

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

Кафедра «Программное обеспечение автоматизированных систем»



УТВЕРЖДАЮ:
Первый проректор
/ Т.Р. Змызгова/
«02» сентября 2022 г.

Рабочая программа учебной дисциплины

ПРИКЛАДНЫЕ ЗАДАЧИ МАШИННОГО ОБУЧЕНИЯ

образовательной программы высшего образования –
программы бакалавриата

09.03.04 – Программная инженерия

Направленность:

Программное обеспечение автоматизированных систем

Форма обучения: очная и заочная

Курган 2022

Рабочая программа дисциплины «Прикладные задачи машинного обучения» составлена в соответствии с учебными планами по программе бакалавриата «Программная инженерия» (Программное обеспечение автоматизированных систем), утвержденными для очной формы обучения «30» августа 2022 года, для заочной формы обучения «30» августа 2022 года.

Рабочая программа дисциплины одобрена на заседании кафедры «Программное обеспечение автоматизированных систем» «01» сентября 2022 года, протокол № 1.

Рабочую программу составили:

Доцент кафедры
«Программное обеспечение
автоматизированных систем»,
к.ф.-м.н, доцент



О.С. Черепанов

Согласовано:

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



В. К. Волк

Начальник управления
образовательной деятельности



И.В. Григоренко

Специалист по учебно-методической
работе учебно-методического отдела



Г.В. Казанкова

1. ОБЪЕМ ДИСЦИПЛИНЫ

Всего: 4 зачетных единиц трудоемкости (144 академических часов)

Очная форма обучения

Вид учебной работы	На всю дисциплину	Семестр
		7
Аудиторные занятия (контактная работа с преподавателем), всего часов	36	36
в том числе:		
Лекции	16	16
Лабораторные работы	16	16
Практические занятия	4	4
Аудиторные занятия в интерактивной форме, часов	-	-
Самостоятельная работа, всего часов	108	108
в том числе:		
Подготовка к зачету	18	18
Другие виды самостоятельной работы	72	72
Контрольная работа	18	18
Вид промежуточной аттестации	зачет	зачет
Общая трудоемкость дисциплины и трудоемкость по семестрам, часов	144	144

Заочная форма обучения

Вид учебной работы	На всю дисциплину	Семестр
		10
Аудиторные занятия (контактная работа с преподавателем), всего часов	10	10
в том числе:		
Лекции	4	4
Лабораторные работы	6	6
Аудиторные занятия в интерактивной форме, часов	-	-
Самостоятельная работа, всего часов	134	134
в том числе:		
Подготовка к экзамену	27	27
Другие виды самостоятельной работы (самостоятельное изучение тем (разделов) дисциплины)	89	89
Контрольная работа	18	18
Вид промежуточной аттестации	экзамен	экзамен
Общая трудоемкость дисциплины и трудоемкость по семестрам, часов	144	144

2. МЕСТО ДИСЦИПЛИНЫ В СТРУКТУРЕ ОБРАЗОВАТЕЛЬНОЙ ПРОГРАММЫ

Дисциплина «Прикладные задачи машинного обучения» относится к элективному модулю «Системы интеллектуальной обработки данных» части, формируемой участниками образовательных отношений, блок 1.

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

- Математический анализ.
- Дискретная математика.
- Математическая логика.
- Теория вероятностей и математическая статистика.
- Вычислительная математика.
- Методы и алгоритмы анализа данных.
- Задачи и методы 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 применение 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 (для 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 (для УК-2),
- способность применять современный математический аппарат и методы компьютерного моделирования в профессиональной деятельности (для ПК-1),

- способность организовывать и проводить работы по исследованию объектов профессиональной деятельности, выявлению, документированию, оценке и сопровождению требований к программному продукту на основе анализа бизнес-процессов предметной области (для ПК-3).

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).

Владеть:

- навыками решения практических задач с использованием различных методов машинного обучения (ПК-3);
- технологиями проектирования программных систем с искусственным интеллектом (ПК-3).

4. СОДЕРЖАНИЕ ДИСЦИПЛИНЫ

4.1. Учебно-тематический план

Очная форма обучения

№	Наименование раздела	Количество часов контактной работы с преподавателем		
		Лекции	Лабораторные работы	Практические занятия
1	Введение в глубокое обучение	2	-	-
2	Компьютерное зрение	6	6	2
	Рубежный контроль 1		2	
3	Обработка естественного языка	8	6	2
	Рубежный контроль 2		2	
Всего:		16	16	4

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. GoogleNet. Inception. SqueezeNet. DenseNet.</p> <p>Разбор примера решения задачи классификации изображений.</p>	2	
<p>Лекция 3. Обнаружение объектов на изображении</p> <p>Постановка задачи детектирования объектов на</p>		

<p>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	2
<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. Self-attention. Трансформеры: ELMo, BERT, GPT. Архитектура трансформеров и методы их обучения.</p> <p>Разбор примера решения задачи машинного перевода.</p>	4	
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. Практические занятия (очная форма обучения)

Номер раздела, темы	Наименование раздела	Наименование практической работы	Норматив времени, час.
2	Компьютерное зрение	Изучение библиотек компьютерного зрения	2
3	Обработка естественного языка	Изучение библиотек NLP	2
Всего:			4

4.4 Контрольная работа

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 работы, выполнение контрольной работы, а также подготовку к рубежным контролям (для очной формы), к зачету (для очной формы) и экзамену (для заочной формы).

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, к практическим работам (для очной формы обучения), а также тщательно прорабатывать материал 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:	50	83
Введение в глубокое обучение	10	23
Компьютерное зрение	20	30
Обработка естественного языка	20	30
Подготовка к лабораторным работам (по 2 ч. на каждую работу)	12	6
Подготовка к практическим занятиям (по 2 ч. на каждую работу)	4	-
Подготовка к рубежным контролям (по 3 ч. на каждый рубеж)	6	-
Подготовка к контрольной работе	18	18
Подготовка к зачету, экзамену	18	27
Всего:	108	107

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 (очная форма обучения).
6. Вопросы к экзамену (заочная форма обучения).

**6.2. Система балльно-рейтинговой оценки
работы студентов по дисциплине**

Наименование		Содержание						
№	Распределение баллов за семестр	Очная форма обучения						
		Распределение баллов за 7 семестр						
	Вид учебной работы:	Посещение лекций	Выполнение и защита результатов лабораторных работ	Выполнение и защита результатов практических работ	Выполнение и защита контрольной работы	Рубежный контроль №1	Рубежный контроль №2	Зачет
1	Распределены баллы по видам учебной работы, сроки сдачи учебной работы (доводятся до сведения студентов на первом учебном занятии)	16 x 8 = 8 б	6 б + 6 б + 6 б + 6 б = 24 б	6 б + 6 б = 12 б	12 б	7 б	7 б	30 б
2	Критерий пересчета баллов в традиционную оценку по итогам работы в семестре и зачета	60 и менее баллов – неудовлетворительно; незачено; 61...73 – удовлетворительно; зачено; 74...90 – хорошо; 91...100 – отлично.						

3	<p>Критерии допуска к промежуточной аттестации, возможности получения, возможность получения бонусных баллов</p>	<p>Для допуска к промежуточной аттестации по дисциплине за семестр обучающийся должен набрать по итогам текущего и рубежного контролей не менее 51 балла. В случае если обучающийся набрал менее 51 балла, то к аттестационным испытаниям он не допускается.</p> <p>Для получения зачета без проведения процедуры промежуточной аттестации обучающемуся необходимо набрать в ходе текущего и рубежных контролей не менее 61 балла. В этом случае итог балльной оценки, получаемой обучающимся, определяется по количеству баллов, набранных им в ходе текущего и рубежных контролей. При этом, на усмотрение преподавателя, балльная оценка обучающегося может быть повышена за счет получения дополнительных баллов за академическую активность.</p> <p>Обучающийся, имеющий право на получение оценки без проведения процедуры промежуточной аттестации, может повысить ее путем сдачи аттестационного испытания. В случае получения обучающимся на аттестационном испытании 0 баллов итог балльной оценки по дисциплине не снижается.</p> <p>За академическую активность в ходе освоения дисциплины, участие в учебной, научно-исследовательской, спортивной, культурно-творческой и общественной деятельности обучающемуся могут быть начислены дополнительные баллы. Максимальное количество дополнительных баллов за академическую активность составляет 30.</p> <p>Основанием для получения дополнительных баллов являются:</p> <ul style="list-style-type: none"> - выполнение дополнительных заданий по дисциплине; дополнительные баллы начисляются преподавателем; - участие в течение семестра в учебной, научно-исследовательской, спортивной, культурно-творческой и общественной деятельности КГУ.
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 51 балла, обучающемуся необходимо набрать недостающее количество баллов за счет выполнения дополнительных заданий, до конца последней (зачетной) недели семестра.</p> <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 и экзамен 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 14 вопросов. Правильный ответ оценивается в 0.5 балла.

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 (экзамене) студенту предлагается ответить на 2 вопроса. Вопросы к зачету (экзамену) 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, зачета и экзамена заносятся преподавателем в зачетно-экзаменационную ведомость, которая сдается в организационный отдел института в день проведения зачета (экзамена), а также выставляются в зачетную книжку студента.

6.4. Примеры оценочных средств для рубежных контролей, зачета и экзамена

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 изображения 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 (очная форма обучения) и экзамена (заочная форма обучения)

1. Успехи глубокого обучение.
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 изображениях. 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 встречаемости.
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.	

9 ДЛЯ СТУДЕНТОВ, ОБУЧАЮЩИХСЯ С ИСПОЛЬЗОВАНИЕМ ДИСТАНЦИОННЫХ ОБРАЗОВАТЕЛЬНЫХ ТЕХНОЛОГИЙ

При использовании электронного обучения и дистанционных образовательных технологий (далее ЭО и ДОТ) занятия полностью или частично проводятся в режиме онлайн. Объем дисциплины и распределение нагрузки по видам работ соответствует п. 4.1. Распределение баллов соответствует п. 6.2, либо может быть изменено в соответствии с решением кафедры, в случае перехода на ЭО и ДОТ в процессе обучения. Решение кафедры об используемых технологиях и системе оценивания достижений обучающихся принимается с учетом мнений ведущего преподавателя и доводится до сведения обучающихся.

Аннотация
рабочей программы учебной дисциплины

ПРИКЛАДНЫЕ ЗАДАЧИ МАШИННОГО ОБУЧЕНИЯ
образовательной программы высшего образования –
программы бакалавриата

09.03.04 Программная инженерия

направленность

Программное обеспечение автоматизированных систем

формы обучения – очная и заочная

Трудоемкость освоения дисциплины – 4 зач. ед. (144 акад. часов)

Семестры: 7-й для очной и 10-й для заочной формы обучения

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

Содержание дисциплины

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