Подписано электронной подписью:
Вержицкий Данил Григорьевич
Должность: Директор КГПИ ФГБОУ ВО «КемГУ»
Дата и время: 2024-02-21 00:00:00
471086fad29a3b30e244c728abc3661ab35c9d5021b0dcf0e75e03a5b6fdf6436

МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ РОССИЙСКОЙ ФЕДЕРАЦИИ
Кузбасский гуманитарно-педагогический институт
федерального государственного бюджетного образовательного учреждения
высшего образования
«Кемеровский государственный университет»

Факультет информатики, математики и экономики

УТВЕРЖДАЮ
Декан А.В. Фомина
10 февраля 2022 г.

Рабочая программа дисциплины
К.М.07.ДВ.01.02 Глубокое обучение

Направление подготовки
01.03.02 Прикладная математика и информатика

Направленность (профиль) подготовки
ИНТЕЛЛЕКТУАЛЬНЫЙ АНАЛИЗ ДАННЫХ

Программа бакалавриата

Квалификация выпускника
бакалавр

Форма обучения
Очная

Год набора 2022

Новокузнецк 2022

Оглавление

1 Цель дисциплины. 3
1.1 Формируемые компетенции 3
1.2 Индикаторы достижения компетенций 3
1.3 Знания, умения, навыки (ЗУВ) по дисциплине 3
2 Объём и трудоёмкость дисциплины по видам учебных занятий. Формы промежуточной аттестации. 4
3. Учебно-тематический план и содержание дисциплины. 4
3.1 Учебно-тематический план 4
3.2. Содержание занятий по видам учебной работы 4
4 Порядок оценивания успеваемости и сформированности компетенций обучающегося в текущей и промежуточной аттестации. 6
5 Материально-техническое, программное и учебно-методическое обеспечение дисциплины. 7
5.1 Учебная литература 7
5.2 Материально-техническое и программное обеспечение дисциплины. 7
5.3 Современные профессиональные базы данных и информационные справочные системы. 8
6 Иные сведения и (или) материалы. 9
6.1. Примерные вопросы и задания / задачи для промежуточной аттестации . 9

1 Цель дисциплины.

В результате освоения данной дисциплины у обучающегося должны быть сформированы компетенции основной профессиональной образовательной программы магистратуры (далее - ОПОП): ПК-1.

Содержание компетенций как планируемых результатов обучения по дисциплине см. таблицы 1 и 2.

1.1 Формируемые компетенции

Таблица 1 - Формируемые дисциплиной компетенции

Наименование вида компетенции *(универсальная, общепрофессиональная, профессиональная)*	Наименование категории (группы) компетенций	Код и название компетенции
профессиональная		ПК-1 Способен проводить аналитические исследования с применением технологий больших данных

1.2 Индикаторы достижения компетенций

Таблица 2 – Индикаторы достижения компетенций, формируемые дисциплиной

Код и название компетенции	Индикаторы достижения компетенции по ОПОП	Дисциплины и практики, формирующие компетенцию ОПОП
ПК-1 Способен проводить аналитические исследования с применением технологий больших данных	ПК 1.1 Способен осуществлять выявление, формирование и согласование требований к результатам аналитических работ с применением технологий больших данных ПК 1.2 Способен планировать и организовывать аналитические работы с использованием технологий больших данных ПК 1.3 Способен подготавливать данные для проведения аналитических работ по исследованию больших данных ПК 1.4 Способен проводить аналитические исследования с применением технологий больших данных в соответствии с требованиями заказчика	К.М.07.01 Машинное обучение К.М.07.02 Аналитика данных К.М.07.03 Технологии работы с большими данными К.М.07.04 Технологии работы с открытыми данными К.М.07.ДВ.01.01 Машинное обучение с подкреплением/ **К.М.07.ДВ.01.02 Глубокое обучение** К.М.09.06(Пд) Преддипломная практика

1.3 Знания, умения, навыки (ЗУВ) по дисциплине

Таблица 3 – Знания, умения, навыки, формируемые дисциплиной

Код и название компетенции	Индикаторы достижения компетенции, закрепленные за дисциплиной	Знания, умения, навыки (ЗУВ), формируемые дисциплиной
ПК-1 Способен проводить аналитические исследования с	ПК 1.4 Способен проводить аналитические исследования с применением технологий больших данных в	**Знать*:*** – современные методы, используемые в машинном обучении, в том числе, глубоком обучении; – тенденции развития, научные и прикладные

Код и название компетенции	Индикаторы достижения компетенции, закрепленные за дисциплиной	Знания, умения, навыки (ЗУВ), формируемые дисциплиной
применением технологий больших данных	соответствии с требованиями заказчика	достижения в области глубокого обучения. **Уметь*:*** – анализировать и модифицировать методы глубокого обучения для определения наиболее подходящего при прогнозе конкретного процесса – использовать методы глубокого обучения для решения научно-исследовательских и прикладных задач. **Владеть*:*** – навыками использования современных методов глубокого обучения для решения прикладных задач

2 Объём и трудоёмкость дисциплины по видам учебных занятий. Формы промежуточной аттестации.

Таблица 4 – Объем и трудоемкость дисциплины по видам учебных занятий

Общая трудоемкость и виды учебной работы по дисциплине, проводимые в разных формах	Объём часов по формам обучения
	ОФО
1 Общая трудоемкость дисциплины	144
2 Контактная работа обучающихся с преподавателем (по видам учебных занятий) (всего)	54
Аудиторная работа (всего):	54
в том числе:	
лекции	18
лабораторные работы	36
Внеаудиторная работа (всего):	
3 Самостоятельная работа обучающихся (всего)	90
4 Промежуточная аттестация обучающегося и объём часов, выделенный на промежуточную аттестацию: - зачет с оценкой 8 семестр	

3. Учебно-тематический план и содержание дисциплины.

3.1 Учебно-тематический план

Таблица 5 - Учебно-тематический план очной формы обучения

№ недели п/п	Разделы и темы дисциплины по занятиям	Общая трудоёмкость *(всего час.)*	Трудоемкость занятий (час.)			Формы текущего контроля и промежуточной аттестации успеваемости
			ОФО			
			Аудиторн. занятия		СРС	
			лекц.	лаб.		
1	Основные понятия глубокого обучения	8	2		6	Реферат
2	Архитектуры нейронных сетей для задач глубокого обучения	90	10	30	50	Лабораторные работы
3	Обзор библиотек глубокого обучения	46	6	6	34	Лабораторные работы
	Всего:	144	18	36	90	Зачет с оценкой 8 семестр

3.2. Содержание занятий по видам учебной работы

Таблица 6 – Содержание дисциплины

№ п/п	Наименование раздела, темы дисциплины	Содержание занятия
Содержание лекционного курса		
1	Основные понятия глубокого обучения	Что такое глубокое обучение (deep learning). Истоки возникновения (связь с биологией). Примеры задач, которые решаются с использованием глубокого обучения: Задачи компьютерного зрения (computer vision): классификация изображений с большим числом категорий, детектирование объектов, семантическая сегментация изображений. Задачи распознавания естественного языка: машинный перевод, генерация текстов естественного языка, грамматический разбор слов. Другие примеры задач (генерация описания модели, задачи планирования). Классификация моделей по способу обучения. Обучение с учителем (supervised learning): многослойные полностью связанные сети, сверточные нейронные сети, рекуррентные нейронные сети. Обучение без учителя (unsupervised learning): автокодировщик, ограниченная машина Больцмана (Restricted Boltzmann Machine, RBM), глубокая машина Больцмана
2	Архитектуры нейронных сетей для задач глубокого обучения	**Многослойные полностью связанные сети** (Fully-Connected Neural Networks, FCNN). Многослойный персептрон (Multiple Layer Perceptron, MLP). Общая структура модели. Слои, функции активации и функции ошибки. Оптимизационная постановка задачи обучения многослойной нейронной сети. Метод обратного распространения ошибки (Back Propagation, BP). Стохастический градиентный спуск (Stochastic Gradient Descent, SGD). Настраиваемые параметры метода. Пример влияния параметров метода на скорость сходимости и результаты работы сети. **Сверточные нейронные сети**. Структура модели. Возможные слои (свертка, pooling, dropout, Local Contrast Normalization, Batch Normalization и другие). Функции активации (сигмоидальные, ReLU). Функции ошибки. Оптимизационная постановка задачи обучения сверточной нейронной сети. Метод обратного распространения ошибки для сверточных нейронных сетей. Пример простейшей сверточной нейронной сети: Структура сети; Влияния параметров метода обучения. Определение числа обучаемых параметров. Оценка объема памяти, необходимой для хранения сети. Принципы построения и оптимизации сверточных сетей. **Рекуррентные нейронные сети** (Recurrent Neural Network, RNN) и их развитие. Общая структура модели. Полностью рекуррентная нейронная сеть. Проблемы обучения рекуррентны сетей. Развертывание рекуррентной сети во времени и адаптация метода обратного распространения ошибки. Примеры простейших сетей: сеть Эльмана, сеть Хопфилда. Пример использования рекуррентных нейронных сетей к задаче распознавания цифр. Двунаправленные рекуррентные нейронные сети. Глубокие двунаправленные рекуррентные нейронные сети. Рекурсивные нейронные сети. Длинные рекуррентные нейронные сети с короткой памятью
3	Обзор библиотек глубокого обучения	Разработка сети, соответствующей логистической регрессии, на примере задачи распознавания рукописных цифр. Структура сети, соответствующая логистической регрессии. Задача распознавания рукописных цифр.

		Открытые библиотеки глубокого обучения: Библиотека Caffe (C/C++, Python).Пример разработки сети, обучения и тестирования сети. Библиотека Torch (Lua). Библиотека TensorFlow (Python).
Содержание лабораторных занятий		
2	Архитектуры нейронных сетей для задач глубокого обучения	Реализация метода обратного распространения ошибки для трехслойного персептрона. Разработка полностью связанной нейронной сети с использованием одной из библиотек глубокого обучения для решения некоторой заданной задачи. Проведение экспериментов с разным количеством скрытых слоев и числом скрытых элементов на каждом слое. Сбор результатов качества работы сетей. Разработка сверточной нейронной сети для решения той же задачи, что и в предыдущей лабораторной работе. Проведение экспериментов с разными конфигурациями сверточных нейронных сетей. Сбор результатов качества работы сетей. Разработка рекуррентных нейронных сетей и их разновидностей для решения той же задачи, что и в предыдущих работах. Проведение экспериментов с разными конфигурациями сетей. Сбор результатов качества работы сетей.
3	Обзор библиотек глубокого обучения	Разработка сети, соответствующей логистической регрессии, на примере задачи распознавания рукописных цифр. Структура сети, соответствующая логистической регрессии. Задача распознавания рукописных цифр. Открытые библиотеки глубокого обучения: Библиотека Caffe (C/C++, Python).Пример разработки сети, обучения и тестирования сети. Библиотека Torch (Lua). Библиотека TensorFlow (Python).

4 Порядок оценивания успеваемости и сформированности компетенций обучающегося в текущей и промежуточной аттестации.

Для положительной оценки по результатам освоения дисциплины обучающемуся необходимо выполнить все установленные виды учебной работы. Оценка результатов работы обучающегося в баллах (по видам) приведена в таблице 7.

Таблица 7 - Шкала и показатели оценивания результатов учебной работы обучающихся по видам в балльно-рейтинговой системе (БРС)

3 семестр

Учебная работа (виды)	Сумма баллов	Виды и результаты учебной работы	Оценка в аттестации (шкала и показатели оценивания)	Баллы
Текущая учебная работа в семестре (Посещение занятий по расписанию и выполнение заданий)	**80**	Реферат	**5 баллов** (пороговое значение) **10 баллов** (максимальное значение)	5-8
		Лабораторные работы	**Отчет по лабораторной работе (18 отчетов):** **2 баллов** (выполнено 51 - 65% заданий) **3 балла** (выполнено 66 - 85% заданий) **4 баллов** (выполнено 86 - 100% заданий)	36-72
Итого по текущей работе в семестре				41 - 80
Промежуточная аттестация (зачет с оценкой)	20	Решение задачи 1.	**5 баллов** (пороговое значение) **10 баллов** (максимальное значение)	5-10
		Решение задачи 2.	**5 баллов** (пороговое значение) **10 баллов** (максимальное значение)	5-10
Итого по промежуточной аттестации (зачет с оценкой)				10-20

Суммарная оценка по дисциплине: Сумма баллов текущей и промежуточной аттестации 51 – 100 б.

В промежуточной аттестации оценка выставляется в ведомость в 100-балльной шкале и в буквенном эквиваленте (таблица 8)

Таблица 8 – Соотнесение 100-балльной шкалы и буквенного эквивалента оценки

Сумма набранных баллов	Уровни освоения дисциплины и компетенций	*Экзамен*		*Зачет*
		Оценка	*Буквенный эквивалент*	*Буквенный эквивалент*
86 - 100	Продвинутый	5	отлично	Зачтено
66 - 85	Повышенный	4	хорошо	
51 - 65	Пороговый	3	удовлетворительно	
0 - 50	Первый	2	неудовлетворительно	Не зачтено

5 Материально-техническое, программное и учебно-методическое обеспечение дисциплины.

5.1 Учебная литература

Основная учебная литература

1. Флах, П. Машинное обучение. Наука и искусство построения алгоритмов, которые извлекают знания из данных / П. Флах. — Москва : ДМК Пресс, 2015. — 400 с. — ISBN 978-5-97060-273-7. — Текст : электронный // Лань : электронно-библиотечная система. — URL: https://e.lanbook.com/book/69955. — Режим доступа: для авториз. пользователей.

Дополнительная учебная литература

1. Гладилин, П. Е. Технологии машинного обучения : учебно-методическое пособие / П. Е. Гладилин, К. О. Боченина. — Санкт-Петербург : НИУ ИТМО, 2020. — 75 с. — Текст : электронный // Лань : электронно-библиотечная система. — URL: https://e.lanbook.com/book/190885. — Режим доступа: для авториз. пользователей.

2. Шалев-Шварц, Ш. Идеи машинного обучения : учебное пособие / Ш. Шалев-Шварц, Бен-ДавидШ. ; перевод с английского А. А. Слинкина. — Москва : ДМК Пресс, 2019. — 436 с. — ISBN 978-5-97060-673-5. — Текст : электронный // Лань : электронно-библиотечная система. — URL: https://e.lanbook.com/book/131686. — Режим доступа: для авториз. пользователей.

3. Баймуратов, И. Р. Методы автоматизации машинного обучения : учебное пособие / И. Р. Баймуратов. — Санкт-Петербург : НИУ ИТМО, 2020. — 40 с. — Текст : электронный // Лань : электронно-библиотечная система. — URL: https://e.lanbook.com/book/190871. — Режим доступа: для авториз. пользователей.

4. Маккинни, У. Python и анализ данных / У. Маккинни ; перевод с английского А. А. Слинкина. — 2-ое изд., испр. и доп. — Москва : ДМК Пресс, 2020. — 540 с. — ISBN 978-5-97060-590-5. — Текст : электронный // Лань : электронно-библиотечная система. — URL: https://e.lanbook.com/book/131721. — Режим доступа: для авториз. пользователей.

5.2 Материально-техническое и программное обеспечение дисциплины.

Учебные занятия по дисциплине проводятся в учебных аудиториях КГПИ ФГБОУ ВО КемГУ:

410 Учебная аудитория (мультимедийная) для проведения: - занятий лекционного типа; - групповых и индивидуальных консультаций; - текущего контроля и промежуточной аттестации; Специализированная (учебная) мебель: доска меловая, кафедра, моноблоки аудиторные. Оборудование: стационарное - компьютер, экран, проектор.	654079, Кемеровская область, г. Новокузнецк, пр-кт Металлургов, д. 19

Используемое программное обеспечение: MSWindows (MicrosoftImaginePremium 3 year по сублицензионному договору № 1212/КМР от 12.12.2018 г. до 12.12.2021 г.), LibreOffice (свободно распространяемое ПО), Яндекс.Браузер (отечественное свободно распространяемое ПО). Интернет с обеспечением доступа в ЭИОС.	
508 Лаборатория компьютерного моделирования Учебная аудитория (мультимедийная) для проведения: - занятий лабораторного типа; - самостоятельной работы; Специализированная (учебная) мебель: доска меловая, кафедра, столы, стулья. Оборудование для презентации учебного материала: стационарное - компьютер преподавателя, проектор, экран. Лабораторное оборудование: стационарное – компьютеры для обучающихся (18 шт.). Используемое программное обеспечение: MSWindows (MicrosoftImaginePremium 3 year по сублицензионному договору № 1212/КМР от 12.12.2018 г. до 12.12.2021 г.), LibreOffice (свободно распространяемое ПО), FoxitReader (свободно распространяемое ПО), Firefox 14 (свободно распространяемое ПО), Яндекс.Браузер (отечественное свободно распространяемое ПО), Opera 12 (свободно распространяемое ПО), Python с расширениями PIL, Py OpenGL, (свободно распространяемое ПО) Интернет с обеспечением доступа в ЭИОС.	654079, Кемеровская область, г. Новокузнецк, пр-кт Металлургов, д. 19

5.3 Современные профессиональные базы данных и информационные справочные системы.

Перечень СПБД и ИСС по дисциплине

1. Информационная система «Общероссийский математический портал», режим доступа : http://www.mathnet.ru/
2. Информационная система «Экспонента» - центр инженерных технологий и моделирования, режим доступа :http://www.exponenta.ru
3. База данных Science Direct (более 1500 журналов издательства Elsevier, среди них издания по математике и информатике), режим доступа :https://www.sciencedirect.com
4. Информационная система «Единое окно доступа к информационным ресурсам» http://window.edu.ru/catalog/
5. Базы данных и аналитические публикации на портале «Университетская информационная система Россия», режим доступа: https://uisrussia.msu.ru/
6. UCI Machine Learning Repository — репозиторий наборов данных для машинного обучения - http://archive.ics.uci.edu/ml/
7. IAPR Education Committee & Resources — коллекция ссылок на образовательные ресурсы по распознаванию образов, машинному обучению, обработке сигналов, обработке изображений и компьютерному зрению, поддерживаемая Международной ассоциацией распознавания образов - http://homepages.inf.ed.ac.uk/rbf/IAPR/index.php
8. Портал по интеллектуальному анализу данных, поддерживаемый Григорием Пятецким-Шапиро - http://www.kdnuggets.com/
9. Профессиональный информационно-аналитический ресурс, посвященный машинному обучению, распознаванию образов и интеллектуальному анализу данных. - http://machinelearning.ru

6 Иные сведения и (или) материалы.

6.1. Примерные вопросы и задания / задачи для промежуточной аттестации

Теоретические вопросы

1. Проблема переобучения при решении задачи восстановления регрессии.
2. Методы борьбы с переобучением: сокращение числа параметров, регуляризация (ридж-регрессия), метод лассо. Трудоемкость методов.
3. Метод ближайших соседей для решения задачи классификации. Теорема об оценке риска в методе ближайшего соседа.
4. Наивный байесовский классификатор.
5. Линейный дискриминантный анализ.
6. Квадратичный дискриминантный анализ.
7. Логистическая регрессия.
8. Нейронные сети. Персептрон Розенблатта. Алгоритм обучения персептрона как метод стохастического градиентного спуска.
9. Нейронные сети для решения задач классификации и восстановления регрессии. Обучение сети. Регуляризация как метод борьбы с переобучением.
10. Понятие о глубоких нейронных сетях.
11. Машина опорных векторов. Ядра и спрямляющие пространства.
12. Деревья решений. Метод CART (classification and regression trees) для решения задач классификации и восстановления регрессии.
13. Отсечения ветвей и выбор финального дерева. Методы обработки пропущенных значений.
14. Ансамбли решающих правил (классификаторов). Простое и взвешенное голосования.
15. Бустинг. Алгоритм AdaBoost. Оценка ошибки предсказания.
16. Бустинг и аддитивные модели. Градиентный бустинг. Алгоритм градиентного бустинга деревьев решений (MART).
17. Баггинг. Алгоритм случайных деревьев (случайный лес.).
18. Обучение без учителя. Кластеризация. Кластеризация методами теории графов.
19. Метод центров тяжести. Метод медиан.
20. Метод нечетких множеств. EM-алгоритм.
21. Введение в глубокое обучение (deep learning): что такое глубокое обучение (deep learning); истоки возникновения (связь с биологией); задачи, которые решаются с

использованием глубокого обучения.

22. Многослойные полностью связанные сети (Fully-Connected Neural Networks, FCNN). Многослойный персептрон (Multiple Layer Perceptron, MLP).

23. Открытые библиотеки глубокого обучения. Принцин разработки сети, соответствующей логистической регрессии, на примере задачи распознавания рукописных цифр.

24. Сверточные нейронные сети: структура модели; возможные слои (свертка, pooling, dropout, Local Contrast Normalization, Batch Normalization и другие), принципы задачи

25. Визуализация фильтров/выходов на промежуточных слоях сети: классификация методов визуализации признаков; открытые библиотеки визуализации; визуализация в библиотеках Caffe, Torch.

26. Рекуррентные нейронные сети (Recurrent Neural Network, RNN) и их развитие: Двунаправленные рекуррентные нейронные сети; Глубокие двунаправленные рекуррентные нейронные сети; Рекурсивные нейронные сети; Длинные рекуррентные нейронные сети с короткой памятью.

27. Обучение без учителя: Автокодировщик; Разверточные нейронные сети; Ограниченная машина Больцмана; Глубокая машина Больцмана; Глубокая доверительная сеть

28. Перенос обучения (transfer learning) глубоких нейронных сетей.

Практические задания

1. Реализовать метод обратного распространения ошибки для трехслойного персептрона.

2. Разработать полностью связанную нейронную сеть с использованием одной из библиотек глубокого обучения для решения некоторой заданной задачи. Провести эксперименты с разным количеством скрытых слоев и числом скрытых элементов на каждом слое. Собрать результаты качества работы сетей.

3. Разработать сверточную нейронную сеть для решения некоторой заданной задачи. Провести эксперименты с разными конфигурациями сверточных нейронных сетей. Собрать результаты качества работы сетей.

4. Разработать рекуррентную нейронную сеть для решения некоторой заданной задачи. Провести эксперименты с разными конфигурациями нейронных сетей. Собрать результаты качества работы сетей.

Составитель (и): Решетникова Е.В., доцент