

Министерство науки и высшего образования Российской Федерации
НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ
ТОМСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ (НИ ТГУ)

Институт прикладной математики и компьютерных наук

УТВЕРЖДЕНО:

Директор
А. В. Замятин

Оценочные материалы по дисциплине

Нейронные сети

по направлению подготовки

01.03.02 Прикладная математика и информатика

Направленность (профиль) подготовки:

Прикладная математика и инженерия цифровых проектов

Форма обучения

Очная

Квалификация

Бакалавр

Год приема

2024

СОГЛАСОВАНО:

Руководитель ОП
Д.Д. Даммер

Председатель УМК
С.П. Сущенко

1. Компетенции и индикаторы их достижения, проверяемые данными оценочными материалами

Целью освоения дисциплины является формирование следующих компетенций:

ОПК-1 Способен применять фундаментальные знания, полученные в области математических и (или) естественных наук, и использовать их в профессиональной деятельности.

ОПК-2 Способен использовать и адаптировать существующие математические методы и системы программирования для разработки и реализации алгоритмов решения прикладных задач.

ОПК-5 Способен разрабатывать алгоритмы и компьютерные программы, пригодные для практического применения.

ПК-2 Способен собирать, обрабатывать и анализировать данные для проведения научно-исследовательских работ в зависимости от проблемной и предметной области, создавать математическую модель исследуемого объекта.

Результатами освоения дисциплины являются следующие индикаторы достижения компетенций:

ИОПК-1.4 Демонстрирует понимание и навыки применения на практике математических моделей и компьютерных технологий для решения практических задач, возникающих в профессиональной деятельности

ИОПК-2.1 Обладает навыками объектно-ориентированного программирования для решения прикладных задач в профессиональной деятельности.

ИОПК-2.2 Проявляет навыки использования основных языков программирования, основных методов разработки программ, стандартов оформления программной документации.

ИОПК-2.3 Демонстрирует умение отбора среди существующих математических методов, наиболее подходящих для решения конкретной прикладной задачи.

ИОПК-2.4 Демонстрирует умение адаптировать существующие математические методы для решения конкретной прикладной задачи.

ИОПК-5.1 Обладает необходимыми знаниями алгоритмов, принципов разработки алгоритмов и компьютерных программ

ИОПК-5.2 Разрабатывает алгоритмы и компьютерные программы для решения задач профессиональной деятельности

ИПК-2.2 Способен строить математическую модель исследуемого объекта и/или процесса в зависимости от проблемной и предметной области

2. Оценочные материалы текущего контроля и критерии оценивания

Элементы текущего контроля:

- тесты;
- контрольная работа.

Тест

1. Какая нейросетевая модель из перечисленных в лучшей степени подходит для прогнозирования временных последовательностей?
 - а) Single-Layer Perceptron
 - б) CNN
 - в) LSTM
 - г) Multi-layer Perceptron
2. Почему модели на сверточных нейронных сетях показывают наилучшие показатели по классификации объектов на изображениях по сравнению с другими моделями?
 - а) Они в высокой степени оптимизированы для обработки векторов с числовыми, а не категориальными признаками

- б) Они обладают широким набором инструментов преобразования признаков пространства, которые может варьировать разработчик в модели
 - в) Они учитывают корреляцию смежных компонент вектора
 - г) Они используют существенно **большее** число настраиваемых параметров, по сравнению с другими моделями
3. Каким главным недостатком обладает рекуррентная нейронная сеть?
- а) Длительная процедура обучения
 - б) Невозможность обучения на категориальных данных
 - в) Сложность запоминания длительных последовательностей
 - г) Использование существенных вычислительных ресурсов
4. Какие меры не приводят к уменьшению переобучения нейросетевой модели?
- а) Установка штрафов за большие значения весов нейронов сети
 - б) Увеличение количества слоев сети
 - в) Добавление шума в выборку
 - г) Уменьшение количества нейронов сети
5. От чего в большей степени зависит успешное решение задачи классификации однослойным персептроном?
- а) от размера выборки
 - б) от размерности признакового пространства
 - в) соотношения разделения выборки на обучающую и тестовую
 - г) от распределения объектов в пространстве признаков

Ключи: 1 в), 2 в), 3 в), 4 б), 5 г).

Критерии оценивания: тест считается пройденным, если обучающий ответил правильно как минимум на половину вопросов.

Контрольная работа

Контрольная работа состоит из 2 теоретических вопросов и 3 задач.

Перечень теоретических вопросов:

1. Опишите три различных метода, с помощью которых можно точно настроить предварительно обученную CNN ImageNet.
2. Опишите, как работает нейронная передача стиля (Neural Style Transfer)
3. Специалист по данным предполагает, что: а) Операция свертки является как линейной, так и инвариантной к сдвигу. б) Операция свертки похожа на корреляцию, за исключением того, что мы переворачиваем фильтр перед применением оператора корреляции. в) Операция свертки достигает максимума только в тех случаях, когда фильтр в основном похож на определенную часть входного сигнала. Прав ли он, предполагая это? Подробно объясните значение этих утверждений.
4. Дано изображение размером $w \times h$ и ядро шириной K . Сколько умножений и сложений требуется для свертки изображения?
5. Верность (Ассигасу) персептрона рассчитывается как количество правильно классифицированных образцов, деленное на общее количество неправильно классифицированных образцов.
6. Каково наиболее распространенное применение слоев Мах-пулинга?
7. Что такое batch-нормализация?
8. Что на самом деле означает термин стохастический в стохастическом градиентном спуске? Использует ли он какой-либо генератор случайных чисел?
9. Объясните, почему при стохастическом градиентном спуске количество эпох, необходимых для преодоления определенного порога потерь, увеличивается по мере уменьшения размера пакета?

10. Как работает обучение с учетом момента? Объясните роль экспоненциального затухания в правиле обновления градиентного спуска.
11. В чем разница между функциями Sigmoid и Softmax?
12. В чем разница между рекуррентной нейронной сетью и нейронной сетью прямого распространения?
13. В чем разница между рекуррентной нейронной сетью и нейронной сетью прямого распространения?
14. Для чего можно использовать рекуррентную нейронную сеть (RNN)?
15. Как выбирается функция потерь?
16. Что такое затухающий градиент в глубоком обучении? Что такое взрывной градиент в глубоком обучении?
17. Из каких частей состоит автоэнкодер?
18. Что такое машина Больцмана?
19. Из каких частей состоит генеративно-состязательная сеть?
20. Истинно ли высказывание: Извлечение признаков в контексте глубокого обучения особенно полезно, когда целевая задача не включает в себя достаточно размеченных данных для успешного обучения CNN, которая хорошо обобщает?
21. Определите термин «тонкая настройка» (Fine-Tuning) предварительно обученной CNN на наборе ImageNet.
22. В каких случаях используется пэддинг (padding) в глубоком обучении?
23. Зачем используются соединения быстрого доступа в ResNet?
24. Как выбирается скорость обучения в уравнении обновления весов?
25. Как можно получить представления признаков, извлекаемых выбранным фильтром в глубокой сверточной модели?
26. Какие функции активации используются в нейронах?
27. Почему в глубоких сверточных сетях количество фильтров не уменьшается по мере продвижения сигналов по сети?
28. Что такое потеря контекста в рекуррентных моделях?
29. Какие гейты используются в ячейке LSTM?
30. Для каких задач используется двунаправленная рекуррентная модель?

Примеры задач:

Задача 1

Построить бинарный классификатор

для набора Оценка вероятности диагностики диабета у человека
<https://www.kaggle.com/datasets/alexteboul/diabetes-health-indicators-dataset> Класс: Diabetes_012. Класс отрицательный – 0 (no diabetes – нет диабета), класс положительный – 1 & 2 (prediabetes – преддиабетическое состояние & diabetes - диабет).

Выполнить загрузку и предварительную обработку данных из наборов. Разделить каждую выборку на обучающую, тестовую и валидационную. Произвести обучение набора нейросетевых архитектур, отличающихся разным набором параметров: число слоёв, количество нейронов в слоях, функции активации в слоях, процедур оптимизации:

Подобрать архитектуры нейронных сетей, которые с одной стороны позволяют получить модели с лучшими метриками качества работы, с другой стороны не являются избыточными и не переобученными.

Вычислить следующие метрики работы:

Для бинарного классификатора: Recall, Precision, Weighted Accuracy, AUC для всех исследованных моделей.

Задача 2

Построить многоклассовый классификатор

для набора Оценка уровня физического развития людей разного возраста: <https://www.kaggle.com/datasets/kukuroo3/body-performance-data> Метка класса: class.

Выполнить загрузку и предварительную обработку данных из наборов. Разделить каждую выборку на обучающую, тестовую и валидационную. Произвести обучение набора нейросетевых архитектур, отличающихся разным набором параметров: число слоёв, количество нейронов в слоях, функции активации в слоях, процедур оптимизации:

Подобрать архитектуры нейронных сетей, которые с одной стороны позволяют получить модели с лучшими метриками качества работы, с другой стороны не являются избыточными и не переобученными.

Вычислить следующие метрики работы:

Для многоклассового классификатора: Recall, Precision, Weighted Accuracy, AUC для всех классов всех исследованных моделей. Вывести ROC-кривые для каждого класса в лучшем классификаторе.

Задача 3

Построить регрессор

для набора Качество вина: <https://archive.ics.uci.edu/ml/datasets/Wine+Quality> предсказываемое значение – качество (Quality), файл winequality-white.csv.

Выполнить загрузку и предварительную обработку данных из наборов. Разделить каждую выборку на обучающую, тестовую и валидационную. Произвести обучение набора нейросетевых архитектур, отличающихся разным набором параметров: число слоёв, количество нейронов в слоях, функции активации в слоях, процедур оптимизации:

Подобрать архитектуры нейронных сетей, которые с одной стороны позволяют получить модели с лучшими метриками качества работы, с другой стороны не являются избыточными и не переобученными.

Вычислить следующие метрики работы:

Для регрессора: MSE, MAE, R2 для всех полученных моделей.

Критерии оценивания:

Результаты контрольной работы определяются оценками «отлично», «хорошо», «удовлетворительно», «неудовлетворительно».

Оценка «отлично» выставляется, если даны правильные ответы на все теоретические вопросы и все задачи решены без ошибок.

Оценка «хорошо» выставляется, если корректные ответы даны на большую часть вопросов, но были отмечены неуверенность в ответе и информация представлена фрагментарно, также задачи были решены правильно с небольшими замечаниями.

Оценка «удовлетворительно» выставляется, если даны правильные ответы на половину теоретических вопросы и одна из задач решена с негрубой ошибкой.

Оценка «неудовлетворительно» выставляется, если корректные ответов меньше половины и задачи были решены неправильно.

3. Оценочные материалы итогового контроля (промежуточной аттестации) и критерии оценивания

Экзаменационный билет состоит из трех частей.

Первая часть представляет собой тест из 5 вопросов. Ответы на вопросы первой части даются путем выбора из списка предложенных.

Вторая часть содержит один вопрос. Ответ на вопрос второй части дается в развернутой форме.

Третья часть содержит 2 вопроса, оформленные в виде практических задач. Ответы на вопросы третьей части предполагают решение задач и краткую интерпретацию полученных результатов.

Перечень теоретических вопросов:

1. Опишите три различных метода, с помощью которых можно точно настроить предварительно обученную CNN ImageNet.
2. Опишите, как работает нейронная передача стиля (Neural Style Transfer)
3. Специалист по данным предполагает, что: а) Операция свертки является как линейной, так и инвариантной к сдвигу. б) Операция свертки похожа на корреляцию, за исключением того, что мы переворачиваем фильтр перед применением оператора корреляции. в) Операция свертки достигает максимума только в тех случаях, когда фильтр в основном похож на определенную часть входного сигнала. Прав ли он, предполагая это? Подробно объясните значение этих утверждений.
4. Дано изображение размером $w \times h$ и ядро шириной K . Сколько умножений и сложений требуется для свертки изображения?
5. Верность (Ассигасу) персептрона рассчитывается как количество правильно классифицированных образцов, деленное на общее количество неправильно классифицированных образцов.
6. Каково наиболее распространенное применение слоев Max-пулинга?
7. Что такое batch-нормализация?
8. Что на самом деле означает термин стохастический в стохастическом градиентном спуске? Использует ли он какой-либо генератор случайных чисел?
9. Объясните, почему при стохастическом градиентном спуске количество эпох, необходимых для преодоления определенного порога потерь, увеличивается по мере уменьшения размера пакета?
10. Как работает обучение с учетом момента? Объясните роль экспоненциального затухания в правиле обновления градиентного спуска.
11. В чем разница между функциями Sigmoid и Softmax?
12. В чем разница между рекуррентной нейронной сетью и нейронной сетью прямого распространения?
13. В чем разница между рекуррентной нейронной сетью и нейронной сетью прямого распространения?
14. Для чего можно использовать рекуррентную нейронную сеть (RNN)?
15. Как выбирается функция потерь?
16. Что такое затухающий градиент в глубоком обучении? Что такое взрывной градиент в глубоком обучении?
17. Из каких частей состоит автоэнкодер?
18. Что такое машина Больцмана?
19. Из каких частей состоит генеративно-состязательная сеть?
20. Истинно ли высказывание: Извлечение признаков в контексте глубокого обучения особенно полезно, когда целевая задача не включает в себя достаточно размеченных данных для успешного обучения CNN, которая хорошо обобщает?
21. Определите термин «тонкая настройка» (Fine-Tuning) предварительно обученной CNN на наборе ImageNet.
22. В каких случаях используется пэддинг (padding) в глубоком обучении?
23. Зачем используются соединения быстрого доступа в ResNet?
24. Как выбирается скорость обучения в уравнении обновления весов?
25. Как можно получить представления признаков, извлекаемых выбранным фильтром в глубокой сверточной модели?
26. Какие функции активации используются в нейронах?
27. Почему в глубоких сверточных сетях количество фильтров не уменьшается по мере продвижения сигналов по сети?
28. Что такое потеря контекста в рекуррентных моделях?
29. Какие гейты используются в ячейке LSTM?
30. Для каких задач используется двунаправленная рекуррентная модель?

Примеры задач:

Задача 1.

Построить бинарный классификатор

для набора Оценка вероятности, того, что клиент откроет банковский депозит в результате маркетинговой акции: <https://archive.ics.uci.edu/ml/datasets/Bank+Marketing>
Класс: атрибут 21 - y - has the client subscribed a term deposit? (binary: 'yes', 'no').

Выполнить загрузку и предварительную обработку данных из наборов. Разделить каждую выборку на обучающую, тестовую и валидационную. Произвести обучение набора нейросетевых архитектур, отличающихся разным набором параметров: число слоёв, количество нейронов в слоях, функции активации в слоях, процедур оптимизации:

Подобрать архитектуры нейронных сетей, которые с одной стороны позволяют получить модели с лучшими метриками качества работы, с другой стороны не являются избыточными и не переобученными.

Вычислить следующие метрики работы:

Для бинарного классификатора: Recall, Precision, Weighted Accuracy, AUC для всех исследованных моделей.

Задача 2.

Построить многоклассовый классификатор для набора Оценка здоровья внутриутробного развития плода: <https://www.kaggle.com/datasets/andrewmvd/fetal-health-classification> Метка класса: fetal_health.

Выполнить загрузку и предварительную обработку данных из наборов. Разделить каждую выборку на обучающую, тестовую и валидационную. Произвести обучение набора нейросетевых архитектур, отличающихся разным набором параметров: число слоёв, количество нейронов в слоях, функции активации в слоях, процедур оптимизации:

Подобрать архитектуры нейронных сетей, которые с одной стороны позволяют получить модели с лучшими метриками качества работы, с другой стороны не являются избыточными и не переобученными.

Вычислить следующие метрики работы:

Для многоклассового классификатора: Recall, Precision, Weighted Accuracy, AUC для всех классов всех исследованных моделей. Вывести ROC-кривые для каждого класса в лучшем классификаторе.

Задача 3.

Построить регрессор для набора Аренда велосипедов: <https://archive.ics.uci.edu/ml/datasets/Bike+Sharing+Dataset>, предсказываемое значение – количество аренд велосипедов в сутки (Area), файл day.csv.

Выполнить загрузку и предварительную обработку данных из наборов. Разделить каждую выборку на обучающую, тестовую и валидационную. Произвести обучение набора нейросетевых архитектур, отличающихся разным набором параметров: число слоёв, количество нейронов в слоях, функции активации в слоях, процедур оптимизации:

Подобрать архитектуры нейронных сетей, которые с одной стороны позволяют получить модели с лучшими метриками качества работы, с другой стороны не являются избыточными и не переобученными.

Вычислить следующие метрики работы:

Для регрессора: MSE, MAE, R2 для всех полученных моделей.

Критерии оценивания:

Результаты экзамена определяются оценками «отлично», «хорошо», «удовлетворительно», «неудовлетворительно».

Оценка «отлично» выставляется, если даны правильные ответы на все теоретические вопросы и все задачи решены без ошибок.

Оценка «хорошо» выставляется, если корректные ответы даны на большую часть вопросов, но были отмечены неуверенность в ответе и информация представлена фрагментарно, также задачи были решены правильно с небольшими замечаниями.

Оценка «удовлетворительно» выставляется, если даны правильные ответы на половину теоретических вопросы и одна из задач решена с негрубой ошибкой.

Оценка «неудовлетворительно» выставляется, если корректные ответов меньше половины и задачи были решены неправильно.

4. Оценочные материалы для проверки остаточных знаний (сформированности компетенций)

Тест

1. Какая нейросетевая модель из перечисленных в лучшей степени подходит для прогнозирования временных последовательностей?
 - а) Single-Layer Perceptron
 - б) CNN
 - в) LSTM
 - г) Multi-layer Perceptron
2. Почему модели на сверточных нейронных сетях показывают наилучшие показатели по классификации объектов на изображениях по сравнению с другими моделями?
 - а) Они в высокой степени оптимизированы для обработки векторов с числовыми, а не категориальными признаками
 - б) Они обладают широким набором инструментов преобразования признакового пространства, которые может варьировать разработчик в модели
 - в) Они учитывают корреляцию смежных компонент вектора
 - г) Они используют существенно **большее** число настраиваемых параметров, по сравнению с другими моделями
3. Каким главным недостатком обладает рекуррентная нейронная сеть?
 - а) Длительная процедура обучения
 - б) Невозможность обучения на категориальных данных
 - в) Сложность запоминания длительных последовательностей
 - г) Использование существенных вычислительных ресурсов
4. Какие меры не приводят к уменьшению переобучения нейросетевой модели?
 - а) Установка штрафов за большие значения весов нейронов сети
 - б) Увеличение количества слоев сети
 - в) Добавление шума в выборку
 - г) Уменьшение количества нейронов сети
5. От чего в большей степени зависит успешное решение задачи классификации однослойным персептроном?
 - а) от размера выборки
 - б) от размерности признакового пространства
 - в) соотношения разделения выборки на обучающую и тестовую
 - г) от распределения объектов в пространстве признаков

Ключи: 1 в), 2 в), 3 в), 4 б), 5 г).

Критерии оценивания: тест считается пройденным, если обучающий ответил правильно как минимум на половину вопросов.

Задачи

Задача 1

Построить бинарный классификатор

для набора Оценка вероятности, того, что клиент откроет банковский депозит в результате маркетинговой акции: <https://archive.ics.uci.edu/ml/datasets/Bank+Marketing>
Класс: атрибут 21 - y - has the client subscribed a term deposit? (binary: 'yes', 'no').

Выполнить загрузку и предварительную обработку данных из наборов. Разделить каждую выборку на обучающую, тестовую и валидационную. Произвести обучение набора нейросетевых архитектур, отличающихся разным набором параметров: число слоёв, количество нейронов в слоях, функции активации в слоях, процедур оптимизации:

Подобрать архитектуры нейронных сетей, которые с одной стороны позволяют получить модели с лучшими метриками качества работы, с другой стороны не являются избыточными и не переобученными.

Вычислить следующие метрики работы:

Для бинарного классификатора: Recall, Precision, Weighted Accuracy, AUC для всех исследованных моделей.

Задача 2

Построить многоклассовый классификатор для набора Оценка здоровья внутриутробного развития плода: <https://www.kaggle.com/datasets/andrewmvd/fetal-health-classification> Метка класса: fetal_health.

Выполнить загрузку и предварительную обработку данных из наборов. Разделить каждую выборку на обучающую, тестовую и валидационную. Произвести обучение набора нейросетевых архитектур, отличающихся разным набором параметров: число слоёв, количество нейронов в слоях, функции активации в слоях, процедур оптимизации:

Подобрать архитектуры нейронных сетей, которые с одной стороны позволяют получить модели с лучшими метриками качества работы, с другой стороны не являются избыточными и не переобученными.

Вычислить следующие метрики работы:

Для многоклассового классификатора: Recall, Precision, Weighted Accuracy, AUC для всех классов всех исследованных моделей. Вывести ROC-кривые для каждого класса в лучшем классификаторе.

Теоретические вопросы:

1. Опишите три различных метода, с помощью которых можно точно настроить предварительно обученную CNN ImageNet.
2. Опишите, как работает нейронная передача стиля (Neural Style Transfer)
3. Специалист по данным предполагает, что: а) Операция свертки является как линейной, так и инвариантной к сдвигу. б) Операция свертки похожа на корреляцию, за исключением того, что мы переворачиваем фильтр перед применением оператора корреляции. в) Операция свертки достигает максимума только в тех случаях, когда фильтр в основном похож на определенную часть входного сигнала. Прав ли он, предполагая это? Подробно объясните значение этих утверждений.
4. Дано изображение размером $w \times h$ и ядро шириной K . Сколько умножений и сложений требуется для свертки изображения?
5. Верность (Accuracy) персептрона рассчитывается как количество правильно классифицированных образцов, деленное на общее количество неправильно классифицированных образцов.
6. Каково наиболее распространенное применение слоев Мах-пулинга?
7. Что такое batch-нормализация?
8. Что на самом деле означает термин стохастический в стохастическом градиентном спуске? Использует ли он какой-либо генератор случайных чисел?
9. Объясните, почему при стохастическом градиентном спуске количество эпох, необходимых для преодоления определенного порога потерь, увеличивается по мере уменьшения размера пакета?

10. Как работает обучение с учетом момента? Объясните роль экспоненциального затухания в правиле обновления градиентного спуска.
11. В чем разница между функциями Sigmoid и Softmax?
12. В чем разница между рекуррентной нейронной сетью и нейронной сетью прямого распространения?
13. В чем разница между рекуррентной нейронной сетью и нейронной сетью прямого распространения?
14. Для чего можно использовать рекуррентную нейронную сеть (RNN)?
15. Как выбирается функция потерь?
16. Что такое затухающий градиент в глубоком обучении? Что такое взрывной градиент в глубоком обучении?
17. Из каких частей состоит автоэнкодер?
18. Что такое машина Больцмана?
19. Из каких частей состоит генеративно-сопоставительная сеть?
20. Истинно ли высказывание: Извлечение признаков в контексте глубокого обучения особенно полезно, когда целевая задача не включает в себя достаточно размеченных данных для успешного обучения CNN, которая хорошо обобщает?
21. Определите термин «тонкая настройка» (Fine-Tuning) предварительно обученной CNN на наборе ImageNet.
22. В каких случаях используется пэддинг (padding) в глубоком обучении?
23. Зачем используются соединения быстрого доступа в ResNet?
24. Как выбирается скорость обучения в уравнении обновления весов?
25. Как можно получить представления признаков, извлекаемых выбранным фильтром в глубокой сверточной модели?
26. Какие функции активации используются в нейронах?
27. Почему в глубоких сверточных сетях количество фильтров не уменьшается по мере продвижения сигналов по сети?
28. Что такое потеря контекста в рекуррентных моделях?
29. Какие гейты используются в ячейке LSTM?
30. Для каких задач используется двунаправленная рекуррентная модель?

Информация о разработчиках

Аксёнов Сергей Владимирович, канд. техн. наук, доцент кафедры теоретических основ информатики НИ ТГУ.