

Министерство науки и высшего образования Российской Федерации
Федеральное государственное бюджетное образовательное учреждение
высшего образования
«Курганский государственный университет»
(КГУ)

Кафедра «Программное обеспечение автоматизированных систем»



УТВЕРЖДАЮ:

Врио ректора

Н.В. Дубив

« 02 » сентября 2019г.

Рабочая программа учебной дисциплины

ПРИКЛАДНЫЕ ЗАДАЧИ МАШИННОГО ОБУЧЕНИЯ

образовательной программы высшего образования –
программы бакалавриата

09.03.03 – Прикладная информатика

Направленность:

Интеллектуальные информационные системы и технологии

Форма обучения: очная и заочная

Курган 2019

Рабочая программа дисциплины «Прикладные задачи машинного обучения» составлена в соответствии с учебными планами по программе бакалавриата «Прикладная информатика» (Программное обеспечение автоматизированных систем), утвержденными для очной формы обучения «29» августа 2019 года, для заочной формы обучения «29» августа 2019 года.

Рабочая программа дисциплины одобрена на заседании кафедры «Программное обеспечение автоматизированных систем» «30» августа 2019 года, протокол № 1.

Рабочую программу составили:

Доцент кафедры
«Программное обеспечение
автоматизированных систем»

О.С. Черепанов

Согласовано:

Заведующий кафедрой
«Программное обеспечение
автоматизированных систем»
к.т.н., доцент

Т.Р. Змызгова

Специалист по учебно-методической
работе Учебно-методического отдела

Г.В. Казанкова

Начальник Управления образовательной
деятельности

С.Н. Сеницын

1. ОБЪЕМ ДИСЦИПЛИНЫ

Всего: 3 зачетных единиц трудоемкости (108 академических часов)

Очная форма обучения

Вид учебной работы	На всю дисциплину	Семестр
		7
Аудиторные занятия (контактная работа с преподавателем), всего часов в том числе:	32	32
Лекции	16	16
Лабораторные работы	16	16
Аудиторные занятия в интерактивной форме, часов	-	-
Самостоятельная работа, всего часов в том числе:	76	76
Подготовка к зачету	18	18
Другие виды самостоятельной работы	40	40
Контрольная работа	18	18
Вид промежуточной аттестации	зачет	зачет
Общая трудоемкость дисциплины и трудоемкость по семестрам, часов	108	108

Заочная форма обучения

Вид учебной работы	На всю дисциплину	Семестр
		8
Аудиторные занятия (контактная работа с преподавателем), всего часов в том числе:	10	10
Лекции	4	4
Лабораторные работы	6	6
Аудиторные занятия в интерактивной форме, часов	-	-
Самостоятельная работа, всего часов в том числе:	98	98
Подготовка к зачету	18	18
Другие виды самостоятельной работы (самостоятельное изучение тем (разделов) дисциплины)	62	62
Контрольная работа	18	18
Вид промежуточной аттестации	зачет	зачет
Общая трудоемкость дисциплины и трудоемкость по семестрам, часов	108	108

2. МЕСТО ДИСЦИПЛИНЫ В СТРУКТУРЕ ОБРАЗОВАТЕЛЬНОЙ ПРОГРАММЫ

Дисциплина «Прикладные задачи машинного обучения» относится к модулю «Системы интеллектуальной обработки данных» вариативной части, формируемой участниками образовательных отношений.

Изучение дисциплины базируется на результатах обучения, сформированных при изучении следующих дисциплин:

- Математический анализ.
- Дискретная математика.
- Математическая логика.
- Теория вероятностей и математическая статистика.
- Вычислительная математика.
- Интеллектуальный анализ данных.
- Основы искусственного интеллекта.
- Нейрокомпьютерные системы
- Основы программирования.
- Объектно-ориентированное программирование.

Результаты обучения по дисциплине необходимы для выполнения выпускной квалификационной работы.

3. ПЛАНИРУЕМЫЕ РЕЗУЛЬТАТЫ ОБУЧЕНИЯ

Целью освоения дисциплины «Прикладные задачи машинного обучения» является формирование теоретических знаний и приобретение практических навыков построения интеллектуальных систем, требующих применения методов машинного обучения.

Задачами дисциплины являются:

- изучение моделей и методов машинного обучения, применяемых в задача компьютерного зрения, обработки естественного языка;
- разработка математических моделей анализа данных;
- овладение студентами программными средствами решения прикладных задач машинного обучения.

Компетенции, формируемые в результате освоения дисциплины:

- Способность осуществлять поиск, критический анализ и синтез информации, применять системный подход для решения поставленных задач (УК-1).
- Способность определять круг задач в рамках поставленной цели и выбирать оптимальные способы их решения, исходя из действующих правовых норм, имеющихся ресурсов и ограничений (УК-2).
- Способность применять системный подход, современные математические методы и технологии для формализации решения прикладных задач (ПК-1),

- Способность оформлять аналитические справки и научно-технические отчеты, готовить обзоры научной литературы и электронных информационно-образовательных ресурсов для профессиональной деятельности, публиковать результаты выполненной работы (ПК-2).

В результате изучения дисциплины обучающийся должен

Знать

- типовые задачи машинного обучения (УК-2, ПК-1);
- модели и методы, применяемые при решения задач машинного обучения (УК-2, ПК-1).

Уметь

- применять различные методы машинного обучения для решения поставленных задач, а также оценивать качество полученного результат (УК-1, УК-2);
- сводит прикладную задачу к основным типам, формализует её и строит математическую модель (УК-1, УК-2, ПК-2).

Владеть:

- навыками решения практических задач с использованием различных методов машинного обучения (ПК-1, ПК-2);
- технологиями проектирования программных систем с искусственным интеллектом (ПК-1, ПК-2).

4. СОДЕРЖАНИЕ ДИСЦИПЛИНЫ

4.1. Учебно-тематический план

Очная форма обучения

№	Наименование раздела	Количество часов контактной работы с преподавателем	
		Лекции	Лабораторные работы
1	Введение в глубокое обучение	2	-
2	Компьютерное зрение	6	6
	Рубежный контроль 1		2
3	Обработка естественного языка	8	6
	Рубежный контроль 2		2
Всего:		16	16

Заочная форма обучения

№	Наименование раздела	Количество часов контактной работы с преподавателем	
		Лекции	Лабораторные работы
1	Введение в глубокое обучение	2	3
2	Компьютерное зрение		
3	Обработка естественного языка	2	3
Всего:		4	6

4.2. Содержание лекционных занятий

Наименование и содержание лекции	Часов контактной работы с преподавателем	
	Очная форма	Заочная форма
Раздел 1. Введение в глубокое обучение		
<p>Лекция 1. Введение в глубокое обучение.</p> <p>Обзор прикладных задач машинного обучения: классификация изображений, обнаружение объектов, изменение изображений, дорисовка и стилизация изображений, генерация изображений по описанию, генерация музыки, машинный перевод, языковые модели.</p> <p>Классификация подходов в искусственном интеллекте. Место глубокого обучения. Глубокие сети прямого распространения. Обзор алгоритмов обучения сетей прямого распространения. Борьба с переобучением нейронных сетей.</p>	2	
Раздел 2. Компьютерное зрение		
<p>Лекция 2. Классификация изображений</p> <p>Постановка задачи. Сверточные нейронные сети (CNN). Операция дискретной свертки. Нелинейность и пуллинг. Полносвязный слой CNN. Разреженная связность. Архитектура серточной нейронной сети. Смысловая интерпретация серточных сетей. Недостатки CNN.</p> <p>Архитектуры серточных нейронных сетей для классификации изображений. LeNet-5. AlexNet. VGG.</p>	2	

<p>GoogleNet. Inception. SqueezeNet. DenseNet.</p> <p>Разбор примера решения задачи классификации изображений.</p>		
<p>Лекция 3. Обнаружение объектов на изображении</p> <p>Постановка задачи детектирования объектов на изображениях. Метрики качества: локализация и классификация. Метрика IoU. Selective Search. Архитектуры нейронных сетей: R-CNN, Fast R-CNN, Faster R-CNN, Yolo, SSD.</p> <p>Разбор примера решения задачи обнаружение объектов на изображениях.</p>	2	2
<p>Лекция 4. Семантическая сегментация изображений</p> <p>Постановка задачи семантической сегментации изображений. Примеры практических задач. Виды сегментаций. Представление данных. Нейросетевой подход. Upsampling. Виды upsampling. FCN. SegNet. Unet. Расширенные свертки. DeepLab. Random crop. Аугментация.</p> <p>Разбор примера решения задачи семантической сегментации изображений.</p>	2	
Раздел 3. Обработка естественного языка		
<p>Лекция 5. Введение в NLP.</p> <p>Типы задач обработки естественного языка. Программные инструменты NLP. Предобработка текста: токенизация, нормализация слов, удаление слов. One-hot encoding. Bag of words. TF-IDF. Коллокации. Pointwise mutual information. Контекстные вложения. Матрица встречаемости. Задача классификации текста.</p> <p>WordToVec. GloVe. FastText.</p> <p>Разбор примера решения задачи классификации текстов.</p>	4	
<p>Лекция 6. Машинный перевод</p> <p>Постановка задачи машинного перевода. Сети Seq2Seq. Encoder-Decoder. Рекуррентные сети. Обучение рекуррентных Encoder-Decoder. Метрики качества машинного перевода. BLEU. Проблемы рекуррентных сетей типа Encoder-Decoder. Beam search. Механизм внимания.</p>	4	2

Self-attention. Трансформеры: ELMo, BERT, GPT. Архитектура трансформеров и методы их обучения. Разбор примера решения задачи машинного перевода.		
Итого:	16	4

4.3. Лабораторные работы

Номер раздела, темы	Наименование раздела	Наименование лабораторной работы	Норматив времени, час.	
			Очная форма обучения	Заочная форма обучения
2	Компьютерное зрение	Реализация системы сегментации изображений	3	3
		Реализация системы обнаружение объектов на изображениях	3	
		Рубежный контроль №1	2	-
3	Обработка естественного языка	Реализация системы классификации текстов	3	3
		Реализация системы машинного перевода	3	
		Рубежный контроль № 2	2	-
Всего:			16	6

4.4 Контрольная работа (для очной и заочной формы обучения)

Основная цель выполнения контрольной работы — ознакомление алгоритмами и методами распознавания и синтеза речи.

Контрольная работа предполагает написание подготовку доклада на тему, утвержденную преподавателем.

5. МЕТОДИЧЕСКИЕ УКАЗАНИЯ ДЛЯ ОБУЧАЮЩИХСЯ ПО ОСВОЕНИЮ ДИСЦИПЛИНЫ

Во время лекций по дисциплин студентам рекомендуется конспектировать теоретический материал, отмечая важные моменты, на которые заострил внимание преподаватель, участвовать в опросах и дискуссиях. Перед лекций необходимо повторить выданный материал, зафиксировать непонятные моменты, чтобы обсудить их на занятии.

Конспект лекций представлен в виде мультимедийных презентаций и включен в состав методического комплекса дисциплины.

Лабораторный практикум включает практические задания по двум разделам дисциплины: «Компьютерное зрение» и «Обработка естественного языка». Все работы выполняются в соответствии с заданием, выданным преподавателем.

Самостоятельная работа студентов включает в себя проработку материала лекционного курса дисциплины, подготовку и выполнение лабораторных работ, выполнение контрольной работы, а также подготовку к рубежным контролям (для очной формы) и к зачету.

Для текущего контроля успеваемости обучения для очной формы обучения используется балльно-рейтинговая система контроля оценки академической активности обучающихся. Для получения высокой оценки настоятельно рекомендуется активно участвовать во время обсуждения материала дисциплины на лекционных и лабораторных занятиях, а также тщательно его прорабатывать при самостоятельной работе.

Рекомендуемый режим самостоятельной работы

Наименование вида самостоятельной работы	Рекомендуемая трудоемкость, акад. час.	
	Очная форма обучения	Заочная форма обучения
Самостоятельное изучение тем дисциплины:	22	56
Введение в глубокое обучение	4	12
Компьютерное зрение	9	22
Обработка естественного языка	9	22
Подготовка к лабораторным работам (по 2 ч. на каждую работу)	12	6
Подготовка к рубежным контролям (по 3 ч. на каждый рубеж)	6	-
Подготовка к контрольной работе	18	18
Подготовка к зачету	18	18
Всего:	76	98

6. ФОНД ОЦЕНОЧНЫХ СРЕДСТВ ДЛЯ АТТЕСТАЦИИ ПО ДИСЦИПЛИНЕ

6.1. Перечень оценочных средств

1. Балльно-рейтинговая система контроля и оценки академической активности студентов в КГУ (для очной формы обучения).
2. Отчеты студентов по лабораторным занятиям.
3. Банк заданий к рубежным контролям № 1, № 2 (для очной формы обучения).
4. Контрольная работа.
5. Вопросы к зачету.

6.2. Система балльно-рейтинговой оценки работы студентов по дисциплине

№	Наименование	Содержание						
		Очная форма обучения						
Распределение баллов за 7 семестр								
1	Распределение баллов за семестры по видам учебной работы, сроки сдачи учебной работы (доводятся до сведения студентов на первом учебном занятии)	Вид учебной работы:	Посещение лекций	Выполнение и защита результатов лабораторных работ	Выполнение и защита контрольной работы	Рубежный контроль №1	Рубежный контроль №2	Зачет
		Балльная оценка:	16 x 8 = 86	86 + 86 + 86 = 326	166	76	76	306
2	Критерий пересчета баллов в традиционную оценку по итогам работы в семестре и экзамена	60 и менее баллов – неудовлетворительно; не зачтено; 61...73 – удовлетворительно; зачтено; 74...90 – хорошо; 91...100 – отлично.						
3	Критерии допуска к промежуточной аттестации, возможности получения, возможность получения бонусных баллов	<p>Для допуска к промежуточной аттестации (зачету) студент должен набрать не менее 50 баллов и выполнить все лабораторные работы и контрольную работу.</p> <p>Для получения «автоматически» оценки «зачтено» студенту необходимо набрать 61 баллов.</p> <p>По согласованию с преподавателем студенту могут быть добавлены дополнительные (бонусные) баллы за активность на консультациях, активное участие в научной и методической работе, оригинальность принятых решений в ходе выполнения лабораторных работ, за участие в значимых учебных и внеучебных мероприятиях кафедры.</p>						

4	<p>Формы и виды учебной работы для неуспевающих (восстановившихся на курсе обучения) студентов для получения недостающих баллов в конце семестра</p>
	<p>В случае если к промежуточной аттестации (зачету) набрана сумма менее 50 баллов, студенту необходимо набрать недостающее количество баллов за счет выполнения дополнительных заданий, до конца последней (зачетной) недели семестра. При этом необходимо проработать материал всех пропущенных лабораторных работ.</p> <p>Формы дополнительных заданий (назначаются преподавателем):</p> <ul style="list-style-type: none">- выполнение и защита пропущенной лабораторной работы (при невозможности дополнительного проведения лабораторной работы преподаватель устанавливает форму дополнительного задания по тематике пропущенной лабораторной работы самостоятельно) – до 8 баллов. <p>Ликвидация академических задолженностей, возникших из-за разности в учебных планах при переводе или восстановлении, проводится путем выполнения дополнительных заданий, форма и объем которых определяется преподавателем.</p>

6.3. Процедура оценивания результатов освоения дисциплины

Рубежные контроли (для очной формы обучения) проводятся в форме письменного тестирования, зачет в виде ответа на вопросы.

Перед проведением рубежного контроля преподаватель прорабатывает со студентами основной материал соответствующих разделов дисциплины в форме краткой лекции-дискуссии.

Варианты заданий для рубежных контролей № 1, № 2 состоят из 10 вопросов. Для определения баллов при проверке рубежных контролей используются интервальные оценки.

На каждую подготовку к рубежному контролю студенту отводится 1 академический час.

Преподаватель оценивает в баллах результаты рубежных контролей каждого студента по количеству правильных ответов и заносит в ведомость учета текущей успеваемости.

На зачете студенту предлагается ответить на 2 вопроса. Вопросы к зачету доводятся до студентов на последней лекции в семестре. Каждый вопрос оценивается в 15 баллов. На подготовку ответа студенту отводится 1 астрономический час.

Результаты текущего контроля успеваемости и зачета заносятся преподавателем в зачетно-экзаменационную ведомость, которая сдается в организационный отдел института в день проведения зачета, а также выставляются в зачетную книжку студента.

6.4. Примеры оценочных средств для рубежных контролей и зачета

6.4.1 Примеры заданий для рубежного контроля №1

1. Какого вида пуллинга в CNN не существует?

- а) Max-pooling;
- б) Mean-pooling;
- в) K-max-pooling;
- г) Static-pooling.

2. Какие нейронные сети используются для классификации изображений:

- а) Fast R-CNN;
- б) BERT;
- в) VGG;
- г) ResNet.

3. Для чего нужен *upsampling* в нейронных сетях?

- а) Для изменения разрешения изображение на выходе сети.
- б) Для ускорения обучения сети.
- в) Для борьбы с переобучением.

г) Все перечисленные варианты.

6.4.2 Примеры заданий для рубежного контроля №2

1. Что не включает процесс предобработки текста в NLP?

- а) Токенизацию;
- б) Нормализацию;
- в) Удаление слов;
- г) Обучение сети.

2. Какая архитектура ИНС в Word2Vec?

- а) многослойный перцептрон;
- б) рекуррентная нейронная сеть;
- в) сверточная нейронная сеть
- г) трансформер.

3. Какое предназначение Decoder в задаче машинного перевода?

- а) генерация перевода;
- б) анализ текста на исходном языке;
- в) токенизация исходного текста;
- г) выполнение процедуры стемминга.

6.4.3 Примерный перечень вопросов для зачета

1. Успехи глубокого обучения.

2. Классификация технологий искусственного интеллекта. Взаимосвязь машинного и глубокого обучения.

3. Глубокие сети прямого распространения. Метод обратного распространения ошибки.

4. Борьба с переобучением нейронных сетей. Модификация данных и функциональная модификация.

5. Борьба с переобучением нейронных сетей. Модификация процесса обучения.

6. Сверточные нейронные сети. Операция свертки. Нелинейность и пуллинг. Полносвязный слой.

7. Обучение сверточных нейронных сетей.

8. Применение сверточных нейронных сетей в задаче классификации изображений. LeNet-5 и AlexNet.

9. Применение сверточных нейронных сетей в задаче классификации изображений. VGG и GoogLeNet.

10. Применение сверточных нейронных сетей в задаче классификации изображений. Inception и ResNet.

11. Постановка задачи детектирования объектов на изображениях. Метрики качества.
12. Сети для детектирования объектов на изображения. Архитектуры R-CNN, Fast R-CNN, Faster R-CNN.
13. Сети для детектирования объектов на изображениях. Архитектуры Yolo, SSD.
14. Постановка задачи семантической сегментации изображений. Виды сегментации. Представление данных.
15. Нейросетевая семантическая сегментация изображений. Upsampling. Виды upsampling. FCN.
16. Нейросетевая семантическая сегментация изображений. Архитектуры SegNet. Unet.
17. Расширенные свертки. DeepLab. Random crop. Аугментация.
18. Типы задач обработки естественного языка. Программные инструменты NLP.
19. Предобработка текста: токенизация, нормализация слов, удаление слов.
20. One-hot encoding. Bag of words. TF-IDF. Коллокации.
21. Pointwise mutual information. Контекстные вложения. Матрица совстречаемости.
22. Вложения. WordToVec. GloVe. FastText.
23. Постановка задачи машинного перевода. Сети Seq2Seq.
24. Encoder-Decoder. Рекуррентные сети. Обучение рекуррентных Encoder-Decoder.
25. Метрики качества машинного перевода. BLEU.
26. Проблемы рекуррентных Encoder-Decoder. Beam search.
27. Механизм внимания. Self-attention.
28. Трансформеры. Архитектура трансформеров и методы их обучения.
29. Архитектура и обучение трансформера ELMO.
30. Архитектура и обучение трансформера BERT.
31. Архитектура и обучение трансформера GPT.

6.5. Фонд оценочных средств

Полный банк заданий для текущего, рубежных контролей и промежуточной аттестации по дисциплине, показатели, критерии, шкалы оценивания компетенций, методические материалы, определяющие процедуры оценивания образовательных результатов, приведены в учебно-методическом комплексе дисциплины.

7. ОСНОВНАЯ УЧЕБНАЯ ЛИТЕРАТУРА

1. Аксенов С.В., Новосельцев В.Б. Организация и использование нейронных сетей (методы и технологии) / Под общ. ред. В.Б. Новосельцева. – Томск: Изд-во НТЛ, 2006. – 128 с.

2. Паттерсон Дж., Гибсон А. Глубокое обучение с точки зрения практика / пер. с англ. А. А. Слинкина. – М.: ДМК Пресс, 2018. – 418 с.

3. Николенко С., Кадури А., Архангельская Е. Глубокое обучение. — СПб.: Питер, 2018. — 480 с.: ил.

4. Гудфеллоу Я., Бенджио И., Курвилль А. Глубокое обучение / пер. с англ. А. А. Слинкина. – 2-е изд., испр. – М.: ДМК Пресс, 2018. – 652 с.

8 МАТЕРИАЛЬНО-ТЕХНИЧЕСКОЕ ОБЕСПЕЧЕНИЕ ДИСЦИПЛИНЫ

8.1 Техническое обеспечение

№	Наименование	Использование
1	Комплект: ноутбук, медиа-проектор, экран	Для демонстрации иллюстративного материала при чтении лекций.
2	Персональный компьютер стандартной комплектации	Используется в качестве инструмента и объекта исследования при выполнении лабораторных и контрольных работ.

8.2 Программное обеспечение

№	Наименование	Использование
1	Интерпретатор Python 3.7+	Программное обеспечения для выполнения лабораторных работ
2	Среда разработки Jupyter	
3	Библиотеки Pytorch, torchtext, numpy, matplotlib, Pillow, razdel.	

Аннотация
рабочей программы учебной дисциплины

ПРИКЛАДНЫЕ ЗАДАЧИ МАШИННОГО ОБУЧЕНИЯ
образовательной программы высшего образования –
программы бакалавриата

09.03.03 Прикладная информатика

направленность

Интеллектуальные информационные системы и технологии

формы обучения – очная и заочная

Трудоемкость освоения дисциплины – 3 зач. ед. (108 акад. часов)

Семестры: 7-й для очной и 8-й для заочной формы обучения

Промежуточная аттестация: зачет в 7-м семестре для очной и зачет в
8-м семестре для заочной формы обучения

Содержание дисциплины

- Раздел 1. Введение в глубокое обучение
- Раздел 2. Компьютерное зрение
- Раздел 3. Обработка естественного языка