

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

Филологический факультет

УТВЕРЖДЕНО:

Декан филологического факультета
И.В. Тубалова

Оценочные материалы по дисциплине

Технологии обработки звучащей речи

по направлению подготовки

45.03.03 Фундаментальная и прикладная лингвистика

Направленность (профиль) подготовки:
«Фундаментальная и прикладная лингвистика»

Форма обучения

Очная

Квалификация

Бакалавр

Год приема

2024

СОГЛАСОВАНО:

Руководитель ОП
А.В. Васильева

Председатель УМК
Ю.А. Тихомирова

Томск – 2024

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

Целью дисциплины является изучение основных принципов и методов автоматической обработки текстов на естественном языке (ЕЯ)

Целью освоения дисциплины является формирование следующих компетенций:

ПК-4. Способен разрабатывать программный код при решении задач автоматической обработки текстов

Результатами освоения дисциплины являются следующие индикаторы достижения компетенций:

ИПК-4.1. Применяет способы формализации и алгоритмизации поставленных задач в сфере автоматической обработки текстов

2. Оценочные материалы текущего контроля и критерии оценивания

К числу форм контроля, оценивающих уровень достижения компетенций по текущей дисциплине, относятся тесты, устные вопросы, выполнение практических заданий (ПК-4, ИПК-4.1), итоговый проект (продукт (ПК-4.1), код (ПК-4), контрольная и самостоятельная работы (ПК-4).

Итоговый проект включает блок вопросов по теоретической части изучаемой дисциплины (ПК-4, ПК-4.1):

Пример практических задач:

1. Запишите предложение: «Катя принесла этого ежика, мяч и юлу к себе домой».

Предобработайте звуковой файл

2. Сделайте разметку файла по звукам и их типам по методике АТР.

3. Сделайте анализ спектрограммы и осциллограммы

4. Сравните форманты между собой и выявите интерференционные признаки

5. Основываясь на гипотезе проверки статистической значимости, сравните звуки в размеченном датасете

6. Напишите код для системы распознавания слитной речи на базе библиотеки

Whisper

Ответы:

Теоретические вопросы (ПК-4, ПК-4.1):

Критерии оценивания ответа на теоретический вопрос

Оценка	Критерии
	1. Понимание и логика высказывания изученного материала 2. Представление взаимосвязей процесса и взаимосвязи теоретических модулей изучаемого предмета 3. Полнота данных ответов; 4. Аргументированность данных ответов; 5. Правильность ответов на вопросы;

«зачтено»	Полно и аргументировано даны ответы по содержанию задания. Обнаружено понимание материала, может обосновать свои суждения, применить знания на практике, привести необходимые примеры не только по учебнику, но и самостоятельно составленные. Изложение материала последовательно и правильно Ответы обучающегося удовлетворяют тем же требованиям, что и для оценки «отлично», но допускается 1-2 ошибки, которые сам же исправляет. Обучающийся демонстрирует знание и понимание основных положений данного задания, но: 1) излагает материал неполно и допускает неточности в определении понятий или формулировке правил; 2) не умеет достаточно глубоко и доказательно обосновать свои суждения и привести свои примеры; 3) излагает материал непоследовательно и допускает ошибки.
«не зачтено»	Демонстрация незнания ответа на соответствующее задание, допускаются ошибки в формулировке определений и правил, искажающие их смысл, беспорядочно и неуверенно излагается материал; отмечаются такие недостатки в подготовке, которые являются серьезным препятствием к успешному овладению последующим материалом.

Критерии оценивания практической работы:

Оценка	Критерии
	1. Понимание и логика алгоритма работы 2. Наличие или отсудив ошибок в коде 3. Полнота решения практических задач 4. Своевременность выполнения; 5. Умения связать практический материал с теоретическим; 6. Понимание базовых формул обработки естественного языка и программирования;
«зачтено»	Основные требования к решению практических задач выполнены. Продемонстрированы умение анализировать алгоритмы и находить оптимальное количество решений, умение работать с информацией, в том числе умение затребовать дополнительную информацию, необходимую для уточнения реализации алгоритма, навыки разработки программного кода; Основные требования к решению практических задач выполнены, но при этом допущены недочеты. В частности, недостаточно раскрыты навыки стиля, недостаточно комментариев Имеются существенные отступления от решения практических задач. В частности отсутствуют навык и умения моделировать решения в соответствии с заданием, представлять различные подходы к разработке алгоритмов, ориентированных на конечный результат
«не зачтено»	Задача не решена, обнаруживается существенное непонимание проблемы

Итоговый проект включает создание корпуса звучащей. К числу тем могут относиться предложения со сложными для русского языка комбинациями звука.

Датасет запись речи респондентов, предобработку звуковых файлов, разметку в программе praat. Разметка должна содержать метку класса. Например, носитель – не носитель языка. Дробление по времени и звукам происходит на основе скриптов, которые формируют структурированный файл csv. Заключительный этап предусматривает использование статистических гипотез, анализа в языке программирования R или Python.

Проект представляется в виде программного кода и презентации, отражающую исследовательскую гипотезу в (фонетика и/или когнитивная лингвистика), этапы анализа и логику проведения исследовательской работы.

Критерии оценивания практической работы:

Оценка	Критерии
	1. Сформулирована гипотеза проекта 2. Собран и структурирован датасет 3. Четкая логика реализации алгоритмов обработки естественного языка в коде; 4. Правильность ответов на вопросы; 5. Наличие структурированной презентации 6. Полнота проекта
«зачтено»	Выполнены все требования к проекту: сформулирована гипотеза, создан прасер сайтов для сбора и структуризации информации, написан код для обработки естественного языка Выполнены все требования к составлению презентаций: дизайн слайдов, логика изложения материала, текст хорошо написан и сформированные идеи ясно изложены и структурированы Существуют незначительные ошибки в проекте, не влияющие на конечный результат. В частности, может быть низкий уровень формальных метрик, неточности в визуализации данных Основные требования к презентациям выполнены, но при этом допущены недочеты. В частности, имеются неточности в изложении материала; отсутствует логическая последовательность в суждениях; не выдержан объем презентации Существенные ошибки в коде, отсекает логика представления проекта, недостаточный объем датасета Имеются существенные отступления от требований к презентациям. В частности: тема освещена лишь частично; допущены фактические ошибки в содержании презентаций или при ответе на дополнительные вопросы.
«не зачтено»	Критические ошибки в коде, гипотеза не подвержена, недостаточное количество или отсутствие обучающей выборки Тема презентации не раскрыта, обнаруживается существенное непонимание проблемы

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

Зачет в пятом семестре состоит из трех частей.

Первая часть представляет собой тест из 20 вопросов, проверяющих ПК-4, ПК-4.1

Ответы на вопросы первой части даются путем выбора из списка предложенных, установления соответствия между объектами или формулирования однословного ответа на вопрос открытого типа.

Примерный тест

1. Какая из нижеперечисленных характеристик не относится к акустическим свойствам речи?
 - a) Частота
 - b) Интенсивность
 - c) Длительность
 - d) Цветность
2. Как называется процесс образования речи?
 - a) Фонетика
 - b) Фонология
 - c) Артикуляция
 - d) Произношение
3. Какая часть речи отвечает за производство звуковых сигналов?
 - a) Голосовой аппарат
 - b) Голосовые связки
 - c) Язык
 - d) Губы
4. Какие органы отвечают за формирование голосовых сигналов?
 - a) Голосовые связки
 - b) Губы
 - c) Небо
 - d) Диафрагма
5. Что такое частота звука?
 - a) Громкость звука
 - b) Сила звука
 - c) Высота звука
 - d) Продолжительность звука
6. Какая из нижеперечисленных характеристик отвечает за громкость звука?
 - a) Частота
 - b) Интенсивность
 - c) Длительность
 - d) Фаза
7. Что такое фонема?
 - a) Звуковая единица языка
 - b) Акустическая характеристика речи
 - c) Орган речи

d) Международный фонетический алфавит

8. Какой из нижеперечисленных звуков не является согласным?

- a) p
- b) i
- c) t
- d) k

9. Какой звук образуется при соприкосновении верхних зубов с нижней губой?

- a) m
- b) s
- c) z
- d) f

10. Какая из нижеперечисленных характеристик отвечает за длительность звука?

- a) Частота
- b) Интенсивность
- c) Амплитуда
- d) Продолжительность

Критерии оценивания тестирования:

Оценка	Критерии
	1. Полнота выполнения тестовых заданий; 2. Своевременность выполнения; 3. Правильность ответов на вопросы; 4. Самостоятельность тестирования
«зачтено»	Выполнено более 85 % заданий предложенного теста, в заданиях открытого типа дан полный, развернутый ответ на поставленный вопрос Выполнено более 70 % заданий предложенного теста, в заданиях открытого типа дан полный, развернутый ответ на поставленный вопрос; однако были допущены неточности в определении понятий, терминов и др. Выполнено более 54 % заданий предложенного теста, в заданиях открытого типа дан неполный ответ на поставленный вопрос, в ответе не присутствуют доказательные примеры, текст со стилистическими и орфографическими ошибками.
«не зачтено»	Выполнено не более 53 % заданий предложенного теста, на поставленные вопросы ответ отсутствует или неполный, допущены существенные ошибки в теоретическом материале (терминах, понятиях).

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

Первая часть. Решение тестовых задач:

Тест №1

1. Что такое акустическая обработка речи?
A) Процесс кодирования речи в цифровой формат
B) Процесс изменения частот и амплитуд звуков речи
C) Процесс удаления шумов из аудиозаписи
D) Процесс преобразования речи в текст
+ A) Процесс кодирования речи в цифровой формат
2. Как называется процесс преобразования речи в текст?
A) Синтез речи
B) Распознавание речи
C) Кодирование речи
D) Анализ речи
+ B) Распознавание речи
3. Какие методы используются для распознавания речи?
A) Методы машинного обучения
B) Методы анализа сигналов
C) Методы статистического моделирования
D) Все вышеперечисленные
+ D) Все вышеперечисленные
4. Какое устройство используется для записи звучащей речи?
A) Микрофон
B) Динамик
C) Акустический фильтр
D) Усилитель звука
+ A) Микрофон
5. Что такое форманты в обработке речи?
A) Частоты, на которых сосредоточена энергия звуков речи
B) Частоты, на которых находятся гармоники звуков речи
C) Частоты, на которых происходит амплитудная модуляция звуков речи
D) Частоты, на которых происходит фазовая модуляция звуков речи
+ A) Частоты, на которых сосредоточена энергия звуков речи
6. Какие методы используются для сжатия аудиоданных?
A) Методы уменьшения частоты дискретизации
B) Методы удаления субъективно незначимых звуков
C) Методы кодирования амплитуды и фазы сигнала
D) Все вышеперечисленные
+ D) Все вышеперечисленные
7. Что такое мел-частотные кепстральные коэффициенты (MFCC)?
A) Частоты, на которых находятся форманты звуков речи
B) Коэффициенты, описывающие спектральные особенности звуков речи
C) Коэффициенты, определяющие длительность звуков речи
D) Коэффициенты, определяющие интенсивность звуков речи
+ B) Коэффициенты, описывающие спектральные особенности звуков речи
8. Какой метод используется для синтеза речи на основе текста?
A) Метод склеивания фраз и слов

- B) Метод морфологического анализа текста
- C) Метод машинного обучения
- D) Метод статистического моделирования
- + A) Метод склеивания фраз и слов

Задача 1. Загрузите датасет размеченных звуков, и сделайте визуализацию звука а, для носителя и не носителя языка.

Ответ

```
abaza <- read.csv("Все форманты_2.csv", encoding = "UTF-8", header = T, sep = ";")
abaza
dim(abaza) #Размерность проверяли
colnames(abaza) #Смотрели какие есть колонки
table(abaza$letter) #Количество классов в колонке букв

a_ltrs <- abaza[abaza$letter=='a',] ##фильтр по букве "a"
print(abaza[abaza$letter=='a',])#извлечение англ а
abaza$letter[283]="a"#изменение англ а на рус а

a_ltrs$time <- as.numeric(a_ltrs$time)
boxplot(a_ltrs$time ~ a_ltrs$class)
```

```
shapiro.test(a_ltrs$time))
colnames(a_ltrs)
library(ggpol)
a_ltrs %>%
  distinct(time, class) %>%
  group_by( class) %>%
  mutate(mean = mean(time)) %>%
  ggplot(aes(time))
```

Задача 2. Проверьте тип распределения формант звука а

Ответ:

```
shapiro.test(a_ltrs$F1)
a_ltrs$F1 <- as.numeric(a_ltrs$F1)
```

Задача 3. Проверьте гипотезу о равенстве форманты 1 и форманты 2 звука а в двух классах, сравнивая их друг с другом

Ответ

```
distinct(word, letter, time, class) %>%
  group_by(class, letter) %>%
  mutate(mean = mean(time)) %>%
  ggplot(aes(letter, time, color = letter, fill = letter))+
  geom_boxjitter(errorbar.draw = TRUE, show.legend = FALSE)
```

```
t.test(duration~vowel, data = abaza[abaza$utterance == "1",])
t.test(duration~vowel, data = abaza[abaza$utterance == "2",])
```

Задача 4. Постройте линейную модель зависимости формант в голосовых звуках после зубных

Ответ

```
baza_mean_f1 %>%
  ggplot(aes(f1, duration, label = vowel))+
  geom_text()+
  geom_smooth(se = FALSE, method = "lm")+
```

```
labs(title = "Гласные после зубных (один спикер)",  
      y = "длительность (мс)",  
      x = "F1 (Hz)")
```

```
summary(lm(duration ~ f1, data = abaza_mean_f1))
```

Задача 6. Разработайте алгоритм для распознавания устной речи.

Ответ:

```
whisper_model = load_model('large-v2')  
whisper_results = whisper_model.transcribe(vocal_target, beam_size=None)  
  
# clear gpu vram  
del whisper_model  
torch.cuda.empty_cache()  
  
device = "cuda"  
alignment_model, metadata = whisperx.load_align_model(  
    language_code=whisper_results["language"], device=device  
)  
result_aligned = whisperx.align(  
    whisper_results["segments"], alignment_model, metadata, vocal_target,  
    device  
)  
  
# clear gpu vram  
del alignment_model  
torch.cuda.empty_cache()
```

Информация о разработчиках

Степаненко Андрей Александрович, старший преподаватель кафедры общей, компьютерной и когнитивной лингвистики