

ГОУ ВПО «ДОНЕЦКИЙ НАЦИОНАЛЬНЫЙ УНИВЕРСИТЕТ»

ФИЗИКО-ТЕХНИЧЕСКИЙ ФАКУЛЬТЕТ

Кафедра компьютерных технологий

УТВЕРЖДАЮ:



Проректор по научно-методической
и учебной работе

Е.И. Скафа

«22» апреля 2020 г.

МП

Рабочая программа учебной дисциплины

«МАШИННОЕ ОБУЧЕНИЕ»

Направления подготовки:	09.04.01 Информатика и вычислительная техника
Магистерская программа:	Информатика и вычислительная техника
Образовательная программа:	академическая магистратура
Квалификация:	магистр
Форма обучения:	<u>очная</u> , <u>очно-заочная</u> , <u>заочная</u>

Донецк 2020



УТВЕРЖДАЮ:

Декан физико-технического факультета

С.А. Фоменко

№2
«17» апреля 2020 г.

МП

Программа учебной дисциплины «**Машинное обучение**» составлена на основании Федерального государственного образовательного стандарта высшего образования (ФГОС ВО) направления подготовки 09.04.01 Информатика и вычислительная техника, утвержденного приказом Министерства образования и науки Российской Федерации от 19 сентября 2017 г. № 918;

Порядка организации учебного процесса в образовательных организациях высшего профессионального образования Донецкой Народной Республики, утвержденного приказом Министерства образования и науки ДНР № 1171 от «10» ноября 2017 г.;

учебного плана и основной образовательной программы Информатика и вычислительная техника, направления подготовки 09.04.01 Информатика и вычислительная техника, разработанных в ГОУ ВПО «Донецкий национальный университет».

Разработчик:

канд. тех. наук, доцент кафедры
компьютерных технологий

Т.В. Ермоленко

Программа учебной дисциплины утверждена на заседании кафедры
компьютерных технологий

Протокол № 12 от «2» апреля 2020 г.

Зав. кафедрой компьютерных технологий

Ермоленко Т.В.

Программа учебной дисциплины одобрена учебно-методической комиссией
физико-технического факультета

Протокол № 5 от «15» апреля 2020 г.

Председатель учебно-методической комиссии
физико-технического факультета

Котенко В.Н

1. ОБЛАСТЬ ПРИМЕНЕНИЯ И МЕСТО ДИСЦИПЛИНЫ В УЧЕБНОМ ПРОЦЕССЕ

Учебная дисциплина «Машинное обучение» относится к базовой части блока «Дисциплины (модули)» и состоит из двух содержательных модулей: модуль 1 – «Методы классификации и регрессии», модуль 2 – «Ансамбли решающих правил. Критерии выбора моделей и методы отбора признаков».

Дисциплина реализуется на физико-техническом факультете кафедрой компьютерных технологий.

Этот курс опирается на математическую подготовку студентов, полученную при изучении дисциплин бакалавриата: «Математика», «Теория вероятности, математическая статистика», «Математическая логика», «Современные информационные системы и технологии», на знания технологий разработки современного программного обеспечения, полученные при изучении дисциплин бакалавриата: «Технологии разработки программного обеспечения» «Объектно-ориентированное программирование», а также на знания технологий искусственного интеллекта, полученные в результате освоения дисциплины магистратуры «Технологии извлечения знаний», закладывает фундамент научно-прикладной подготовки будущих исследователей в области инженерии знаний.

Полученные знания используются студентами при изучении дисциплины «Распознавание речи», а также во время выполнения научно-исследовательской работы и при написании магистерской диссертации.

2. СТРУКТУРА ДИСЦИПЛИНЫ

<i>Характеристика учебной дисциплины</i>		
Направление подготовки	09.04.01 Информатика и вычислительная техника	
Магистерская программа	Информатика и вычислительная техника	
Программа подготовки	академическая магистратура	
Квалификация	магистр	
Количество содержательных модулей	2	
Дисциплина базовой / вариативной части образовательной программы	дисциплина вариативной части	
Формы контроля	1 модульный контроль, 1 экзамен	
Показатели	очная форма обучения	заочная форма обучения
Количество зачетных единиц (кредитов)	3,5	3,5
Год подготовки	2	2
Семестр	3	3
Количество часов	126	126
- лекционных	14	2
- практических, семинарских		
- лабораторных	28	6
- самостоятельной работы	84	118
в т.ч. индивидуальное задание		
Недельное количество часов,	9	15
в т.ч. аудиторных	3	1,2

3. ОПИСАНИЕ ДИСЦИПЛИНЫ

Цели и задачи

Цель – формирование у магистрантов знаний основных математических методов и алгоритмов машинного обучения, развить интуицию для лучшего понимания основных идей, лежащих за этими методами, привить навыки работы с программным обеспечением, реализующим алгоритмы машинного обучения, и умение применять полученные знания при проектировании и реализации интеллектуальных информационных систем.

Задачи – освоение логических, метрических и вероятностных методов классификации и регрессии, алгоритмов построения разделяющих гиперплоскостей и нейросетей, критериев выбора моделей и методов отбора признаков, методов построения композиций классификаторов.

Требования к результатам освоения дисциплины:

Процесс изучения дисциплины «Интеллектуальные системы» направлен на формирование элементов следующих компетенций в соответствии с ФГОС ВО РФ по направлению подготовки 09.04.01 Информатика и вычислительная техника, утвержденного приказом Министерства образования и науки Российской Федерации от 19 сентября 2017 г. № 918, и основной образовательной программы высшего профессионального образования направления подготовки 09.04.01 Информатика и вычислительная техника (магистерская программа: Информатика и вычислительная техника):

а) универсальных (УК):

– способен осуществлять критический анализ проблемных ситуаций на основе системного подхода, вырабатывать стратегию действий (УК-1);

б) общепрофессиональных (ОПК):

– способен самостоятельно приобретать, развивать и применять математические, естественнонаучные, социально-экономические и профессиональные знания для решения нестандартных задач, в том числе в новой или незнакомой среде и в междисциплинарном контексте (ОПК-1);

– способен разрабатывать оригинальные алгоритмы и программные средства, в том числе с использованием современных интеллектуальных технологий, для решения профессиональных задач (ОПК-2);

– способен применять на практике новые научные принципы и методы исследований (ОПК-4);

– способен разрабатывать и модернизировать программное и аппаратное обеспечение информационных и автоматизированных систем (ОПК-5).

в) профессиональных (ПК):

производственно-технологическая деятельность:

– способен осуществлять технологическую поддержку подготовки технических публикаций (ПК-3);

– способен осуществлять управление развитием инфокоммуникационной системы организации (ПК-6);

организационно-управленческая деятельность:

– способен осуществлять руководство разработкой комплексных проектов на всех стадиях и этапах выполнения работ (ПК-20);

научно-исследовательская деятельность:

– способен осуществлять экспертный анализ эргономических характеристик программных продуктов и/или аппаратных средств. (ПК-21).

В результате изучения учебной дисциплины студент должен

Знать:

- постановку основных задач машинного обучения;
- возможности алгоритмов машинного обучения и способы их применения для решения практических задач управления и распознавания образов;
- логические, метрические и вероятностные методы классификации и регрессии;
- основные понятия и принципы работы искусственных нейронных сетей;
- композиции классификаторов, методы бустинга;
- критерии выбора моделей и методы отбора признаков;

Уметь:

- выполнять грамотную постановку задач, возникающих в практической деятельности, для их решения с помощью методов машинного обучения;
- обосновать применение того или иного алгоритма машинного обучения для решения конкретной задачи;
- проводить анализ работы методов машинного обучения с выявлением их сильных и слабых сторон;
- анализировать результаты обучения алгоритма машинного обучения, предлагать пути повышения его точности;
- программно реализовывать алгоритмы машинного обучения;

Владеть:

- базовым инструментарием машинного обучения;
- навыками построения различного рода классификаторов.

4. Содержание дисциплины и формы организации учебного процесса

Порядковый номер и тема	Краткое содержание темы
1	2
<i>Содержательный модуль 1.</i> Инженерия знаний. Системы, основанные на знаниях	
<i>Тема 1.</i> Основные понятия и примеры прикладных задач.	Постановка задач обучения по прецедентам. Объекты и признаки. Типы шкал. Типы задач. Основные понятия: модель алгоритмов, метод обучения, функция потерь и функционал качества, принцип минимизации эмпирического риска, обобщающая способность, скользящий контроль.

1	2
<p>Тема 2. Методы классификации и регрессии.</p>	<p>Метрические и логические методы классификации и регрессии: метод эталонных образов, метод ближайших соседей kNN и его обобщения; подбор числа k по критерию скользящего контроля; логическая закономерность; параметрические семейства закономерностей; переборные алгоритмы синтеза конъюнкций; решающее дерево; вывод критериев ветвления; энтропийный критерий, критерий Джини. Линейный классификатор: разделяющая поверхность; метод решающих функций и опорных векторов; линейные модели регрессии и классификации; метод наименьших квадратов; полиномиальная регрессия. Байесовская теория классификации: наивный байесовский классификатор; непараметрические и параметрические методы оценивания плотности; смеси. Нейронные сети: модель биологического нейрона; функции активации; алгоритм обратного распространения ошибок; свёрточные сети, рекуррентные сети.</p>
<p>Тема 3. Прогнозирование временных рядов.</p>	<p>Задача прогнозирования временных рядов. Примеры приложений. Экспоненциальное скользящее среднее. Адаптивная авторегрессионная модель. Адаптивная селективная модель. Адаптивная композиция моделей.</p>
<p>Тема 4. Нечеткие знания.</p>	<p>Нечеткие множества и отношения. Теория приближенных рассуждений. Мягкая экспертная система.</p>
<p align="center">Содержательный модуль 2. Ансамбли решающих правил. Критерии выбора моделей и методы отбора признаков</p>	
<p>Тема 4. Линейные композиции, бустинг.</p>	<p>Основные понятия: базовый алгоритм (алгоритмический оператор), корректирующая операция. Взвешенное голосование. Алгоритм AdaBoost. Экспоненциальная аппроксимация пороговой функции потерь. Процесс последовательного обучения базовых алгоритмов. Теорема о сходимости бустинга. Базовые алгоритмы в бустинге. Алгоритм AnyBoost. Алгоритм ComBoost.</p>
<p>Тема 5. Критерии выбора моделей и методы отбора признаков.</p>	<p>Внутренние и внешние критерии. Эмпирические, статистические и аналитические критерии. Сложность задачи отбора признаков. Полный перебор. Метод добавления и удаления, шаговая регрессия. Поиск в глубину, метод ветвей и границ. Анализ главных компонент и факторный анализ.</p>

Тематический план

Содержательный модуль 1												
Названия содержательных модулей и тем	Количество часов											
	Очная форма обучения						Заочная форма обучения					
	всего	в т.ч.					всего	в т.ч.				
		лекции	практические	лабораторные	самостоятельная работа	индивидуальная работа		лекции	практические	лабораторные	самостоятельная работа	индивидуальная работа
Тема 1. Основные понятия и примеры прикладных задач.	4	2			2							
Тема 2. Методы классификации и регрессии.	40	4		12	24							
Тема 3. Прогнозирование временных рядов.	20	2		8	10							
Итого по содержательному модулю 1	64	8		20	36							
Содержательный модуль 2												
Тема 4. Линейные композиции, бустинг.	32	4			28							
Тема 5. Критерии выбора моделей и методы отбора признаков.	30	2		8	20							
Итого по содержательному модулю 2	62	6		8	48							
Всего часов	126	14		28	84							

Курс дисциплины «Машинное обучение» предусматривает следующие **формы организации учебного процесса**:

- лекции;
- практические занятия;
- лабораторные занятия;
- самостоятельная работа студента.

По источнику передачи и восприятия учебной информации используются словесные (лекция, беседа), наглядные (слайды, иллюстрации, коды программ), практические (исследования, упражнения, лабораторные работы) методы.

По характеру познавательной деятельности студентов используются объяснительно-иллюстративные и репродуктивные методы, проблемное преподавание, частично-поисковый и исследовательский методы.

В зависимости от основной дидактической цели и задач используются методы устного изложения знаний, закрепление учебного материала, самостоятельной работы студентов по осмыслению и усвоению нового материала, работы по применению знаний на практике и выработке умений и навыков, проверки и оценки знаний, умений и навыков.

Используются следующие методы контроля:

- устный контроль (экспресс-опрос на лекциях);
- защита лабораторных работ;
- модульная контрольная работа.
- итоговый контроль (экзаменационные билеты).

5. МЕТОДИЧЕСКИЕ РЕКОМЕНДАЦИИ ДЛЯ ПРОВЕДЕНИЯ ЛЕКЦИОННЫХ, ПРАКТИЧЕСКИХ И ЛАБОРАТОРНЫХ ЗАНЯТИЙ

Темы лекционных занятий

<i>№ п/п</i>	<i>Название темы</i>	<i>Количество часов</i>
1	Основные понятия и примеры прикладных задач.	2
2	Метрические и логические методы классификации и регрессии.	1
3	Линейные и вероятностные методы классификации и регрессии.	1
4	Искусственные нейронные сети.	2
5	Прогнозирование временных рядов.	2
6	Алгоритмы Adaboost, Anyboost.	2
7	Голосование по большинству и старшинству.	2
8	Отбор информативных признаков.	1
9	Синтез информативных признаков.	1
	ВСЕГО	14

Темы лабораторных занятий

<i>№ п/п</i>	<i>Название темы</i>	<i>Количество часов</i>
1	Построение модели парной и множественной линейной регрессии средствами языка R.	8
2	Построение наивного байесовского классификатора.	4
3	Анализ временных рядов средствами языка R.	8
4	Анализ данных с помощью метода главных компонент средствами языка R.	8
	ВСЕГО	28

6. Методические рекомендации по организации самостоятельной работы студентов

Самостоятельная работа студентов по курсу «Машинное обучение» предусматривает:

- систематическое посещение лекционных занятий, ведение конспекта лекций;
- повседневное изучение лекционного материала и содержания технической литературы, рекомендуемые этой программой и рабочим учебным планом;
- добросовестную подготовку к лабораторным и практическим занятиям;
- своевременное и качественное оформление отчётов по лабораторным работам.
- самостоятельную разработку алгоритмов и текстов программ лабораторных работ.

Организация самостоятельной работы студентов

№ п/п	Название темы	Количество часов
1	2	3
1	Изучение темы: Обучение по прецедентам. Обучение без учителя. Примеры практических задач	2
2	Изучение темы: Метрические методы классификации и регрессии. В том числе вопросов, не освещаемых на лекциях: – метод окна Парзена с постоянной и переменной шириной окна; – непараметрическая регрессия. Локально взвешенный метод наименьших квадратов. Ядерное сглаживание; – робастная непараметрическая регрессия. Алгоритм LOWESS.	4
3	Изучение темы: Логические методы классификации и регрессии.	4
	Подготовка к лабораторной работе №1: Изучение возможностей языка R для построения моделей парной и множественной линейной регрессии и проведение их анализа.	4
4	Изучение темы: Вероятностные методы классификации и регрессии. В том числе вопросов, не освещаемых на лекциях: – непараметрическое оценивание плотности: ядерная оценка плотности Парзена-Розенблатта. – параметрическое оценивание плотности: нормальный дискриминантный анализ; – выбор числа компонентов смеси. Пошаговая стратегия. Априорное распределение Дирихле; – смесь многомерных нормальных распределений. Сеть радиальных базисных функций (RBF) и применение ЕМ-алгоритма для её настройки; – сравнение RBF-сети и SVM с гауссовским ядром.	6
	Подготовка к лабораторной работе №2: Изучение байесовского подхода к классификации.	4
5	Изучение темы: Нейронные сети. В том числе вопросов, не освещаемых на лекциях: – быстрые методы стохастического градиента; – методы постепенного усложнения сети, оптимальное прореживание нейронных сетей; – функции активации ReLU и PReLU. Проблема «паралича» сети. – обучение рекуррентных сетей: Backpropagation Through Time (BPTT).	6
6	Изучение темы: Прогнозирование временных рядов. В том числе вопросов, не освещаемых на лекциях: – прогнозирование временных рядов с помощью рекуррентной сети LSTM	4
7	Подготовка к лабораторной работе №3: Изучение возможностей языка R для прогнозирования временных рядов.	2
8	Изучение темы: Ансамбли моделей классификаторов. В том числе вопросов, не освещаемых на лекциях: – бэггинг и метод случайных подпространств; – случайный лес; – градиентный бустинг; – алгоритм LogitBoost; – смесь алгоритмов.	14

1	2	3
9	Изучение темы: Методы голосования. В том числе вопросов, не освещаемых на лекциях: – простое голосование: алгоритм ComBoost.	14
	Изучение темы: Критерии выбора модели. В том числе вопросов, не освещаемых на лекциях: – критерий непротиворечивости моделей; – критерий регуляризации.	8
	Изучение темы: Отбор информативных признаков. В том числе вопросов, не освещаемых на лекциях: – отбор признаков методом стохастического поиска.	8
10	Подготовка к лабораторной работе №4: Изучение метода анализа главных компонент. Изучение возможностей языка R для проведения анализа данных с помощью анализа главных компонент.	4
	ВСЕГО	84

7. ИНДИВИДУАЛЬНЫЕ ЗАДАНИЯ

Индивидуальные задания не предусмотрены.

8. КОНТРОЛЬНЫЕ ВОПРОСЫ К ПРОМЕЖУТОЧНОЙ АТТЕСТАЦИИ

1. Функциональная и вероятностная постановка задачи обучения по прецедентам.
2. Типы признаков. Понятие модели алгоритмов и метода обучения. Обобщающая способность метода. Эмпирические оценки обобщающей способности.
3. Функция потерь и функционал качества. Наиболее употребительные функции потерь.
4. Суть метода ближайшего соседа. Какими недостатками он обладает?
5. Алгоритм k ближайших соседей и его модификация – k взвешенных ближайших соседей. Какие недостатки простейших метрических алгоритмов типа kNN?
6. Отступ объекта относительно алгоритма. Типы объектов в порядке убывания отступа. Алгоритм STOLP для отбора эталонных объектов.
7. Критерии информативности, позволяющие называть предикаты закономерностями. Приведите эвристическое, статистическое, энтропийное определения информативности.
8. Бинаризация количественных признаков. Разбиение диапазона значений признака на информативные зоны.
9. Алгоритмы синтеза конъюнкций.
10. Решающие списки, разновидности решающих списков. Жадный алгоритм их построения.
11. Решающие деревья, синтез решающих деревьев, критерии расщепления. Преимущества метода деревьев решений.
12. Основные этапы алгоритма построения решающего дерева ID3, его преимущества и недостатки. Алгоритм CART. Редукция решающих деревьев.
13. Решающая функция и дискриминантная поверхность. Каким свойством обладает гиперплоскость, построенная методом SVM? Зачем в SVM используются ядра?
14. Общий вид регрессионной модели. Модель множественной линейной регрессии. Могут ли зависимости в регрессионной модели быть нелинейными?
15. Подход для нахождения оптимальных оценок параметров регрессионной модели. Оценки качества регрессионной модели.

16. Общая идея метода потенциальных функций. Каковы основные этапы алгоритма метода потенциальных функций?
17. Байесовский подход к распознаванию образов.
18. Параметрические и непараметрические методы оценивания плотности вероятностей. Смеси.
19. Модель нейрона, ее ограничения. Модель многослойного персептрона.
20. Нейронная сеть Хопфилда, ее ограничения.
21. Основные правила обучения нейросетей. Алгоритм обратного распространения ошибки.
22. Общая архитектура сверточных нейросетей. Двумерная свёртка. Техники свёртки, «фильтр» и «ядро» свёртки. Слои пулинга.
23. Глубокие нейросети. Подход к обучению глубоких нейросетей.
24. Схема однослойной рекуррентной сети. Отличие рекуррентной сети Элмана от предыдущих концепций рекуррентных сетей.
25. Общая архитектура рекуррентной сети. Особенности архитектуры LSTM-сети, преимущества LSTM-сети.
26. Алгоритм обучения нейросети методами Хебба.
27. Принцип работы сети Кохонена.
28. Составляющие временного ряда. Модель временного ряда. Типы моделей временного ряда. Трансформация данных ряда, виды трансформации данных.
29. Стационарный временной ряд, тесты для определения стационарности временного ряда. Единичный корень.
30. Коэффициент автокорреляции. Его статистическая оценка. Коррелограмма. Частичная автокорреляция.
31. Наиболее часто используемые на практике критерии проверки «наличия-отсутствия» тренда?
32. Критерии оценки информативности модели ряда. Этапы алгоритма анализа временного ряда.
33. Модель AR(p). Модель MA(p). Модель ARIMA. Определение порядков моделей.

9. ОБРАЗЕЦ МОДУЛЬНОГО КОНТРОЛЯ

ГОУ ВПО «ДОНЕЦКИЙ НАЦИОНАЛЬНЫЙ УНИВЕРСИТЕТ»

Физико-технический факультет

<i>Направление подготовки:</i>	09.04.01 Информатика и вычислительная техника
<i>Магистерская программа:</i>	Информатика и вычислительная техника
<i>Программа подготовки:</i>	академическая магистратура
<i>Семестр</i>	3
<i>Учебная дисциплина</i>	Машинное обучение

МОДУЛЬНАЯ КОНТРОЛЬНАЯ РАБОТА

ВАРИАНТ №1

1. Производственные системы. Структура производственной системы и стратегии вывода.
2. Композиционное правило вывода в нечеткой логике.

Утверждено на заседании кафедры компьютерных технологий,
протокол № ____ от “__” _____ 20__ г.

Зав. кафедрой _____

Преподаватель _____

Критерии оценивания модульного контроля

<i>Номер задания</i>	<i>Количество баллов</i>
Задание 1	5
Задание 2	5
<i>Всего</i>	<i>10</i>

10. ОБРАЗЕЦ ЭКЗАМЕНАЦИОННОГО БИЛЕТА

Теоретические вопросы к экзамену

1. Композиция алгоритмов Взвешенное голосование.
2. Алгоритма AdaBoost. Достоинства и недостатки AdaBoost.
3. Итерации обобщённого алгоритма бустинга AnyBoost.
4. Алгоритм построения композиции классификаторов для голосования по большинству. Преобразование простого голосования во взвешенное.
5. Алгоритм построения композиции для голосования по старшинству. Стратегии выбора параметра алгоритма голосования по старшинству.
6. Типы критериев задачах отбора признаков и моделей. Критерий средней ошибки на контрольных данных, критерий скользящего контроля.
7. Bootstrap.
8. Алгоритм полного перебора всевозможных наборов признаков.
9. Алгоритм Add. Каковы его достоинства и недостатки?
10. Алгоритм Add-Del. Каковы его достоинства и недостатки?
11. Стратегии обхода дерева возможных наборов признаков.
12. Критерии выбора модели, основанные на оценках обобщающей способности.
13. Что строит алгоритм PCA, выполняя обработку всей матрицы наблюдений? Зачем используют PCA? Что такое ковариационная матрица, матрица счетов, матрица нагрузок?
14. Что из себя представляют главные компоненты? Являются ли они коррелированными? Из каких соображений выбирают количество главных компонент, которым стоит ограничиться? Критерия broken stick model.
15. Общая архитектура рекуррентной сети. Особенности архитектуры LSTM-сети, преимущества LSTM-сети.
16. Принцип работы сети Кохонена.
17. Общая архитектура сверточных нейросетей. Двумерная свёртка. Техники свёртки, «фильтр» и «ядро» свёртки. Слой пулинга.
18. Параметрические и непараметрические методы оценивания плотности вероятностей. Смеси.
19. Подход для нахождения оптимальных оценок параметров регрессионной модели. Оценки качества регрессионной модели.
20. Схема однослойной рекуррентной сети. Отличие рекуррентной сети Элмана от предыдущих концепций рекуррентных сетей.
21. Алгоритм k ближайших соседей и его модификация – k взвешенных ближайших соседей. Какие недостатки простейших метрических алгоритмов типа kNN?
22. Отступ объекта относительно алгоритма. Типы объектов в порядке убывания отступа. Алгоритм STOLP для отбора эталонных объектов.
23. Решающие деревья. Основные этапы алгоритма построения решающего дерева ID3, его преимущества и недостатки. Алгоритм CART. Редукция решающих деревьев.

Образец экзаменационного билета

ГОУ ВПО «ДОНЕЦКИЙ НАЦИОНАЛЬНЫЙ УНИВЕРСИТЕТ»

Физико-технический факультет

Направление подготовки: **09.04.01 Информатика и вычислительная техника**
 Магистерская программа: **Информатика и вычислительная техника**
 Программа подготовки: **академическая магистратура**
 Семестр: **3**
 Учебная дисциплина: **Машинное обучение**

БИЛЕТ №1

1. Композиция алгоритмов Взвешенное голосование.
2. Общая архитектура рекуррентной сети. Особенности архитектуры LSTM-сети, преимущества LSTM-сети.
3. Рассмотрим задачу классификации с двумя классами 0 и 1. Пусть пространство признаков двумерное. Объекты каждого класса имеют нормальное распределение с математическим ожиданием (0, 0.5) и (0.4, 1) соответственно и матрицей ковариации

$$\Sigma = \begin{pmatrix} 1 & 2 \\ 2 & 3 \end{pmatrix}$$

Априорные вероятности классов равны $P\{Y=0\}=1/3$ и $P\{Y=1\}=2/3$

Найти: уравнение разделяющей поверхности байесова классификатора; собственное разложение матрицы Σ .

Перейти к новым координатам, оси которых совпадают с собственными векторами матрицы Σ . Выписать уравнение разделяющей поверхности байесова классификатора в новых координатах.

Утверждено на заседании кафедрой компьютерных технологий,
 протокол № ____ от “__” _____ 20__ г.

Зав. кафедрой _____

Преподаватель _____

Критерии оценивания экзамена

Номер задания	Количество баллов
Задание 1	10
Задание 2	10
Задание 3	30
Всего	50 баллов

11. КРИТЕРИИ ОЦЕНИВАНИЯ

Распределение баллов, которые могут получить студенты в процессе изучения дисциплины

Форма контроля	Максимальное количество баллов
Лабораторная работа №1	20
Лабораторная работа №2	20
Лабораторная работа №3	20
Лабораторная работа №4	20
Модульный контроль	10
Организационно-учебная работа студента	10

Согласно модульному принципу организации учебного процесса, содержание дисциплины «Машинное обучение» включает в себя два содержательных модуля. Каждый содержательный модуль состоит из теоретического материала и лабораторных работ, выполнение которых требует овладения теорией в указанном в модуле объеме.

К модульному контролю студент должен защитить 3 лабораторные работы, каждая из которых оценивается в 20 баллов.

На модульном контроле студент имеет возможность получить 10 баллов, ответив на 2 теоретических вопроса, каждый из которых оценивается в 5 баллов.

К концу семестра студент должен защитить еще 1 лабораторную работу, которая оценивается в 20 баллов.

Дополнительно 10 баллов студент может получить в течение семестра, посещая лекции и участвуя в устных опросах.

Экзамен сдают студенты с целью повышения рейтинга. На экзамене студент может получить 50 баллов, ответив на два теоретических вопроса, каждый из которых оценивается в 10 баллов, и решив задачу, оцениваемую в 30 баллов.

Шкала соответствия баллов национальной шкале

Оценка по шкале ECTS	Оценка по 100-балльной шкале	Оценка по государственной шкале (экзамен, дифференцированный зачет)	Оценка по государственной шкале (зачет)
A	90-100	5 (отлично)	зачтено
B	80-89	4 (хорошо)	зачтено
C	75-79	4 (хорошо)	зачтено
D	70-74	3 (удовлетворительно)	зачтено
E	60-69	3 (удовлетворительно)	зачтено
FX	35-59	2 (неудовлетворительно) с возможностью повторной сдачи	не зачтено
F	0-34	2 (неудовлетворительно) с возможностью повторной сдачи при условии обязательного набора дополнительных баллов	не зачтено

Оценка за овладение курса выставляется по следующим критериям:

– Оценку «отлично» заслуживает студент, который обнаружил глубокие знания при ответах на теоретические вопросы по темам курса, а также выполнил лабораторные работы в полном объеме и набрал более 90 баллов.

– Оценку «хорошо» заслуживает студент, сдавший все лабораторные работы и сделавший ошибки в теоретических или практических ответах при их защите, которые могут быть интерпретированы как малосущественные для вопросов, которые рассматривались. Студент должен набрать более 75 баллов.

– Оценку «удовлетворительно» заслуживает студент, который сдал лабораторные работы, отвечающий на вопросы неполно и с ошибками, но при этом набрал более 60 баллов.

– Оценку «неудовлетворительно» заслуживает студент, который не выполнил большинства лабораторных работ и набрал менее 60 баллов.

12. МАТЕРИАЛЬНО-ТЕХНИЧЕСКОЕ ОБЕСПЕЧЕНИЕ УЧЕБНОГО ПРОЦЕССА

Лекционные занятия проводятся в аудитории, оснащенной мультимедийной техникой и доской. Лабораторные занятия проводятся в компьютерном классе, оборудованном компьютерами с лицензионным программным обеспечением, доступом к сети Интернет, столами, доской.

13. РЕКОМЕНДОВАННАЯ ЛИТЕРАТУРА

№ п/п	Наименование	Кол-во экземпляров в библиотеке ДонНУ	Наличие электронной версии в ЭБС
<i>Основная литература</i>			
1.	Ермоленко Т.В. Методы машинного обучения [Электронный ресурс]: учебное пособие / Т.В. Ермоленко – Донецк: ДонНУ, 2017. – Электронные данные (1 файл).	-	+
2.	Ермоленко Т.В. Введение в машинное обучение [Электронный ресурс]: учебное-методическое пособие / Т.В. Ермоленко – Донецк: ДонНУ, 2017. – Электронные данные (1 файл).	-	+
<i>Дополнительная литература</i>			
3.	Коэлья Л.П. Построение систем машинного обучения на языке Python [Электронная книга] / Л.П. Коэлья, В. Ричарт. – М.: Издательство «ДМК Пресс», 2016. – 302 с. URL: http://mirknig.su/knigi/programming/10054-postroenie-sistem-mashinnogo-obucheniya-na-yazyke-python-2-e-izdanie.html (в свободном доступе)	-	-
4.	Шитиков В.К., Мاستицкий С.Э. Классификация, регрессия и другие алгоритмы Data Mining с использованием R [Электронный ресурс]. / В.К. Шитиков, С.Э. Мастицкий – 2017. – 351 с. – Электронные данные, URL: https://github.com/ranalytics/data-mining (в свободном доступе)	-	-

14. ИНФОРМАЦИОННЫЕ РЕСУРСЫ

1. К.В. Воронцов. Машинное обучение (курс лекций). URL: <http://www.machinelearning.ru/wiki/index.php?title=ML> (дата обращения 03.01.2017).
2. Портал статей по применению ИТ и машинному обучению - http://habrahabr.ru/hub/machine_learning/ (дата обращения 03.01.2017).
3. Портал искусственного интеллекта URL: <http://www.aiportal.ru/downloads> (дата обращения 03.01.2017).

15. ПРОГРАММНОЕ ОБЕСПЕЧЕНИЕ

1. Язык R – свободная программная среда для статистических вычислений и графики (open source, лицензия GNU GPL 2).
2. RStudio – интегрированная среда разработки (IDE) для R (open source, лицензия GNU GPL 2).

Рабочая программа рассмотрена и переутверждена на заседании кафедры компьютерных технологий с изменениями (без изменений) на 2020 год.

Протокол № 12 от «2» апреля 2020 г.

Заведующий кафедрой

Ермоленко Т.В.

Рабочая программа рассмотрена и переутверждена на заседании кафедры компьютерных технологий с изменениями (без изменений) на 2021 год.

Протокол № ____ от «____» _____ 2021 г.

Заведующий кафедрой

Рабочая программа рассмотрена и переутверждена на заседании кафедры компьютерных технологий с изменениями (без изменений) на 2022 год.

Протокол № ____ от «____» _____ 2022 г.

Заведующий кафедрой

Рабочая программа рассмотрена и переутверждена на заседании кафедры компьютерных технологий с изменениями (без изменений) на 2023 год.

Протокол № ____ от «____» _____ 2023 г.

Заведующий кафедрой