

Документ подписан простой электронной подписью  
Информация о владельце:  
ФИО: Смирнов Сергей Николаевич  
Должность: врио ректора  
Дата подписания: 19.09.2022 11:28:14  
Уникальный программный ключ:  
69e375c64f7e975d4e8830e7b4fcc2ad1bf35f08

Приложение № 4  
к Положению об основной  
образовательной программе  
высшего образования

Министерство науки и высшего образования Российской Федерации  
ФГБОУ ВО «Тверской государственный университет»

Утверждаю:  
Руководитель ООП  
\_\_\_\_\_ Цветков В.П.  
«\_\_» \_\_\_\_\_ 20\_\_ г.

Рабочая программа дисциплины (с аннотацией)

Машинное обучение

Направление подготовки

02.03.01 Математика и компьютерные науки

Направленность (профиль)

Математическое и компьютерное моделирование

Для студентов 3 курса, очная

Составитель: к.ф.-м.н., ст.преп. Лебедев Д.Ю.

Тверь, 2019

## **I. Аннотация**

### **1. Цель и задачи дисциплины**

Целью освоения дисциплины является: изучение основ теории машинного обучения и основных алгоритмов машинного обучения, овладение навыками практического решения задач анализа данных.

Задачами освоения дисциплины являются: изучение принципов построения основных алгоритмов машинного обучения, особенностей применения их на практике, их недостатков, а также методики выбора подходящего алгоритма для решения задач анализа и прогнозирования поведения социально-экономических и природных систем с использованием языка программирования R. Изложить последние достижения в области практического использования методов машинного обучения, дать представление о современных проблемах машинного обучения.

### **2. Место дисциплины в структуре ООП**

Дисциплина входит в вариативную часть ООП направления подготовки «Математика и компьютерные науки».

Для успешного освоения дисциплины необходимы знания и умения, приобретенные в результате обучения дисциплинам: алгебра и теория чисел, дискретная математика и математическая логика, основы программирования, структуры и алгоритмы компьютерной обработки данных.

Список дисциплин, для успешного освоения которых необходимы знания и умения, сформированные данной дисциплиной: учебная практика, управление в динамических системах, численные методы в математическом моделировании, фрактальные методы в исследовании социально-экономических и природных систем, катастрофы в теории гравитирующих конфигураций, математические методы гравитации и космологии, все виды производственной практики, подготовка к сдаче и сдача государственного экзамена, подготовка к защите и защита выпускной квалификационной работы.

**3. Объем дисциплины: 5 зачетных единиц, 180 академических часов, в том числе:**

**контактная аудиторная работа:** лабораторные работы 32 часа.

**самостоятельная работа:** 121 час, в том числе контроль 27 часов.

### **4. Планируемые результаты обучения по дисциплине, соотнесенные с**

планируемыми результатами освоения образовательной программы

Планируемые результаты освоения образовательной программы (формируемые компетенции)	Планируемые результаты обучения по дисциплине
ПК-1. Способен анализировать и прогнозировать поведение социально-экономических и природных систем на основе их математических и компьютерных моделей	ПК-1.1. Составляет и реализует комплексы программ для вычисления основных параметров математических и компьютерных моделей социально-экономических и природных систем

ПК-1.2. Анализирует и прогнозирует поведение социально-экономических и природных систем при изменении значений управляющих параметров математических и компьютерных моделей этих систем

### 5. Форма промежуточной аттестации и семестр прохождения

Итоговой формой аттестации в 6-м семестре обучения является экзамен.

### 6. Язык преподавания русский.

## II. Содержание дисциплины, структурированное по темам (разделам) с указанием отведенного на них количества академических часов и видов учебных занятий

Учебная программа – наименование разделов и тем	Всего (час.)	Контактная работа (час.)	Самостоятельная работа, в том числе Контроль (час.)
		Лабораторные работы	
Тема 1. Базовое программирование с R.	22	4	18
Тема 2. Манипуляции с данными в R. Библиотека dplyr.	22	4	18
Тема 3. Средства визуализации в R. Базовая графика и библиотека ggplot2.	22	4	18
Тема 4. Введение в машинное обучение. Линейная регрессия. Построение и интерпретация моделей линейной регрессии от одной и многих переменных.	22	4	18
Тема 5. Кластерный анализ.	22	4	18

Тема 6. Алгоритмы классификации.	22	4	18
Тема 7. Оценка качества алгоритмов.	22	4	18
Тема 8. Машинное обучение в прикладных задачах.	22	4	22
<b>ИТОГО</b>	<b>180</b>	<b>32</b>	<b>148</b>

### **III. Образовательные технологии**

Учебная программа – наименование разделов и тем	Вид занятия	Образовательные технологии
Тема 1. Базовое программирование с R.	лекция  практическое	Дискуссионные технологии, дистанционные образовательные технологии, проблемная лекция, разбор конкретных ситуаций
Тема 2. Манипуляции с данными в R. Библиотека dplyr.	лекция  практическое	Дискуссионные технологии, дистанционные образовательные технологии, проблемная лекция, разбор конкретных ситуаций
Тема 3. Средства визуализации в R. Базовая графика и библиотека ggplot2.	лекция  практическое	Дискуссионные технологии, дистанционные образовательные технологии, проблемная лекция, разбор конкретных ситуаций
Тема 4. Введение в машинное обучение. Линейная регрессия. Построение и интерпретация моделей линейной регрессии от одной и многих переменных.	лекция  практическое	Дискуссионные технологии, дистанционные образовательные технологии, проблемная лекция, разбор конкретных ситуаций
Тема 5. Кластерный анализ.	лекция  практическое	Дискуссионные технологии, дистанционные образовательные технологии, проблемная лекция, разбор конкретных ситуаций

Тема 6. Алгоритмы классификации.	лекция практическое	Дискуссионные технологии, дистанционные образовательные технологии, проблемная лекция, разбор конкретных ситуаций
Тема 7. Оценка качества алгоритмов.	лекция практическое	Дискуссионные технологии, дистанционные образовательные технологии, проблемная лекция, разбор конкретных ситуаций
Тема 8. Машинное обучение в прикладных задачах.	лекция практическое	Дискуссионные технологии, дистанционные образовательные технологии, проблемная лекция, разбор конкретных ситуаций

#### **IV. Оценочные материалы для проведения текущей и промежуточной аттестации**

##### ***Оценочные материалы для проведения текущей аттестации***

##### **Задания для практических (семинарских) занятий**

**Задание 1 (ПК-1.1, ПК-1.2):** Используя данные библиотеки nycflights13 вычислить:

а). суммарное расстояние, которое налетали судна каждой авиакомпании (группировка по названию) и вывести топ-3 компании.

б). используя метаданные самолетов, вывести общее количество перелетов в разрезе производитель, модель (группировка). Отсортировать результаты по убыванию.

в). прирост (убыль) количества вылетов по каждой авиакомпании относительно прошлого месяца.

г). информацию по первому вылету воздушного судна.

д) кумулятивную (накопительную) сумму количества всех полетов по неделям с 2013-1-1 с использованием `cumsum()`

**Задание 2 (ПК-1.1, ПК-1.2):** Рассчитать коэффициенты модели линейной регрессии от одной переменной с помощью функции `lm()` для зависимостей  $y$  от  $x$  по данным, представленным в таблице.

**Задание 3 (ПК-1.1, ПК-1.2):** Реализовать алгоритм k-means для разделения набора данных без меток на кластеры по исходным данным с

использованием R. Оценить результат работы алгоритма по выбранной метрике оценки качества кластеризации.

#### **Задание 4 (ПК-1.1, ПК-1.2):.**

Обучите несколько классификаторов, чтобы предсказать требуемое поле из набора данных. Рассчитайте следующие метрики, чтобы проверить качество моделей: доля правильных ответов (accuracy), точность (precision), полнота (recall). Постройте precision-recall и ROC-кривые для ваших моделей.

### **V. Учебно-методическое и информационное обеспечение дисциплины**

#### **1) Рекомендуемая литература**

##### **а) Основная литература:**

1. Классификация, регрессия и другие алгоритмы Data Mining с использованием R. Шитиков В. К., Мастицкий С. Э. Режим доступа: <https://ranalytics.github.io/data-mining/>
2. Мастицкий С. Э. (2016) Визуализация данных с помощью ggplot2. - М.: ДМК Пресс, 222 с
3. Самсонов Т.Е. Визуализация и анализ географических данных на языке R. Режим доступа: <https://tsamsonov.github.io/r-geo-course/>

##### **б) Дополнительная литература:**

1. Наглядная статистика. Используем R! А. Б. Шипунов. Режим доступа: <https://cran.r-project.org/doc/contrib/Shipunov-rbook.pdf>

#### **2) Программное обеспечение**

##### **а) Лицензионное программное обеспечение**

- Microsoft Windows 10 Enterprise Акт приема-передачи № 369 от 21 июля 2017;
- MS Office 365 pro plus Акт приема-передачи № 369 от 21 июля 2017
- Kaspersky Endpoint Security 10 для Windows – Акт на передачу прав №2129 от 25 октября 2016 г

##### **б) Свободно распространяемое программное обеспечение**

- RStudio
- Google Chrome
- WinDjView

### 3) Современные профессиональные базы данных и информационные справочные системы

1. ЭБС «ZNANIUM.COM» [www.znanium.com](http://www.znanium.com) ;
2. ЭБС «ЮРАИТ» [www.biblio-online.ru](http://www.biblio-online.ru) ;
3. ЭБС «Университетская библиотека онлайн» <https://biblioclub.ru/> ;
4. ЭБС IPRbooks [http://www.iprbookshop.ru /](http://www.iprbookshop.ru/);
5. ЭБС «Лань» <http://e.lanbook.com>;
6. ЭБС BOOK.ru <https://www.book.ru/>
7. ЭБС ТвГУ <http://megapro.tversu.ru/megapro/Web>
8. Научная электронная библиотека eLIBRARY.RU (подписка на журналы) [https://elibrary.ru/projects/subscription/rus\\_titles\\_open.asp?](https://elibrary.ru/projects/subscription/rus_titles_open.asp?) ;
9. Репозиторий ТвГУ <http://eprints.tversu.ru>

### 4) Перечень ресурсов информационно-телекоммуникационной сети «Интернет», необходимых для освоения дисциплины: <https://cyberleninka.ru/> научная электронная библиотека «Киберленинка».

## VI. Методические материалы для обучающихся по освоению дисциплины

### *Методические рекомендации по организации самостоятельной работы студентов*

Методические рекомендации по организации самостоятельной работы по дисциплине «Теория баз данных» см. в личном кабинете электронной образовательной среды (LMS).

### Требования к рейтинг-контролю для студентов очной формы обучения.

Баллы за текущую работу студентов распределяются между двумя модулями (периодами обучения) следующим образом:

Модуль (период обучения)	Максимальная сумма баллов в модуле	Максимальная сумма баллов за работу на практических занятиях	Посещаемость	Максимальный балл за рейтинговую контрольную работу
1	30	10	5	15
2	30	10	5	15

Правила формирования рейтинговой оценки и шкалу пересчета рейтинговых баллов в оценку на экзамене см. в «Положении о рейтинговой системе обучения в ТвГУ»:

[https://www.tversu.ru/sveden/files/Pologhenie\\_o\\_reytingovoy\\_sisteme\\_obucheniya\(1\).pdf](https://www.tversu.ru/sveden/files/Pologhenie_o_reytingovoy_sisteme_obucheniya(1).pdf)

## VII. Материально-техническое обеспечение

Учебный процесс по данной дисциплине проводится в аудиториях, оснащенных мультимедийными средствами обучения. Для организации

**самостоятельной работы студентов необходимо наличие персональных компьютеров с доступом в Интернет.**

<b>Наименование специальных* помещений и помещений для самостоятельной работы</b>	<b>Оснащенность специальных помещений и помещений для самостоятельной работы</b>	<b>Перечень лицензионного программного обеспечения. Реквизиты подтверждающего документа</b>
<p>Учебная аудитория для проведения занятий лекционного типа, занятий семинарского типа, курсового проектирования (выполнения курсовых работ), групповых и индивидуальных консультаций, текущего контроля и промежуточной аттестации, Учебная аудитория. Математический кабинет № 213 (170002, Тверская обл., г.Тверь, пер. Садовый, дом 35)</p>	<p>Набор учебной мебели, Меловая доска, Переносной ноутбук, Компьютер:(процессор Core i5-2400+монитор LC E2342T (10шт.) Графопроектор, мультимедийный комплект учебного класса (вариант № 1) Проектор Casio XJ-M140, кронштейн, кабель, удлинитель, настенный проекц. экран Lumien 180*180.</p>	<p>Google Chrome бесплатно Kaspersky Endpoint Security 10 для Windows Акт на передачу прав №2129 от 25 октября 2016 г PostgreSQL 9.6 бесплатно Rstudio бесплатно WinDjView бесплатно Microsoft Windows 10 Enterprise Акт приема-передачи № 369 от 21 июля 2017; MS Office 365 pro plus Акт приема-передачи № 369 от 21 июля 2017</p>
<p>Учебная аудитория для проведения занятий лекционного типа, занятий семинарского типа, курсового проектирования (выполнения курсовых работ), групповых и индивидуальных консультаций, текущего контроля и промежуточной аттестации, Учебная аудитория № 208 (170002, Тверская обл., г.Тверь, пер. Садовый, дом 35)</p>	<p>Набор учебной мебели, Меловая доска, Переносной ноутбук, Проектор EpsonEB-W41 с креплением KROMAX Projector-10 Настенный экран Lumien Eco Picture 180x180 см Matte White</p>	<p>Google Chrome бесплатно Kaspersky Endpoint Security 10 для Windows Акт на передачу прав №2129 от 25 октября 2016 г PostgreSQL 9.6 бесплатно Rstudio бесплатно WinDjView бесплатно Microsoft Windows 10 Enterprise Акт приема-передачи № 369 от 21 июля 2017; MS Office 365 pro plus Акт приема-передачи № 369 от 21 июля 2017</p>

### **VIII. Сведения об обновлении рабочей программы дисциплины**

№п.п.	Обновленный раздел рабочей программы дисциплины	Описание внесенных изменений	Реквизиты документа, утвердившего изменения
1			
2			