

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_2021_118031
Актуализировано: 04.04.2021

Рабочая программа дисциплины
Вероятностно-статистические методы в машинном обучении

	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

Шатров Анатолий Викторович

ФИО

Цели и задачи дисциплины

Цель дисциплины	Курс призван познакомить студентов с некоторыми элементами современного анализа данных. Студенты получают представление об основных классах задач машинного обучения в вероятностной постановке, более подробно знакомятся с алгоритмами для решения задач классификации и кластеризации, регрессионного анализа, вероятностных графических моделей, тематическим моделированием. Дисциплина базируется на методах математической статистики.
Задачи дисциплины	<ul style="list-style-type: none"> - ознакомление слушателей с задачами, принципами, методами и подходами машинного обучения в вероятностной постановке; - приобретение слушателями теоретических знаний, и практических умений и навыков в области исследования задач анализа данных и их решения методами машинного обучения в вероятностной постановке.

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; основные вероятностно-статистические методы в машинном обучении	применять вероятностно-статистические методы и методы системного подхода для критического анализа проблемных ситуаций; формировать стратегию действий с обоснованием принимаемых решений	навыками построения стратегии действий для разрешения проблемных ситуаций; навыками системного анализа стратегии действий, позволяющими выделять ошибочные элементы

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 программного кода; основные принципы отладки программного кода; основные виды диагностических данных и способы их представления; типовые метрики	использовать современные средства обработки данных; использовать основные, наиболее употребительные методы математической обработки результатов исследований	способностью осуществлять оценку качества разработанных процедур отладки программного кода, процедур сбора диагностических данных, измерения требуемых характеристик программного обеспечения

программного обеспечения, основные методы измерения и оценки характеристик программного обеспечения		
---	--	--

Структура дисциплины
Тематический план

№ п/п	Наименование разделов дисциплины	Шифр формируемых компетенций
1	Вероятность и статистика	УК-1
2	Классификация и кластеризация	ПК-7, УК-1
3	Вероятностные графические и тематические модели	ПК-7, УК-1
4	Подготовка и прохождение промежуточной аттестации	ПК-7, УК-1

Формы промежуточной аттестации

Зачет	Не предусмотрен (Очная форма обучения)
Экзамен	2 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	1	2	216	6	113	54	18	0	36	103			2

Содержание дисциплины

Очная форма обучения

Код занятия	Наименование тем занятий	Трудоемкость, академических часов
Раздел 1 «Вероятность и статистика»		43.00
Лекции		
Л1.1	Условная вероятность. Распределения случайных величин	2.00
Л1.2	Проверка статистических гипотез	2.00
Лабораторные занятия		
Р1.1	Вероятность. Выборки. Случайные величины	2.00
Р1.2	Распределения случайных величин	2.00
Р1.3	Проверка статистических гипотез	2.00
Р1.4	Корреляция	2.00
Самостоятельная работа		
С1.1	Подготовка к лекциям	6.00
С1.2	Выполнение заданий лабораторных работ	12.50
Контактная внеаудиторная работа		
КВР1.1	Контактная внеаудиторная работа	12.50
Раздел 2 «Классификация и кластеризация»		76.00
Лекции		
Л2.1	Наивный байесовский классификатор	2.00
Л2.2	EM-алгоритм: оценка плотности распределения и вероятностный алгоритм кластеризации	4.00
Л2.3	Метод стохастического градиента	2.00
Лабораторные занятия		
Р2.1	Задача рубрикации новостных статей	4.00
Р2.2	Смесь нормальных распределений. Метод GaussianMixture	2.00
Р2.3	Ядерная оценка плотности распределения (KDEClassifier)	2.00
Р2.4	Вероятностный алгоритм кластеризации (EM-алгоритм)	4.00
Р2.5	Метод стохастического градиента	4.00
Самостоятельная работа		
С2.1	Подготовка к лекциям	6.00
С2.2	Выполнение заданий лабораторных работ	24.00
Контактная внеаудиторная работа		
КВР2.1	Контактная внеаудиторная работа	22.00
Раздел 3 «Вероятностные графические и тематические модели»		70.00
Лекции		
Л3.1	Вероятностные графические модели	2.00
Л3.2	Тематическое моделирование	4.00
Лабораторные занятия		
Р3.1	Марковские модели (скрытые)	4.00
Р3.2	Conditional Random Fields	4.00
Р3.3	Topic Modelling	4.00

Самостоятельная работа		
С3.1	Проработка материала лекций	6.00
С3.2	Выполнение заданий лабораторных работ	24.00
Контактная внеаудиторная работа		
КВР3.1	Контактная внеаудиторная работа	22.00
Раздел 4 «Подготовка и прохождение промежуточной аттестации»		27.00
Э4.1	Подготовка к сдаче экзамена	24.50
КВР4.1	Консультация перед экзаменом	2.00
КВР4.2	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) Кухаренко, Б. Г. Интеллектуальные системы и технологии : учебное пособие / Б.Г. Кухаренко. - Москва : Альтаир|МГАВТ, 2015. - 115 с. - Б. ц. - URL: <http://biblioclub.ru/index.php?page=book&id=429758/> (дата обращения: 24.03.2020). - Режим доступа: ЭБС Университетская библиотека ONLINE. - Текст : электронный.

2) Программные системы статистического анализа: обнаружение закономерностей в данных с использованием системы R и языка Python : учебное пособие / В.М. Волкова, М.А. Семенова, Е.С. Четвертакова, С.С. Вожов. - Новосибирск : Новосибирский государственный технический университет, 2017. - 74 с. : ил., табл. - Библиогр.: с. 48. - ISBN 978-5-7782-3183-2 : Б. ц. - URL: <http://biblioclub.ru/index.php?page=book&id=576496/> (дата обращения: 24.03.2020). - Режим доступа: ЭБС Университетская библиотека ONLINE. - Текст : электронный.

3) Прикладная математическая статистика : учебное пособие. - Томск : ТУСУР, 2016. - 113 с. - Б. ц. - URL: <http://biblioclub.ru/index.php?page=book&id=480889/> (дата обращения: 24.03.2020). - Режим доступа: ЭБС Университетская библиотека ONLINE. - Текст : электронный.

Учебная литература (дополнительная)

1) Кобзарь, А. И. Прикладная математическая статистика / А.И. Кобзарь. - 2-е изд., испр. - Москва : Физматлит, 2012. - 816 с. - ISBN 978-5-9221-1375-5 : Б. ц. - URL: <http://biblioclub.ru/index.php?page=book&id=82617/> (дата обращения: 24.03.2020). - Режим доступа: ЭБС Университетская библиотека ONLINE. - Текст : электронный.

Учебно-методические издания

1) Многомерные статистические методы : методические указания по выполнению лабораторных работ. для студентов-магистров очной и заочной форм обучения по направлению подготовки 09.04.03 «прикладная информатика». профиль «прикладная информатика в экономике». - Сочи : СГУ, 2018. - 64 с. - Б. ц. - URL: <https://e.lanbook.com/book/147679> (дата обращения: 15.05.2020). - Режим доступа: ЭБС Лань. - Текст : электронный.

2) Петров, А. А. Вероятностное и статистическое моделирование : учебно-методическое пособие (лекционные материалы и материалы к лабораторным работам) для студентов, магистрантов и аспирантов всех форм обучения / А. А. Петров, М. В. Куркина. - Ханты-Мансийск : ЮГУ, 2017. - 86 с. - Б. ц. - URL: <https://e.lanbook.com/book/149000> (дата обращения: 15.05.2020). - Режим доступа: ЭБС Лань. - 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
Ноутбук Lenovo B590 4G
ПРОЕКТОР NEC V302H FULL 3D

Специализированное оборудование

Перечень используемого оборудования
ГРАФИЧЕСКАЯ РАБОЧАЯ СТАНЦИЯ DEPO Race G560S
ПЕРСОНАЛЬНЫЙ КОМПЬЮТЕР TP CORP OPTIMA

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=118031