

МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ РОССИЙСКОЙ ФЕДЕРАЦИИ
федеральное государственное бюджетное образовательное учреждение высшего
образования «Вятский государственный университет»
(ВятГУ)
г. Киров

Утверждаю
Директор/Декан Бушмелева Н. А.



Номер регистрации
РПД_3-09.04.03.03_2020_112505
Актуализировано: 22.03.2021

Рабочая программа дисциплины
Анализ больших данных

	наименование дисциплины
Квалификация выпускника	Магистр
Направление подготовки	09.04.03 шифр
	Прикладная информатика наименование
Направленность (профиль)	3-09.04.03.03 шифр
	Машинное обучение и анализ данных наименование
Формы обучения	Очная наименование
Кафедра-разработчик	Кафедра прикладной математики и информатики (ОРУ) наименование
Выпускающая кафедра	Кафедра прикладной математики и информатики (ОРУ) наименование

Сведения о разработчиках рабочей программы дисциплины

Котельников Евгений Вячеславович

ФИО

Цели и задачи дисциплины

Цель дисциплины	Сформировать понимание внутреннего устройства, механики работы, области применимости существующих решений, осветить сильные и слабые стороны, научить практическим навыкам анализа больших массивов информации.
Задачи дисциплины	<ul style="list-style-type: none"> - ознакомление с существующими методами и системами обработки больших данных; - формирование представления об области применимости методов и систем обработки больших данных; - формирование понимания преимуществ и недостатков методов и систем обработки больших данных; - формирование культуры работы с информацией; - формирование навыка проведения научных исследований.

Перечень планируемых результатов обучения по дисциплине, соотнесенных с планируемыми результатами освоения образовательной программы

Компетенция УК-1

Способен осуществлять критический анализ проблемных ситуаций на основе системного подхода, выработать стратегию действий		
Знает	Умеет	Владеет
основы системного подхода для критического анализа проблемных ситуаций и способы разрешения этих ситуаций; принципы декомпозиции задач анализа больших данных; основные методологические подходы к формированию стратегии действий, направленных на решение поставленной задачи	применять методы системного подхода для критического анализа проблемных ситуаций; находить решение элементарных (атомарных) задач при анализе больших данных; формировать стратегию действий с обоснованием принимаемых решений	навыками построения стратегии действий для разрешения проблемных ситуаций; навыками получения комплексных решений путем композиции решений элементарных (атомарных) задач анализа больших данных; навыками системного анализа стратегии действий, позволяющими выделять ошибочные элементы

Компетенция ПК-5

Способен осуществлять оценку качества формализации и алгоритмизации поставленных задач, оценку качества и эффективности программного кода, редактирование программного кода и контроль версий программного обеспечения		
Знает	Умеет	Владеет
методы и приемы формализации задач анализа больших данных; методы и приемы алгоритмизации задач анализа больших данных; программные продукты для графического отображения	использовать методы и приемы формализации задач; использовать методы и приемы алгоритмизации поставленных задач; использовать программные продукты для графического отображения алгоритмов;	навыком оценки качества формализации и алгоритмизации поставленных задач, оценки качества и эффективности программного кода, редактирования программного кода и

алгоритмов; стандартные алгоритмы анализа больших данных и области их применения; критерии качества и эффективности анализа больших данных	применять стандартные алгоритмы анализа больших данных; применять коллективную среду разработки программного обеспечения и систему контроля версий	контроль версий программного обеспечения
--	--	--

Компетенция ПК-6

Способен осуществлять анализ и согласование архитектуры программного обеспечения, оценку качества проектирования программного обеспечения, структуры базы данных, программных интерфейсов

Знает	Умеет	Владеет
критерии качества проектирования программного обеспечения, структуры базы данных, программных интерфейсов; принципы построения архитектуры программного обеспечения и вида архитектуры программного обеспечения, методологии и средства проектирования программного обеспечения; методы и средства проектирования программных интерфейсов	применять методологии и средства проектирования программного обеспечения, применять методы и средства проектирования баз данных; применять методы и средства проектирования программных интерфейсов	навыком оценки качества проектирования программного обеспечения, структуры данных, программных интерфейсов

Компетенция ПК-7

Способен осуществлять оценку качества разработанных процедур отладки программного кода, процедур сбора диагностических данных, измерения требуемых характеристик программного обеспечения, тестовых наборов данных в соответствии с выбранной методикой, оценку результатов проверки работоспособности программного обеспечения, рефакторинг и оптимизацию кода

Знает	Умеет	Владеет
критерии качества и эффективности моделей анализа больших данных; основные виды диагностических данных и способы их представления	использовать современные средства обработки и анализа больших данных; производить подготовку наборов данных и проверку работоспособности программного обеспечения на их основе	способностью осуществлять оценку качества разработанных процедур отладки программного кода, процедур сбора диагностических данных, измерения требуемых характеристик программного обеспечения

Компетенция ПК-8

Способен осуществлять исследование и изучение мировых практик выполнения аналитических работ, выявление проблем и сложностей в существующих практиках,

описание методик проведения аналитических работ, их апробацию на выбранных проектах и доработку		
Знает	Умеет	Владеет
<p>существующие мировые практики выполнения анализа больших данных их достоинства и недостатки; способы описания методик проведения аналитических работ</p>	<p>выявлять проблемы и сложности в существующих практиках выполнения анализа больших данных; планировать аналитические работы, планировать ресурсы, управлять проектами</p>	<p>навыком исследования и изучения мировых практик выполнения анализа больших данных, выявления проблем и сложностей в существующих практиках; навыком описания методик решения задач анализа больших данных, их апробации на выбранных проектах и доработки</p>

Структура дисциплины
Тематический план

№ п/п	Наименование разделов дисциплины	Шифр формируемых компетенций
1	Пакетная и потоковая обработка данных	ПК-5, ПК-8, УК-1
2	Хранение данных	ПК-6, ПК-7
3	Подготовка и прохождение промежуточной аттестации	ПК-5, ПК-6, ПК-7, ПК-8, УК-1

Формы промежуточной аттестации

Зачет	Не предусмотрен (Очная форма обучения)
Экзамен	3 семестр (Очная форма обучения)
Курсовая работа	Не предусмотрена (Очная форма обучения)
Курсовой проект	Не предусмотрена (Очная форма обучения)

Трудоемкость дисциплины

Форма обучения	Курсы	Семестры	Общий объем (трудоемкость)		Контактная работа, час	в том числе аудиторная контактная работа обучающихся с преподавателем, час				Самостоятельная работа, час	Курсовая работа (проект), семестр	Зачет, семестр	Экзамен, семестр
			Часов	ЗЕТ		Всего	Лекции	Семинарские, практические занятия	Лабораторные занятия				
Очная форма обучения	2	3	216	6	113	54	18	0	36	103			3

Содержание дисциплины

Очная форма обучения

Код занятия	Наименование тем занятий	Трудоемкость, академических часов
Раздел 1 «Пакетная и потоковая обработка данных»		128.00
Лекции		
Л1.1	Распределенные файловые системы	4.00
Л1.2	Модель вычислений MapReduce	4.00
Л1.3	Устройство MapReduce и Hadoop	2.00
Л1.4	Потоковая обработка данных	4.00
Лабораторные занятия		
Р1.1	Устройство MapReduce и Hadoop	8.00
Р1.2	Часто применяемые методы в обработке больших данных	8.00
Р1.3	Модель вычислений Spark Streaming	8.00
Самостоятельная работа		
С1.1	Подготовка к лабораторным занятиям	50.00
Контактная внеаудиторная работа		
КВР1.1	Контактная внеаудиторная работа	40.00
Раздел 2 «Хранение данных»		61.00
Лекции		
Л2.1	BigTable-подобные хранилища	2.00
Л2.2	Дунамо-подобные хранилища	2.00
Лабораторные занятия		
Р2.1	BigTable-подобные хранилища	6.00
Р2.2	Дунамо-подобные хранилища	6.00
Самостоятельная работа		
С2.1	Подготовка к лабораторным занятиям	28.50
Контактная внеаудиторная работа		
КВР2.1	Контактная внеаудиторная работа	16.50
Раздел 3 «Подготовка и прохождение промежуточной аттестации»		27.00
ЭЗ.1	Подготовка к сдаче экзамена	24.50
КВР3.1	Консультация перед экзаменом	2.00
КВР3.2	Сдача экзамена	0.50
ИТОГО		216.00

Содержание дисциплины данной рабочей программы используется при обучении по индивидуальному учебному плану, при ускоренном обучении, при применении дистанционных образовательных технологий и электронном обучении (при наличии).

Методические указания для обучающихся по освоению дисциплины

Успешное освоение дисциплины предполагает активное, творческое участие обучающегося на всех этапах ее освоения путем планомерной, повседневной работы. Обучающийся обязан посещать лекции, семинарские, практические и лабораторные занятия (при их наличии), получать консультации преподавателя и выполнять самостоятельную работу.

Изучение дисциплины следует начинать с проработки настоящей рабочей программы, методических указаний и разработок, указанных в программе, особое внимание уделить целям, задачам, структуре и содержанию дисциплины.

Главной задачей каждой лекции является раскрытие сущности темы и анализ ее основных положений. Тематика лекций определяется настоящей рабочей программой дисциплины.

Лекции – это систематическое устное изложение учебного материала. На них обучающийся получает основной объем информации по каждой конкретной теме. Лекции обычно носят проблемный характер и нацелены на освещение наиболее трудных и дискуссионных вопросов.

Предполагается, что обучающиеся приходят на лекции, предварительно проработав соответствующий учебный материал по источникам, рекомендованным программой. Часто обучающимся трудно разобраться с дискуссионными вопросами, дать однозначный ответ. Преподаватель, сравнивая различные точки зрения, излагает свой взгляд и нацеливает их на дальнейшие исследования и поиск научных решений. После лекции желательно вечером перечитать и закрепить полученную информацию, тогда эффективность ее усвоения значительно возрастает. При работе с конспектом лекции необходимо отметить материал, который вызывает затруднения для понимания, попытаться найти ответы на затруднительные вопросы, используя предлагаемую литературу. Если самостоятельно не удалось разобраться в материале, сформулируйте вопросы и обратитесь за помощью к преподавателю.

Целью семинарских занятий является проверка уровня понимания обучающимися вопросов, рассмотренных на лекциях и в учебной литературе.

Целью практических и лабораторных занятий является формирование у обучающихся умений и навыков применения теоретических знаний в реальной практике решения задач; восполнение пробелов в пройденной теоретической части курса.

Семинарские, практические и лабораторные занятия в равной мере направлены на совершенствование индивидуальных навыков решения теоретических и прикладных задач, выработку навыков интеллектуальной работы, а также ведения дискуссий. Для успешного участия в семинарских, практических и лабораторных занятиях обучающемуся следует тщательно подготовиться.

Основной формой подготовки обучающихся к практическим (лабораторным) занятиям является самостоятельная работа с учебно-методическими материалами, научной литературой, статистическими данными и т.п.

Изучив конкретную тему, обучающийся может определить, насколько хорошо он в ней разобрался. Если какие-то моменты остались непонятными, целесообразно составить список вопросов и на занятии задать их преподавателю. Практические (лабораторные) занятия предоставляют обучающемуся возможность творчески раскрыться, проявить инициативу и развить навыки публичного ведения дискуссий и общения.

Самостоятельная работа обучающихся включает в себя выполнение различного рода заданий (изучение учебной и научной литературы, материалов лекций, систематизацию прочитанного материала, подготовку контрольной работы, решение

задач, подготовка докладов, написание рефератов, публикация тезисов, научных статей, подготовка и защита курсовой работы / проекта и другие), которые ориентированы на глубокое усвоение материала изучаемой дисциплины.

Обучающимся рекомендуется систематически отводить время для повторения пройденного материала, проверяя свои знания, умения и навыки.

Внутренняя система оценки качества освоения дисциплины включает входной контроль уровня подготовленности обучающихся, текущий контроль успеваемости, промежуточную аттестацию, направленную на оценивание промежуточных и окончательных результатов обучения по дисциплине (в том числе результатов курсового проектирования (выполнения курсовых работ) при наличии).

При проведении промежуточной аттестации обучающегося учитываются результаты текущего контроля, проводимого в течение освоения дисциплины.

Процедура оценивания результатов освоения дисциплины осуществляется на основе действующих локальных нормативных актов ФГБОУ ВО «Вятский государственный университет», с которыми обучающиеся ознакамливаются на официальном сайте университета www.vyatsu.ru.

Учебно-методическое обеспечение дисциплины, в том числе учебно-методическое обеспечение самостоятельной работы обучающегося по дисциплине

Учебная литература (основная)

1) Чубукова, И. А. Data Mining / И.А. Чубукова. - 2-е изд., испр. - Москва : Интернет-Университет Информационных Технологий, 2008. - 383 с. - (Основы информационных технологий). - ISBN 978-5-94774-819-2 : Б. ц. - URL: <http://biblioclub.ru/index.php?page=book&id=233055/> (дата обращения: 24.03.2020). - Режим доступа: ЭБС Университетская библиотека ONLINE. - Текст : электронный.

2) Технологии анализа данных: Data Mining, Visual Mining, Text Mining, OLAP : учеб. пособие / А. А. Барсегян, М. С. Куприянов, В. В. Степаненко, И. И. Холод. - 2-е изд. - СПб. : БХВ-Петербург, 2008. - 375 с. + 1 эл. опт. диск (CD-ROM). - Библиогр.: с.368-372. - Предм. указ.: с. 372-375. - ISBN 978-5-94157-991-4 : 368.00 р. - Текст : непосредственный.

3) Лазицкас, Е. А. Базы данных и системы управления базами данных : учебное пособие / Е.А. Лазицкас. - Минск : РИПО, 2016. - 267 с. - ISBN 978-985-503-558-0 : Б. ц. - URL: <http://biblioclub.ru/index.php?page=book&id=463305/> (дата обращения: 24.03.2020). - Режим доступа: ЭБС Университетская библиотека ONLINE. - Текст : электронный.

Учебная литература (дополнительная)

1) Петрунин, Юрий Юрьевич. Информационные технологии анализа данных. Data analysis : учеб. пособие по дисциплине "Информатика" для студентов высших учебных заведений, обучающихся по управленческим и экономическим специальностям и направлениям / Ю. Ю. Петрунин ; МГУ. - М. : Ун-т книжный дом (КДУ), 2008. - 291 с. : ил., табл. ; 20 см. - Библиогр.: с. 290-291. - ISBN 978-5-98227-416-8 : 264.00 р. - Текст : непосредственный.

Учебно-методические издания

1) Сидорова, Н. П. Базы данных: практикум по проектированию реляционных баз данных : учебное пособие / Н.П. Сидорова. - Москва|Берлин : Директ-Медиа, 2020. - 93 с. : ил. - Библиогр.: с. 85. - ISBN 978-5-4499-0799-8 : Б. ц. - URL: <http://biblioclub.ru/index.php?page=book&id=575080/> (дата обращения: 24.03.2020). - Режим доступа: ЭБС Университетская библиотека ONLINE. - Текст : электронный.

Электронные образовательные ресурсы

1) Портал дистанционного обучения ВятГУ [электронный ресурс] / - Режим доступа: <http://mooc.do-kirov.ru/>

- 2) Раздел официального сайта ВятГУ, содержащий описание образовательной программы [электронный ресурс] / - Режим доступа: https://www.vyatsu.ru/php/programms/eduPrograms.php?Program_ID=3-09.04.03.03
- 3) Личный кабинет студента на официальном сайте ВятГУ [электронный ресурс] / - Режим доступа: <https://new.vyatsu.ru/account/>
- 4) Единое окно доступа к образовательным ресурсам <http://window.edu.ru/>

Электронные библиотечные системы (ЭБС)

- ЭБС «Научная электронная библиотека eLIBRARY» (<http://elibrary.ru/defaultx.asp>)
- ЭБС «Издательства Лань» (<http://e.lanbook.com/>)
- ЭБС «Университетская библиотека online» (www.biblioclub.ru)
- Внутренняя электронно-библиотечная система ВятГУ (<http://lib.vyatsu.ru/>)
- ЭБС «ЮРАЙТ» (<https://urait.ru>)

Современные профессиональные базы данных и информационные справочные системы

- ГАРАНТ
- КонсультантПлюс
- Техэксперт: Нормы, правила, стандарты
- Роспатент (<https://www1.fips.ru/elektronnye-servisy/informatsionno-poiskovaya-sistema>)
- Web of Science® (<http://webofscience.com>)

Материально-техническое обеспечение дисциплины

Демонстрационное оборудование

Перечень используемого оборудования
ПРОЕКТОР NEC V302H FULL 3D

Специализированное оборудование

Перечень используемого оборудования
ПЕРСОНАЛЬНЫЙ КОМПЬЮТЕР ICL S253.MI (МОНОБЛОК)

Перечень информационных технологий, используемых при осуществлении образовательного процесса по дисциплине, в том числе лицензионное и свободно распространяемое ПО (включая ПО отечественного производства)

№ п.п	Наименование ПО	Краткая характеристика назначения ПО
1	Программная система с модулями для обнаружения текстовых заимствований в учебных и научных работах «Антиплагиат.ВУЗ»	Программный комплекс для проверки текстов на предмет заимствования из Интернет-источников, в коллекции диссертация и авторефератов Российской государственной библиотеки (РГБ) и коллекции нормативно-правовой документации LEXPRO
2	Microsoft Office 365 ProPlusEdu ALNG SubsVL MVL AddOn toOPP	Набор веб-сервисов, предоставляющий доступ к различным программам и услугам на основе платформы Microsoft Office, электронной почте бизнес-класса, функционалу для общения и управления документами
3	Office Professional Plus 2016	Пакет приложений для работы с различными типами документов: текстами, электронными таблицами, базами данных, презентациями
4	Windows Professional	Операционная система
5	Kaspersky Endpoint Security для бизнеса	Антивирусное программное обеспечение
6	Справочная правовая система «Консультант Плюс»	Справочно-правовая система по законодательству Российской Федерации
7	Электронный периодический справочник ГАРАНТ Аналитик	Справочно-правовая система по законодательству Российской Федерации
8	Security Essentials (Защитник Windows)	Защита в режиме реального времени от шпионского программного обеспечения, вирусов.
9	МойОфис Стандартный	Набор приложений для работы с документами, почтой, календарями и контактами на компьютерах и веб браузерах
10	Python	Язык программирования
11	Anaconda	дистрибутив языков программирования Python и R с набором приложений. По умолчанию в Anaconda Navigator доступны следующие приложения: JupyterLab Jupyter Notebook QtConsole Spyder Glue Orange RStudio Visual Studio Code

Обновленный список программного обеспечения данной рабочей программы находится по адресу:
https://www.vyatsu.ru/php/list_it/index.php?op_id=112505