

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

01.04.02 Прикладная математика и информатика

Направленность (профиль) подготовки:
Информационная безопасность

Форма обучения
Очная

Квалификация
Магистр

Год приема
2024

СОГЛАСОВАНО:
Руководитель ОП
А.Ю. Матророва

Председатель УМК
С.П. Сущенко

Томск – 2024

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

Целью освоения дисциплины является формирование следующих компетенций:

ПК-5 Способен управлять получением, хранением, передачей, обработкой больших данных.

УК-1 Способен осуществлять критический анализ проблемных ситуаций на основе системного подхода, вырабатывать стратегию действий.

Результатами освоения дисциплины являются следующие индикаторы достижения компетенций:

ИПК-5.1 Осуществляет мониторинг и оценку производительности обработки больших данных

ИПК-5.2 Использует методы и инструменты получения, хранения, передачи, обработки больших данных

ИПК-5.3 Разрабатывает предложения по повышению производительности обработки больших данных

ИУК-1.1 Выявляет проблемную ситуацию, на основе системного подхода осуществляет её многофакторный анализ и диагностику.

ИУК-1.2 Осуществляет поиск, отбор и систематизацию информации для определения альтернативных вариантов стратегических решений в проблемной ситуации.

ИУК-1.3 Предлагает и обосновывает стратегию действий с учетом ограничений, рисков и возможных последствий.

2. Оценочные материалы текущего контроля и критерии оценивания

Элементы текущего контроля:

- тесты;
- контрольная работа;
- лабораторные работы.

Пример теста (ИУК-1.1, ИУК-1.2)

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

Темы лабораторных работ по дисциплине (ИПК-5.1, ИПК-5.2, ИПК-5.3, ИУК-1.1, ИУК-1.2, ИУК-1.3)

1. Построение нейросетевого регрессора

Выбрать признаки, используемые при обучении, и, если необходимо, выполнить их предобработку. В работе необходимо исследовать работу архитектур и алгоритмов обучения с разными значениями параметров структуры и обучения (гиперпараметров) сетей и выбрать наилучшие значения последних.

Написать короткий отчет по работе, включив в него программу с комментариями, значения качества моделей. Выбрать лучшую модель.

2. Обучение нейросетевого классификатора

Написать программу на Python, которая обучает нейросетевой классификатор, с помощью библиотек scikit-learn и keras (опционально PyTorch).

Написать короткий отчет по работе, включив в него программу с комментариями, значения качества моделей. Выбрать лучшую модель.

Выборка для классификатора Covertypes Data Set (<https://archive.ics.uci.edu/ml/datasets/Covertypes>).

3. Исследование архитектуры нейронной сети

Для выполнения работы необходимо взять нейронную сеть, полученную в предыдущем задании. Выполнить повторное обучение полученной архитектуры и визуализировать поверхность функции потерь в пространстве двух выбранных весов, а также последовательность изменения этих весовых коэффициентов на каждой итерации обучения. Получить важность весов с помощью исследования чувствительности сети. Исследовать возможности по редукции сети с учетом полученной важности.

Написать короткий отчет по работе, включив в него программу с комментариями, рисунки настройки модели и значения метрик полной и редуцированных архитектур.

4. Кластеризация данных SOM

В качестве обучающих данных рекомендуется взять набор из лабораторной работы 2. Написать программу на языке Python, выполняющую задачу кластеризации данных с использованием самоорганизующихся архитектур SOM, требуется подобрать архитектуру сети (выполнить регуляцию размеров двумерной и трехмерной сети, а также произвести исследование функций близости и их параметров), привести метрики качества модели и визуализировать процесс обучения модели. Результаты работы привести в отчете.

5. Классификатор изображений

Написать программу на Python, которая обучает бинарный классификатор изображений на основе сверточных нейронных сетей.

Для этого нужно подготовить две папки с изображениями, принадлежащими двум классам (изображения можно загрузить из сети Интернет). Разделить выборку на обучающую и тестовую.

Построить набор сверточных нейронных сетей, отличающихся количеством слоев, чередованием слоев, наличием прореживания и обучить их на подготовленном наборе изображений.

Выбрать сеть, позволяющую классифицировать изображения с достаточным качеством, и не обладающую избыточностью.

Написать короткий отчет по работе, включив в него программу с комментариями, графики настройки модели и значения метрик моделей (точность, полнота).

6. Визуализация процесса функционирования нейронной сети

Для выполнения задания взять нейронную сеть, полученную в предыдущем задании.

1. Выполнить визуализацию активаций сверточных слоев (промежуточных активаций) для нескольких объектов разных классов.

2. Осуществить визуализацию фильтров сверточных слоев сети.

3. Произвести визуализацию тепловых карт активации классов с помощью метода CAM.

Написать короткий отчет по работе, включив в него программу с комментариями и результаты визуализации.

7. Исследование рекуррентной нейронной сети (прогноз температуры)

Написать программу на Python, которая обучает регрессионные модели, предсказывающие температуру на завтрашний день в одном из городов и построенных на наборе данных с помощью трёх алгоритмов: простая рекуррентная нейронная сеть, LSTM (или GRU), и стек LSTM (или стек GRU).

Выбрать признаки, используемые при обучении, и, если необходимо, выполнить их предобработку. Разделить выборку на обучающую и тестовую.

В работе необходимо исследовать работу нейронных сетей, посмотреть возможности настройки сети с прореживанием, определить количество дней до прогнозного дня, для которых достигается наивысшая точность прогнозирования

Написать короткий отчет по работе, включив в него программу с комментариями, графики настройки модели и значения метрик моделей (коэффициент детерминации, среднюю квадратичную и среднюю абсолютную ошибки).

Выбрать наилучшую модель из полученных регрессоров.

8. Генерация музыкальных композиций с помощью рекуррентных моделей

Подготовить набор временных последовательностей нот не менее 20 музыкальных композиций одного стиля (направления, автора). Обучить набор рекуррентных моделей для исследования этих временных последовательностей и написать генератор мелодий, позволяющий создавать мелодию, похожую на исследуемые.

Написать короткий отчет по работе, включив в него программу с комментариями, графики настройки моделей и значения метрик моделей.

Критерии оценивания лабораторных работ:

«Зачтено» – студент в целом хорошо разбирается в задаче, знает и использует методы решения практически самостоятельно или при небольшой подсказке

преподавателя, отвечает на вопросы, возможно с негрубыми ошибками. Представляет работу на защите в целом хорошо, с несущественными замечаниями.

«Не зачтено» – студент слабо разбирается в задаче, не знает или знает плохо методы решения, не отвечает, либо отвечает, но с грубыми ошибками на вопросы преподавателя.

Контрольная работа (ИПК-5.1, ИПК-5.2, ИПК-5.3, ИУК-1.1, ИУК-1.2, ИУК-1.3)
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 (Accuracy) перцептрона 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 вопросов и одна из задач решена с негрубой ошибкой.

Оценка «неудовлетворительно» выставляется, если корректных 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, проверяющих ИУК-1.1, ИУК-1.2. Ответы на вопросы первой части даются путем выбора из списка предложенных.

Вторая часть содержит один вопрос, проверяющий ИПК-5.1, ИПК-5.3, ИУК-1.3. Ответ на вопрос второй части дается в развернутой форме.

Третья часть содержит 2 вопроса, проверяющих ИПК-5.2, ИУК-1.3 и 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 (Accuracy) перцептрона 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-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?

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 теоретические 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 (ИУК-1.1, ИУК-1.2)

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 (ИПК-5.1, ИПК-5.2, ИУК-1.3)

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 (ИУК-1.1, ИУК-1.2, ИПК-5.3)

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) перцептрона 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?

Информация о разработчиках

Аксёнов Сергей Владимирович, к.т.н., кафедра теоретических основ информатики (ТОИ) Института прикладной математики и компьютерных наук (ИПМKN) Национальный исследовательский Томский государственный университет (НИ ТГУ), доцент каф. ТОИ