Государственное бюджетное образовательное учреждение
высшего образования Московской области
«Университет «Дубна»
(государственный университет «Дубна»)

Институт системного анализа и управления
Кафедра распределенных информационных вычислительных систем

УТВЕРЖДАЮ
Проректор по учебно-
методической работе
_____ / Деникин А.С./
«15» 06 2021 г.

Рабочая программа дисциплины

Машинное обучение и анализ данных

Направление подготовки (специальность)
01.03.02 Прикладная математика и информатика

Уровень высшего образования
Бакалавриат

Направленность (профиль) программы (специализация)
Математическое моделирование

Форма обучения
очная

Дубна, 2021

Преподаватель:
Доктор физ.-мат наук Ососков Г.А.
кафедра распределённых информационно-вычислительных систем

подпись

Рабочая программа разработана в соответствии с требованиями ФГОС ВО по направлению подготовки высшего образования **01.03.02 Прикладная математика и информатика**

Программа рассмотрена на заседании кафедры **распределенных информационных вычислительных систем**

Протокол заседания № 12 от «11» июня 2021 г.

Заведующий кафедрой д.т.н. профессор Кореньков В.В.

подпись

СОГЛАСОВАНО

Заведующий выпускающей кафедрой распределённых информационно-вычислительных систем Кореньков В.В.

подпись

Эксперт (рецензент):

Помощник директора лаборатории информационных технологий имени М.Г. Мещерякова Объединенного института ядерных исследований по международному сотрудничеству и работе с кадрами, к.ф.-м.н., с.н.с., Айрян Э.А.

подпись

Содержание

1	Цели и задачи освоения дисциплины	4
2	Место дисциплины в структуре ОПОП ВО	4
3	Планируемые результаты обучения по дисциплине (модулю)	4
4	Объем дисциплины (модуля)	6
5	Содержание дисциплины (модуля)	7
6	Перечень учебно-методического обеспечения по дисциплине (модулю)	9
7	Фонд оценочных средств по дисциплине (модулю)	9
8	Ресурсное обеспечение	9
Приложение. Фонд оценочных средств		13

1 Цели и задачи освоения дисциплины

Цель дисциплины — приобретение студентами теоретических знаний и устойчивых навыков практической работы со средствами и технологиями машинного обучения.

В задачи дисциплины входит подготовка студентов в области машинного обучения и обеспечение их профессиональными знаниями для системного подхода к созданию программ глубокого обучения на языке Python с использованием библиотек Keras-Tensorflow и PyTorch и современных методов распараллеливания алгоритмов.

2 Место дисциплины в структуре ОПОП ВО

Дисциплина «Машинное обучение и анализ данных» относится к части образовательной программы, формируемой участниками образовательных отношений, является обязательной дисциплиной по направлению 01.03.02 Прикладная математика и информатика. Дисциплина преподается в 6 семестре, на 3 курсе. Форма промежуточной аттестации – зачет.

Перечень дисциплин с указанием разделов (тем), предшествующих дисциплине «Прикладная математика и информатика»:

Семестр	Дисциплина	Разделы
1,2	Математический анализ	Пределы, непрерывность функции, выпуклость, ряды, сходимость,
		Производные, численные методы вычисления производных; функции многих переменных, частные производные.
1,2	Линейная алгебра и аналитическая геометрия	Матрицы и определители, методы их нахождения, системы линейных уравнений, собственные векторы
		Виды уравнений прямой, кривые второго порядка
3	Теория вероятностей и математической статистики	классическая вероятность, условная вероятность и независимость, теорема Байеса и основы байесовского подхода, непрерывные распределения, свертка распределений, математическое ожидание, дисперсия, ковариация, моменты высших порядков, условное математическое ожидание и условное распределение.

3 Планируемые результаты обучения по дисциплине (модулю)

После изучения данной дисциплины студенты приобретают знания, умения и владения (навыки), соответствующие результатам основной профессиональной образовательной программы.

Формируемые компетенции *(код и наименование)*	**Индикаторы достижения компетенций** *(код и формулировка)*	**Планируемые результаты обучения по дисциплине (модулю)**
УК-1 Способен осуществлять поиск, критический анализ и синтез информации, применять системный подход для решения поставленных задач	УК-1.1. Анализирует задачу, выделяя ее базовые составляющие	Знать специфику системного подхода
		Знать специфику логических методов анализа и синтеза
		Уметь проводить анализ задачи, выделять ее базовые составляющие и формулировать результаты, которые необходимо достигнуть
		Уметь применять системный подход для анализа и решения поставленных задач
	УК-1.2. Определяет и ранжирует информацию, требуемую для решения поставленной задачи	Владеть навыками работы с научной и учебной литературой
		Владеть методами критического анализа и синтеза информации
ПК-1 Способен создавать и исследовать новые математические модели в естественных науках, промышленности и бизнесе, с учетом возможностей современных информационных технологий, программирования и компьютерной техники	ПК-1.1. Использует математические модели в естественных науках, промышленности и бизнесе	Знать классические математические модели, применяемые в естественных науках, промышленности и бизнесе
		Уметь использовать математические модели в профессиональной деятельности
		Владеть навыками разработки и исследования новых математических моделей в естествознании, промышленности и бизнесе
	ПК-1.2. Исследует математические модели с применением информационных технологи, программирования и компьютерной техники	Уметь создавать, исследовать и анализировать математические модели, применяемые в естественных науках, промышленности и бизнесе
		Владеть практическим опытом применения языков программирования и пакетов прикладных программ для проведения математического моделирования при помощи компьютерной техники
ПК-2 Способен использовать современные методы разработки и реализации алгоритмов математических моделей на базе языков программирования и пакетов	ПК-2.1. Применяет современные методы разработки и реализации алгоритмов математических	Знать современные методы разработки и реализации алгоритмов математических моделей на базе языков программирования и пакетов прикладных программ моделирования

прикладных программ моделирования	моделей на базе языков программирования и пакетов прикладных программ для решения задач профессиональной деятельности	Уметь разрабатывать и реализовывать алгоритмы математических моделей на базе языков программирования и пакетов прикладных программ моделирования
		Владеть практическим опытом разработки и реализации алгоритмов на базе языков программирования и пакетов прикладных программ моделирования для решения задач профессиональной деятельности
ПК-4 Способен демонстрировать базовые знания математических и естественных наук, программирования и информационных технологий	ПК-4.1. Обладает базовыми знаниями в области математических, естественных наук, программирования и информационных технологий, применяет их в профессиональной деятельности	Знать основные понятия математического аппарата и информационных технологий
		Уметь находить, формулировать и решать стандартные задачи в собственной научно-исследовательской деятельности в математике и информатике
		Владеть навыками научно-исследовательской деятельности в математике и информатике
ПК-5 Способен работать в составе научно-исследовательского и производственного коллектива, готовить презентации, оформлять научно-технические отчеты по результатам выполненной работы	ПК-5.1. Решает задачи профессиональной деятельности в составе научно-исследовательского и производственного коллектива	Знать правила работы в составе научно-исследовательского и производственного коллектива, методы и способы решения задач профессиональной деятельности
		Уметь работать в составе научно-исследовательского и производственного коллектива и решать задачи профессиональной деятельности
		Владеть навыками работы в составе научно-исследовательского и производственного коллектива, алгоритмами решения задач профессиональной деятельности
	ПК-5.2. Представляет результаты работы в виде презентаций, научно-технических отчетов, статей и докладов	Уметь в составе команды разработчиков участвовать в процессах разработки программного обеспечения

4 Объем дисциплины (модуля)

Объем дисциплины составляет 3 зачетных единицы, всего 108 академических часа.

5 Содержание дисциплины (модуля)

Очная **форма обучения**

Наименование разделов и тем дисциплины (модуля)	**Всего** (академ. часы)	в том числе:						
		Контактная работа (работа во взаимодействии с преподавателем)						Самостоятельная работа обучающегося
		Лекции	Практические (семинарские) занятия	Лабораторные занятия	КРП	…	**Всего**	
6 семестр								
Наука о данных: искусственный интеллект, машинное и глубокое обучение и прикладные задачи.	12	2	4				6	6
Принципы работы искусственных нейронных сетей (ИНС). Типы ИНС Метод обратного распространения ошибки	12	2	4				6	6
Проблемы обучения глубоких ИНС и методы их решения для различных типов глубоких ИНС. Многослойные персептроны, автоэнкодеры, сверточные и рекуррентные ИНС.	12	2	4				6	6
LSTM сети с долгой краткосрочной памятью, нейросети-трансформеры, обучение с подкреплением, подходы к работе с естественными языками.	16	2	6				8	8
Примеры с применением ИНС в практике ОИЯИ, для идентификации частиц, предсказания временных рядов	12	2	2				4	8
Применение прямоточных и сверточных автоэнкодеров для понижения размерности обучающей выборки и шумоподавления.	16	2	6				8	8
Методы обучения глубоких ИНС в условиях малой обучающей выборки.	16	2	6				8	8
Методы глубокого трекинга в экспериментальной физике высоких энергий	12	2	2				4	8
Промежуточная аттестация: зачет		X						
Итого за семестр	***108***	**16**	**34**				**50**	**58**

Содержание дисциплины (модуля)

№	Содержание раздела
Раздел 1	Введение. Эра Больших Данных и необходимость новых подходов, основанных на применении искусственного интеллекта. Соотношение ИИ, машинного и глубокого обучения. Методы машинного обучения и его роль в хранении и распределенной обработке данных экспериментальной физики и других областей научных исследований (биологии, медицины, экологии и др.).
Раздел 2	Принципы работы искусственных нейронных сетей (ИНС). Биологические предпосылки, описание искусственного нейрона и схемы объединения нейронов в ИНС. Типы ИНС по способам их обучения. Метод обратного распространения ошибки и его математическое описание. Проблемы и способы и минимизации функции ошибки ИНС. Проблема переобучения, метрики качества обучения, кросс-валидация для выбора параметров ИНС
Раздел 3	Экономические и технологические предпосылки появления глубокого обучения. Проблемы обучения глубоких ИНС и методы их решения для различных типов глубоких ИНС. Многослойные персептроны, автоэнкодеры, Сверточные ИНС, Рекуррентные ИНС, LSTM сети с долгой краткосрочной памятью, нейросети-трансформеры, обучение с подкреплением, подходы к работе с естественными языками.
Раздел 4	Примеры с применением ИНС в практике ОИЯИ. Применение прямоточных ИНС с одним скрытым слоем для идентификации частиц в эксперименте СВМ. Применение авторегрессионых нейросетей для предсказания временных рядов. Применение прямоточных и сверточных автоэнкодеров для понижения размерности обучающей выборки и шумоподавления. Применение глубокого обучения для моделирование и предсказание атмосферных загрязнений.
Раздел 5	Примеры с применением ИНС в практике ОИЯИ. Применение прямоточных ИНС с одним скрытым слоем для идентификации частиц в эксперименте СВМ. Применение авторегрессионых нейросетей для предсказания временных рядов. Применение прямоточных и сверточных автоэнкодеров для понижения размерности обучающей выборки и шумоподавления. Применение глубокого обучения для моделирование и предсказание атмосферных загрязнений.
Раздел 6	Методы обучения глубоких ИНС в условиях малой обучающей выборки. Способы аугментации обучающей выборки. Метод переноса обучения. ИНС со специальной структурой для обучения с одной попытки, сиамские нейросети, применение составной (триплетной) функции ошибки сети.
Раздел 7	Методы глубокого трекинга в экспериментальной физике высоких энергий. Локальный и глобальный подходы. Примеры реконструкции траекторий элементарных частиц в различных экспериментах в фиксированной мишенью и коллайдерного типа

Практическая подготовка при изучении дисциплины реализуется непосредственно в университете.

6 Перечень учебно-методического обеспечения по дисциплине (модулю)

Для обеспечения реализации программы дисциплины (модуля) разработаны:

– методические материалы к лекционным и практическим (семинарским) занятиям;
– методические рекомендации для преподавателя;
– методические материалы по организации самостоятельной работы обучающихся;
– методическое обеспечение инновационных форм учебных занятий и проч.
– методические материалы по организации изучения дисциплины (модуля) с применением электронного обучения, дистанционных образовательных технологий;
– методические рекомендации для обучающихся с ограниченными возможностями здоровья и инвалидов по освоению программы дисциплины (модуля).

Методические материалы по дисциплине (модулю) и образовательной программе в целом представлены на официальном сайте образовательной организации (раздел «Сведения об образовательной организации» - Образование - Образовательные программы).

7 Фонд оценочных средств по дисциплине (модулю)

Для аттестации обучающихся на соответствие их персональных достижений поэтапным требованиям образовательной программы по дисциплине (модулю) разработаны фонды оценочных средств, позволяющие оценить результаты обучения (знания, умения, навыки) и сформированные (формируемые) компетенции. Эти фонды включают теоретические вопросы, типовые практические задания, контрольные работы, домашние работы и иные оценочные материалы, используемые при проведении процедур текущего контроля успеваемости и промежуточной аттестации.

Фонды оценочных средств представлены в приложении к рабочей программе.

При необходимости обучающиеся с ограниченными возможностями здоровья и инвалиды обеспечиваются оценочными материалами в формах, адаптированных к ограничениям их здоровья и восприятия информации:

Для лиц с нарушениями зрения:

– в печатной форме увеличенным шрифтом,
– в форме электронного документа,
– в форме аудиофайла,
– в печатной форме на языке Брайля.

Для лиц с нарушениями слуха:

– в печатной форме,
– в форме электронного документа.

Для лиц с нарушениями опорно-двигательного аппарата:

– в печатной форме,
– в форме электронного документа,
– в форме аудиофайла

8 Ресурсное обеспечение

Перечень литературы

Основная литература

1. *Плас Дж. В.* Python для сложных задач: наука о данных и машинное обучение / Плас Дж.Вандер. - СПб.: Питер, 2018. - 576 с.: ил. - (Бестселлеры O'Reilly). - ISBN 978-5-496-03068-7.

2. *Флах П.* Машинное обучение. Наука и искусство построения алгоритмов, которые извлекают знания из данных / Флах Петер; перевод с английского А. А. Слинкина. - М.: ДМК Пресс, 2015. - 400 с.: ил. - Библиогр.:с.376.-Предм.указ.:с.387. - ISBN 978-1-107-09639-4.
3. *Николенко С. И.* Глубокое обучение. Погружение в мир нейронных сетей / Николенко С. И., Кадурин А. А., Архангельская Е. О. - СПб.: Питер, 2020. - 480 с.: ил. - (Библиотека программиста). - ISBN 978-5-4461-1537-2.
4. *Шолле Ф.* Глубокое обучение на Python / Шолле Франсуа. - СПб.: Питер, 2018. - 400 с.: ил. - (Библиотека программиста). - ISBN 978-5-4461-0770-4.
5. The official home of the Python Programming Language: https://www.python.org/
6. *М. Саммерфилд.* Программирование на Python 3. Подробное руководство. -СПБ.: Символ-Плюс. 2015.-608 с.
7. Профессиональный информационно-аналитический ресурс, посвященный машинному обучению, распознаванию образов и интеллектуальному анализу данных: http://www.machinelearning.ru
8. Айвазян С.В., Мхитарян В.С. Прикладная статистика в задачах и упражнениях. –М.: ЮНИТИ-ДАНА . 2001. -270 с.
9. Библиотека научных вычислений: http://www.numpy.org/
10. Python Data Analysis Library: https://pandas.pydata.org/
11. Библиотека scikit-learn: https://scikit-learn.org/stable/
12. Keras: The Python Deep Learning library: https://keras.io/

Дополнительная литература

1. Траск Э. Грокаем глубокое обучение / Траск Эндрю. - СПб.: Питер, 2019. - 352 с.: ил. - (Библиотека программиста). - ISBN 978-5-4461-1334-7.
2. Силен Д. Основы Data Science и Big Data. Python и наука о данных / Силен Дэви, Мейсман Арно, Али Мохамед. - СПб.: Питер, 2018. - 338 с.: ил. - (Библиотека программиста). - ISBN 9785496025171.
3. Шарден Б. Крупномасштабное машинное обучение вместе с Python: учитесь быстро развивать мощные модели машинного обучения и развертывать крупномасштабные приложения прогнозирования / Шарден Бастиан, Массарон Лука, Боскетти Альберто; перевод с английского А. В. Логунова. - М.: ДМК Пресс, 2018. - 358 с.: ил. - Предм.указ.:с.350. - ISBN 978-5-97060-506-6.
4. Брюс П. Практическая статистика для специалистов Data Science. 50 важнейших понятий / Брюс Питер, Брюс Эндрю. - СПб.: БХВ-Петербург, 2018. - 304 с.: ил. - Библиогр.:с.295.-Предм.указ.:с.297. - ISBN 978-5-9775-3974-6.
5. Вьюгин В. В. Математические основы машинного обучения и прогнозирования / Вьюгин Владимир Вячеславович; МФТИ. Лаборатория структурных методов анализа данных в предсказательном моделировании (ПреМоЛаб); Институт проблем передачи информации им. А. А. Харкевича РАН. - М.: МЦНМО, 2013. - 304 с. - Лит.:с.301. - ISBN 978-5-4439-0111-4.
6. *Джонс М.Т.* Программирование искусственного интеллекта в приложениях. Пер.с англ. А.И.Осипова. - М.: Пресс, 2006. - 312с. ISBN 5-94074-275-0. - ISBN 1-58450-278-9.

Профессиональные базы данных и информационные справочные системы
Электронно-библиотечные системы и базы данных

1. Электронно-библиотечная система «Лань» http://e.lanbook.com/
2. Электронно-библиотечная система «ЮРАЙТ». biblio-online.ru

3. Электронно-библиотечная система (ЭБС) «Университетская библиотека онлайн» www.biblioclub.ru
4. Электронно-библиотечная система Znanium.com https://new.znanium.com/

Профессиональные ресурсы сети «Интернет»

1. Электронные ресурсы издательства «Elsevier» на платформе «ScienceDirect» www.sciencedirect.com
2. Базы данных компании EBSCO Publishing: http://search.ebscohost.com/
3. БД российских научных журналов на Elibrary.ru (РУНЭБ): http://elibrary.ru/defaultx.asp.
4. http://www.scopus.com/home.url
5. Web of Science webofknowledge.com
6. Национальная электронная библиотека (НЭБ) http://нэб.рф/

Необходимое программное обеспечение

Практические занятия будут проводится на гетерогенной вычислительной платформе HybriLIT (http://hlit.jinr.ru/) в рамках экосистемы для ML/DL.

Необходимое материально-техническое обеспечение

Специализированный компьютерный класс (например, ауд. 1-307, 1-321, 1-322, 1-318, 1-211, 1-219, 1-215), подключенный к сети Интернет и к локальной сети университета, обеспечивающей доступ к программному обеспечению и ЛМС системы MOODLE для проведения семинарских занятий.

Обучающиеся из числа инвалидов и лиц с ограниченными возможностями здоровья могут использовать специализированное программное и материально-техническое обеспечение:

– обучающиеся с нарушениями опорно-двигательного аппарата при необходимости могут использовать адаптивные технические средства: функцию «сенсорная клавиатура», «управление указателем мыши с клавиатуры», специально оборудованные джойстики, увеличенные выносные кнопки, клавиатуры с большими клавишами или накладки «Клавита»;

– обучающиеся с ограничениями по зрению могут прослушать доступный аудиоматериал или прочитать тексты, увеличив шрифт на экране монитора компьютера. Рекомендуется использовать экранную лупу и другие визуальные вспомогательные средства, чтобы изменить шрифт текста, межстрочный интервал, синхронизацию с речью и т.д., программы экранного доступа (скринридеры для прочтения текстовой информации через синтезированную речь) и/или включить функцию «экранного диктора» на персональном компьютере с операционной системой Windows 7, 8, 10, Vista, XP. Студенты с полным отсутствием зрения могут использовать тексты, напечатанные шрифтом Брайля, а для набора текста на компьютере – клавиатуры Брайля;

– обучающиеся с ограничениями по слуху могут воспользоваться индивидуальными техническими средствами (аппараты «Глобус», «Монолог», индивидуальными слуховыми аппаратами, компьютерной аудиогарнитурой, наушниками и др.) при прослушивании необходимой информации, а также услугами сурдопереводчика.

При необходимости обучающиеся с ограниченными возможностями здоровья и инвалиды обеспечиваются печатными и (или) электронными образовательными ресурсами (образовательная программа, учебники, учебные пособия и др.) в формах, адаптированных к ограничениям их здоровья и восприятия информации:

Для лиц с нарушениями зрения:

– в печатной форме увеличенным шрифтом,
– в форме электронного документа,
– в форме аудиофайла,
– в печатной форме на языке Брайля.

Для лиц с нарушениями слуха:

– в печатной форме,
– в форме электронного документа.

Для лиц с нарушениями опорно-двигательного аппарата:

– в печатной форме,
– в форме электронного документа,
– в форме аудиофайла.

Фонд оценочных средств по дисциплине (модулю)

Машинное обучение и анализ данных

Направление подготовки (специальность)

01.03.02 Прикладная математика и информатика

Уровень высшего образования

Бакалавриат

Направленность (профиль) программы (специализация)

Математическое моделирование

Форма обучения

очная

Для 2021 года набора

Перечень компетенций выпускников образовательной программы с указанием результатов обучения (знаний, умений, владений), характеризующих этапы их формирования, описание показателей и критериев оценивания компетенций на различных этапах их формирования

Полный перечень компетенций выпускников образовательной программы 01.03.02 Прикладная математика и информатика (профиль Математическое моделирование) с указанием результатов обучения (знаний, умений, владений), характеризующих этапы их формирования, описание показателей и критериев оценивания компетенций на различных этапах их формирования приведен в картах компетенций образовательной программы.

Перечень компетенций выпускников образовательной программы 01.03.02 Прикладная математика и информатика (профиль Математическое моделирование), в формировании которых участвует данная дисциплина представлен в разделе 3 рабочей программы дисциплины.

Указание результатов обучения (знаний, умений, владений), характеризующих этапы формирования компетенций, описание показателей и критериев оценивания компетенций на различных этапах их формирования представлено ниже.

Описание шкал оценивания

Критерии оценивания ответов студентов на зачете:

Зачтено	Студент показывает высокий уровень теоретических знаний по изучаемой дисциплине, выполнены все задания (лабораторные работы)
Не зачтено	Студент показывает слабый уровень теоретических знаний и не выполнил задания (лабораторные работы).

Описание шкал оценивания для различных заданий, выполняемых в рамках текущего контроля, представлено в методических материалах, определяющих процедуры оценивания знаний, умений, навыков и (или) опыта деятельности, характеризующих этапы формирования компетенций.

Критерии и процедуры оценивания результатов обучения по дисциплине, характеризующих этапы формирования компетенций

УК-1. Способен осуществлять поиск, критический анализ и синтез информации, применять системный подход для решения поставленных задач

РЕЗУЛЬТАТ ОБУЧЕНИЯ по дисциплине (модулю)	**КРИТЕРИИ ОЦЕНИВАНИЯ РЕЗУЛЬТАТА ОБУЧЕНИЯ по дисциплине (модулю) ШКАЛА оценивания**		**ПРОЦЕДУРЫ ОЦЕНИВАНИЯ**
	не зачтено	**зачтено**	
Знать специфику системного подхода	Отсутствие знаний	Демонстрирует свободное и уверенное знание предметной области. Не допускает ошибок.	*Устное собеседование*
Знать специфику логических методов анализа и синтеза	Отсутствие знаний	Демонстрирует свободное и уверенное знание предметной области. Не допускает ошибок.	*Устное собеседование*
Уметь проводить анализ задачи, выделять ее базовые составляющие и формулировать результаты, которые необходимо достигнуть	Отсутствие умений	Демонстрирует свободное и уверенное умение. Не допускает ошибок.	*Выполнение практического контрольного задания*
Уметь применять системный подход для анализа и решения поставленных задач	Отсутствие умений	Демонстрирует свободное и уверенное умение. Не допускает ошибок.	*Выполнение практического контрольного задания*
Владеть навыками работы с научной и учебной литературой	Отсутствие умений	Демонстрирует свободное и уверенное владение. Не допускает ошибок.	*Выполнение практического контрольного задания*
Владеть методами критического анализа и синтеза информации	Отсутствие умений	Демонстрирует свободное и уверенное владение. Не допускает ошибок.	*Выполнение практического контрольного задания*

ПК-1 Способен создавать и исследовать новые математические модели в естественных науках, промышленности и бизнесе, с учетом возможностей современных информационных технологий, программирования и компьютерной техники

РЕЗУЛЬТАТ ОБУЧЕНИЯ по дисциплине (модулю)	**КРИТЕРИИ ОЦЕНИВАНИЯ РЕЗУЛЬТАТА ОБУЧЕНИЯ по дисциплине (модулю) ШКАЛА оценивания**		**ПРОЦЕДУРЫ ОЦЕНИВАНИЯ**
	не зачтено	**зачтено**	
Знать классические математические модели, применяемые в естественных науках, промышленности и бизнесе	Отсутствие знаний	Демонстрирует свободное и уверенное знание.	*Устное собеседование*
Уметь использовать математические модели в профессиональной деятельности	Отсутствие умений	Демонстрирует свободное и уверенное умение.	*Выполнение практического контрольного задания*
Владеть навыками разработки и исследования новых математических моделей в естествознании, промышленности и бизнесе	Отсутствие владений	Демонстрирует свободное и уверенное владение.	*Выполнение практического контрольного задания*
Уметь создавать, исследовать и анализировать математические модели, применяемые в естественных науках, промышленности и бизнесе	Отсутствие умений	Демонстрирует свободное и уверенное умение.	*Выполнение практического контрольного задания*
Владеть практическим опытом применения языков программирования и пакетов прикладных программ для проведения математического моделирования при помощи компьютерной техники	Отсутствие владений	Демонстрирует свободное и уверенное владение.	*Выполнение практического контрольного задания*

ПК-2 Способен использовать современные методы разработки и реализации алгоритмов математических моделей на базе языков программирования и пакетов прикладных программ моделирования

РЕЗУЛЬТАТ ОБУЧЕНИЯ по дисциплине (модулю)	**КРИТЕРИИ ОЦЕНИВАНИЯ РЕЗУЛЬТАТА ОБУЧЕНИЯ по дисциплине (модулю) ШКАЛА оценивания**		**ПРОЦЕДУРЫ ОЦЕНИВАНИЯ**
	не зачтено	**зачтено**	
Знать современные методы разработки и реализации алгоритмов математических моделей на базе языков программирования и пакетов прикладных программ моделирования	Отсутствие знаний	Демонстрирует свободное и уверенное знание.	*Устное собеседование*
Уметь разрабатывать и реализовывать алгоритмы математических моделей на базе языков программирования и пакетов прикладных программ моделирования	Отсутствие умений	Демонстрирует свободное и уверенное умение.	*Выполнение практического контрольного задания*
Владеть практическим опытом разработки и реализации алгоритмов на базе языков программирования и пакетов прикладных программ моделирования для решения задач профессиональной деятельности	Отсутствие владений	Демонстрирует свободное и уверенное владение.	*Выполнение практического контрольного задания*

ПК-4 Способен демонстрировать базовые знания математических и естественных наук, программирования и информационных технологий

РЕЗУЛЬТАТ ОБУЧЕНИЯ по дисциплине (модулю)	**КРИТЕРИИ ОЦЕНИВАНИЯ РЕЗУЛЬТАТА ОБУЧЕНИЯ по дисциплине (модулю) ШКАЛА оценивания**		**ПРОЦЕДУРЫ ОЦЕНИВАНИЯ**
	не зачтено	**зачтено**	
Знать основные понятия математического аппарата и информационных технологий	Отсутствие знаний	Демонстрирует свободное и уверенное знание.	*Устное собеседование*

Уметь находить, формулировать и решать стандартные задачи в собственной научно-исследовательской деятельности в математике и информатике	Отсутствие умений	Демонстрирует свободное и уверенное умение.	*Выполнение практического контрольного задания*
Владеть навыками научно-исследовательской деятельности в математике и информатике	Отсутствие владений	Демонстрирует свободное и уверенное владение.	*Выполнение практического контрольного задания*

ПК-5 Способен работать в составе научно-исследовательского и производственного коллектива, готовить презентации, оформлять научно-технические отчеты по результатам выполненной работы

РЕЗУЛЬТАТ ОБУЧЕНИЯ по дисциплине (модулю)	**КРИТЕРИИ ОЦЕНИВАНИЯ РЕЗУЛЬТАТА ОБУЧЕНИЯ по дисциплине (модулю) ШКАЛА оценивания**		**ПРОЦЕДУРЫ ОЦЕНИВАНИЯ**
	не зачтено	**зачтено**	
Знать правила работы в составе научно-исследовательского и производственного коллектива, методы и способы решения задач профессиональной деятельности	Отсутствие знаний	Демонстрирует свободное и уверенное знание.	*Устное собеседование*
Уметь работать в составе научно-исследовательского и производственного коллектива и решать задачи профессиональной деятельности	Отсутствие умений	Демонстрирует свободное и уверенное умение.	*Выполнение практического контрольного задания*
Владеть навыками работы в составе научно-исследовательского и производственного коллектива, алгоритмами решения задач профессиональной деятельности	Отсутствие владений	Демонстрирует свободное и уверенное владение.	*Выполнение практического контрольного задания*
Уметь в составе команды разработчиков участвовать в процессах разработки программного обеспечения	Отсутствие умений	Демонстрирует свободное и уверенное умение.	*Выполнение практического контрольного задания*

Типовые контрольные задания или иные материалы, необходимые для оценки результатов обучения, характеризующих этапы формирования компетенций и (или) для итогового контроля сформированности компетенции

Перечень вопросов к зачету

№	Вопрос	Индикаторы достижения компетенции	Результаты обучения
1	Привести классификацию задач машинного обучения, примеры.	УК-1.1	Знает специфику системного подхода
2	Дать краткое описание особенностей задач, решаемых методами машинного обучения.	УК-1.1	Знает специфику логических методов анализа и синтеза
3	Дайте краткое описание возможностей библиотеки Scikit-Learn.	ПК-2.1	Умеет разрабатывать и реализовывать алгоритмы математических моделей на базе языков программирования и пакетов прикладных программ моделирования
4	Дайте краткую характеристику библиотек для предобработки данных, визуализации данных и построения моделей, их обучения и анализа полученных результатов.	УК-1.1	Умеет применять системный подход для анализа и решения поставленных задач
5	Привести примеры линейных моделей.	УК-1.2	Владеет навыками работы с научной и учебной литературой
6	Привести алгоритмы для решения задачи классификации.	ПК-1.1 ПК-5.2	Знает классические математические модели, применяемые в естественных науках, промышленности и бизнесе. Умеет в составе команды разработчиков участвовать в процессах разработки программного обеспечения.
7	Дать определение проблемы переобучения и методов ее обнаружения и разрешения (регуляризация).	УК-1.2	Владеет методами критического анализа и синтеза информации

8	Привести примеры алгоритмов машинного обучения - решающие деревья и их композиции.	ПК-1.1	Умеет использовать математические модели в профессиональной деятельности
9	Привести описание метода К-ближайших соседей.	ПК-2.1	Владеет практическим опытом разработки и реализации алгоритмов на базе языков программирования и пакетов прикладных программ моделирования для решения задач профессиональной деятельности
10	Дать описание метода опорных векторов (SVM) и привести основные этапы его вывода, указать виды ядер в методе опорных векторов.	ПК-4.1	Знает основные понятия математического аппарата и информационных технологий
11	Привести байесовские модели в задачах машинного обучения	ПК-1.1	Владеет навыками разработки и исследования новых математических моделей в естествознании, промышленности и бизнесе
12	Привести критерии оценки эффективности моделей машинного обучения в задачах классификации и регрессии.	ПК-1.2	Умеет создавать, исследовать и анализировать математические модели, применяемые в естественных науках, промышленности и бизнесе
13	Привести математическую модель биологического нейрона, привести примеры функций активации.	ПК-4.1	Умеет находить, формулировать и решать стандартные задачи в собственной научно-исследовательской деятельности в математике и информатике
14	Дать определения основных объектов нейронных сети (слои, функции активации, функция потерь, оптимизаторы), дать доказательство алгоритма обратного распространения ошибки.	ПК-4.1	Владеет навыками научно-исследовательской деятельности в математике и информатике
15	Дать определения основных элементов сверточной нейронной сети: свертка (2D и 3D), pooling.	УК-1.2	Владеет навыками работы с научной и учебной литературой

16	Привести примеры различных архитектур нейронных сетей.	УК-1.1 ПК-5.1	Умеет проводить анализ задачи, выделять ее базовые составляющие и формулировать результаты, которые необходимо достигнуть. Знает правила работы в составе научно-исследовательского и производственного коллектива, методы и способы решения задач профессиональной деятельности.
17	Дайте описание библиотек и фрейворков для построения нейросетевых моделей, их обучения: *TensorFlow, Keras, PyTorch.*	ПК-1.2 ПК-5.1	Владеет практическим опытом применения языков программирования и пакетов прикладных программ для проведения математического моделирования при помощи компьютерной техники. Умеет работать в составе научно-исследовательского и производственного коллектива и решать задачи профессиональной деятельности.
18	Привести описание возможностей библиотеки компьютерного зрения opencv.	ПК-2.1 ПК-5.1	Знает современные методы разработки и реализации алгоритмов математических моделей на базе языков программирования и пакетов прикладных программ моделирования. Владеет навыками работы в составе научно-исследовательского и производственного коллектива, алгоритмами решения задач профессиональной деятельности.

Материалы для текущего контроля

Формы текущего контроля: контроль посещаемости, отчеты по практическим занятиям (лабораторные работы).

Тематика практических работ.

№	Тема работы
ПР1	Вводное занятие по работе на гетерогенной вычислительной платформе ЛИТ ОИЯИ *HybriLIT* в рамках экосистемы для ML/DL.

	Работа в *Jupyter Notebook* – интерактивной web-среде для разработки приложений на Python, анализа и визуализации данных, решения задач математического моделирования, в том числе на базе методов машинного и глубокого обучения. **Основы** *Python*: базовые типы данных, типы коллекций, управляющие структуры и функции. Язык программирования *Python*, библиотеки для анализа и визуализации данных: *NumPy, matplotlib, pandas* (основы). **Библиотека** *NumPy*: работа с массивами (создание, индексация, срезы, слияние и разбиение), выполнение вычислений над массивами, агрегирование, операции над массивами, сортировка массивов, структурированные массивы. **Примеры:** • Создание одномерных и многомерных массивов (np.arange(),np.linspace(), np.random.normal() и т.д.): • Атрибуты массивов (ndim, shape, size) • Срезы: напечатать строку, столбец и т.д. Маскирование и «прихотливая» идексация. Визуализация с помощью **библиотек** *matplotlib* и *seaborn*: графики, гистограммы и т.д. **Примеры:** • простые линейные графики: цвета и стили линий, настройка графиков, подписи к осям, метки на графиках; • простые диаграммы рассеяния; • визуализация погрешностей; графики плотности и контурные графики.
ПР2	Работа с данными с **библиотекой** *Pandas*: объекты библиотеки, операции над данными, иерархическая индексация, объединение набор данных, агрегирование и группировка, сводные таблицы, работа с временными рядами. **Примеры:** • Объекты библиотеки *Pandas:*объект DataFrame, объект *Series;* • объект *Series* как специализированный словарь; • индексация и выборка данных; • чтение данных из CSV файла; сводные таблицы на примере данных о рождаемости.
ПР3	Машинное обучение с **библиотекой** *Scikit-learn.* Лабораторная работа № 1: задача о конверсии сайта.
ПР4	Машинное обучение с **библиотекой** *Scikit-learn.* Линейные модели.
ПР5	Алгоритмы машинного обучения: задачи классификации (наивная байесовская классификация, гауссов байесовский классификатор, полиномиальный баесовский классификатор).
ПР6	Алгоритмы машинного обучения: линейная регрессия, регрессия по комбинации базисных функций, вопросы регуляризации.
ПР7	Алгоритмы машинного обучения: метод опорных векторов (*Support*

	Vector Machines).
ПР8	Задачи машинного обучения: деревья принятия решений и случайные леса.
ПР9	Задачи машинного обучения: кластеризация, методы понижения размерности. Метод главных компонент.
ПР10	Фреймворки и библиотеки TensorFlow, Keras: библиотека глубокого обучения *Python*. **На примерах** рассматривается процесс создания нейронной сети в фреймворке Keras: • создания модели с использованием функционального API; • подготовку данных для обучения, создание генератора данных; • обучение нейросети и оценка качества получившейся модели.
ПР11	Фреймворки и библиотеки TensorFlow, Keras: библиотека глубокого обучения *Python*. Рассматриваются необходимые приёмы повышения точности результатов: • аугментация данных для обучения; • регуляризация сети с помощью слоев Dropout;
ПР12	Фреймворки и библиотеки TensorFlow, Keras: библиотека глубокого обучения *Python*. *Fine-tuning* нейросетевых моделей на основе предварительно обученной сети.
ПР13	Примеры на использование сверточных нейронных сетей для решения прикладных задач.
ПР14	Машинное и глубокое обучение для задач анализа текстов. Рассматривается workflow типичного проекта Data Science, работающего с текстовыми документами: формулировка проблемы, генерация данных, анализ данных на основе соответствующих функций: несколько способов извлечения и описания семантической информации и будет показано как включить / дополнить ее дополнительной несемантической (что может помочь улучшить результаты). Далее мы рассмотрим, сконструируем и применим несколько стандартных моделей машинного обучения (ML) для описания наших данных: мы приведем их к задачам классификации и регрессии. Затем мы анализируем эффективность методов ML, а также роль, влияние и актуальность наших семантических и несематических признаков. Далее мы покажем, как применять методы глубокого обучения для решения той же проблемы - мы рассмотрим простые модели DNN (глубокая нейронная сеть) и CNN (сверточная нейронная сеть). В конце мы сопоставляем наши результаты по ML и DL, обсуждаем их плюсы и минусы: эффективность, необходимые вычислительные ресурсы, возможные пути их улучшения.
ПР15	Компьютерное зрение с *Python*: работа с *opencv*.
ПР16	Компьютерное зрение с *Python*: глубокое обучение в технологиях компьютерного зрения.

Примерные темы домашних заданий

№	Тема домашнего задания	Индикаторы достижения компетенции	Результаты обучения
Д1	Задачи на работу с многомерными массивами (библиотека *NumPy*), задачи на построение графиков функций (построение функций активации), диаграмм. Интерактивное управление в Jupyter Notebooks: библиотека *IPywidgets* – задача на нахождение точек пересечения графиков функций в зависимости от параметров.	УК-1.1	Знает специфику системного подхода. Знает специфику логических методов анализа и синтеза.
Д2	Задачи на работу с данными с библиотекой *Pandas.*	УК-1.1	Умеет применять системный подход для анализа и решения поставленных задач
Д3	Лабораторная работа 1. Корреляция и линейная регрессия (часть 1): исследовать зависимость количества скачиваний программного продукта от количества посещений сайта: средствами Python, Python+NumPy, Scikit-Learn.	ПК-2.1	Умеет разрабатывать и реализовывать алгоритмы математических моделей на базе языков программирования и пакетов прикладных программ моделирования
Д4	Лабораторная работа 2. Классификация (часть 1): изучение метода k ближайших соседей для задач классификации.	УК-1.2	Владеет навыками работы с научной и учебной литературой
Д5	Лабораторная работа 2. Классификация (часть 2): изучение метода опорных векторов и наивного байесовского классификатора.	ПК-1.1 ПК-5.2	Знает классические математические модели, применяемые в естественных науках, промышленности и бизнесе. Умеет в составе команды разработчиков участвовать в процессах разработки программного обеспечения.
Д6	Лабораторная работа 3. Сравнение алгоритмов, метрики. Задача бинарной классификации на наборе данных с Kaggle: *pima-indians-diabetes.*	УК-1.2	Владеет методами критического анализа и синтеза информации
Д7	Лабораторная работа 4.	ПК-1.1	Умеет использовать математические

	Применение методов машинного обучения для экономических задач.		модели в профессиональной деятельности
Д8	Лабораторная работа 5. Задачи машинного обучения: деревья принятия решений и случайные леса.	ПК-2.1	Владеет практическим опытом разработки и реализации алгоритмов на базе языков программирования и пакетов прикладных программ моделирования для решения задач профессиональной деятельности
Д9	Задачи машинного обучения: кластеризация, методы понижения размерности. Метод главных компонент. Моделирование данных определенной структуры (Scikit-Learn)	ПК-4.1	Знает основные понятия математического аппарата и информационных технологий
Д10	Лабораторная работа 6. Нейроконтроллер для персонажей компьютерных игр. Построение модели в Keras, обучение модели, анализ полученных результатов.	ПК-1.1	Владеет навыками разработки и исследования новых математических моделей в естествознании, промышленности и бизнесе
Д11	Лабораторная работа 7. Решение задачи классификации рукописных цифр с помощью многослойного перцептрона.	ПК-1.2	Умеет создавать, исследовать и анализировать математические модели, применяемые в естественных науках, промышленности и бизнесе
Д12	Задачи на работу с изображениями (библиотека opencv)	ПК-4.1	Умеет находить, формулировать и решать стандартные задачи в собственной научно-исследовательской деятельности в математике и информатике
Д13	Лабораторная работа 8. Нейронные сети. Сравнительный анализ использования различных алгоритмов оптимизации.	ПК-4.1 ПК-5.1	Владеет навыками научно-исследовательской деятельности в математике и информатике. Владеет навыками работы в составе научно-исследовательского и производственного коллектива, алгоритмами решения задач профессиональной деятельности.
Д14	Лабораторная работа 9. Решение задачи классификации изображений с помощью сверточной нейронной сети. Исходные данные: DataSet CIFAR-10	УК-1.2 ПК-2.1	Владеет навыками работы с научной и учебной литературой. Знает современные методы разработки и реализации алгоритмов математических моделей на базе языков программирования и пакетов прикладных программ моделирования.

Д15	Лабораторная работа 10. Распознавание рукописных цифр Решение задачи классификации рукописных цифр. Исходные данные: DataSet MNIST + собственный набор данных	УК-1.1 ПК-5.1	Умеет проводить анализ задачи, выделять ее базовые составляющие и формулировать результаты, которые необходимо достигнуть. Знает правила работы в составе научно-исследовательского и производственного коллектива, методы и способы решения задач профессиональной деятельности.
Д16	Лабораторная работа 11. Детектор масок: детекция наличия или отсутствия маски на лице человека	ПК-1.2 ПК-5.1	Владеет практическим опытом применения языков программирования и пакетов прикладных программ для проведения математического моделирования при помощи компьютерной техники. Умеет работать в составе научно-исследовательского и производственного коллектива и решать задачи профессиональной деятельности.

Пример домашней работы

Задача. (*С.В. Айвазян, В.С. Мхитарян.* Прикладная статистика в задачах и упражнениях. – М.: ЮНИТИ-ДАНА . 2001. -270 с.**)**

Специально исследование показало, что склонность фирм к утаиванию части своих доходов

(и, соответственно, - к уклонению от уплаты части налогов) в существенной мере определяется двумя показателями

- x_1 – соотношением «быстроты активов» и текущих пассивов;
- x_2 – соотношением прибыли и процентных ставок

(оба показателя оцениваются по определенной методике в шкале от 300 до 900 балов);

Обучающая выборка:

№	Фирмы, уклоняющиеся от налогов		Фирмы, не уклоняющиеся от налогов	
	x_1	x_2	x_1	x_2
1	740	68	750	590
2	670	600	360	600
3	560	550	720	750
4	540	520	540	710
5	590	540	570	700
6	590	700	520	670
7	470	600	590	790
8	560	540	670	700
9	540	630	620	730
10	500	600	690	840
11	-	-	610	680
12	-	-	550	730
13	-	-	590	750

Требуется:

- Предложить (с обоснованием) и реализовать метод, с помощью которого можно определить, к какому классу (1 или 2) следует отнести фирму с показателями:
$$x_1 = 740, x_2 = 590.$$
- **предложить (с обоснованием) и вычислить расстояние, измеряющее степень различия двух анализируемых классов**

Алгоритмы классификации :

- **SVM** (*C-Support Vector Classification* с различными значениями параметра *C*)
- Gaussian Naive Bayes algorithm for classification (*GaussianNB*)
- Метод K-ближайших соседей (*Nearest Neighbors Classification, KNeighborsClassifier):*

провести исследование в зависимости от параметров)

Примеры лабораторных работ

Лабораторная работа № 1

Задачи машинного обучения: Задача о конверсии сайта (корреляция и линейная регрессия).

Цель работы: На примере построения математической модели зависимости количества регистрации (и скачиваний) программного продукта от количества посещений сайта (данные работы [1]) освоить работу с моделями библиотеки *scikit-learn.*

Метод решения:

Для построения линейной модели предлагается использовать метод наименьших квадратов [2].

Библиотеки: для визуализации данных: *matplotlib* и *seaborn* [3], для построения линейной модели *Scikil-Lean* [4].

Подготовка и анализ данных:

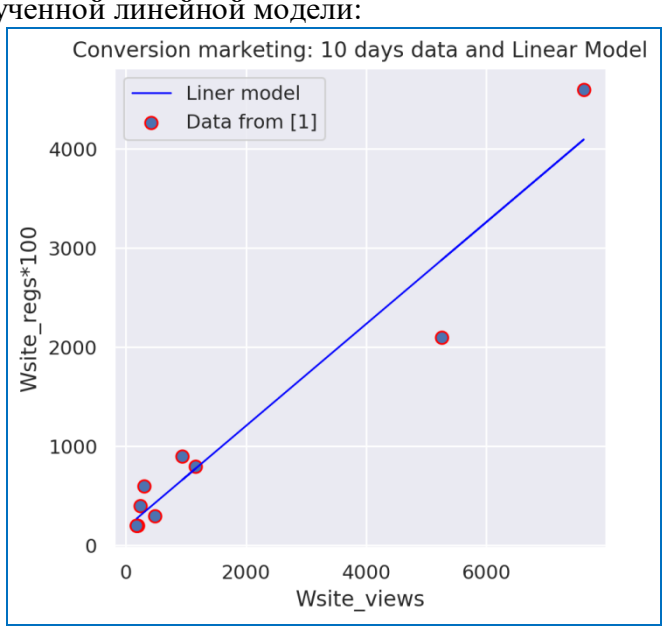
- Создать CSV-файл с данными: /data/**website_stat.csv**
- Напечатать таблицу, данные по посещаемости и регистрации на сайте, сводные статистические показатели
- Построить зависимость зарегистрировавшихся на сайте от количества посещений

Реализации:

1. Реализация средствами *Python* и *Numpy*.
2. Реализации с использованием *Python*, *Numpy* и *Scikil-Lean*.

Отчет должен содержать:

1. Постановку задачи.
2. Описание метода решения (вывод формул).
3. Описание двух реализаций.
4. Построенные графики зависимостей:
   Построить зависимость зарегистрировавшихся на сайте от количества посещений и полученной линейной модели:



5. Ответить на вопросы:
   - При количестве посещений сайта в 8000 человек, сколько планируется получить зарегистрировавшихся (или скачавших) программный продукт;
   - Для обеспечения 500 скачиваний продукта, сколько человек должны зайти на сайт?
6. Анализ полученных результатов и выводы.
7. Список использованной литературы.

Литература:

[1] Конверсия посетителей сайта:
https://habr.com/ru/company/nerepetitor/blog/250633/

[2] *Каханер Д., Моулер К., Нэш С.* Численные методы и программное обеспечение. М.: Мир. 1998.

[3] Библиотека *seaborn*: Statistical Data Visualization:
https://seaborn.pydata.org/

[4] Библиотека *scikit-learn*: Регрессия:
https://scikit-learn.org/stable/index.html

Лабораторная работа № 6
Нейроконтроллер для персонажей компьютерных игр

Цель работы: на примере задачи разработки *Нейроконтроллера для персонажей компьютерных игр* (пример взят из [1]) освоить работу с нейросетями, изучить метод обратного распространения ошибки и освоить *workflow* с нейронными сетями в фрейворке *Keras.*

Данные (обучающая и тестовые выборки) представлены в виде таблицы:

- здоровье (числовой параметр, принимающий значения от 0 до 2)
- наличие ножа (числовой параметр от или 1)
- наличие пистолета (числовой параметр от или 1)
- наличие врагов (числовой параметр - количество врагов)
- последний столбец показывает принадлежность классу, т.е. действие
  - уворачиваться
  - атаковать
  - прятаться
  - бежать

Требуется: создания нейронной сети в фреймворка Keras:

- создания модели в Keras, моделирующего поведение персонажа игры, в зависимости от значения 4 входных парамеров ;
- обучение нейросети и оценка качества получившейся модели.

Литература:

[1] *Джонс М.Т.* Программирование искусственного интеллекта в приложениях. Пер.с англ. А.И.Осипова. - М.: Пресс, 2006. - 312с. ISBN 5-94074-275-0. - ISBN 1-58450-278-9.
[2] *Keras*: The Python Deep Learning library:
https://keras.io/

Методические материалы, определяющие процедуры оценивания знаний, умений, навыков и (или) опыта деятельности, характеризующих этапы формирования компетенций

Для контроля знаний студентов по данной дисциплине необходимо проводить текущий и промежуточный контроль. Текущий контроль выполняется в виде приема допусков и защит практических и индивидуальных работ, устного опроса на практических занятиях. Промежуточный контроль проводится в виде зачета, на котором обсуждаются теоретические вопросы курса. Практическая часть зачитывается по результатам работ, выполненным в семестре, на основе балльно-рейтинговой системы. При рейтинговой системе все знания, умения и навыки, приобретаемые студентами в результате изучения дисциплины, оцениваются в баллах.

Оценка качества работы в рейтинговой системе является накопительной и используется для оценивания системной работы студентов в течение всего периода обучения. Максимально возможная сумма баллов по всем видам заданий приравнивается 100 %. При этом также учитывается посещаемость занятий – 0,5 балла за посещение одного занятия продолжительностью 1 академический час, при этом максимальный балл за посещаемость равен 17. Каждое практическое задание оценивается в зависимости от

сложности задания и определяется преподавателем в начале семестра, проставляется в журнале успеваемости и доводится до сведения студентов.

Чтобы быть допущенным к зачету, студент обязан проделать и сдать все практические работы. Работы, пропущенные без уважительной причины, а также работы, не защищенные студентом в течение 2-х недель после проведения практической части, оцениваются преподавателем с понижающим коэффициентом 0,5.

Преподаватель может использовать «штрафы» в виде уменьшения набранных баллов за пропуск практических занятий, за нарушение сроков выполнения учебной работы, за систематический отказ отвечать на семинарских занятиях и т.д.

Результаты работы студентов фиксируются преподавателем в журнале успеваемости.