

Министерство науки и высшего образования Российской Федерации
федеральное государственное бюджетное образовательное учреждение высшего образования
«Новгородский государственный университет имени Ярослава Мудрого»
Институт электронных и информационных систем
Кафедра прикладной математики и информатики



УТВЕРЖДАЮ
Директор ИЭИС
Электронный
Информационный

С.И. Эминов

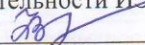
2019 г.

РАБОЧАЯ ПРОГРАММА
учебной дисциплины
Машинное обучение

по направлению подготовки (специальности)
01.04.02 Прикладная математика и информатика
Направленность (профиль) Прикладной анализ данных

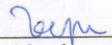
СОГЛАСОВАНО

Начальник отдела обеспечения
деятельности ИЭИС

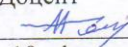
 Лысухо П.В.
« 05 » апреля 2019 г.

Разработал

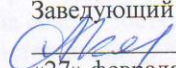
Старший преподаватель

 Гарбарь С.В.
«18» февраля 2019 г.

Доцент

 Токмачёв М.С.
«18» февраля 2019 г.

Принято на заседании кафедры
Протокол №7 от «27» февраля 2019 г.
Заведующий кафедрой

 Колногоров А.В.
«27» февраля 2019 г.

1 Цели и задачи освоения учебной дисциплины

Цели освоения учебной дисциплины «Машинное обучение» — изучение основ теории, необходимых для успешного освоения методов машинного обучения, соответствующих алгоритмов и инструментов. В рамках освоения данного модуля студенты получают знания, умения и навыки, необходимые при практическом решении задач в профессиональной деятельности, проектировании и разработке информационных систем.

Задачи УМ

- знакомство с некоторыми методами машинного обучения и анализа данных;
- получение навыков применения различных методов анализа данных и машинного обучения.

2 Место учебной дисциплины в структуре ОПОП

Учебная дисциплина относится к обязательной части учебного плана основной профессиональной образовательной программы направления подготовки 01.04.02 Прикладная математика и информатика и направленности (профилю) Прикладной анализ данных (далее – ОПОП).

В качестве входных требований выступают сформированные ранее компетенции обучающихся, приобретенные ими в рамках программы бакалавриата.

Освоение учебной дисциплины является компетентностным ресурсом для дальнейшего изучения следующих дисциплин (модулей, практик):

- Интеллектуальный анализ данных
- Подготовка к защите и защиты выпускной квалификационной работы.

3 Требования к результатам освоения учебной дисциплины

Перечень компетенций, которые формируются в процессе освоения учебной дисциплины:

Универсальные компетенции

УК-1 Способен осуществлять критический анализ проблемных ситуаций на основе системного подхода, вырабатывать стратегию действий

Общепрофессиональные компетенции:

ОПК-1 Способен решать актуальные задачи фундаментальной и прикладной математики

ОПК-2 Способен совершенствовать и реализовывать новые математические методы решения прикладных задач

Таблица 1 — Результаты освоения учебной дисциплины

<i>Код и наименование компетенции</i>	<i>Результаты освоения учебной дисциплины (индикаторы достижения компетенций)</i>		
УК-1 Способен осуществлять критический анализ проблемных ситуаций на основе системного подхода, вырабатывать стратегию действий	Знать УК-1.1 Анализирует проблемную ситуацию как систему, выявляя ее составляющие и связи между ними;	Уметь УК-1.2 Осуществляет поиск вариантов решения поставленной проблемной ситуации на основе доступных источников информации;	Владеть УК-1.3 Разрабатывает и содержательно аргументирует стратегию решения проблемной ситуации на основе системного подхода

<i>Код и наименование компетенции</i>	<i>Результаты освоения учебной дисциплины (индикаторы достижения компетенций)</i>			
ОПК-1 Способен решать актуальные задачи фундаментальной и прикладной математики	ОПК-1.1 Знает основы, современные и актуальные проблемы и задачи фундаментальной и прикладной математики;	ОПК-1.2 Умеет решать стандартные профессиональные задачи с применением знаний и методов фундаментальной и прикладной математики, математического и имитационного моделирования, методов машинного обучения;	ОПК-1.3 Владеет навыками теоретического исследования, математического и имитационного моделирования в области прикладного анализа данных	
ОПК-2 Способен совершенствовать и реализовывать новые математические методы решения прикладных задач	ОПК-2.1 Знает современные математические методы решения прикладных задач;	ОПК-2.2 Умеет совершенствовать и реализовывать новые математические методы при решении задач в профессиональной области;	ОПК-2.3 Владеет навыками выбора и применения математических методов решения задач прикладного анализа данных	

4 Структура и содержание учебной дисциплины

4.1 Трудоемкость учебной дисциплины

4.1.1 Трудоемкость учебной дисциплины для очной формы обучения:

Таблица 2 — Трудоемкость учебной дисциплины для очной формы обучения

<i>Части учебной дисциплины</i>	<i>Всего</i>	<i>Распределение по семестрам</i>
		<i>2 семестр</i>
1. Трудоемкость учебной дисциплины в зачетных единицах (ЗЕТ)	6	6
2. Контактная аудиторная работа в академических часах (АЧ)	45	45
3. Курсовая работа/курсовой проект (АЧ) (при наличии)	—	—
4. Внеаудиторная СРС в академических часах (АЧ)	171	171
5. Промежуточная аттестация (экзамен) (АЧ)	36	36

4.2 Содержание учебной дисциплины

Раздел №1. Основные модели машинного обучения.

1.1 Введение в машинное обучение. Основные понятия. Постановка задачи анализа данных. Методика анализа данных. Методы сбора и подготовки исходного набора данных. Постановка задачи обучения по прецедентам. Классификация и восстановление регрессии.

1.2 Библиотеки Python для задач машинного обучения. Введение в Python, Jupyter Notebook. Pandas. NumPy. SciPy. Matplotlib. Визуальный анализ данных. Применение различных средств визуализации. Библиотека Sklearn. Встроенные датасеты. Метрики качества. Подбор параметров по сетке.

1.3 Линейные модели в задачах регрессии и классификации. Постановка задач классификации и регрессии. Основные алгоритмы и методы решения. Градиентный спуск для линейной регрессии. Линейная классификация. Функции потерь в задачах классификации. Проблема переобучения. Регуляризация. Масштабирование признаков. Спрямоляющие пространства. Работа с категориальными признаками.

1.4 Метрические модели и метод опорных векторов. Метод k ближайших соседей.

Настройка параметров и метрики. Метод опорных векторов.

1.5 Логические модели машинного обучения. Композиции простых алгоритмов. Деревья решений. Критерии останова и стрижка деревьев. Случайные леса. Градиентный бустинг. Борьба с переобучением в градиентном бустинге. Градиентный бустинг для регрессии и классификации. Градиентный бустинг над решающими деревьями.

1.6 Нейронные сети. Однослойная нейронная сеть. Многослойная нейронная сеть.

Функция ошибки. Оптимизация параметров нейронной сети. Регуляризация и прореживание нейронной сети.

4.3 Трудоемкость разделов учебной дисциплины и контактной работы

Таблица 3 - Трудоемкость разделов учебной дисциплины

№	Наименование разделов (тем) учебной дисциплины, УЭМ, наличие КП/КР	Контактная работа (в АЧ)				Внеауд. СРС (в АЧ)	Формы текущего контроля
		Аудиторная			В т.ч. СРС		
		ЛЕК	ПЗ	ЛР			
Раздел №1. Основные модели машинного обучения							
1.1	Введение в машинное обучение. Основные понятия. Постановка задачи анализа данных. Методика анализа данных. Методы сбора и подготовки исходного набора данных. Постановка задачи обучения по прецедентам. Классификация и восстановление регрессии.	1				21	
1.2	Библиотеки Python для задач машинного обучения. Введение в Python, Jupyter Notebook. Pandas. NumPy. SciPy. Matplotlib. Визуальный анализ данных. Применение различных средств визуализации. Библиотека Sklearn. Встроенные датасеты. Метрики качества. Подбор параметров по сетке.			6	2	30	Защита лабораторной работы №1
1.3	Линейные модели в задачах регрессии и классификации. Постановка задач классификации и регрессии. Основные алгоритмы и методы решения. Градиентный спуск для линейной регрессии. Линейная классификация. Функции потерь в задачах классификации. Проблема переобучения. Регуляризация. Масштабирование признаков. Спрямоляющие пространства. Работа с категориальными признаками.	2		6	2	30	Защита лабораторной работы №2
1.4	Метрические модели и метод опорных векторов. Метод k ближайших соседей. Настройка параметров и метрики. Метод опорных векторов.	2		6	1	20	Защита лабораторной работы №3
1.5	Логические модели машинного обучения. Композиции простых алгоритмов. Деревья решений. Критерии останова и стрижка деревьев. Случайные леса. Градиентный бустинг. Борьба с переобучением в градиентном бустинге. Градиентный бустинг для регрессии и классификации. Градиентный бустинг над решающими	2		12	2	40	Защита лабораторной работы №4 Защита лабораторной работы №5 Защита лабораторной

	деревьями.						работы №6
1.6	Нейронные сети. Однослойная нейронная сеть. Многослойная нейронная сеть. Функция ошибки. Оптимизация параметров нейронной сети. Регуляризация и прореживание нейронной сети.	2		6	2	30	Защита лабораторной работы №7
	Промежуточная аттестация	Экзамен, 36 часов					
	ИТОГО	9		36		171	

4.4 Лабораторные работы и курсовые работы/курсовые проекты

4.4.1 Перечень тем лабораторных работ:

- Первичный анализ данных (Лабораторная работа №1)
- Линейная регрессия, регуляризация (Лабораторная работа №2)
- Метрические методы классификации, кросс-валидация, подбор параметров (Лабораторная работа №3)
- Случайные леса (Лабораторная работа №4)
- Градиентный бустинг (Лабораторная работа №5)
- Анализ текстов (Лабораторная работа №6)
- Нейронные сети (Лабораторная работа №7)

Курсовые работы/курсовые проекты не предусмотрены учебным планом.

5 Методические рекомендации по организации освоения учебной дисциплины

Таблица 4 - Методические рекомендации по организации лекций

№	Темы лекционных занятий (форма проведения)	Трудоемкость в АЧ
1.	Введение в машинное обучение. Основные понятия. Постановка задачи анализа данных (<i>информационная лекция</i>)	1
2.	Линейные модели в задачах регрессии и классификации (<i>проблемная лекция</i>)	2
3.	Метрические модели и метод опорных векторов (<i>проблемная лекция</i>)	2
4.	Логические модели машинного обучения. Композиции простых алгоритмов (<i>проблемная лекция</i>)	2
5.	Нейронные сети (<i>проблемная лекция</i>)	2
	ИТОГО	18

6 Фонд оценочных средств учебной дисциплины

Фонд оценочных средств представлен в Приложении А.

7 Условия освоения учебной дисциплины

7.1 Учебно-методическое обеспечение

Учебно-методического обеспечение учебной дисциплины представлено в Приложении Б.

7.2 Материально-техническое обеспечение

Таблица 5 — Материально-техническое обеспечение учебной дисциплины

№	Требование к материально-техническому обеспечению	Наличие материально-технического оборудования и программного обеспечения
1.	Наличие специальной аудитории	Компьютерный класс с компьютерами под управлением ОС Windows
2.	Мультимедийное оборудование	Проектор, компьютер, экран, интерактивная доска
3.	Программное обеспечение	Пакет Anaconda

Приложение А (обязательное)
Фонд оценочных средств учебной дисциплины
Машинное обучение

1 Структура фонда оценочных средств

Фонд оценочных средств состоит из двух частей:

а) открытая часть — общая информация об оценочных средствах (название оценочных средств, проверяемые компетенции, баллы, количество вариантов заданий, методические рекомендации для применения оценочных средств и пр.), которая представлена в данном документе, а также те вопросы и задания, которые могут быть доступны для обучающегося;

б) закрытая часть — фонд вопросов и заданий, которая не может быть заранее доступна для обучающихся (экзаменационные билеты, вопросы к контрольной работе и пр.) и которая хранится на кафедре.

2 Перечень оценочных средств текущего контроля и форм промежуточной аттестации

Таблица А.1 — Перечень оценочных средств

№	Оценочные средства для текущего контроля	Разделы (темы) учебной дисциплины	Баллы	Проверяемые компетенции
1.	Защита лабораторной работы №1	Тема 1.2	35	УК-1, ОПК-1, ОПК-2
2.	Защита лабораторной работы №2	Тема 1.3	35	УК-1, ОПК-1, ОПК-2
3.	Защита лабораторной работы №3	Тема 1.4	35	УК-1, ОПК-1, ОПК-2
4.	Защита лабораторной работы №4	Тема 1.5	35	УК-1, ОПК-1, ОПК-2
5.	Защита лабораторной работы №5	Тема 1.5	35	УК-1, ОПК-1, ОПК-2
6.	Защита лабораторной работы №6	Тема 1.5	35	УК-1, ОПК-1, ОПК-2
7.	Защита лабораторной работы №7	Тема 1.6	35	УК-1, ОПК-1, ОПК-2
<i>Промежуточная аттестация</i>				
	Экзамен		50	
	ИТОГО		300	

3 Рекомендации к использованию оценочных средств

Таблица А.2 — Защита лабораторной работы №1

Критерии оценки	Количество вариантов заданий
В соответствии с методическими указаниями	10

Таблица А.3 — Защита лабораторной работы №2

Критерии оценки	Количество вариантов заданий
В соответствии с методическими указаниями	10

Таблица А.4 — Защита лабораторной работы №3

Критерии оценки	Количество вариантов заданий
В соответствии с методическими указаниями	10

Таблица А.5 — Защита лабораторной работы №4

Критерии оценки	Количество вариантов заданий
В соответствии с методическими указаниями	10

Таблица А.6 — Защита лабораторной работы №5

Критерии оценки	Количество вариантов заданий
В соответствии с методическими указаниями	10

Таблица А.7 — Защита лабораторной работы №6

<i>Критерии оценки</i>	<i>Количество вариантов заданий</i>
В соответствии с методическими указаниями	10

Таблица А.8 — Защита лабораторной работы №7

<i>Критерии оценки</i>	<i>Количество вариантов заданий</i>
В соответствии с методическими указаниями	10

Таблица А.9 — Экзамен №7

<i>Критерии оценки</i>	<i>Количество вариантов заданий</i>
Правильность ответов на вопросы билета	15

Пример вопросов экзаменационного билета:**НовГУ****ИЭИС****1. Подбор параметров по сетке****2. Случайные леса**

Все материалы для проведения промежуточного контроля хранятся на кафедре.

Приложение Б
(обязательное)

**Карта учебно-методического обеспечения
Учебной дисциплины Машинное обучение**

Таблица Б.1 – Основная литература

Библиографическое описание издания (автор, наименование, вид, место и год издания, кол. стр.)	Кол. экз. в библ. НовГУ	Наличие в ЭБС
Печатные источники		
Чубукова И. А. Data Mining : учеб. пособие / И. А. Чубукова. - 2-е изд., испр. - М. : Интернет-университет Информ. Технологий : БИНОМ. Лаборатория знаний, 2008. - 382 с.	12	
Дюк В. Data Mining / В. Дюк, А. Самойленко. - СПб. : Питер, 2001. - 366 с.	15	

Таблица Б.2 – Дополнительная литература

Библиографическое описание издания (автор, наименование, вид, место и год издания, кол. стр.)	Кол. экз. в библ. НовГУ	Наличие в ЭБС
Печатные источники		
Айвазян С. А. Прикладная статистика и основы эконометрики = Applied statistics and essentials of econometrics : учеб. для вузов / С. А. Айвазян, В. С. Мхитарян ; Гос. ун-т Высш. шк. экономики. - М. : ЮНИТИ, 1998. - 1022 с.	3	
Электронные ресурсы		
MachineLearning.ru — Профессиональный информационно-аналитический ресурс, посвященный машинному обучению, распознаванию образов и интеллектуальному анализу данных. — Режим доступа: http://www.machinelearning.ru		
2		

Зав. кафедрой _____
подпись И.О.Фамилия

«_____» _____ 20__ г.

Рабочая программа актуализирована на 20__/20__ учебный год.
 Протокол № ____ заседания кафедры от «____» _____ 20__ г.
 Разработчик: _____
 Зав. кафедрой _____

Рабочая программа актуализирована на 20__ / 20__ учебный год.
 Протокол № ____ заседания кафедры от « ____ » _____ 20__ г.
 Разработчик: _____
 Зав. кафедрой _____

Рабочая программа актуализирована на 20__/20__ учебный год.
 Протокол № ____ заседания кафедры от «____» _____ 20__ г.
 Разработчик: _____
 Зав. кафедрой _____

Таблица В.1 — Перечень изменений, внесенных в рабочую программу:

[illegible]