

МИНИСТЕРСТВО ОБРАЗОВАНИЯ И НАУКИ РФ
филиал федерального государственного бюджетного образовательного учреждения
высшего образования «Мурманский арктический государственный университет»
в г. Апатиты

РАБОЧАЯ ПРОГРАММА ДИСЦИПЛИНЫ (МОДУЛЯ)

Б1.В.ДВ.1.2 Методы и технологии анализа данных

(название дисциплины (модуля) в соответствии с учебным планом)

**основной профессиональной образовательной программы
по направлению подготовки**

**09.03.02 Информационные системы и технологии
направленность (профиль): «Программно-аппаратные комплексы»**

(код и наименование направления подготовки
с указанием направленности (профиля) (наименования магистерской программы))

высшее образование – бакалавриат

уровень профессионального образования: высшее образование – бакалавриат / высшее образование –
специалитет, магистратура / высшее образование – подготовка кадров высшей квалификации

бакалавр

квалификация

очная

форма обучения

2018

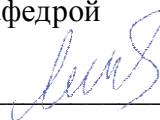
год набора

Составитель:

Тоичкин Н.А., канд. техн. наук,
доцент кафедры информатики и
вычислительной техники

Утверждено на заседании кафедры
информатики и вычислительной техники
(протокол № 9 от «30» мая 2018 г.)

Зав. кафедрой



Яковлев С.Ю.

1. ЦЕЛЬ ОСВОЕНИЯ ДИСЦИПЛИНЫ (МОДУЛЯ) - освоение обучающимися современных технологий, методов и алгоритмов для обработки и анализа данных и их применение в различных задачах практической деятельности.

В результате освоения дисциплины, обучающийся должен:

знать:

- понятие информации и данных;
- основные свойства информации и способы ее измерения, тенденции роста информации в современном мире;
- общую характеристику процессов накопления;
- понятие базы данных и различных моделей данных;
- основы работы в электронных таблицах;
- принципы разработки реляционной СУБД;
- принципы создания хранилищ и витрин данных;
- правила Кодда для хранилищ данных;
- принципы архитектуры OLAP;
- постановку различных задач Data Mining и алгоритмы их решения.

уметь:

- осуществлять математическую и информационную постановку задач по обработке информации;
- использовать алгоритмы обработки информации для различных приложений;
- использовать методы оперативной аналитической обработки информации;
- выполнять постановку различных задач Data Mining.

владеть:

- инструментальными средствами обработки информации;
- информационными технологиями анализа данных.

2. ПЕРЕЧЕНЬ ПЛАНИРУЕМЫХ РЕЗУЛЬТАТОВ ОБУЧЕНИЯ ПО ДИСЦИПЛИНЕ (МОДУЛЮ), СООТНЕСЕННЫХ С ПЛАНИРУЕМЫМИ РЕЗУЛЬТАТАМИ ОСВОЕНИЯ ОБРАЗОВАТЕЛЬНОЙ ПРОГРАММЫ

В результате освоения дисциплины формируются следующие компетенции:

- способность проводить предпроектное обследование объекта проектирования, системный анализ предметной области, их взаимосвязей (ПК-1);
- способность разрабатывать средства реализации информационных технологий (методические, информационные, математические, алгоритмические, технические и программные) (ПК-12).

3. УКАЗАНИЕ МЕСТА ДИСЦИПЛИНЫ (МОДУЛЯ) В СТРУКТУРЕ ОБРАЗОВАТЕЛЬНОЙ ПРОГРАММЫ

Данная дисциплина относится к дисциплинам по выбору вариативной части цикла Б1 основной профессиональной образовательной программы для направления подготовки 09.03.02 Информационные системы и технологии, направленность (профиль) Программно-аппаратные комплексы.

Для освоения данной дисциплины обучающиеся используют знания, умения, навыки, способы деятельности и установки, которые они получили в процессе изучения дисциплин: «Информатика», «Развитие информационного общества»,

В свою очередь, «Методы и технологии анализа данных» представляет собой методологическую базу для усвоения обучающимся содержания дисциплин: «Информационные технологии», «Инженерная и компьютерная графика», «Защита информации», «Методы и средства проектирования информационных систем и технологий», «Операционные системы», «Инструментальные средства информационных систем» а также для выполнения междисциплинарной курсовой работы.

4. ОБЪЕМ ДИСЦИПЛИНЫ (МОДУЛЯ) В ЗАЧЕТНЫХ ЕДИНИЦАХ С УКАЗАНИЕМ КОЛИЧЕСТВА АКАДЕМИЧЕСКИХ ИЛИ АСТРОНОМИЧЕСКИХ ЧАСОВ, ВЫДЕЛЕННЫХ НА КОНТАКТНУЮ РАБОТУ ОБУЧАЮЩИХСЯ С ПРЕПОДАВАТЕЛЕМ (ПО ВИДАМ УЧЕБНЫХ ЗАНЯТИЙ) И НА САМОСТОЯТЕЛЬНУЮ РАБОТУ ОБУЧАЮЩИХСЯ

Общая трудоемкость дисциплины составляет 7 зачетных единиц или 252 часа.
(из расчета 1 ЗЕТ= 36 часов).

Курс	Семестр	Трудоемкость в ЗЕТ	Общая трудоемкость (час.)	Контактная работа			Всего контактных часов	Из них в интерактивной форме	Кол-во часов на СРС	Курсовые работы	Кол-во часов на контроль	Форма контроля
				ЛК	ПР	ЛБ						
1	2	3	108	16	18	-	34	8	74	-	-	-
2	3	4	144	16	18	-	34	8	74	-	36	экзамен
Итого:		7	252	32	36	-	68	16	148	-	36	экзамен

В интерактивной форме часы используются в виде: групповой дискуссии, заслушивании и обсуждении подготовленных обучающимися докладов по тематике дисциплины.

5. СОДЕРЖАНИЕ ДИСЦИПЛИНЫ, СТРУКТУРИРОВАННОЕ ПО ТЕМАМ (РАЗДЕЛАМ) С УКАЗАНИЕМ ОТВЕДЕННОГО НА НИХ КОЛИЧЕСТВА АКАДЕМИЧЕСКИХ ИЛИ АСТРОНОМИЧЕСКИХ ЧАСОВ И ВИДОВ УЧЕБНЫХ ЗАНЯТИЙ

№ п/п	Наименование раздела, темы	Контактная работа			Всего контактных часов	Из них в интерактивной форме	Кол-во часов на СРС	Кол-во часов на контроль
		ЛК	ПР	ЛБ				
1.	Данные, информация, знание. Свойства и меры информации.	4	-	-	4	-	19	-
2.	Работа в электронной таблице Calc.	4	6	-	10	2	19	-
3.	Информационный процесс накопления данных.	4	6	-	10	2	18	-
4.	Разработка баз данных в СУБД Access.	4	6	-	10	4	18	-
	Всего за 3 семестр:	16	18	-	34	8	74	-
5.	Хранилища данных. Технология OLAP.	4	4	-	12	2	19	-
6.	Классификация и регрессия	4	4	-	12	2	19	-
7.	Поиск ассоциативных правил	4	5	-	12	2	18	-
8.	Кластеризация	4	5	-	14	2	18	-
	Экзамен	-	-	-	-	-	-	36
	Всего за 4 семестр:	16	18	-	34	8	74	36

№ п/п	Наименование раздела, темы	Контактная работа			Всего контактных часов	Из них в интерактивной форме	Кол-во часов на СРС	Кол-во часов на контроль
		ЛК	ПР	ЛБ				
	Итого:	32	36	-	68	16	148	36

Содержание дисциплины:

Тема 1. Данные, информация, знание. Свойства и меры информации.

Информация и ее адекватность. Меры информации. Общая характеристика процессов сбора, передачи, обработки и накопления информации. Качество информации. Системы классификации информации. Кодирование информации в информационных системах. Сокращение избыточности данных.

Тема 2. Работа в электронной таблице Calc.

Создание ЭТ. Форматирование ячеек таблицы. Различные типы данных в ЭТ. Формулы, встроенные функции. Построение диаграмм.

Тема 3. Информационный процесс накопления данных.

Общая характеристика процессов накопления. Выбор хранимых данных. Базы данных. Реляционная модель БД. Реляционная структура данных. Целостность реляционных данных.

Тема 4. Разработка баз данных в СУБД Access.

Интерфейс Access. Однотабличная база данных. Связи между таблицами, типы связей. Ключи первичный и вторичный. Многотабличные база данных Создание: таблиц, форм, запросов, отчетов.

Тема 5. Хранилища данных. Технология OLAP.

Тест FASMI. Что такое хранилище данных. Понятие гиперкуба. Структура OLAP куба. Компоненты OLAP. Архитектура OLAP. Данные в ХД: детализированные и агрегированные данные, метаданные.

Тема 6. Классификация и регрессия.

Постановка задачи. Классификационные правила. Методы построения правил классификации. Алгоритм построения 1 - правил. Метод Naïve Bayes. Деревья решений. Методы построения деревьев решений. Математические функции. Методы построения математических функций. Прогнозирование временных рядов. Методы прогнозирования временных рядов.

Тема 7. Поиск ассоциативных правил.

Постановка задачи. Сиквенциальный анализ. Разновидности задачи поиска ассоциативных правил. Представление результатов. Алгоритм Apriori

Тема 8. Кластеризация.

Постановка задачи. Меры близости, основанные на расстояниях, используемые в алгоритмах кластеризации. Представление результатов. Алгоритмы кластеризации. Иерархические алгоритмы. Неиерархические алгоритмы. Адаптивные методы кластеризации.

6. ПЕРЕЧЕНЬ УЧЕБНО-МЕТОДИЧЕСКОГО ОБЕСПЕЧЕНИЯ ДЛЯ САМОСТОЯТЕЛЬНОЙ РАБОТЫ ОБУЧАЮЩИХСЯ ПО ДИСЦИПЛИНЕ (МОДУЛЮ)

Основная литература:

1. Чубукова И. А. Data Mining. М.: Интернет-Университет Информационных Технологий, 2008, 383 с. Режим доступа: https://biblioclub.ru/index.php?page=book_red&id=233055&sr=1
2. OpenOffice.org Calc: курс. Москва: Интернет-Университет Информационных Технологий, 2011, 479 с. Режим доступа: https://biblioclub.ru/index.php?page=book_red&id=234807&sr=1

Дополнительная литература:

1. Архипенков С. Я., Голубев Д., Максименко О. Хранилища данных: от концепции до внедрения. М.: Диалог-МИФИ, 2002, 528 с. Режим доступа: https://biblioclub.ru/index.php?page=book_red&id=89285&sr=1
2. Быкова В. В. Искусство создания базы данных в Microsoft Office Access 2007: учебное пособие, Красноярск: Сибирский федеральный университет, 2011, 260 с. Режим доступа: https://biblioclub.ru/index.php?page=book_red&id=229161&sr=1

7. МАТЕРИАЛЬНО-ТЕХНИЧЕСКОЕ ОБЕСПЕЧЕНИЕ ДИСЦИПЛИНЫ (МОДУЛЯ)

В образовательном процессе используются:

- учебные аудитории для проведения занятий лекционного типа, занятий семинарского типа, курсового проектирования (выполнения курсовых работ), групповых и индивидуальных консультаций, текущего контроля и промежуточной аттестации (мебель аудиторная (столы, стулья, доска аудиторная), комплект мультимедийного оборудования, включающий мультимедиапроектор, экран, переносной ноутбук для демонстрации презентаций; учебно-наглядные пособия; обеспечивающие тематические иллюстрации);
- помещения для самостоятельной работы (оснащены компьютерными столами, стульями, доской аудиторной, компьютерной техникой с возможностью подключения к сети «Интернет» и обеспечением доступа в электронную информационно-образовательную среду университета);
- помещения для хранения и профилактического обслуживания учебного оборудования (оснащены наборами инструментов, оборудованием, расходными материалами для монтажа, ремонта и обслуживания информационно-телекоммуникационной сети филиала и вычислительной техники);
- лаборатория информационных технологий (оснащена компьютерными столами, стульями, мультимедийным проектором, экраном проекционным, компьютерной техникой с возможностью подключения к сети «Интернет» и обеспечением доступа в электронную информационно-образовательную среду университета).

7.1. ПЕРЕЧЕНЬ ЛИЦЕНЗИОННОГО ПРОГРАММНОГО ОБЕСПЕЧЕНИЯ

1. Microsoft Windows
2. Microsoft Office / LibreOffice
3. Microsoft Access – система управления базами данных.
4. Deductor Academic – аналитическая платформа.

7.2. ЭЛЕКТРОННО-БИБЛИОТЕЧНЫЕ СИСТЕМЫ

1. ЭБС «Издательство Лань» [Электронный ресурс]: электронная библиотечная система / ООО «Издательство Лань». – Режим доступа: <https://e.lanbook.com/>;
2. ЭБС «Электронная библиотечная система ЮРАЙТ» [Электронный ресурс]: электронная библиотечная система / ООО «Электронное издательство ЮРАЙТ». – Режим доступа: <https://biblio-online.ru/>;
3. ЭБС «Университетская библиотека онлайн» [Электронный ресурс]: электронно-периодическое издание; программный комплекс для организации онлайн-

доступа к лицензионным материалам / ООО «НексМедиа». – Режим доступа: <https://biblioclub.ru/>.

7.3. СОВРЕМЕННЫЕ ПРОФЕССИОНАЛЬНЫЕ БАЗЫ ДАННЫХ

1. Электронная база данных SCOPUS.
2. Электронная база данных РИНЦ.

7.4. ИНФОРМАЦИОННЫЕ СПРАВОЧНЫЕ СИСТЕМЫ

1. Справочно-правовая информационная система Консультант Плюс
<http://www.consultant.ru/>
2. Электронный справочник "Информιο" для высших учебных заведений
<http://www.informio.ru/>

8. ИНЫЕ СВЕДЕНИЯ И МАТЕРИАЛЫ НА УСМОТРЕНИЕ ВЕДУЩЕЙ КАФЕДРЫ

Не предусмотрено.

9. ОБЕСПЕЧЕНИЕ ОБРАЗОВАНИЯ ДЛЯ ЛИЦ С ОВЗ

Для обеспечения образования инвалидов и лиц с ограниченными возможностями здоровья реализация дисциплины может осуществляться в адаптированном виде, с учетом специфики освоения и дидактических требований, исходя из индивидуальных возможностей и по личному заявлению обучающегося.