

**МИНОБРНАУКИ РОССИИ**  
ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ БЮДЖЕТНОЕ ОБРАЗОВАТЕЛЬНОЕ УЧРЕЖДЕНИЕ  
ВЫСШЕГО ОБРАЗОВАНИЯ  
**«ВОРОНЕЖСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ»**  
**(ФГБОУ ВО «ВГУ»)**

УТВЕРЖДАЮ  
Заведующий кафедрой  
теоретической и прикладной лингвистики



проф. А.А. Кретов

19.06.2018 г.

**РАБОЧАЯ ПРОГРАММА УЧЕБНОЙ ДИСЦИПЛИНЫ**

**Б1.В.ДВ.4.2 Компьютерные технологии обработки текстов**

- 1. Шифр и наименование направления подготовки/специальности:** 44.04.01.  
Педагогическое образование
  - 2. Профиль подготовки/специализации:** Преподавание иностранных языков с использованием он-лайн технологий
  - 3. Квалификация (степень) выпускника:** магистр
  - 4. Форма обучения:** заочная
  - 5. Кафедра, отвечающая за реализацию дисциплины:** кафедра теоретической и прикладной лингвистики
  - 6. Составители программы:** Подтележникова Елена Николаевна, канд. филол. наук, доцент кафедры теоретической и прикладной лингвистики
  - 7. Рекомендована:** Рекомендована: Научно-методическим советом факультета РГФ, протокол № 10 от 19.06.2018 г.
  - 8. Учебный год:** 2019-2020
- Семестр(-ы):** 2 курс  
Сессия 2, 3

**9. Цели и задачи учебной дисциплины:** Изучить формальные, программно-реализуемые подходы к изучению структур и закономерностей естественных языков, ознакомится с основными прикладными практическими задачами компьютерной лингвистики

Основные задачи дисциплины:

- Изучить основные принципы и методов обработки естественного языка. Получение навыков разработки и интеграции программных систем обработки естественного языка в программные продукты.
- Знакомство с принципами проектирования программных систем, ориентированных на обработку естественно-языковых текстов.

**10. Место учебной дисциплины в структуре ООП:** Дисциплина относится к дисциплинам по выбору вариативной части. Основой для освоения дисциплины являются знания, умения, навыки и компетенции, полученные в результате освоения программ подготовки специалистов и в бакалавриата по направлению «Лингвистика», «Филология».

**11. Планируемые результаты обучения по дисциплине (знания, умения, навыки), соотнесенные с планируемыми результатами освоения образовательной программы (компетенциями выпускников):**

Компетенция		Планируемые результаты обучения
Код	Название	
ОК-1	способность к абстрактному мышлению, анализу, синтезу, совершенствовать и развивать свой интеллектуальный и общекультурный уровень	<b>знать</b> основные законы и подходы, объясняющие развитие сознания, научной мысли и образования; особенности и область применения основных мыслительных операций; основные понятия математической статистики как абстрактного языка научной деятельности; область применения операций анализа и синтеза в интеллектуальной деятельности; <b>уметь</b> применять мыслительные операции для качественного и количественного анализа эмпирических данных; осуществлять синтез и выделение существенных свойств социальных объектов; <b>владеть</b> навыками всестороннего и критического анализа принятых решений и собственной деятельности в учебных ситуациях.
ОК-4	способность формировать ресурсно-информационные базы для осуществления практической деятельности в различных сферах	<b>ЗНАТЬ:</b> степень развития современной цивилизации, характеризующуюся увеличением роли информации, возрастанием доли информационных средств коммуникаций, информационных продуктов и услуг; знать отличительные черты информационного общества. <b>УМЕТЬ:</b> пользоваться программными средствами при создании различных ресурсно-информационных баз в практической деятельности. <b>ВЛАДЕТЬ:</b> навыками сбора, обработки и анализа информации.
ОК-5	способность самостоятельно приобретать и использовать, в том числе с помощью информационных технологий, новые знания и умения, непосредственно не	<b>ЗНАТЬ:</b> иметь знания в области организации самостоятельной работы с учебной и научной литературой, поиска необходимых информационных источников (ресурсов) в локальных и глобальных сетях <b>УМЕТЬ:</b> использовать новые знания и умения, приобретенные в смежных предметных областях, для

	связанные со сферой профессиональной деятельности	решения задач в своей профессиональной деятельности. <b>ВЛАДЕТЬ:</b> современными методиками, инновационными и информационными технологиями приобретения и использования новых знаний и умений, необходимых для решения задач, связанных или не связанных со сферой профессиональной деятельности.
ПК-2	способностью формировать образовательную среду и использовать профессиональные знания и умения в реализации задач инновационной образовательной политики	<b>знать:</b> современные тенденции развития образовательной системы; теоретические положения, характеризующие образовательную среду и инновационную деятельность; виды инноваций в образовании; критерии инновационных процессов в образовании; <b>уметь:</b> поставить цели инновационной деятельности в образовательном учреждении; обосновать необходимость внесения запланированных изменений в образовательное учреждение; представить результаты работы в соответствии с принятыми решениями; внедрять инновационные приемы в педагогический процесс с целью создания условий для эффективной мотивации обучающихся; <b>владеть:</b> технологией планирования, организации и управления инновационной деятельностью в образовательном учреждении; - технологиями проведения опытно-экспериментальной работы, участия в инновационных процессах

**12. Объем дисциплины в зачетных единицах/часах в соответствии с учебным планом — 2 ЗЕТ / 72 часа.**

**Форма промежуточной аттестации:** зачет, контрольная работа

**13. Виды учебной работы:**

Вид учебной работы	Трудоемкость (часы)			
	Всего	По семестрам		
		Сессия 2	Сессия 3	.....
Аудиторные занятия				
в том числе: лекции				
практические				
лабораторные	8	2	6	
Самостоятельная работа	60	34	26	
Форма промежуточной аттестации (зачет)	зачёт		зачёт	
Форма промежуточной аттестации (контрольная работа)	4		4	
Итого:	72	36	36	

**13.1 Содержание дисциплины:**

п/п	Наименование раздела дисциплины	Содержание раздела дисциплины
	<b>1. Лабораторные работы</b>	

1.1	Задачи компьютерной лингвистики в изучении естественного языка (ЕЯ).	История научного направления. Классификация языков Хомского. Особенности ЕЯ. Знакомство с направлениями современных исследований.
1.2	Лингвистический процессор. Алгоритмы лингвистического разбора и анализа текста. Парсеры ЕЯ-предложений.	Лингвистический процессор - функциональная структура. Методы морфологического анализа, используемые в лингвистических процессорах. Морфологические словари. Алгоритмы синтаксического и семантического анализа для автоматических систем обработки текстов. Парсеры ЕЯ. Прикладные системы- спэлчекеры, текстовые редакторы, системы профессионального редактирования.
1.3	Формальные методы исследования структуры ЕЯ текста	Статистические методы анализа структур ЕЯ текста на морфологическом, синтаксическом, семантическом уровнях. Метод позиционных статистик. Приложение методов для задач дешифровки ЕЯ текстов на неизвестных языках. Марковские цепи.
1.4	Формальные методы классификации полнотекстовых документов	Математическая постановка задачи распознавания образов и классификации. Формальные методы определения сходства ЕЯ документов на различных уровнях лингвистического анализа (морфологическом, синтаксическом, семантическом): кластерный анализ, деревья принятия решений, векторные методы, Байесовский классификатор. Применение методов классификации для задач определения авторства текстов.

### 13.2 Темы (разделы) дисциплины и виды занятий:

№ п/п	Наименование темы (раздела) дисциплины	Виды занятий (часов)				Всего
		Лекции	Практические	Лабораторные	Самостоятельная работа	
1	Задачи компьютерной лингвистики в изучении естественного языка (ЕЯ).			2	15	17
2	Лингвистический процессор. Алгоритмы лингвистического разбора и анализа текста. Парсеры ЕЯ-предложений.			2	15	17
3	Формальные методы исследования структуры ЕЯ текста			2	15	17
4	Формальные методы классификации полнотекстовых документов			2	15	17
	Итого:			8	60	68

### 14. Методические указания для обучающихся по освоению дисциплины:

Для изучения разделов данной учебной дисциплины необходимо вспомнить и систематизировать знания, полученные ранее.

При изучении материала учебной дисциплины по учебнику нужно, прежде всего, уяснить существо каждого излагаемого там вопроса. Главное - это понять изложенное в учебнике, а не «заучить».

Изучать материал рекомендуется по темам конспекта лекций и по главам (параграфам) учебника (учебного пособия). Сначала следует прочитать весь материал темы (параграфа), особенно не задерживаясь на том, что показалось не совсем понятным: часто это становится понятным из последующего. Затем надо вернуться к местам, вызвавшим затруднения и внимательно разобраться в том, что было неясно.

Особое внимание при повторном чтении необходимо обратить на формулировки соответствующих определений, формулы и т.п. (они обычно

бывают набраны в учебнике курсивом); в точных формулировках, как правило, существенно каждое слово и очень полезно понять, почему данное положение сформулировано именно так. Однако не следует стараться заучивать формулировки; важно понять их смысл и уметь изложить результат своими словами.

Закончив изучение раздела, полезно составить краткий конспект, по возможности не заглядывая в учебник (учебное пособие).

При изучении учебной дисциплины особое внимание следует уделить приобретению навыков решения профессионально-ориентированных задач. Для этого, изучив материал данной темы, надо сначала обязательно разобраться в решениях соответствующих задач, которые рассматривались на практических занятиях, приведены в учебно-методических материалах, пособиях, учебниках, ресурсах Интернета, обратив особое внимание на методические указания по их решению. Затем необходимо самостоятельно решить несколько аналогичных задач из сборников задач, приводимых в разделах рабочей программы, и после этого решать соответствующие задачи из сборников тестовых заданий и контрольных работ.

Закончив изучение раздела, нужно проверить умение ответить на все вопросы программы курса по этой теме (осуществить самопроверку).

Все вопросы, которые должны быть изучены и усвоены, в программе перечислены достаточно подробно. Однако очень полезно составить перечень таких вопросов самостоятельно (в отдельной тетради) следующим образом:

– начав изучение очередной темы программы, выписать сначала в тетради последовательно все перечисленные в программе вопросы этой темы, оставив справа широкую колонку;

– по мере изучения материала раздела (чтения учебника, учебно-методических пособий, конспекта лекций) следует в правой колонке указать страницу учебного издания (конспекта лекции), на которой излагается соответствующий вопрос, а также номер формулы, которые выражают ответ на данный вопрос.

В результате в этой тетради будет полный перечень вопросов для самопроверки, который можно использовать и при подготовке к экзамену. Кроме того, ответив на вопрос или написав соответствующую формулу (уравнение), можете по учебнику (конспекту лекций) быстро проверить, правильно ли это сделано, если в правильности своего ответа Вы сомневаетесь. Наконец, по тетради с такими вопросами Вы можете установить, весь ли материал, предусмотренный программой, Вами изучен.

Следует иметь в виду, что в различных учебных изданиях материал может излагаться в разной последовательности. Поэтому ответ на какой-нибудь вопрос программы может оказаться в другой главе, но на изучении курса в целом это, конечно, никак не скажется.

Указания по выполнению тестовых заданий и контрольных работ приводятся в учебно-методической литературе, в которых к каждой задаче даются конкретные методические указания по ее решению и приводится пример решения.

### **13. Перечень основной и дополнительной литературы, ресурсов Интернет, необходимых для освоения дисциплины:**

#### **а) основная литература:**

№ п/п	Источник
1.	Автоматическая обработка текстов на естественном языке и компьютерная лингвистика: учеб. пособие / Большакова Е.И., Клышинский Э.С., Ландэ Д.В., Носков А.А., Пескова О.В.,

	Ягунова Е.В. – Москва: МИЭМ, 2011. 272 с.
2.	Захаров В.П., Богданова С.Ю. Корпусная лингвистика: учебник для студентов гуманитарных вузов. – Иркутск: ИГЛУ, 2011. 161 с.
3.	Методы когнитивного анализа семантики слова: компьютерно-корпусный подход / под общ. ред. В.И. Заботкиной М.: Языки славянской культуры. 2015. 344с.

**б) дополнительная литература:**

№ п/п	Источник
4.	<b>Corpora in Cognitive Linguistics: Corpus-Based Approaches to Syntax and Lexis Mouton Select. <u>Anatol Stefanowitsch</u> (Editor), <u>Stefan Th. Gries</u> (Editor). 2007</b>
5.	Encyclopedia of Language and Linguistics. Volume 1-14 / Ed. by Keith Brown. – 2nd Edition. – Amsterdam: Elsevier, 2006.

**в) информационные электронно-образовательные ресурсы:**

№ п/п	Источник
6.	Каталог лингвистических программ и ресурсов <a href="http://www.rvb.ru/soft/catalogue/catalogue.html">http://www.rvb.ru/soft/catalogue/catalogue.html</a>
7.	The Visual Thesaurus. URL: <a href="http://www.visualthesaurus.com/app/view">http://www.visualthesaurus.com/app/view</a>
8.	New BYU <u>Wikipedia Corpus</u> (with virtual corpora) See <u>overview and tutorials</u>
9.	e-library.ru
10.	WordNet. URL: <a href="http://wordnetweb.princeton.edu/perl/webwn">http://wordnetweb.princeton.edu/perl/webwn</a> Сайты образовательных порталов <a href="http://online-learning.harvard.edu">http://online-learning.harvard.edu</a> <a href="http://online-learning.com">http://online-learning.com</a> <a href="https://alison.com">https://alison.com</a> <a href="https://www.coursera.org">https://www.coursera.org</a> <a href="https://www.khanacademy.org">https://www.khanacademy.org</a> <a href="http://www.ontariolearn.com/en/">http://www.ontariolearn.com/en/</a> <a href="http://oedb.org">http://oedb.org</a> <a href="https://www.edx.org">https://www.edx.org</a> <a href="http://www.onlineeducation.net">http://www.onlineeducation.net</a> <a href="http://www.ashford.edu/degrees/online/education.htm">http://www.ashford.edu/degrees/online/education.htm</a> <a href="http://www.liberty.edu/online/">http://www.liberty.edu/online/</a> <a href="http://www.nu.edu/OnlineEducation.html">http://www.nu.edu/OnlineEducation.html</a> <a href="http://www.devry.edu/online-education.html">http://www.devry.edu/online-education.html</a> <a href="http://www.fullsail.edu/online-education">http://www.fullsail.edu/online-education</a>

**16. Перечень учебно-методического обеспечения для самостоятельной работы:**

№ п/п	Источник
1.	Захаров В.П., Богданова С.Ю. Корпусная лингвистика: учебник для студентов гуманитарных вузов. – Иркутск: ИГЛУ, 2011. 161 с.

**17. Информационные технологии, используемые для реализации учебной дисциплины, включая программное обеспечение и информационно-справочные системы (при необходимости):**

Программное обеспечение (Microsoft Office).

**18. Материально-техническое обеспечение дисциплины:**

/ауд. 14/ Проектор Benq MW523 (1 шт.)

Экран проекционный (1 шт.)

Компьютер Asus H81m-Plus (11 шт.)

/ауд. 12/ Компьютер Arbyte Tempo/AOC (12 шт.),

Проектор Benq MW523 (1 шт.),

Сканер Canon Canoscan LiDE 120 (5 шт.)

Экран проекционный (1 шт.)

/ауд. 52/ DVD+VHS Philips DVP 3100 V (1 шт.)

Домашний кинотеатр Aleks DR9000 (1 шт.)

Мультимедиа-проектор Epson EB-X18 (1 шт.)

Настенный экран 180\*180 (1 шт.)

Телевизор LCD Samsung (1 шт.)

**19. Фонд оценочных средств:**

**19.1. Перечень компетенций с указанием этапов формирования и планируемых результатов обучения**

Код и содержание компетенции (или ее части)	Планируемые результаты обучения (показатели достижения заданного уровня освоения компетенции посредством формирования знаний, умений, навыков)	Этапы формирования компетенции (разделы (темы) дисциплины или модуля и их наименование)	ФОС (средства оценивания)
ОК-1 способность к абстрактному мышлению, анализу, синтезу, способностью совершенствовать и развивать свой интеллектуальный и общекультурный уровень	<b>знать</b> основные законы и подходы, объясняющие развитие сознания, научной мысли и образования; особенности и область применения основных мыслительных операций; основные понятия математической статистики как абстрактного языка научной деятельности; область применения операций анализа и синтеза в интеллектуальной деятельности;	Компьютерные информационные ресурсы Методика использования ресурсов Специализированные образовательные порталы Информационный поиск	
	<b>уметь</b> применять мыслительные операции для качественного и количественного анализа эмпирических данных; осуществлять синтез и выделение существенных свойств социальных объектов;	Компьютерные информационные ресурсы Методика использования ресурсов Специализированные образовательные порталы Информационный поиск	Лабораторная работа
	<b>владеть</b> навыками всестороннего и критического анализа принятых решений и собственной деятельности в учебных ситуациях.	Компьютерные информационные ресурсы Методика использования	Лабораторная работа

		ресурсов Специализированные образовательные порталы Информационный поиск	
ОК-4 способность формировать ресурсно-информационные базы для осуществления практической деятельности в различных сферах	<b>ЗНАТЬ:</b> степень развития современной цивилизации, характеризующуюся увеличением роли информации, возрастанием доли информационных средств коммуникаций, информационных продуктов и услуг; знать отличительные черты информационного общества. <b>УМЕТЬ:</b> пользоваться программными средствами при создании различных ресурсно-информационных баз в практической деятельности. <b>ВЛАДЕТЬ:</b> навыками сбора, обработки и анализа информации.	Компьютерные информационные ресурсы Методика использования ресурсов Специализированные образовательные порталы Информационный поиск	Лабораторная работа
ОК-5 способность самостоятельно приобретать и использовать, в том числе с помощью информационных технологий, новые знания и умения, непосредственно не связанные со сферой профессиональной деятельности	<b>ЗНАТЬ:</b> иметь знания в области организации самостоятельной работы с учебной и научной литературой, поиска необходимых информационных источников (ресурсов) в локальных и глобальных сетях <b>УМЕТЬ:</b> использовать новые знания и умения, приобретенные в смежных предметных областях, для решения задач в своей профессиональной деятельности. <b>ВЛАДЕТЬ:</b> современными методиками, инновационными и информационными технологиями приобретения и использования новых знаний и умений, необходимых для решения задач, связанных или не связанных со сферой профессиональной деятельности.	Компьютерные информационные ресурсы Методика использования ресурсов Специализированные образовательные порталы Информационный поиск	Лабораторная работа
ПК-2 способностью формировать образовательную среду и использовать профессиональные знания и умения в реализации задач инновационной образовательной политики	<b>знать:</b> современные тенденции развития образовательной системы; теоретические положения, характеризующие образовательную среду и инновационную деятельность; виды инноваций в образовании; критерии инновационных процессов в образовании; <b>уметь:</b> поставить цели инновационной деятельности в образовательном учреждении; обосновать необходимость внесения запланированных изменений в образовательное учреждение; представить результаты работы в соответствии с принятыми решениями; внедрять инновационные приемы в педагогический процесс с целью создания условий для эффективной мотивации обучающихся; <b>владеть:</b> технологией планирования, организации и управления инновационной деятельностью в образовательном	Компьютерные информационные ресурсы Методика использования ресурсов Специализированные образовательные порталы Информационный поиск	Лабораторная работа Контрольная работа



	учреждении; - технологиями проведения опытно-экспериментальной работы, участия в инновационных процессах		
Промежуточная аттестация			КИМ

## 19.2 Описание критериев и шкалы оценивания компетенций (результатов обучения) при промежуточной аттестации:

Для оценивания результатов обучения на экзамене используются следующие показатели (ЗУНы из 19.1):

знать основные законы и подходы, объясняющие развитие сознания, научной мысли и образования; особенности и область применения основных мыслительных операций; основные понятия математической статистики как абстрактного языка научной деятельности; область применения операций анализа и синтеза в интеллектуальной деятельности; степень развития современной цивилизации, характеризующуюся увеличением роли информации, возрастанием доли информационных средств коммуникаций, информационных продуктов и услуг; знать отличительные черты информационного общества; иметь знания в области организации самостоятельной работы с учебной и научной литературой, поиска необходимых информационных источников (ресурсов) в локальных и глобальных сетях; современные тенденции развития образовательной системы; теоретические положения, характеризующие образовательную среду и инновационную деятельность; виды инноваций в образовании; критерии инновационных процессов в образовании;

уметь применять мыслительные операции для качественного и количественного анализа эмпирических данных; осуществлять синтез и выделение существенных свойств социальных объектов; пользоваться программными средствами при создании различных ресурсно-информационных баз в практической деятельности; использовать новые знания и умения, приобретенные в смежных предметных областях, для решения задач в своей профессиональной деятельности; поставить цели инновационной деятельности в образовательном учреждении; обосновать необходимость внесения запланированных изменений в образовательное учреждение; представить результаты работы в соответствии с принятыми решениями; внедрять инновационные приемы в педагогический процесс с целью создания условий для эффективной мотивации обучающихся;

владеть навыками всестороннего и критического анализа принятых решений и собственной деятельности в учебных ситуациях; навыками сбора, обработки и анализа информации; современными методиками, инновационными и информационными технологиями приобретения и использования новых знаний и умений, необходимых для решения задач, связанных или не связанных со сферой профессиональной деятельности; технологией планирования, организации и управления инновационной деятельностью в образовательном учреждении; - технологиями проведения опытно-экспериментальной работы, участия в инновационных процессах

Для оценивания результатов обучения на экзамене используется 2-балльная шкала: «зачет», «незачет»

Соотношение показателей, критериев и шкалы оценивания результатов обучения.

Критерии оценивания компетенций	Уровень сформированности компетенций	Шкала оценок
<p>Обучающийся демонстрирует достаточные способности</p> <p>знать основные законы и подходы, объясняющие развитие сознания, научной мысли и образования; особенности и область применения основных мыслительных операций; основные понятия математической статистики как абстрактного языка научной деятельности; область применения операций анализа и синтеза в интеллектуальной деятельности; степень развития современной цивилизации, характеризующуюся увеличением роли информации, возрастанием доли информационных средств коммуникаций, информационных продуктов и услуг; знать отличительные черты информационного общества; иметь знания в области организации самостоятельной работы с учебной и научной литературой, поиска необходимых информационных источников (ресурсов) в локальных и глобальных сетях; современные тенденции развития образовательной системы; теоретические положения, характеризующие образовательную среду и инновационную деятельность; виды инноваций в образовании; критерии инновационных процессов в образовании;</p> <p>уметь применять мыслительные операции для качественного и количественного анализа эмпирических данных; осуществлять синтез и выделение существенных свойств социальных объектов; пользоваться программными средствами при создании различных ресурсно-информационных баз в практической деятельности; использовать новые знания и умения, приобретенные в смежных предметных областях, для решения задач в своей профессиональной деятельности; поставить цели инновационной деятельности в образовательном учреждении; обосновать необходимость внесения запланированных изменений в образовательное учреждение; представить результаты работы в соответствии с принятыми решениями; внедрять инновационные приемы в педагогический процесс с целью создания условий для эффективной мотивации обучающихся;</p> <p>владеть навыками всестороннего и критического анализа принятых решений и собственной деятельности в учебных ситуациях; навыками сбора, обработки и анализа информации; современными методиками, инновационными и информационными технологиями приобретения и использования новых знаний и умений, необходимых для решения задач, связанных или не связанных со сферой профессиональной деятельности; технологией планирования, организации и управления инновационной деятельностью в образовательном учреждении; - технологиями проведения опытно-экспериментальной работы, участия в инновационных процессах</p>	Повышенный уровень	Зачет
<p>Ответ на контрольно-измерительный материал не соответствует одному (двум) из перечисленных показателей, но обучающийся демонстрирует достаточные способности</p> <p>знать основные законы и подходы, объясняющие развитие</p>	Базовый уровень	Зачет

<p>сознания, научной мысли и образования; особенности и область применения основных мыслительных операций; основные понятия математической статистики как абстрактного языка научной деятельности; область применения операций анализа и синтеза в интеллектуальной деятельности; степень развития современной цивилизации, характеризующуюся увеличением роли информации, возрастанием доли информационных средств коммуникаций, информационных продуктов и услуг; знать отличительные черты информационного общества; иметь знания в области организации самостоятельной работы с учебной и научной литературой, поиска необходимых информационных источников (ресурсов) в локальных и глобальных сетях; современные тенденции развития образовательной системы; теоретические положения, характеризующие образовательную среду и инновационную деятельность; виды инноваций в образовании; критерии инновационных процессов в образовании;</p> <p>уметь применять мыслительные операции для качественного и количественного анализа эмпирических данных; осуществлять синтез и выделение существенных свойств социальных объектов; пользоваться программными средствами при создании различных ресурсно-информационных баз в практической деятельности; использовать новые знания и умения, приобретенные в смежных предметных областях, для решения задач в своей профессиональной деятельности; поставить цели инновационной деятельности в образовательном учреждении; обосновать необходимость внесения запланированных изменений в образовательное учреждение; представить результаты работы в соответствии с принятыми решениями; внедрять инновационные приемы в педагогический процесс с целью создания условий для эффективной мотивации обучающихся;</p> <p>владеть навыками всестороннего и критического анализа принятых решений и собственной деятельности в учебных ситуациях; навыками сбора, обработки и анализа информации; современными методиками, инновационными и информационными технологиями приобретения и использования новых знаний и умений, необходимых для решения задач, связанных или не связанных со сферой профессиональной деятельности; технологией планирования, организации и управления инновационной деятельностью в образовательном учреждении; - технологиями проведения опытно-экспериментальной работы, участия в инновационных процессах</p>		
<p>Ответ на контрольно-измерительный материал не соответствует любым двум (трем) из перечисленных показателей, обучающийся дает неполные ответы на дополнительные вопросы. Демонстрирует частичные способности знать основные законы и подходы, объясняющие развитие сознания, научной мысли и образования; особенности и область применения основных мыслительных операций; основные понятия математической статистики как абстрактного языка научной деятельности; область применения операций анализа и синтеза в интеллектуальной деятельности; степень развития современной цивилизации, характеризующуюся увеличением роли информации, возрастанием доли информационных средств коммуникаций, информационных продуктов и услуг; знать отличительные черты</p>	<p>Пороговый уровень</p>	<p>Зачет</p>

<p>информационного общества; иметь знания в области организации самостоятельной работы с учебной и научной литературой, поиска необходимых информационных источников (ресурсов) в локальных и глобальных сетях; современные тенденции развития образовательной системы; теоретические положения, характеризующие образовательную среду и инновационную деятельность; виды инноваций в образовании; критерии инновационных процессов в образовании;</p> <p>уметь применять мыслительные операции для качественного и количественного анализа эмпирических данных; осуществлять синтез и выделение существенных свойств социальных объектов; пользоваться программными средствами при создании различных ресурсно-информационных баз в практической деятельности; использовать новые знания и умения, приобретенные в смежных предметных областях, для решения задач в своей профессиональной деятельности; поставить цели инновационной деятельности в образовательном учреждении; обосновать необходимость внесения запланированных изменений в образовательное учреждение; представить результаты работы в соответствии с принятыми решениями; внедрять инновационные приемы в педагогический процесс с целью создания условий для эффективной мотивации обучающихся;</p> <p>владеть навыками всестороннего и критического анализа принятых решений и собственной деятельности в учебных ситуациях; навыками сбора, обработки и анализа информации; современными методиками, инновационными и информационными технологиями приобретения и использования новых знаний и умений, необходимых для решения задач, связанных или не связанных со сферой профессиональной деятельности; технологией планирования, организации и управления инновационной деятельностью в образовательном учреждении; - технологиями проведения опытно-экспериментальной работы, участия в инновационных процессах</p>		
<p>Ответ на контрольно-измерительный материал не соответствует любым трем (четырем) из перечисленных показателей. Обучающийся демонстрирует отрывочные, фрагментарные знания по предмету: не понимает сути вопроса</p>	<p>–</p>	<p>незачет</p>

### **19.3 Типовые контрольные задания или иные материалы, необходимые для оценки знаний, умений, навыков и (или) опыта деятельности, характеризующие этапы формирования компетенций в процессе освоения образовательной программы**

#### **19.3.1 Перечень вопросов к зачету:**

1. История научного направления.
2. Классификация языков Хомского.
3. Особенности ЕЯ.
4. Знакомство с направлениями современных исследований.
5. Лингвистический процессор - функциональная структура. Методы морфологического анализа, используемые в лингвистических процессорах.
6. Морфологические словари. Алгоритмы синтаксического и семантического анализа для автоматических систем обработки текстов.

7. Парсеры ЕЯ. Прикладные системы- спэлчекеры, текстовые редакторы, системы профессионального редактирования.

8. Статистические методы анализа структур ЕЯ текста на морфологическом, синтаксическом, семантическом уровнях.

9. Метод позиционных статистик. Приложение методов для задач дешифровки ЕЯ текстов на неизвестных языках. Марковские цепи.

10. Математическая постановка задачи распознавания образов и классификации. Формальные методы определения сходства ЕЯ документов на различных уровнях лингвистического анализа (морфологическом, синтаксическом, семантическом): кластерный анализ, деревья принятия решений, векторные методы, Байесовский классификатор.

11. Применение методов классификации для задач определения авторства текстов.

12. Автоматическое извлечение знаний из ЕЯ текстов. Формирование онтологии предметной области по тексту. Построение семантической модели текста. Семантическая классификация и кластеризация текстов.

### 19.3.3. Пример задания для выполнения лабораторной/контрольной работы

Лабораторная работа № 4 datamining и классификация документов.

**Цель:** усвоить основы интеллектуального анализа данных и применение наивного байесовского классификатора для задачи категоризации текстовых документов.

*Основные теоретические сведения.*

**Интеллектуальный анализ данных (DataMining)** — это процесс обнаружения в множестве данных ранее неизвестных, полезных и доступных интерпретации знаний, необходимых для принятия решений в различных сферах человеческой деятельности. Английский термин «DataMining» не имеет однозначного перевода на русский язык (интеллектуальный анализ данных, добыча данных, вскрытие данных, информационная проходка, извлечение данных/информации) поэтому в большинстве случаев используется в оригинале.

Методы DataMining разделяются на две группы:

1. **статистические** (дескриптивный анализ, корреляционный и регрессионный анализ, факторный анализ, дисперсионный анализ, компонентный анализ, дискриминантный анализ, анализ временных рядов)
2. **кибернетические** (искусственные нейронные сети, байесовские сети, эволюционное программирование, генетические алгоритмы, ассоциативная память, нечеткая логика, деревья решений, системы обработки экспертных знаний)

Визуальные инструменты DataMining позволяют проводить анализ данных предметными специалистами (аналитиками), не владеющими соответствующими математическими знаниями.

*Задачи, решаемые иад*

1. **Классификация** — отнесение входного вектора (объекта, события, наблюдения) к одному из заранее известных классов.
2. **Кластеризация** — разделение множества входных векторов на группы (кластеры) по степени «похожести» друг на друга.
3. **Сокращение описания** — для визуализации данных, лаконизма моделей, упрощения счета и интерпретации, сжатия объемов собираемой и хранимой информации.
4. **Ассоциация** — поиск повторяющихся образцов. Например, поиск «устойчивых связей в корзине покупателя» (*marketbasketanalysis*) — вместе с пивом часто покупают орешки.
5. **Прогнозирование** – предсказание следующего состояния системы по наблюдаемым.

6. **Анализ отклонений** — выявление нетипичной сетевой активности позволяет обнаружить вредоносные программы.
7. ...

В литературе можно встретить еще ряд классов задач. Базовыми задачами являются первые три. Остальные задачи сводятся к ним тем или иным способом.

#### *Алгоритмы решения задач над*

Для задач классификации характерно «обучение с учителем», при котором построение (обучение) модели производится по выборке содержащей входные и выходные векторы.

Для задач кластеризации и ассоциации применяется «обучение без учителя», при котором построение модели производится по выборке, в которой нет выходного параметра. Значение выходного параметра («относится к кластеру ...», «похож на вектор ...») подбирается автоматически в процессе обучения.

Для задач сокращения описания характерно *отсутствие разделения на входные и выходные векторы*. Начиная с классических работ К. Пирсона по методу главных компонент, основное внимание здесь уделяется аппроксимации данных.

### **19.4. Методические материалы, определяющие процедуры оценивания знаний, умений, навыков и (или) опыта деятельности, характеризующих этапы формирования компетенций**

Оценка знаний, умений и навыков, характеризующая этапы формирования компетенций в рамках изучения дисциплины осуществляется в ходе текущей и промежуточной аттестаций.

Текущая аттестация проводится в соответствии с Положением о текущей аттестации обучающихся по программам высшего образования Воронежского государственного университета. Текущая аттестация проводится в форме *тестирования*. Критерии оценивания приведены выше.

Промежуточная аттестация проводится в соответствии с Положением о промежуточной аттестации обучающихся по программам высшего образования.

Контрольно-измерительные материалы промежуточной аттестации включают в себя теоретические вопросы, позволяющие оценить уровень полученных знаний и/или практическое(ие) задание(я), позволяющее(ие) оценить степень сформированности умений и(или) навыков, и(или) опыт деятельности лингвиста. При оценивании используются количественные шкалы оценок. Критерии оценивания приведены выше.

