лабораторные работы, академ. час.		0	0
в форме практической подготовки		0	0
Практические занятия, академ. час.		32	32
в форме практической подготовки		0	0
Курсовая работа, академ. час.		36	36
в форме практической подготовки		0	0
Консультации, академ. час.		0	0
в форме практической подготовки		0	0
Самостоятельная работа, академ. час.		96	96
в форме практической подготовки		0	0
Контроль, академ. час.		36	36
в форме практической подготовки		0	0

Содержание учебной дисциплины

Раздел 1 Введение в науку о данных;

Тема 1.1 Основы анализа данных (Принципы анализа данных, классический и информационный подходы. Структурированные данные. Источники данных. Сбор данных. Требования к данным. Подготовка данных к анализу);

Тема 1.2 Концепция Big Date (Большие данные) (Предпосылки формирования тренда больших данных; определение термина «большие данные». Сферы использования больших данных. Модель MapReduce; пути развития передовых аналитических инструментов; подходы к построению моделей для эффективного использования);

Тема 1.3 Консолидация, визуализация и подготовка данных (Организация хранилища данных. Хранилища данных и анализ. Визуализаторы общего назначения. Многомерная модель данных. OLAP-анализ. Качество данных. Сбор и интеграция данных. Виды источников данных. Методы интеграции данных. Преобразование и обогащение данных. Методы трансформации данных.);

Раздел 2 Интеллектуальный анализ данных и машинное обучение;

Тема 2.1 Методы и инструменты интеллектуального анализа данных (Задачи интеллектуального анализа данных (ИАД) и их классификация. Модели и методы ИАД. Цели, задачи и методы машинного обучения.);

Тема 2.2 Кластерный анализ (Введение в кластеризацию. Алгоритмы k-means, g-means. Иерархический кластерный анализ. Сети и карты Кохонена);

Тема 2.3 Классификация и регрессия (Метод опорных векторов. Метод k-ближайших соседей. Байесовский классификатор. Основы логистической регрессии. Правила классификации. Деревья решений, случайный лес. Простая линейная регрессия. Искусственные нейронные сети).

5 Перечень тем лекций

№ раздела / темы дисциплины	Темы лекций	Трудоемкость, <i>академ. час</i>	
		всего	в форме практической подготовки
Раздел 1.	Введение в науку о данных		
Тема 1.1.	Основы анализа данных	2	
Тема 1.2.	Концепция Big Date (Большие данные)	2	
Тема 1.3.	Консолидация, визуализация и подготовка данных	2	
Раздел 2.	Интеллектуальный анализ данных и машинное обучение		
Тема 2.1.	Методы и инструменты интеллектуального анализа данных	2	
Тема 2.2.	Кластерный анализ	4	
Тема 2.3.	Классификация и регрессия	4	
Итого:		16	0

6 Перечень тем практических занятий (семинаров)

№ раздела / темы дисциплины	Темы практических занятий (семинаров)	Трудоемкость, <i>академ. час</i>	
		всего	в форме практической подготовки
Тема 1.3.	Консолидация и визуализация данных	6	
Тема 1.3.	Подготовка данных	4	

Тема 2.2.	Кластерный анализ	8	
Тема 2.3.	Классификация и регрессия	14	
Итого:		32	0

7 Перечень тем лабораторных работ

№ раздела / темы дисциплины	Темы лабораторных работ	Трудоемкость, <i>академ. час</i>	
		всего	в форме практической подготовки
	<i>Отсутствуют</i>		
Итого:		0	0

8 Перечень тем курсовых работ (проектов)

№ раздела / темы дисциплины	Темы курсовых работ (проектов)	Трудоемкость, <i>академ. час</i>	
		всего	в форме практической подготовки
Раздел 1; Раздел 2.	<p>1) Построение описательной модели и анализ ассоциаций по данным о клиентах организации предоставляющей услуги (мобильный оператор, интернет магазин, розничная торговая сеть, банк, государственные органы и т.п.).</p> <p>2) Построение описательной модели и сегментация клиентов организации предоставляющей услуги (мобильный оператор, интернет магазин, розничная торговая сеть, банк, страховая компания и т.п.).</p> <p>3) Построение модели оценки качества конечного продукта на основе физико-химических свойств, протекания технологического процесса и т.п.</p> <p>4) Построение предсказательной модели поведения клиента организации</p>	36	

	<p>предоставляющей услуги (банк, издательство, мобильный оператор и т.д.).</p> <p>5) Построение скоринговой модели для оценки риска невозврата кредита заемщиком.</p> <p>6) Построение диагностической модели по данным предметной области (медицина, техническая диагностика и т.п.).</p> <p>7) Построение рекомендательной системы для посетителей интернет ресурса.</p>		
Итого:		36	0

9 Виды самостоятельной работы

№ раздела / темы дисциплины	Виды самостоятельной работы	Трудоемкость, <i>академ. час</i>	
		всего	в форме практической подготовки
Раздел 1.	1. Изучение лекционного материала; 2. Подготовка к практическому занятию; 3. Прохождение тестирования.	24	
Раздел 2.	1. Изучение лекционного материала; 2. Подготовка к практическому занятию; 3. Прохождение тестирования.	72	
<i>Курсовая работа</i>	<i>Выполнение курсовой работы</i>	36	0
<i>Контроль</i>	<i>Подготовка к экзамену</i>	36	
Итого:		168	0

10 Учебно-методическое и информационное обеспечение учебной дисциплины

а) литература:

1 Станкевич, Л. А. Интеллектуальные системы и технологии : учебник и практикум для вузов / Л. А. Станкевич. – Москва :

Издательство Юрайт, 2022. – 397 с. – ISBN 978-5-534-02126-4. – URL: <https://urait.ru/bcode/489694> (дата обращения: 23.03.2022);

2 Гасанов, Э. Э. Интеллектуальные системы. Теория хранения и поиска информации : учебник для вузов / Э.Э. Гасанов, В.Б. Кудрявцев. – 2-е изд., испр. и доп. – Москва : Юрайт, 2022. – 271 с. – ISBN 978-5-534-08684-3. – URL: <https://urait.ru/bcode/491100> (дата обращения: 23.03.2022);

3 Парфенов, Ю. П. Постреляционные хранилища данных : учебное пособие для вузов / Ю.П. Парфенов, Н.В. Папуловская. – Москва : Юрайт, 2022. – 121 с. – ISBN 978-5-534-09837-2. – URL: <https://urait.ru/bcode/492609> (дата обращения: 23.03.2022);

4 Бессмертный, И. А. Системы искусственного интеллекта : учебное пособие для вузов. – 2-е изд., испр. и доп. – Москва : Юрайт, 2022. – 157 с. – ISBN 978-5-534-07467-3. – URL: <https://urait.ru/bcode/490657> (дата обращения: 23.03.2022);

5 Загорулько, Ю. А. Искусственный интеллект. Инженерия знаний : учебное пособие для вузов / Ю.А. Загорулько, Г.Б. Загорулько. – Москва : Юрайт, 2022. – 93 с. – ISBN 978-5-534-07198-6. – URL: <https://urait.ru/bcode/494205> (дата обращения: 23.03.2022);

6 Назаров, Д. М. Интеллектуальные системы: основы теории нечетких множеств : учебное пособие для вузов / Д.М. Назаров, Л.К. Конышева. – 3-е изд., испр. и доп. – Москва : Юрайт, 2022. – 186 с. – ISBN 978-5-534-07496-3. – URL: <https://urait.ru/bcode/492333> (дата обращения: 23.03.2022);

7 Чубукова, И. А. Data Mining : учебное пособие / И. А. Чубакова. – Москва : ИНТУИТ, 2016. – ISBN 978-5-94774-819-2. – URL: <http://www.studentlibrary.ru/book/ISBN9785947748192.html> (дата обращения: 23.03.2022).

б) ресурсы информационно-телекоммуникационной сети «Интернет»:

1 Консультант студента. Электронная библиотека технического ВУЗа : электронно-библиотечная система / ООО «Политехресурс». – Москва, [200 –]. – URL: <http://www.studentlibrary.ru>. – Режим доступа: для авторизир. пользователей;

2 ЛАНЬ : электронно-библиотечная система : [коллекция «Инженерно-технические науки»] / ООО «Издательство Лань». – Санкт-Петербург, [200 –]. – URL: <http://e.lanbook.com>. – Режим доступа: для авторизир. пользователей;

3 НАУЧНАЯ ЭЛЕКТРОННАЯ БИБЛИОТЕКА eLIBRARY.RU : база данных / ООО «НЭБ». – Москва, [200 –]. – URL: <http://elibrary.ru>. – Режим доступа: по подписке;

4 Образовательная платформа ЮРАЙТ / ООО «Электронное издательство Юрайт». – Москва, [200 –]. – URL: <https://urait.ru>. – Режим доступа: для авторизир. пользователей;

5 Университетская библиотека онлайн : электронно-библиотечная система / ООО «Директ-Медиа». – Москва, [200 –]. – URL: <http://www.biblioclub.ru>. – Режим доступа: для авторизир. пользователей;

6 Электронная библиотека // Научно-техническая библиотека СибГИУ : сайт. – Новокузнецк, [200 –]. – URL: <https://library.sibsiu.ru/LibrELibraryFullText.asp>. – Режим доступа: для авторизир. пользователей;

7 Электронно-библиотечная система IPRbooks : [коллекции: «Дошкольная педагогика. Педагогика школы», «Педагогика. Образование»] / ООО Компания «Ай Пи Ар Медиа». – Москва, [200 –]. – URL: <http://www.iprbookshop.ru>. – Режим доступа: для авторизир. пользователей;

8 Электронный каталог : сайт / Научно-техническая библиотека СибГИУ. – Новокузнецк, [199 –]. – URL: <http://libr.sibsiu.ru>.

в) лицензионное и свободно распространяемое программное обеспечение:

- Adobe Acrobat Reader;
- Deductor Academic;
- Loginom Academic;
- Microsoft Office 2010;
- Microsoft Visual Studio Community;
- Microsoft Windows 7;
- Orange.

г) базы данных и информационно-справочные системы:

1 КонсультантПлюс : справочно-правовая система / ООО «Информационный центр АНВИК». – Новокузнецк, [199 –]. – Режим доступа: компьютерная сеть библиотеки Сиб. гос. индустр. ун-та.;

2 Система ГАРАНТ : электронный периодический справочник / ООО «Правовой центр «Гарант». – Кемерово, [200 –]. – Режим доступа: компьютерная сеть Сиб. гос. индустр. ун-та.;

3 Техэксперт : информационно-справочная система / ООО «Группа компаний «Кодекс». – Кемерово, [200 –]. – Режим доступа: компьютерная сеть Сиб. гос. индустр. ун-та.;

4 Электронный реферативный журнал (ЭлРЖ) : база данных / ВИНТИ РАН. – Москва, [200 –]. – Режим доступа: компьютерная сеть библиотеки Сиб. гос. индустр. ун-та.

11 Материально-техническое обеспечение учебной дисциплины

Материально-техническое обеспечение учебной дисциплины включает учебные аудитории, оснащенные оборудованием, компьютерной техникой, и техническими средствами обучения, в том числе:

- учебную аудиторию для проведения занятий лекционного типа,

оборудованную учебной доской, экраном и мультимедийным проектором;

- учебную аудиторию для проведения занятий семинарского типа (практических занятий);

- учебную аудиторию (помещения) для групповых и индивидуальных консультаций, текущего контроля и промежуточной аттестации;

- помещения для самостоятельной работы, оснащенные компьютерной техникой с возможностью подключения к сети «Интернет» и обеспечением доступа в электронную информационно-образовательную среду, научно-техническую библиотеку СибГИУ;

- учебную аудиторию (помещения) для групповых и индивидуальных консультаций, текущего контроля и промежуточной аттестации

Рабочая программа дисциплины составлена в соответствии с требованиями ФГОС ВО по направлению подготовки (специальности) 09.04.03 «Прикладная информатика».

Составитель(и):

доцент Кораблина Татьяна Валентиновна (кафедра прикладных информационных технологий и программирования).

Рабочая программа дисциплины рассмотрена и утверждена на заседании кафедры.

Приложение А

Аннотация

рабочей программы дисциплины «Математические и инструментальные методы анализа данных»

по направлению подготовки (специальности)

09.04.03 «Прикладная информатика»

(направленность (профиль): «Системы корпоративного управления»)

форма обучения – Очная форма

1 Цели и задачи освоения учебной дисциплины

Целями учебной дисциплины являются:

- формирование целостного представления о методах и алгоритмах сбора, хранения и анализа больших объемов данных.

Задачами учебной дисциплины являются:

- знакомство с основными понятиями в области обработки «больших данных»;
- знакомство с методами и инструментами получения, хранения и анализа данных;
- получение практических навыков обработки и анализа данных с использованием специализированного программного обеспечения.

2 Место учебной дисциплины в структуре ООП по направлению подготовки (специальности)

Учебная дисциплина относится к учебным дисциплинам части, формируемой участниками образовательных отношений **Блока 1 «Дисциплины (модули)»** ООП по направлению подготовки (специальности) 09.04.03 «Прикладная информатика».

Учебная дисциплина опирается на базовые знания и компетенции, полученные в процессе получения предыдущего образования.

Учебная дисциплина дополняет знания, умения и навыки, получаемые по одновременно изучаемым и последующим дисциплинам:

- Python для анализа данных;
- Анализ и оптимизация бизнес-процессов;
- Машинное обучение для задач управления;
- Разработка и реализация проектов 1;
- Разработка и реализация проектов 2;
- Научно-исследовательская работа;
- Выполнение и защита выпускной квалификационной работы.

3 Планируемые результаты обучения по учебной дисциплине

Процесс изучения учебной дисциплины направлен на формирование следующих компетенций:

– Профессиональные компетенции

Наименование категории (группы) ПК	Код и наименование ПК	Код и наименование индикатора достижения ПК	Планируемые результаты обучения
	ПК-2: Способен разрабатывать и применять алгоритмы интеллектуального анализа больших объемов данных для управления технологическими системами	ПК-2.2 Собирает данные из различных источников и осуществляет их подготовку для анализа	<p>– знать: методы сбора и интеграции данных из различных источников, их подготовки, первичной обработки и визуализации .</p> <p>– уметь: осуществлять сбор, первичную обработку и визуализацию данных при решении профессиональных задач.</p> <p>– владеть: навыками сбора, подготовки, первичной обработки и визуализации данных с использованием специализированных аналитических платформ.</p>
		ПК-2.3 Выбирает и использует методы и инструментальные средства для проведения анализа больших объемов данных, строит модели на основе данных	<p>– знать: теоретические и прикладные основы анализа больших объемов данных и построения моделей на основе данных.</p> <p>– уметь: использовать типовые алгоритмы анализа и обработки больших объемов данных.</p> <p>– владеть: навыками</p>

			построения моделей на основе данных в специализированных аналитических платформах.
--	--	--	--

4 Объем учебной дисциплины

Семестр / курс		ИТОГО	1 семестр
Форма промежуточной аттестации			экзамен, зачет с оценкой по КР
Трудоёмкость	<i>академ. час.</i>	216	216
	<i>зачетных единиц</i>	6	6
Лекции, <i>академ. час.</i>		16	16
в форме практической подготовки		0	0
Лабораторные работы, <i>академ. час.</i>		0	0
в форме практической подготовки		0	0
Практические занятия, <i>академ. час.</i>		32	32
в форме практической подготовки		0	0
Курсовая работа, <i>академ. час.</i>		36	36
в форме практической подготовки		0	0
Консультации, <i>академ. час.</i>		0	0
в форме практической подготовки		0	0
Самостоятельная работа, <i>академ. час.</i>		96	96
в форме практической подготовки		0	0
Контроль, <i>академ. час.</i>		36	36
в форме практической подготовки		0	0

5 Краткое содержание учебной дисциплины

В структуре учебной дисциплины выделяются следующие основные разделы (темы):

Раздел 1 Введение в науку о данных;

Тема 1.1 Основы анализа данных (Принципы анализа данных, классический и информационный подходы. Структурированные данные. Источники данных. Сбор данных. Требования к данным. Подготовка данных к анализу);

Тема 1.2 Концепция Big Date (Большие данные) (Предпосылки формирования тренда больших данных; определение термина «большие данные». Сферы использования больших данных. Модель MapReduce; пути развития передовых аналитических инструментов; подходы к построению моделей для эффективного использования);

Тема 1.3 Консолидация, визуализация и подготовка данных (Организация хранилища данных. Хранилища данных и анализ. Визуализаторы общего назначения. Многомерная модель данных. OLAP-анализ. Качество данных. Сбор и интеграция данных. Виды источников данных. Методы интеграции данных. Преобразование и обогащение данных. Методы трансформации данных.);

Раздел 2 Интеллектуальный анализ данных и машинное обучение;

Тема 2.1 Методы и инструменты интеллектуального анализа данных (Задачи интеллектуального анализа данных (ИАД) и их классификация. Модели и методы ИАД. Цели, задачи и методы машинного обучения.);

Тема 2.2 Кластерный анализ (Введение в кластеризацию. Алгоритмы k-means, g-means. Иерархический кластерный анализ. Сети и карты Кохонена);

Тема 2.3 Классификация и регрессия (Метод опорных векторов. Метод k-ближайших соседей. Байесовский классификатор. Основы логистической регрессии. Правила классификации. Деревья решений, случайный лес. Простая линейная регрессия. Искусственные нейронные сети).

6 Составитель(и):

доцент Кораблина Татьяна Валентиновна (кафедра прикладных информационных технологий и программирования).