МИНОБРНАУКИ РОССИИ

Федеральное государственное бюджетное
образовательное учреждение высшего образования
«Астраханский государственный университет имени В. Н. Татищева»
(Астраханский государственный университет им. В. Н. Татищева)

СОГЛАСОВАНО	УТВЕРЖДАЮ
Руководитель ОПОП	И.о. заведующего кафедрой ИБ
_____ И.М. Ажмухамедов	_____ Т.Г. Гурская
_____	протокол заседания кафедры № 9
«06» июня 2024 г.	«06» июня 2024 г.

РАБОЧАЯ ПРОГРАММА ДИСЦИПЛИНЫ (МОДУЛЯ)

«Системы искусственного интеллекта в информационной безопасности»

Составитель(и)	**Демина Р.Ю., к.т.н, доц., доцент кафедры ИБ**
Направление подготовки / специальность	**10.03.01 Информационная безопасность**
Направленность (профиль) ОПОП	**Организация и технология защиты информации**
Квалификация (степень)	**бакалавр**
Форма обучения	**очная**
Год приёма	**2021**
Курс	4
Семестр(ы)	**7**

Астрахань – 2024

1. ЦЕЛИ И ЗАДАЧИ ОСВОЕНИЯ ДИСЦИПЛИНЫ (МОДУЛЯ)

1.1. Целями освоения дисциплины (модуля) «Системы искусственного интеллекта в информационной безопасности» являются овладение студентами основными методами теории интеллектуальных систем, приобретение навыков по использованию интеллектуальных систем в области информационной безопасности, изучение основных методов представления знаний и моделирования рассуждений.

1.2. Задачи освоения дисциплины (модуля): «Системы искусственного интеллекта в информационной безопасности»

- изучение теоретических моделей рассуждений, поведения, обучения в когнитивных науках, постановки проблем математического и информационного моделирования сложных систем;
- умение планировать процесс моделирования и вычислительного эксперимента;
- овладение навыками постановки задач и обработки результатов компьютерного моделирования.

2. МЕСТО ДИСЦИПЛИНЫ (МОДУЛЯ) В СТРУКТУРЕ ОПОП

2.1. Учебная дисциплина (модуль) «Системы искусственного интеллекта в информационной безопасности» относится к части, формируемой участниками образовательных отношений и осваивается в 7 семестре.

2.2. Для изучения данной учебной дисциплины (модуля) необходимы следующие знания, умения, навыки, формируемые предшествующими учебными дисциплинами (модулями):

- теория вероятностей и математическая статистика;
- дискретная математика.

Знания: основные понятия теории вероятностей, математической статистики, дискретной математики.

Умения: решать типовые задачи теории вероятностей, математической статистики и дискретной математики.

Навыки: владеть методами оценки репрезентативности выборки и составления деревьев решений.

2.3. Последующие учебные дисциплины (модули) и (или) практики, для которых необходимы знания, умения, навыки, формируемые данной учебной дисциплиной (модулем):

- преддипломная практика.

3. ПЛАНИРУЕМЫЕ РЕЗУЛЬТАТЫ ОБУЧЕНИЯ ПО ДИСЦИПЛИНЕ (МОДУЛЮ)

Процесс освоения дисциплины (модуля) направлен на формирование элементов следующей(их) компетенции(ий) в соответствии с ФГОС ВО и ОПОП ВО по данному направлению подготовки / специальности:

профессиональные (ПК):

- ПК 3;
- ПК 4.

Таблица 1 – Декомпозиция результатов обучения

Код и наименование компетенции	Планируемые результаты обучения по дисциплине (модулю)		
	Знать (1)	Уметь (2)	Владеть (3)
ПК - 3. Способен осуществлять внедрение систем защиты информации для обеспечения информационной безопасности автоматизированных систем	ИПК – 3.1.1: знать основные алгоритмы искусственного интеллекта, применяемые для решения задач в сфере информационной безопасности	ИПК – 3.2.1: уметь решать задачи информационной безопасности методами искусственного интеллекта	ИПК – 3.3.1: владеть навыками оценки обученной модели.
ПК-4. Способен администрировать средства защиты информации в компьютерных системах и сетях	ИПК – 4.1.1: знать основные функции программной среды WEKA	ИПК – 4.2.1: уметь решать задачи информационной безопасности с использованием программной среды WEKA	ИПК – 4.3.1: владеть навыками представления объектов обучающего множества в формате csv и/или arff-файлов.

4. СТРУКТУРА И СОДЕРЖАНИЕ ДИСЦИПЛИНЫ (МОДУЛЯ)

Объём дисциплины (модуля) составляет 3 зачётных единицы, в том числе 54 часа, выделенных на контактную работу обучающихся с преподавателем (из них 18 часов(а) – лекции, 0 часов(а) – практические, семинарские занятия, 36 часов(а) – лабораторные работы), и 54 часов(а) – на самостоятельную работу обучающихся.

Таблица 2 – Структура и содержание дисциплины (модуля)

Раздел, тема дисциплины (модуля)	Семестр	Контактная работа (в часах)			Самост. работа		Форма текущего контроля успеваемости, форма промежуточной аттестации *[по семестрам]*
		Л	ПЗ	ЛР	КР	СР	
Раздел I. Жизненный цикл модели машинного обучения в задачах защиты информации	7	**8**		**16**		**30**	**Опросы, лабораторные работы**
Тема 1. Введение в системы искусственного интеллекта в информационной безопасности		2		4		7	Опрос
Тема 2. Обучающее множество в задачах защиты информации		2		4		8	Лабораторная работа 1
Тема 3. Моделирование и		2		4		7	Лабораторная

Раздел, тема дисциплины (модуля)	Семестр	Контактная работа (в часах)			Самост. работа		Форма текущего контроля успеваемости, форма промежуточной аттестации *[по семестрам]*
		Л	ПЗ	ЛР	КР	СР	
прогнозирование в задачах защиты информации							работа 2
Тема 4. Оценка и оптимизация моделей в сфере информационной безопасности		2		4		8	Лабораторная работа 3
Раздел II. Примеры прикладных задач из сферы информационной безопасности		**10**		**20**		**24**	
Тема 5. Атаки на системы искусственного интеллекта		2		4		5	Лабораторная работа 4
Тема 6. Задача обнаружения вредоносного программного обеспечения		2		4		5	Реферат
Тема 7. Задача поиска аномалий пользовательского трафика		2		4		5	Лабораторная работа 5
Тема 8. Задача распознавания лиц		2		4		5	Реферат
Тема 9. Задача распознавания «дипфейков»		2		4		4	Лабораторная работа 6
Итого		**18**		**36**		**54**	**Экзамен**

Примечание: Л – лекция; ПЗ – практическое занятие, семинар; ЛР – лабораторная работа; КР – курсовая работа; СР – самостоятельная работа.

Таблица 3 – Матрица соотнесения разделов, тем учебной дисциплины (модуля) и формируемых компетенций

Раздел, тема дисциплины (модуля)	Кол-во часов	Код компетенции		Общее количество компетенций
		ПК-3	ПК-4	
Раздел I. Жизненный цикл модели машинного обучения в задачах защиты информации	**54**	+	+	**2**
Тема 1. Введение в системы искусственного интеллекта в информационной безопасности	13	+	+	2
Тема 2. Обучающее множество в задачах защиты информации	14	+	+	2
Тема 3. Моделирование и прогнозирование в задачах защиты информации	13	+	+	2
Тема 4. Оценка и оптимизация	14	+	+	2

Раздел, тема дисциплины (модуля)	Кол-во часов	Код компетенции		Общее количество компетенций
		ПК-3	ПК-4	
моделей в сфере информационной безопасности				
Раздел II. Примеры прикладных задач из сферы информационной безопасности	**54**	+	+	2
Тема 5. Атаки на системы искусственного интеллекта	11	+	+	2
Тема 6. Задача обнаружения вредоносного программного обеспечения	11	+	+	2
Тема 7. Задача поиска аномалий пользовательского трафика	11	+	+	2
Тема 8. Задача распознавания лиц	11	+	+	2
Тема 9. Задача распознавания «дипфейков»	10	+	+	2
Итого	**108**	**108**	**108**	

Краткое содержание каждой темы дисциплины (модуля)

Раздел I. ***Жизненный цикл модели машинного обучения в задачах защиты информации***

Тема 1. Введение в системы искусственного интеллекта в информационной безопасности. Принятие решений на основе данных: традиционный подход, подход с машинным обучением (преимущества и сложности). Сбор и подготовка данных. Обучение модели. Оценка производительности. Оптимизация производительности. Способы повышения эффективности.

Тема 2. Обучающее множество в задачах защиты информации. Важность качества обучающего множества с точки зрения информационной безопасности. Специфика сбора данных в сфере информационной безопасности. Сбор данных: входные признаки, целевая переменная, объем и репрезентативность обучающей выборки. Подготовка данных к моделированию: категориальные признаки, отсутствующие данные, проектирование признаков, нормализация данных. Визуализация данных: мозаичные диаграммы, диаграмма размаха, графики плотности, диаграммы рассеяния.

Тема 3. Моделирование и прогнозирование в задачах защиты информации. Основы моделирования с машинным обучением в сфере информационной безопасности: поиск связи между входными данными и целевой переменной, методы моделирования, обучение с учителем и без. Специфика сложности классификации и предсказаний в сфере информационной безопасности. Классификация: построение классификатора и получение предсказаний, классификация сложных нелинейных данных, классификация в случае множества классов. Регрессия: построение регрессора и генерация прогнозов, регрессия для сложных нелинейных данных.

Тема 4. Оценка и оптимизация моделей в сфере информационной безопасноти. Нюансы проверки моделей в сфере информационной безопасности. Оценка прогностической точности на новых данных: проблема переобучения, скользящий контроль, перекрестная проверка. Оценка моделей классификации: таблица сопряженности, ROC-кривые, оценка многоклассовых классификаторов. Оценка моделей регрессии: показатели эффективности

регрессионных моделей, исследование остатков. Оптимизация модели путем подбора параметров: параметры настройки алгоритмов и сеточный поиск.

Раздел II. Примеры прикладных задач из сферы информационной безопасности.

Тема 5. Атаки на системы искусственного интеллекта. Атаки на обучающее множество. Возможные негативные последствия. Состязательные атаки. Детектирование состязательных атак.

Тема 6. Задача обнаружения вредоносного программного обеспечения. Методы обнаружения вредоносного программного обеспечения. Признаки, извлекаемые из вредоносных файлов. Алгоритмы обучения, применяемые для обучения антивирусных классификаторов.

Тема 7. Задача поиска аномалий пользовательского трафика. Проблема ботов. Проблема DDos-атак. Задача определения нелегитимного трафика.

Тема 8. Задача распознавания лиц. Признаки, извлекаемые из изображений лиц. Алгоритмы распознавания. Проблема распознавания. Распознавание по походке.

Тема 9. Задача распознавания «дипфейков». Морально-этические аспекты создания «дипфейков»: легальное и зловредное применение. Морфинг. «Дипфейки». Методы распознавания искусственно сгенерированных изображений.

5. МЕТОДИЧЕСКИЕ УКАЗАНИЯ ПО ПРЕПОДАВАНИЮ И ОСВОЕНИЮ ДИСЦИПЛИНЫ (МОДУЛЯ)

5.1. Указания для преподавателей по организации и проведению учебных занятий по дисциплине (модулю)

Для организаций лекционных занятий необходим проектор и доска. Для проведения лабораторных занятий требуется компьютерный класс с установленными средами разработки на языках программирования высокого уровня и программный пакет WEKA.

5.2. Указания для обучающихся по освоению дисциплины (модулю)

Таблица 4 – Содержание самостоятельной работы обучающихся

Вопросы, выносимые на самостоятельное изучение	Кол-во часов	Форма работы
Раздел I. Жизненный цикл модели машинного обучения в задачах защиты информации	**30**	**Опросы, лабораторные работы**
Тема 1. Введение в системы искусственного интеллекта в информационной безопасности	7	Подготовка к опрос
Тема 2. Обучающее множество в задачах защиты информации	8	Выполнение лабораторной работы 1
Тема 3. Моделирование и прогнозирование в задачах защиты информации	7	Выполнение лабораторной работы 2
Тема 4. Оценка и оптимизация моделей в сфере информационной безопасности	8	Выполнение лабораторной работы 3
Раздел II. Примеры прикладных задач из сферы информационной безопасности	**24**	
Тема 5. Атаки на системы искусственного интеллекта	5	Выполнение лабораторной работы 4
Тема 6. Задача обнаружения вредоносного программного обеспечения	5	Презентация реферата
Тема 7. Задача поиска аномалий пользовательского трафика	5	Выполнение лабораторной работы 5
Тема 8. Задача распознавания лиц	5	Презентация реферата

Вопросы, выносимые на самостоятельное изучение	Кол-во часов	Форма работы
Раздел I. Жизненный цикл модели машинного обучения в задачах защиты информации	**30**	**Опросы, лабораторные работы**
Тема 9. Задача распознавания «дипфейков»	4	Выполнение лабораторной работы 6

5.3. Виды и формы письменных работ, предусмотренных при освоении дисциплины (модуля), выполняемые обучающимися самостоятельно

Лабораторные работы. На ЯВУ или с помощью программной среды WEKA необходимо провести эксперимент, предложенный в рамках лабораторной работы. Для отчета требуется предоставить обученную модель и отчет с результатами эксперимента.

Опрос. Студентам требуется подготовиться к ответам на вопросы по прошедшей теме занятия.

Реферат. Студентам требуется проанализировать отечественную и зарубежную литературу и представить реферат на тему используемых подходов к решению задач информационной безопасности.

6. ОБРАЗОВАТЕЛЬНЫЕ И ИНФОРМАЦИОННЫЕ ТЕХНОЛОГИИ

При реализации различных видов учебной работы по дисциплине могут использоваться электронное обучение и дистанционные образовательные технологии.

6.1. Образовательные технологии

Таблица 5 – Образовательные технологии, используемые при реализации учебных занятий

Раздел, тема дисциплины (модуля)	Форма учебного занятия		
	Лекция	Практическое занятие, семинар	Лабораторная работа
Раздел I. Жизненный цикл модели машинного обучения в задачах защиты информации			
Тема 1. Введение в системы искусственного интеллекта в информационной безопасности	Обзорная лекция, лекция-презентация	Не предусмотрено	Фронтальный опрос
Тема 2. Обучающее множество в задачах защиты информации	Обзорная лекция, лекция-презентация	Не предусмотрено	Анализ ситуаций и имитационных моделей
Тема 3. Моделирование и прогнозирование в задачах защиты информации	Интерактивная лекция	Не предусмотрено	Анализ ситуаций и имитационных моделей
Тема 4. Оценка и оптимизация моделей в сфере информационной безопасности	Интерактивная лекция	Не предусмотрено	Анализ ситуаций и имитационных моделей
Раздел II. Примеры прикладных задач из сферы информационной безопасности			
Тема 5. Атаки на системы искусственного интеллекта	Интерактивная лекция	Не предусмотрено	Анализ ситуаций и имитационных моделей
Тема 6. Задача обнаружения вредоносного программного обеспечения	Интерактивная лекция	Не предусмотрено	Групповая дискуссия после заслушивания рефератов

Тема 7. Задача поиска аномалий пользовательского трафика	Интерактивная лекция	Не предусмотрено	Анализ ситуаций и имитационных моделей
Тема 8. Задача распознавания лиц	Интерактивная лекция	Не предусмотрено	Групповая дискуссия после заслушивания рефератов
Тема 9. Задача распознавания «дипфейков»	Интерактивная лекция	Не предусмотрено	Анализ ситуаций и имитационных моделей

6.2. Информационные технологии

При организации учебной и внеучебной работы используются возможности сети Интернет, учебные пособия и литература в электронном виде, презентации. Отправка отчетов и рефератов на проверку возможна на электронный адрес (kafedra_ib_agu@mail.ru).

При реализации различных видов учебной и внеучебной работы используются следующие информационные технологии: виртуальная обучающая среда (или система управления обучением LMS Moodle) или иные информационные системы, сервисы и мессенджеры.

6.3. Программное обеспечение, современные профессиональные базы данных и информационные справочные системы

6.3.1. Программное обеспечение

Наименование программного обеспечения	Назначение
Adobe Reader	Программа для просмотра электронных документов
Платформа дистанционного обучения LMS Moodle	Виртуальная обучающая среда
Mozilla FireFox	Браузер
Microsoft Office 2013, Microsoft Office Project 2013, Microsoft Office Visio 2013	Офисная программа
7-zip	Архиватор
Microsoft Windows 7 Professional	Операционная система
Kaspersky Endpoint Security	Средство антивирусной защиты
Microsoft Visual Studio	Среда разработки
Python	Среда разработки
Weka	Среда моделирования

6.3.2. Современные профессиональные базы данных и информационные справочные системы

– Электронный каталог Научной библиотеки АГУ на базе MARK SQL НПО «Информ-систем»: https://library.asu.edu.ru.

– Электронный каталог «Научные журналы АГУ»: http://journal.asu.edu.ru/.

– Универсальная справочно-информационная полнотекстовая база данных периодических изданий ООО «ИВИС»: http://dlib.eastview.com/

- Электронно-библиотечная система elibrary. http://elibrary.ru
- Справочная правовая система КонсультантПлюс: http://www.consultant.ru
- Информационно-правовое обеспечение «Система ГАРАНТ»: http://garant-astrakhan.ru

7. ФОНД ОЦЕНОЧНЫХ СРЕДСТВ ДЛЯ ПРОВЕДЕНИЯ ТЕКУЩЕГО КОНТРОЛЯ И ПРОМЕЖУТОЧНОЙ АТТЕСТАЦИИ ПО ДИСЦИПЛИНЕ (МОДУЛЮ)

7.1. Паспорт фонда оценочных средств

При проведении текущего контроля и промежуточной аттестации по дисциплине (модулю) «Системы искусственного интеллекта в информационной безопасности» проверяется сформированность у обучающихся компетенций, указанных в разделе 3 настоящей программы. Этапность формирования данных компетенций в процессе освоения образовательной программы определяется последовательным освоением дисциплин (модулей) и прохождением практик, а в процессе освоения дисциплины (модуля) – последовательным достижением результатов освоения содержательно связанных между собой разделов, тем.

Таблица 6 – Соответствие разделов, тем дисциплины (модуля), результатов обучения по дисциплине (модулю) и оценочных средств

Контролируемый раздел, тема дисциплины (модуля)	Код контролируемой компетенции	Наименование оценочного средства
Раздел I. Жизненный цикл модели машинного обучения в задачах защиты информации	ПК-3, ПК-4	Лабораторные работы, опрос
Тема 1. Введение в системы искусственного интеллекта в информационной безопасности	ПК-3, ПК-4	Опрос
Тема 2. Обучающее множество в задачах защиты информации	ПК-3, ПК-4	Лабораторная работа 1
Тема 3. Моделирование и прогнозирование в задачах защиты информации	ПК-3, ПК-4	Лабораторная работа 2
Тема 4. Оценка и оптимизация моделей в сфере информационной безопасности	ПК-3, ПК-4	Лабораторная работа 3
Раздел II. Примеры прикладных задач из сферы информационной безопасности	ПК-3, ПК-4	Лабораторные работы, реферат
Тема 5. Атаки на системы искусственного интеллекта	ПК-3, ПК-4	Лабораторная работа 4
Тема 6. Задача обнаружения вредоносного программного обеспечения	ПК-3, ПК-4	Реферат
Тема 7. Задача поиска аномалий пользовательского трафика	ПК-3, ПК-4	Лабораторная работа 5
Тема 8. Задача распознавания лиц	ПК-3, ПК-4	Реферат
Тема 9. Задача распознавания «дипфейков»	ПК-3, ПК-4	Лабораторная работа 6

7.2. Описание показателей и критериев оценивания компетенций, описание шкал оценивания

Таблица 7 – Показатели оценивания результатов обучения в виде знаний

Шкала оценивания	Критерии оценивания
5	демонстрирует глубокое знание теоретического материала, умение

Шкала оценивания	Критерии оценивания
«отлично»	обоснованно излагать свои мысли по обсуждаемым вопросам, способность полно, правильно и аргументированно отвечать на вопросы, приводить примеры
4 «хорошо»	демонстрирует знание теоретического материала, его последовательное изложение, способность приводить примеры, допускает единичные ошибки, исправляемые после замечания преподавателя
3 «удовлетвори тельно»	демонстрирует неполное, фрагментарное знание теоретического материала, требующее наводящих вопросов преподавателя, допускает существенные ошибки в его изложении, затрудняется в приведении примеров и формулировке выводов
2 «неудовлетво рительно»	демонстрирует существенные пробелы в знании теоретического материала, не способен его изложить и ответить на наводящие вопросы преподавателя, не может привести примеры

Таблица 8 – Показатели оценивания результатов обучения в виде умений и владений

Шкала оценивания	Критерии оценивания
5 «отлично»	демонстрирует способность применять знание теоретического материала при выполнении заданий, последовательно и правильно выполняет задания, умеет обоснованно излагать свои мысли и делать необходимые выводы
4 «хорошо»	демонстрирует способность применять знание теоретического материала при выполнении заданий, последовательно и правильно выполняет задания, умеет обоснованно излагать свои мысли и делать необходимые выводы, допускает единичные ошибки, исправляемые после замечания преподавателя
3 «удовлетвори тельно»	демонстрирует отдельные, несистематизированные навыки, испытывает затруднения и допускает ошибки при выполнении заданий, выполняет задание по подсказке преподавателя, затрудняется в формулировке выводов
2 «неудовлетво рительно»	не способен правильно выполнить задания

7.3. Контрольные задания и иные материалы, необходимые для оценки результатов обучения по дисциплине (модулю)

Раздел I. Жизненный цикл модели машинного обучения в задачах защиты информации

Тема 1. Введение в системы искусственного интеллекта в информационной безопасности

Опрос

1. Исследование предметной области. Формулировка задачи.
2. Объекты, признаки, виды признаков, целевая переменная.
3. Обучающее и проверочное множества.
4. Обучение модели
5. Алгоритмы обучения
6. Оценка качество модели
7. Оптимизация модели.

Тема 2. Обучающее множество в задачах защиты информации

Лабораторная работа 1.

Цель: разработать классификатор, который мог бы отличать корректные почтовые сообщения от спама.

Задачи:

1. Найдите или возьмите предложенный датасет.
2. Визуализируйте аналитику по датасету. В каких пропорциях представлены объекты в обучающем множестве?
3. Выберите подходящий алгоритм распознавания.
4. Обучите классификатор с использованием перекрестной проверки.
5. Для каждого классификатора вычислите:
   - матрицу сопряженности
   - верность(точность) и частоту ошибок
   - TP, TN, FP, FN - частоту ошибок и правильного распознавания каждого класса
6. Повторите шаг 3, построив модель с использованием других алгоритмов.
7. Для классификаторов, обученных с использованием разных алгоритмов, постройте график покрытия. Определите по нему наилучший классификатор.

Тема 3. Моделирование и прогнозирование в задачах защиты информации

Лабораторная работа 2

Цель: разработать регрессор, который мог бы предсказывать какое-либо фактическое значение.

Задачи:

1. Найдите или возьмите предложенный датасет.
2. Выберите подходящий алгоритм прогнозирования.
3. Оцените качество прогнозирования с использованием квадратного корня из среднеквадратичной ошибки и с использованием анализа остатков.
4. Сделайте вывод, какие именно признаки оказывают наиболее существенное влияние на целевую переменную.

Тема 4. Оценка и оптимизация моделей в сфере информационной безопасности

Лабораторная работа 3.

Цель: Обучить кластеризатор, который смог бы определять категории фейковых новостей.

Задачи:

1. Найдите или возьмите предложенный датасет.
2. Выберите подходящий алгоритм кластеризации.
3. На какие кластеры были разбиты данные? Сколько кластеров получилось? В чем вы не согласны с предложенным разбиением?

Раздел II. Машинное обучение в задачах информационной безопасности

Тема 5. Атаки на системы искусственного интеллекта

Лабораторная работа 4.

Цель: сымитировать состязательную атаку на модель компьютерного зрения и реализовать метод детектирования.

Задачи:

1. Разработать модель компьютерного зрения, которая бы классифицировала объекты, изображенные на фотографиях.
2. Провести состязательную атаку на собственную модель
3. Разработать и проверить способы детектирования состязательных атак.

Тема 6. Задача обнаружения вредоносного программного обеспечения

Реферат 1. Различные подходы к обнаружению вредоносного программного обеспечения с использованием методов машинного обучения.

Тема 7. Задача поиска аномалий пользовательского трафика

Лабораторная работа 5.

Цель: Выявить основные закономерности в предложенной статистике по посещаемости web-сайта.

Задачи:

4. Определить основные тенденции в посещаемости сайта в зависимости от времени суток, от дня недели, от дня месяца.
5. Проверить как выявленные тенденции соответствуют часовому поясу пользователей.
6. Выявить аномалии в трафике.

Тема 8. Задача распознавания лиц

Реферат 2. Различные подходы к распознаванию лиц на изображениях и в видео.

Тема 9. Задача распознавания «дипфейков»

Лабораторная работа 6.

Цель: искусственно сгенерировать изображение и разработать метод детектирования.

Задачи:

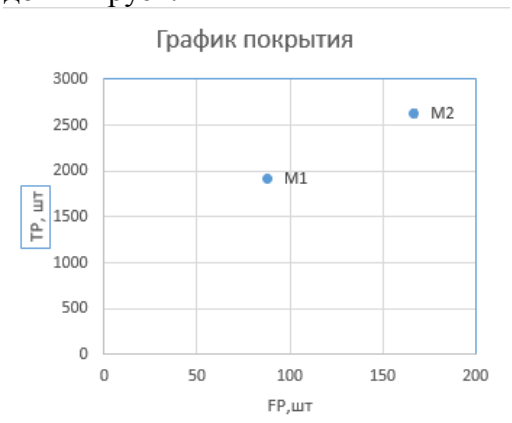
1. Искусственно сгенерировать изображение/видео.
2. Разработать и проверить способы детектирования искусственно сгенерированных изображений.

Перечень вопросов и заданий,
выносимых на экзамен / зачёт / дифференцированный зачёт

Таблица 9 – Примеры оценочных средств с ключами правильных ответов

№ п/п	Тип задания	Формулировка задания	Правильный ответ	Время выполнения (в минутах)
ПК-3				
1.	Задание закрытого типа	Что такое нормализация данных? 1. Усреднение данных 2. Преобразование категориальных признаков в численные 3. Преобразование численных признаков в категориальные 4. Подгонка под единую шкалу	3	1
2.		Укажите соответствие между типами входных/целевых признаков и диаграммой, которую целесообразно использовать для визуализации 1. Входной признак-категориальный, целевая переменная-категориальная 2. Входной признак-категориальный, Целевая переменная- числовая 3. Входной признак- числовой, Целевая переменная-категориальная 4. Входной признак- числовой,	1-d 2-c 3-b 4-a	5

№ п/п	Тип задания	Формулировка задания	Правильный ответ	Время выполнения (в минутах)
		Целевая переменная- числовая a. Диаграмма рассеяния b. Диаграмма размаха c. График плотности d. Мозаичная диаграмма		
3.		Если вам необходимо, рассортировать содержимое корзины с фруктами, то какую задачу вы будете решать? 1. Понижения размерности 2. Регрессии 3. Классификации 4. Кластеризации	4	2
4.		Для оценки эффективности регрессора применяют: 1. Точность 2. Верность 3. Долю истинно положительных результатов 4. Квадратный корень из среднеквадратичной ошибки 5. Частотой ошибки	3	3
5.		Какой алгоритм основан на гипотезе «Набор слабых обучающих алгоритмов спсобен создать сильный обучающий алгоритм»? 1. Бустинг 2. Случайный лес 3. Нейронные сети 4. Наивный Байес	1	3
6.	Задание открытого типа	Что делать в случае, если в обучающем множестве отсутствуют какие-либо данные	Существует несколько стратегий: • создать новую категорию для отсутствующих данных • удалить экземпляры с отсутствующими данными • подставить значение предшествующего экземпляра • заместить отсутствующее значение	5

№ п/п	Тип задания	Формулировка задания	Правильный ответ	Время выполнения (в минутах)
			средним значением столбца • заместить с помощью модели МО	
7.		Что такое точность классификации?	Точность-это доля правильно распознанных экземпляров.	2
8.		Какие признаки называются категориальными?	Признаки называются категориальными, если их можно отнести к какой-либо группе, но при этом не важен порядок	2
9.		В чем выражается проблема переобучения?	Модель эффективно работает только с теми данными, на которых была обучена	3
10.		В проверочном множестве 800 объектов: 300 объектов класса *a*, 500 объектов класса *b*. Правильно были распознаны 275 объектов класса *a* и 480 объектов класса *b*. Рассчитайте верность (accuracy) классификатора.	0,94	3
ПК-4				
11.	Задание закрытого типа	На рисунке представлен график покрытия, изображающий сравнение двух моделей. Какая модель доминирует?  1) M1	3	3

№ п/п	Тип задания	Формулировка задания	Правильный ответ	Время выполнения (в минутах)
		2) M2 3) Ни одна		
12.		ROC-кривая описывает… 1) зависимость ложных срабатываний от размера обучающего множества 2) чувствительность модели к разным порогам классификации 3) качество модели при тестировании с использованием различных проверочных множеств 4) частоту истино положительных ответов модели	2	4
13.		Как связаны между собой AUC и ROC 1) ROC-«кривая ошибок», а AUC – площадь под ней 2) ROC- «кривая правильных ответов» , а AUC- площадь над ней 3) ROC – график, а AUC- оптимальное пороговое значение 4) AUC-«кривая ошибок», а ROC- площадь под ней 5) AUC - «кривая правильных ответов», а ROC - площадь над ней 6) AUC – график, а ROC - оптимальное пороговое значение	1	3
14.		Выберите верное утверждение: 1) Чем ближе к 1 индекс корреляции, тем выше качество модели множественной регрессии. 2) Чем ближе к 0 коэффициент детерминации, тем выше качество модели множественной регрессии. 3) Независимость остатков проверяется с помощью критерия Дарбина – Уотсона. 4) Качество регрессора характеризуется фактом обоснованной зависимости	1, 3	5

№ п/п	Тип задания	Формулировка задания	Правильный ответ	Время выполнения (в минутах)
		остатков от целевой переменной.		
15.		Для решения каких задач информационной безопасности обычно используются методы машинного обучения? 1) Распознавание лиц 2) Составление модели нарушителя 3) Прогнозирование числа кибер-атак 4) Обнаружение аномалий в сетевом трафике 5) Построение модели угроз	1,4	5
16.		Укажите соответствие между мерами схожести и формулами по которым они вычисляются 1) Евклидово расстояние 2) Квадрат евклидова расстояния 3) Расстояние городских кварталов (манхэттенское расстояние) 4) Расстояние Чебышева a) $\rho(x,x') = \max(\lvert x_i - x_i' \rvert)$ b) $\rho(x,x') = \sum_i^n (x_i - x_i')^2$ c) $\rho(x,x') = \sqrt{\sum_i^n (x_i - x_i')^2}$ d) $\rho(x,x') = \sum_i^n \lvert x_i - x_i' \rvert$	1) -c 2) -b 3) -d 4)-a	5
17.	Задания открытого типа	Какая информация хранится в матрицах сопряженности по результатам тестирования классификатора?	Информация о правильно и неправильно распознанных объектах каждого класса: FP, TP, FN, TN	5
18.		Перечислите основные типы алгоритмов кластеризации	• Иерархический • k-средних • c-средних • Выделение связных компонент • Минимальное покрывающее	5

№ п/п	Тип задания	Формулировка задания	Правильный ответ	Время выполнения (в минутах)
			дерево • Послойная кластеризация	

7.4. Методические материалы, определяющие процедуры оценивания результатов обучения по дисциплине (модулю)

Таблица 10 – Технологическая карта рейтинговых баллов по дисциплине (модулю)

№ п/п	Контролируемые мероприятия	Количество мероприятий / баллы	Максимальное количество баллов	Срок представления
Основной блок				
1.	*Выполнение лабораторной работы*	5/7	35	2 недели
2.	*Ответ во время опроса*	1/3	3	2 недели
3.	*Реферат*	2/6	12	2 недели
Всего			**50**	-
Блок бонусов				
4.	*Посещение занятий*	0,25/24	6	
5.	*Своевременное выполнение всех заданий*	0,5/8	4	
Всего			**10**	-
Дополнительный блок**				
6.	*Экзамен*		40	
Всего			**40**	-
ИТОГО			**100**	-

Таблица 11 – Система штрафов (для одного занятия)

Показатель	Балл
Опоздание на занятие	-0,5
Нарушение учебной дисциплины	-5
Неготовность к занятию	-1
Пропуск занятия без уважительной причины	-2

Таблица 12 – Шкала перевода рейтинговых баллов в итоговую оценку за семестр по дисциплине (модулю)

Сумма баллов	Оценка по 4-балльной шкале
90–100	5 (отлично)
85–89	4 (хорошо)
75–84	
70–74	
65–69	3 (удовлетворительно)
60–64	
Ниже 60	2 (неудовлетворительно)

При реализации дисциплины (модуля) в зависимости от уровня подготовленности обучающихся могут быть использованы иные формы, методы контроля и оценочные средства, исходя из конкретной ситуации.

7. УЧЕБНО-МЕТОДИЧЕСКОЕ И ИНФОРМАЦИОННОЕ ОБЕСПЕЧЕНИЕ ДИСЦИПЛИНЫ (МОДУЛЯ)

8.

8.1. Основная литература

1. Кольер, Р. Машинное обучение в Elastic Stack / Р. Кольер, К. Монтонен, Б. Азарми; пер. с англ. В. С. Яценкова. - Москва : ДМК Пресс, 2021. - 380 с. - ISBN 978-5-93700-107-8. - Текст : электронный // ЭБС "Консультант студента" : [сайт]. - URL : https://www.studentlibrary.ru/book/ISBN9785937001078.html (дата обращения: 21.12.2022). - Режим доступа : по подписке.Андреева Г. М. Социальная психология: учебник. М.: Аспект Пресс, 2002. 364 с. (23 экз.).

2. Машинное обучение с использованием библиотеки H2O / Д. Кук - Москва : ДМК Пресс, 2018. - ISBN 978-5-97060-508-0. - Текст : электронный // ЭБС "Консультант студента" : [сайт]. - URL : https://www.studentlibrary.ru/book/ISBN9785970605080.html (дата обращения: 21.12.2022). - Режим доступа : по подписке.

3. Машинное обучение с использованием библиотеки H2O / Д. Кук - Москва : ДМК Пресс, 2018. - ISBN 978-5-97060-508-0. - Текст : электронный // ЭБС "Консультант студента" : [сайт]. - URL : https://www.studentlibrary.ru/book/ISBN9785970605080.html (дата обращения: 21.12.2022). - Режим доступа : по подписке.

4. Python и машинное обучение / С. Рашка - Москва : ДМК Пресс, 2017. - ISBN 978-5-97060-409-0. - Текст : электронный // ЭБС "Консультант студента" : [сайт]. - URL : https://www.studentlibrary.ru/book/ISBN9785970604090.html (дата обращения: 21.12.2022). - Режим доступа : по подписке.

5. Python и машинное обучение / С. Рашка - Москва : ДМК Пресс, 2017. - ISBN 978-5-97060-409-0. - Текст : электронный // ЭБС "Консультант студента" : [сайт]. - URL : https://www.studentlibrary.ru/book/ISBN9785970604090.html (дата обращения: 21.12.2022). - Режим доступа : по подписке.

6. Горбаченко, В. И. Машинное обучение : учебное пособие / В. И. Горбаченко, К. Е. Савенков, М. А. Малахов. — Москва : Ай Пи Ар Медиа, 2023. — 217 с. — ISBN 978-5-4497-1860-0. — Текст : электронный // Цифровой образовательный ресурс IPR SMART : [сайт]. — URL: https://www.iprbookshop.ru/125886.html (дата обращения: 28.11.2022). — Режим доступа: для авторизир. пользователей. - DOI: https://doi.org/10.23682/125886

7. Павлова, А. И. Искусственные нейронные сети : учебное пособие / А. И. Павлова. — Москва : Ай Пи Ар Медиа, 2021. — 190 с. — ISBN 978-5-4497-1165-6. — Текст : электронный // Цифровой образовательный ресурс IPR SMART : [сайт]. — URL: https://www.iprbookshop.ru/108228.html (дата обращения: 21.12.2022). — Режим доступа: для авторизир. пользователей. - DOI: https://doi.org/10.23682/108228

8. Барский, А. Б. Введение в нейронные сети : учебное пособие / А. Б. Барский. — 3-е изд. — Москва, Саратов : Интернет-Университет Информационных Технологий (ИНТУИТ), Ай Пи Ар Медиа, 2020. — 357 с. — ISBN 978-5-4497-0309-5. — Текст : электронный // Цифровой образовательный ресурс IPR SMART : [сайт]. — URL: https://www.iprbookshop.ru/89426.html (дата обращения: 21.12.2022). — Режим доступа: для авторизир. пользователей

8.2. Дополнительная литература

1. Паттерсон, Дж. , Гибсон А. Глубокое обучение с точки зрения практика / Паттерсон Дж. , Гибсон А. , пер. с анг. А. А. Слинкина. - Москва : ДМК Пресс, 2018. - 418 с. - ISBN 978-5-97060-481-6. - Текст : электронный // ЭБС "Консультант студента" : [сайт]. - URL : https://www.studentlibrary.ru/book/ISBN9785970604816.html (дата обращения: 21.12.2022). - Режим доступа : по подписке.

2. (Манро), Р. Машинное обучение с участием человека / Монарх Р. (Манро) ; перевод В. И. Бахур. — Москва : ДМК Пресс, 2022. — 498 с. — ISBN 978-5-97060-934-7. —

Текст : электронный // Цифровой образовательный ресурс IPR SMART : [сайт]. — URL: https://www.iprbookshop.ru/125122.html (дата обращения: 20.10.2022). — Режим доступа: для авторизир. Пользователей

3. Яхъяева, Г. Э. Нечеткие множества и нейронные сети : учебное пособие / Г. Э. Яхъяева. — 4-е изд. — Москва : Интернет-Университет Информационных Технологий (ИНТУИТ), Ай Пи Ар Медиа, 2020. — 315 с. — ISBN 978-5-4497-0665-2. — Текст : электронный // Цифровой образовательный ресурс IPR SMART : [сайт]. — URL: https://www.iprbookshop.ru/97552.html (дата обращения: 21.12.2022). — Режим доступа: для авторизир. Пользователей

4. Барский, А. Б. Искусственный интеллект и логические нейронные сети : учебное пособие / А. Б. Барский. — Санкт-Петербург : Интермедия, 2019. — 360 с. — ISBN 978-5-4383-0155-4. — Текст : электронный // Цифровой образовательный ресурс IPR SMART : [сайт]. — URL: https://www.iprbookshop.ru/95270.html (дата обращения: 21.12.2022). — Режим доступа: для авторизир. пользователей

8.3. Интернет-ресурсы, необходимые для освоения дисциплины (модуля)

1. Электронно-библиотечная система (ЭБС) ООО «Политехресурс» «Консультант студента». Многопрофильный образовательный ресурс «Консультант студента» является электронной библиотечной системой, предоставляющей доступ через сеть Интернет к учебной литературе и дополнительным материалам, приобретенным на основании прямых договоров с правообладателями. Каталог в настоящее время содержит около 15000 наименований. www.studentlibrary.ru.

2. Kaggle. Система организации конкурсов по исследованию данных, а также социальная сеть специалистов по обработке данных и машинному обучению. www.kaggle.com

9. МАТЕРИАЛЬНО-ТЕХНИЧЕСКОЕ ОБЕСПЕЧЕНИЕ ДИСЦИПЛИНЫ (МОДУЛЯ)

Для проведения лекционных занятий необходима мультимедийная аудитория, оснащенная компьютерной презентационной техникой.

Для проведения лабораторных занятий необходима аудитория, оснащенная компьютерами.

Рабочая программа дисциплины (модуля) при необходимости может быть адаптирована для обучения (в том числе с применением дистанционных образовательных технологий) лиц с ограниченными возможностями здоровья, инвалидов. Для этого требуется заявление обучающихся, являющихся лицами с ограниченными возможностями здоровья, инвалидами, или их законных представителей и рекомендации психолого-медико-педагогической комиссии. Для инвалидов содержание рабочей программы дисциплины (модуля) может определяться также в соответствии с индивидуальной программой реабилитации инвалида (при наличии).