

Документ подписан простой электронной подписью
 Информация о владельце:
 ФИО: Косенок Сергей Михайлович
 Должность: ректор
 Дата подписания: 21.06.2025 16:46:17
 Уникальный программный идентификатор:
 e3a68f3eaa1e674574b5414998099d3d6b1dc1836

Оценочные материалы для промежуточной аттестации по дисциплине

Квантитативная лингвистика и новые информационные технологии

Код, направление	45.04.02 Лингвистика
подготовки	
Направленность (профиль)	Теория перевода и межкультурная / межъязыковая коммуникация
Форма обучения	очная
Кафедра-разработчик	Лингвистики и переводоведения
Выпускающая кафедра	Лингвистики и переводоведения

Типовая контрольная работа (тест) по дисциплине «Квантитативная лингвистика и новые информационные технологии в лингвистике»

Время на выполнение теста: 45 минут
 За каждый правильный ответ дается 1 балл

- Выберите из списка определение, соответствующее виду контент-анализа:
 структурный контент-анализ.
 - интерпретация количественных данных, с выявлением определенных смыслов.
 - констатация наличия/отсутствия/количества и информации.
 - анализ структурных компонентов текста.
- Метод лингвистических переменных – это:
 - метод построения терм-множеств;
 - формульное представление количественных подсчетов;
 - метод анализа деятельности нейронов при компьютерной томографии;
- Количественные методы в лингвистике включают:
 - когнитивное моделирование;
 - статистический анализ;
 - дефиниционный анализ;
 - компонентный анализ.
- Исключите лишнее:
 - Средняя арифметическая, среднее линейное отклонение, мода, медиана;
 - t-критерий Стьюдента, среднее квадратичное отклонение, дисперсия, дизъюнкция;
 - Semantic Web, Web 3.0., социальные сети.
- Для современного человека преобладающей является...
 - звуковая информация;
 - визуальная (символьная) информация;
 - вкусовая и тактильная информация;
 - визуальная (образная) информация.
- Адекватность информации — это ...
 - степень соответствия информации объективной реальности окружающего
 - мира;
 - степень соответствия информации, полученной потребителем, тому,

- d. что автор вложил в ее содержание;
 - e. достаточность информации для принятия решения;
 - f. степень соответствия информации текущему моменту времени.
7. *Машинный синтаксис – это ...*
- a. правила строения имен;
 - b. правила построения слов в более сложные структуры;
 - c. соотношение слова и его значения;
 - d. правила перевода письменного символа в устный.
8. *Естественный язык – это ...*
- a. знаковая система, используемая человеком с момента рождения;
 - b. знаковая система, используемая человеком в непринужденной обстановке;
 - c. знаковая система, созданная для естественных наук;
 - d. знаковая система, стихийно возникшая и закрепившаяся в обществе.
9. *Если дисперсия одного признака зависит от значения другого признака, то между этими признаками:*
- a. имеется функциональная связь
 - b. имеется статистическая связь
 - c. нет правильного ответа
 - d. связь отсутствует
 - e. имеется корреляционная связь
10. *Волапюк – это...*
- a. специализированный язык науки;
 - b. родной язык одного из малочисленных племен;
 - c. неспециализированный искусственный язык;
 - d. система символического кодирования.
11. *Какие из следующих приложений не являются текстовыми редакторами?*
- a. MS Excel;
 - b. Corel WordPerfect;
 - c. MS Works;
 - d. Adobe InCopy.
12. *Microsoft Word не включает...*
- a. функции настольных издательских систем;
 - b. функцию удалённого доступа;
 - c. функцию редактирования графических объектов;
 - d. шаблоны типовых таблиц.
13. *К устройствам ввода данных не относится*
- a. сканер;
 - b. принтер;
 - c. клавиатура;
 - d. цифровой фотоаппарат.
14. *OCR —это ...*
- a. система автоматического распознавания символов;
 - b. система переводческой памяти;
 - c. система машинного перевода;
 - d. функция текстового процессора.
15. *Реферат – это...*
- a. связный текст, который кратко выражает тему, предмет, цель, методы и результаты исследования;
 - b. процесс составления содержания документа (книги, статьи, патента на изобретение и др.);
 - c. краткое изложение содержания документа, дающее общее представление о его теме;

- d. краткий текст, выполняющий сигнальную функцию (информирует о том, что есть публикация на определенную тему).
16. Слово, относящееся к основному содержанию текста и повторяющееся в нем несколько раз, в автоматическом реферировании называется ...
- лейтмотивом;
 - термином;
 - символом;
 - ключевым словом.
17. Метод автоматического аннотирования, при котором важные слова выделяются в заголовке, подзаголовке, начале и конце текста, называется ...
- статистическим;
 - логико-семантическим;
 - позиционным;
 - функциональным.
18. Совокупность специально отобранных текстов, размеченных по различным лингвистическим параметрам и обеспеченных системой поиска, называется ...
- базой данных;
 - словарем;
 - информационным массивом;
 - корпусом.
19. Разметка бывает ...
- морфологической; синтаксической; семантической и просодической;
 - полнотекстовой и фрагментной;
 - синхронической и диахронической;
 - звуковой, письменной, смешанной.
20. УНК – это...
- корпус естественного языка, представительный по отношению ко всему языку;
 - универсальный национальный код;
 - собрание текстов, которое существует в Интернете;
 - собрание текстов, размеченных по различным лингвистическим параметрам и обеспеченных системой поиска.
21. Требования к корпусам
- полнота, адекватность, актуальность, компьютерная поддержка;
 - устойчивость, тиражируемость, адаптируемость, оптимальность временных параметров, комфорт пользователя;
 - репрезентативность, полнота, экономичность, структуризация, компьютерная поддержка;
 - полнота, экономичность, достоверность, структуризация, компьютерная поддержка.
22. Корпусный менеджер ...
- обеспечивает сортировку результатов поиска, статистические подсчеты;
 - составление списков слов на основе корпуса;
 - это специальная программа поиска по корпусу;
 - это человек, составляющий корпуса и управляющий ими;
 - это специальная программа подготовки текстов к их включению в корпус.
23. ПОД – это ...
- вид информационно-поисковой системы;
 - специальная программа поиска по корпусу;
 - поисковый образ документа;
 - поисковая оценка данных.
24. Одна из основных проблем компьютерного анализа речи состоит в том, что ...
- невозможно создать искусственный интеллект;

- b. компьютер не умеет работать со смыслом;
 - c. у компьютера нет дополнительных источников информации (ситуация,
 - d. контекст, прошлый опыт в данной области и т.п.);
 - e. разработчики не желают делиться своими профессиональными секретами.
25. *Что не относится к понятию термина?*
- a. слово (словосочетание) метаязыка науки, а также областей конкретной
 - b. практической деятельности человека;
 - a. понятие задается через свойства, реализуемые в системе;
 - b. использование основывается не на интуиции, а на четких определениях;
 - c. сопоставляется, как правило, несколько значений.
26. *Что не относится к процессу и понятию машинного перевода?*
- a. междисциплинарность;
 - b. использование машинных средств;
 - c. принципиальное сходство этапов понимания и синтеза текста;
 - d. учет языковых и экстралингвистических знаний.
27. *Типовая парадигма лексемы в автоматическом морфологическом анализе – это ...*
- a. последовательность букв от начала словоформы, общая для всех словоформ;
 - b. элементы, описывающие формоизменение конкретной лексемы,
 - c. совокупность наборов машинных окончаний;
 - d. совпадение основ разных слов.
28. *Требования к системам МП включают ...*
- a. устойчивость, тиражируемость, адаптируемость, оптимальность временных параметров, комфорт пользователя;
 - b. полнота, адекватность, актуальность, достоверность;
 - c. репрезентативность, полнота, экономичность, адекватность, компьютерная поддержка;
 - d. репрезентативность, полнота, экономичность, структуризация, компьютерная поддержка.
29. *Сущность когнитивно-интеллектуального подхода в компьютерном обучении состоит в том, что ...*
- a. программы ориентированы на обучающегося, дают свободу выбора уровня и типа действий;
 - b. программы построены по формуле стимул – реакция;
 - c. обучающемуся отводится роль объекта обучения;
 - d. в нем используются программы-тренажеры обучению языку с помощью компьютера.
30. *Программа для машинного перевода текстов:*
- a. Abbyy FineReader
 - b. Prompt
 - c. Adobe Photoshop
 - d. Mozilla Firefox

Примерная контрольная работа (практическое задание) по дисциплине
«Квантитативная лингвистика и новые информационные технологии в лингвистике»

1. Дайте определение следующим понятиям: выборка, переменная, статистическая значимость, нормальное распределение, коэффициент корреляций, мода, медиана, среднее линейное отклонение, коэффициент корреляций.
2. Составить выборку по заданным параметрам;
3. Рассчитать частоту употребления в тексте: а) частей речи; б) ключевых слов; в) квазисинонимов;

4. Произвести атрибуцию текста;
5. Интерпретировать количественный результат, полученный методом количественной лингвистики.

Типовые вопросы к экзамену:

1. Понятие «количественная лингвистика». Предмет, объект, цели и задачи количественной лингвистики.
2. Математический аппарат в лингвистике. Математическая (комбинаторная и количественная) лингвистика.
3. Количественные методы в лингвистических исследованиях.
4. Языковой закон или лингвистический закон.
5. Основные области применения структурно-вероятностной модели языка.
6. Методика проведения лингвостатистического анализа. Базовые концепты методики лингвостатистического анализа.
7. Выборки. Способы формирования выборок. Типы вероятностных выборок. Частота. Выборочная частота. Генеральная и выборочная совокупности. Виды ГЛС.
8. Среднее арифметическое. Вычисление среднего арифметического. Определение среднего квадратического отклонения. Решение лингвистических задач с использованием среднего квадратического отклонения.
9. Выявление внешней вариативности поведения определенной статистической характеристики. Понятия «коэффициент вариации», «размах вариации». Функции коэффициента вариации. Дисперсия.
10. Установление тесноты связи между данностями. Применение коэффициента ассоциации. Метод четырех полей.
11. Формула показателя корреляции рангов при определении степени родства данностей.
12. Система организации слов в словаре. Частотный словарь. Основные научные и практические направления, в которых применяются данные теории частоты слов.
13. Частота как характеристика употребительности слова в тексте. Модель текста «ранг-частота». Закон Ципфа. Уточнение закона Ципфа: закон Ципфа – Мандельброта. Области применения закона Ципфа. Статистические закономерности в развитии лексики.
14. Автоматическое аннотирование и реферирование текстов: основы технологии.
15. Исходные понятия корпусной лингвистики. Требования к корпусу текстов.
16. Виды лингвистических корпусов. Универсальный национальный корпус.
17. Национальный корпус русского языка.
18. Типы информационно-поисковых систем. Информационно-поисковые языки.
19. Стратегии машинного перевода. Автоматизированные и полуавтоматизированные системы перевода.
20. Области использования машинного перевода.
21. Важнейшие системы машинного перевода. Память переводов.
22. Перспективы развития систем машинного перевода (когнитивные и прагматические аспекты).
23. Политическая лингвистика и контент-анализ (принципы, результаты).
24. Этапы развития компьютерных обучающих средств. Особенности компьютерного обучения.
25. Программные средства обработки письменных текстов: автоматические конкордансы, морфологические и синтаксические анализаторы, их применение в области корпусной лингвистики, частотной лексикографии, проектирования информационно-поисковых систем.

26. Технологии распознавания печатного и рукописного текстов: известные на настоящий момент программные продукты: компьютерные программы и онлайн-сервисы OCR; применение технологий OCR в современных электронных устройствах.
27. Анализ и синтез звучащей речи: сферы применения, обзор компьютерных программ и онлайн-инструментов.
28. Методы автоматического реферирования и аннотирования текста. Программы и системы автоматического аннотирования и реферирования