

МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ РОССИЙСКОЙ ФЕДЕРАЦИИ
Федеральное государственное автономное образовательное учреждение высшего образования
«НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ
ТОМСКИЙ ПОЛИТЕХНИЧЕСКИЙ УНИВЕРСИТЕТ»

УТВЕРЖДАЮ

Директор ИШИТР

 Д.М. Сонькин

« 01 » 09 2020 г

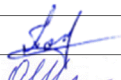
РАБОЧАЯ ПРОГРАММА ДИСЦИПЛИНЫ
ПРИЕМ 2020 г.
ФОРМА ОБУЧЕНИЯ очная

DATA MINING		
Направление подготовки/ специальность	09.03. 01 Информатика и вычислительная техника	
Образовательная программа (направленность (профиль))	Программирование вычислительных и телекоммуникационных систем	
Специализация	Информационно-коммуникационные технологии	
Уровень образования	высшее образование – бакалавриат	
Курс	4	8
Трудоемкость в кредитах (зачетных единицах)	3	
Виды учебной деятельности	Временной ресурс	
Контактная (аудиторная) работа, ч	Лекции	11
	Практические занятия	0
	Лабораторные занятия	33
	ВСЕГО	44
Самостоятельная работа, ч		64
ИТОГО, ч		108

Вид промежуточной аттестации

Экзамен	Обеспечивающее подразделение	ОИТ ИШИТР
----------------	------------------------------	------------------

Заведующий кафедрой -
руководитель отделения на
правах кафедры
Руководитель ООП
Преподаватель

	Шерстнев В. С.
	Погребной А.В.
	Шефер О.В.

2020 г.

1. Цели освоения дисциплины

Целями освоения дисциплины является формирование у обучающихся определенного состава компетенций для подготовки к профессиональной деятельности.

Код компетенции	Наименование компетенции	Составляющие результатов освоения (дескрипторы компетенций)	
		Код	Наименование
ОПК(У)-9	Способен осваивать методики использования программных средств для решения практических задач	ОПК(У)-9.1В1	Имеет навыки использования программных средств для решения практических задач
		ОПК(У)-9.1У1	Умеет использовать программные средства для решения практических задач
		ОПК(У)-9.1З1	Знает методики использования программных средств для решения практических задач
ПК(У)-3	Способен оптимизировать функционирование БД	ПК(У)-3.1В1	Владеет навыками выбора основных статистических показателей работы БД
		ПК(У)-3.1В2	Владеет навыками анализа полученных статистических данных, формирование выводов об эффективности работы БД
		ПК(У)-3.1У1	Умеет обрабатывать статистические данные, применять методы статистических расчётов
		ПК(У)-3.1У2	Умеет самостоятельно вести поиск информации, необходимой для выполнения профессиональных задач по управлению БД
		ПК(У)-3.1З1	Знает основные понятия статистики
		ПК(У)-3.1З2	Знает методы статистических исследований результатов испытаний

2. Место дисциплины в структуре ООП

Дисциплина относится к вариативной части Блока 1 учебного плана образовательной программы.

3. Планируемые результаты обучения по дисциплине

После успешного освоения дисциплины будут сформированы результаты обучения:

Планируемые результаты обучения по дисциплине		Компетенция
Код	Наименование	
РД-1	Владение навыками применения методик использования программных средств для решения практических задач	ОПК(У)-9
РД-2	Умение проведение мониторинга работы БД, сбора статистической информации о работе БД.	ПК(У)-3

Оценочные мероприятия текущего контроля и промежуточной аттестации представлены в календарном рейтинг-плане дисциплины.

4. Структура и содержание дисциплины

Основные виды учебной деятельности

Разделы дисциплины	Формируемый результат обучения по дисциплине	Виды учебной деятельности	Объем времени, ч.
Раздел 1. Введение в Data Mining	РД-1	Лекции	2
		Практические занятия	0
		Лабораторные занятия	6
		Самостоятельная работа	8
Раздел 2. Методы и алгоритмы Data Mining	РД-2	Лекции	9
		Практические занятия	0
		Лабораторные занятия	27
		Самостоятельная работа	56

Содержание разделов дисциплины:

Раздел 1. Введение в Data Mining

Деятельность любого предприятия (коммерческого, производственного, медицинского, научного и т.д. теперь сопровождается регистрацией и записью всех подробностей его деятельности. Без продуктивной переработки потока сырых данных образуют никому не нужную свалку.

Data Mining, как процесс обнаружения в сырых данных ранее неизвестных, нетривиальных, практически полезных и доступных интерпретации знаний, необходимых для принятия решений в различных сферах человеческой деятельности.

Темы лекций:

1. Понятие Data Mining. Data Mining как часть рынка информационных технологий.

Названия лабораторных работ:

1. Предварительная обработка средствами Excel и Statgraphics

Раздел 2. Методы и алгоритмы Data Mining ...

Основы анализа данных. Данные. Набор данных и их атрибутов. Шкалы. Корреляционный анализ статистический метод анализа данных, предназначенный для исследования взаимосвязи выборок. Основной показатель – выборочный коэффициент корреляции. Корреляционный анализ является составной частью любого статистического исследования. Регрессионный анализ – статистический метод анализа данных, предназначенный для исследования зависимости одной переменной от одной или нескольких независимых переменных.

Темы лекций:

2. Описательная статистика. Анализ данных в Excel.
3. Корреляционный анализ. Регрессионный анализ.
4. Методы кластерного анализа. Иерархические методы. Итеративные методы.
5. Компонентный анализ. Метод главных компонент.

6. Методы поиска ассоциативных правил.

Названия лабораторных работ

2. Проведение корреляционного и регрессионного анализа средствами Excel и STATGRAPHICS.
3. Провести кластерный анализ данных с использованием Excel и STATGRAPHICS.
4. Провести компонентный анализ с использованием Excel и STATGRAPHICS.
5. Поиск ассоциативных правил средствами Deductor.
6. Визуализация данных с использованием Excel и STATGRAPHICS.

5. Организация самостоятельной работы студентов

Самостоятельная работа студентов при изучении дисциплины предусмотрена в следующих видах и формах:

- Работа с лекционным материалом, поиск и обзор литературы и электронных источников информации по индивидуальным заданиям;
- Изучение тем, вынесенных на самостоятельную проработку;
- Подготовка к лабораторным работам;
- Подготовка к выполнению контрольных заданий.

6. Учебно-методическое и информационное обеспечение дисциплины

6.1. Учебно-методическое обеспечение

Основная литература

1. Юре, Л. Анализ больших наборов данных / Л. Юре, Р. Ананд, Д. У. Джеффри ; перевод с английского А. А. Слинкин. — Москва : ДМК Пресс, 2016. — 498 с. — ISBN 978-5-97060-190-7. — Текст : электронный // Лань : электронно-библиотечная система. — URL: <https://e.lanbook.com/book/93571> (дата обращения: 28.05.2020). — Режим доступа: для авториз. пользователей.
2. Пселтис, Э. Д. Поточная обработка данных. Конвейер реального времени / Э. Д. Пселтис ; перевод с английского А. А. Слинкин. — Москва : ДМК Пресс, 2018. — 218 с. — ISBN 978-5-97060-606-3. — Текст : электронный // Лань : электронно-библиотечная система. — URL: <https://e.lanbook.com/book/105840> (дата обращения: 28.05.2020). — Режим доступа: для авториз. пользователей.
3. Мельниченко, А. С. Математическая статистика и анализ данных : учебное пособие / А. С. Мельниченко. — Москва : МИСИС, 2018. — 45 с. — ISBN 978-5-906953-62-9. — Текст : электронный // Лань : электронно-библиотечная система. — URL: <https://e.lanbook.com/book/108035> (дата обращения: 28.05.2020). — Режим доступа: для авториз. пользователей.
4. Стефанова, И. А. Обработка данных и компьютерное моделирование : учебное пособие / И. А. Стефанова. — Санкт-Петербург : Лань, 2020. — 112 с. — ISBN 978-5-8114-4010-8. — Текст : электронный // Лань : электронно-библиотечная система. — URL: <https://e.lanbook.com/book/126939> (дата обращения: 28.05.2020). — Режим доступа: для авториз. пользователей.

Дополнительная литература

1. Ганичева, А. В. Прикладная статистика : учебное пособие / А. В. Ганичева. — Санкт-Петербург : Лань, 2017. — 172 с. — ISBN 978-5-8114-2450-4. — Текст : электронный // Лань : электронно-библиотечная система. — URL: <https://e.lanbook.com/book/91890> (дата обращения: 28.05.2020). — Режим доступа: для авториз. пользователей.
2. Степанов, П. Е. Планирование эксперимента : учебно-методическое пособие / П. Е. Степанов. — Москва : МИСИС, 2017. — 22 с. — Текст : электронный // Лань : электронно-библиотечная система. — URL: <https://e.lanbook.com/book/108113> (дата обращения: 28.05.2020). — Режим доступа: для авториз. пользователей.
3. Берестнева, О. Г. Прикладная математическая статистика : учебное пособие / О. Г. Берестнева, О. В. Марухина, Г. Е. Шевелев; Национальный исследовательский Томский политехнический университет (ТПУ), Институт кибернетики (ИК), Кафедра прикладной математики (ПМ). — Томск: Изд-во ТПУ, 2012. — URL: <http://www.lib.tpu.ru/fulltext2/m/2012/m341.pdf> (дата обращения 28.05.2020). — Режим доступа: из корпоративной сети ТПУ. — Текст : электронный.
4. Гмурман, В. Е. Руководство к решению задач по теории вероятностей и математической статистике : учебное пособие для бакалавров / В. Е. Гмурман. — 11-е изд. — Москва: Юрайт, 2013. — URL: <http://www.lib.tpu.ru/fulltext2/m/2013/FN/fn-2433.pdf> (дата обращения 28.05.2020). — Режим доступа: из корпоративной сети ТПУ. — Текст : электронный.

6.2. Информационное и программное обеспечение

Internet-ресурсы:

1. Электронно-библиотечная система «Лань» – <https://e.lanbook.com/>
2. Электронно-библиотечная система «Консультант студента» – <http://www.studentlibrary.ru/>
3. Wikipedia about Data Mining- http://en.wikipedia.org/wiki/Data_mining
4. Data Mining Tutorials - <http://www.eruditionhome.com/datamining/tut.html>
5. INTUIT.ru Учебный курс - Data Mining - <http://www.intuit.ru/departement/database/datamining/>

Профессиональные базы данных и информационно-справочные системы доступны по ссылке: <https://www.lib.tpu.ru/html/irs-and-pdb>.

Используемое лицензионное программное обеспечение (в соответствии с **Перечнем лицензионного программного обеспечения ТПУ**):

1. Google Chrome;
2. MathWorks MATLAB Full Suite R2017b;
3. Microsoft Office 2007 Standard Russian Academic;
4. Microsoft Visual Studio 2019 Community;
5. PTC Mathcad 15 Academic Floating;
6. Document Foundation LibreOffice.

7. Особые требования к материально-техническому обеспечению дисциплины

В учебном процессе используется следующее лабораторное оборудование для практических и лабораторных занятий:

№	Наименование специальных помещений	Наименование оборудования
1.	Аудитория для проведения учебных занятий всех типов, курсового проектирования, консультаций, текущего контроля и промежуточной аттестации 634034, Томская область, г. Томск, Советская улица, 84/3, 413	Комплект учебной мебели на 24 посадочных мест; Тумба стационарная - 1 шт.; Компьютер - 1 шт.; Проектор - 1 шт.; Телевизор - 1 шт.
2.	Аудитория для проведения учебных занятий всех типов, курсового проектирования, консультаций, текущего контроля и промежуточной аттестации (компьютерный класс) 634034, Томская область, г. Томск, Советская улица, 84/3, 411	Комплект учебной мебели на 12 посадочных мест Компьютер - 23 шт.; Принтер - 1 шт.

Рабочая программа составлена на основе Общей характеристики образовательной программы «Программирование вычислительных и телекоммуникационных систем» по специализации «Информационно-коммуникационные технологии» направления 09.03.01 Информатика и вычислительная техника (приема 2020 г., очная форма обучения).

Разработчик:

Должность	Подпись	ФИО
Доцент ОИТ		О. В. Шефер

Программа одобрена на заседании Отделения информационных технологий ИШИТР (протокол от «01» сентября 2020г. №19).

Заведующий кафедрой - руководитель отделения
на правах кафедры

 В.С. Шерстнев
подпись