

Паспорт
оценочных материалов для проведения текущего контроля и
промежуточной аттестации обучающихся по дисциплине (модулю)
Прикладные задачи анализа данных

Перечень оценочных материалов и индикаторов достижения компетенций, сформированность которых они контролируют¹

Наименование оценочного средства	Коды индикаторов достижения формируемых компетенции	Номер приложения ²
Выполнение практических работ	ИД-1 _{ПК-1} ИД-2 _{ПК-1} ИД-3 _{ПК-1} ИД-4 _{ПК-1}	1
Экзамен	ИД-1 _{ПК-1} ИД-2 _{ПК-1} ИД-3 _{ПК-1} ИД-4 _{ПК-1}	2

Утверждено на заседании кафедры «Вычислительная техника»
протокол №3 от «11» октября 2021 года
Заведующий кафедрой _____ К.В.Святов

¹ Перечисляются все оценочные материалы, указанные в рабочей программе дисциплины.

² Указывается порядковый номер приложения, в котором размещены оценочные средства. Нумерация изменяется в зависимости от имеющихся оценочных средств.

Выполнение практических работ

1. Процедура выполнения практических работ

Количество проводимых практических работ в течение всего периода освоения дисциплины	8 работ
Формат проведения результатов	Электронный
Методические рекомендации (при необходимости)	Рекомендуется представлять выполненную практическую работу в виде ноутбука в формате .ipynb (или ссылки на ноутбук в Google Colab) Приведенный код должен быть аккуратно оформлен, прокомментирован для удобства его проверки, сопровождаться текстовыми блоками с обоснованиями использованных методов и анализом полученных результатов.

2. Шкала оценивания с учетом срока сдачи³

Показатель качества исполнения работы	Балл (по 100-балльной шкале)
Практическая работа выполнена в полном объеме и в срок; результаты работы скрипта/приложения корректны на тестовых примерах или проведен требуемый вычислительный эксперимент; результаты работы представлены преподавателю; исполнитель может объяснить действия команд и внести простые изменения в алгоритм по требованию преподавателя.	90-100
Практическая работа выполнена практически в полном объеме и в срок; результаты работы скрипта/приложения корректны на тестовых примерах или проведен требуемый вычислительный эксперимент; результаты работы представлены преподавателю; исполнитель может объяснить действия команд программы, возникают проблемы с быстрой модификацией алгоритмов по требованию преподавателя.	60-89
Работа выполнена не в полном объеме (скрипт/приложение работает некорректно); исполнитель может объяснить действия команд программы, однако не может модифицировать код по требованию преподавателя.	40-59
Работа не выполнена или выполнена не в полном объеме (не проведены	< 40

³ За несвоевременную сдачу обучающемуся могут быть начислены штрафные баллы.

Показатель качества исполнения работы	Балл (по 100-балльной шкале)
заданные вычислительные эксперименты); результаты работы не представлены преподавателю или представлены с существенным нарушением срока; исполнитель не может объяснить действия команд программы и не может внести простые изменения в алгоритм по требованию преподавателя.	

3. Перечень практических работ

Практическая работа 1. Прогнозирование временных рядов

Задание:

1. Построить график временного ряда.
2. Проверить временной ряд на автокорреляцию.
3. Проверить временной ряд на стационарность.
4. Выбрать подходящие подходы и построить точечный и интервальный прогнозы.

Практическая работа 2. Анализ поведения пользователей

Задание: имеется массив данных, характеризующих отклик пользователя на письмо: id_пользователя, дата и время получения письма, признак конверсии (0 – нет, 1 – да), а также ряд дополнительных признаков в зависимости от варианта. Спрогнозировать наилучшую маркетинговую активность:

1. Подготовить данные (обнаружить и обработать аномальные и пропущенные данные, провести фильтрацию и трансформацию данных).
2. Разделить выборку на тестовую и обучающую.
3. Агрегировать информацию по каждому пользователю.
4. Выбрать и обучить модель, которая позволит спрогнозировать, откроет письмо пользователь или нет.
5. Оценить качество построенной модели.

Практическая работа 3. Классификация изображений

Задание: Набор данных: изображения с соревнования по машинному обучению под названием "Dogs vs Cats" на [kaggle.com](https://www.kaggle.com).

1. Импортировать нейронную сеть (согласно варианту) из библиотеки Keras.
2. Настроить сеть.
3. Провести тонкую настройку нижнего свёрточного слоя нейронной сети.
4. Применить аугментацию данных при помощи ImageDataGenerator.
5. Обучить сеть.
6. Провести оценку точности модели.

Практическая работа 4. Распознавание лиц

Задание:

1. Подключить OpenCV к проекту.
2. Реализовать обнаружение лиц в кадре.

3. Реализовать извлечение лица из кадра.
4. Обучить модель распознавания лиц сокурсников.
5. Написать скрипт для распознавания лиц сокурсников в кадре, используя обученную модель.

Практическая работа 5. Детекция объектов

Задание: Напишите скрипт, позволяющий находить объект в кадре и определять его смещение. Решите следующие задачи:

- поиск границ объекта;
- детектирование структурных элементов;
- определение положения объекта.

Практическая работа 6. Анализ тональности текста

Задание: на основе одного из открытых русскоязычных наборов данных разработать скрипт, содержащий обученную модель для анализа тональности.

1. Провести предварительную подготовку текстов (лемматизация, стеммирование и др.).
2. Векторизовать подготовленный текст.
3. Выбрать и обучить классификатор.
4. Оценить точность модели.

Практическая работа 7. Генерация текстов

Задание: написать скрипт, содержащий обученную модель для автоматической генерации новостей с заданной тематикой (согласно вариантам).

1. Проработать устройство генератора.
2. Ознакомиться с особенностями новостей по заданной тематике.
3. Выбрать технологию обработки данных о заданной тематике.
4. Выбрать подход к реализации автоматической генерации текста.
5. Написать скрипт.
6. Сгенерировать при помощи скрипта 5 текстов новостей.
7. Переслать тексты новостей 3 однокурсников для оценки текстов.
8. Оценить тексты новостей однокурсников.
9. Подготовить ноутбук, содержащий отчет о проделанной работе.

Практическая работа 8. Рекомендательные системы

Задание: на основе набора данных из варианта разработать скрипт, содержащий 3 реализованных алгоритма рекомендательных систем (Content-based, Neighborhood подход в коллаборативной фильтрации, Latent factor подход в коллаборативной фильтрации), сравнить их скорость и качество получаемых рекомендаций.

Экзамен

1. Процедура проведения

Общее количество вопросов к экзамену	26 вопросов
Количество вопросов в билете	2 вопросов
Наличие задач в билете	нет
Формат проведения	Устно и письменно
Методические рекомендации (при необходимости)	Экзамен по билетам проводится, как правило, для тех обучающихся, которые не смогли своей текущей работой в семестре продемонстрировать высокую активность и хорошие результаты. Цель подготовки к ответам на вопросы билетов – закрепление базовых теоретических понятий дисциплины, а также типовых практико-ориентированных знаний и навыков в рассматриваемой области.

2. Шкала оценивания с учетом текущего контроля работы обучающегося в семестре

Критерии оценки уровня сформированности компетенций по дисциплине	Оценка
Обучающийся при текущей работе в семестре проявил высокую активность и продемонстрировал высокие результаты выполнения заданий (>92 по 100-балльной шкале); зачет рекомендуется выставить на основе краткого итогового собеседования по тем вопросам, которые вызвали у обучающегося наибольший интерес	Отлично
Обучающийся при текущей работе в семестре проявил активность, но продемонстрировал нестабильные результаты выполнения заданий; ответы на вопросы на зачете представляются достаточно полными	Хорошо
Обучающийся при текущей работе в семестре проявил низкую активность и продемонстрировал невысокие результаты выполнения заданий (<80 по 100-балльной шкале); ответы на вопросы на зачете не обладают полнотой, но обучающийся в целом понимает основные положения теоретического курса и имеет общее представление о способах решения практических задач в рассматриваемой области	Удовлетворительно
Обучающийся при текущей работе в семестре проявил	Неудовлетворительно

Критерии оценки уровня сформированности компетенций по дисциплине	Оценка
низкую активность и продемонстрировал невысокие результаты выполнения заданий; ответы на вопросы на зачете демонстрируют непонимание обучающимся основных положений теоретического курса и отсутствие общего представления о способах решения практических задач в рассматриваемой области	

3. Вопросы

1. Прогнозированием временных рядов.
2. Анализ поведения пользователей. Аудиторные метрики. Прогнозирование оттока пользователей.
3. Прикладные задачи анализа медиаинформации
4. Задачи компьютерного зрения
5. «Низкоуровневое» зрение: арифметические операции, эквализация гистограммы, блендинг, цветовые пространства, каскады Хаара — детектор лиц, сегментация
6. Линейная фильтрация изображений: скользящее среднее — свертка, детекция границ, корреляция
7. Классификация изображений. Прикладные задачи классификации изображений
8. Распознавание лиц
9. Детекция объектов
10. Стилизация изображений
11. Прикладные задачи анализа текстов на естественном языке
12. Первичная обработка текстовых данных
13. Этапы работы с текстом
14. Извлечение признаков из текста
15. Языковые модели
16. Классификация спама
17. Распределение текстов по топикам (задача со многими классами)
18. Анализ тональности текста
19. Аннотирование
20. Генерация текстов
21. Прикладные задачи рекомендательных систем
22. Задача ранжирования. Метрики качества ранжирования.
23. Рекомендательные системы. Подходы к построению.
24. Признаковые описания пользователей и объектов. Обучение с учителем.
25. Коллаборативная фильтрация.
26. Факторизационные машины.