Федеральное государственное автономное образовательное учреждение высшего образования
"Национальный исследовательский университет
"Высшая школа экономики"

Факультет гуманитарных наук
Школа лингвистики
Общеуниверситетская кафедра высшей математики

Рабочая программа дисциплины
«Математические основания компьютерной лингвистики»

для образовательной программы «Компьютерная лингвистика»
направления подготовки 45.04.03 «Фундаментальная и прикладная лингвистика»
подготовки магистра

Разработчики программы:
Щуров И.В., к.ф.-м.н., ischurov@hse.ru
Сахарова Н.Е., nsakharova@hse.ru

Одобрена на заседании Школы лингвистики «30» мая 2016 г.
Руководитель Школы лингвистики Е.В. Рахилина _____ [подпись]

Рекомендована Академическим советом образовательной программы
«06» июня 2016 г., Протокол № 7

Утверждена «06» июня 2016 г.
Академический руководитель образовательной программы
Бонч-Осмоловская А. А. _____ [подпись]

Москва, 2016

Настоящая программа не может быть использована другими подразделениями университета и другими вузами без разрешения подразделения-разработчика программы.

1. Область применения и нормативные ссылки

Настоящая программа учебной дисциплины устанавливает минимальные требования к знаниям и умениям студента и определяет содержание и виды учебных занятий и отчетности.

Программа предназначена для преподавателей, ведущих данную дисциплину, учебных ассистентов и студентов направления 45.04.03 «Фундаментальная и компьютерная лингвистика» подготовки магистра, изучающих дисциплину «Математические основания компьютерной лингвистики». Настоящая программа учебной дисциплины устанавливает минимальные требования к знаниям и умениям студента и определяет содержание и виды учебных занятий и отчетности.

Программа разработана в соответствии с:

- образовательным стандартом НИУ ВШЭ, в отношении которого установлена категория «Национальный исследовательский университет$

- образовательной программой «Теория языка и компьютерная лингвистика» подготовки магистра;

- рабочим учебным планом НИУ ВШЭ на 2016/2017 по направлению подготовки 45.04.03 «Фундаментальная и прикладная лингвистика», утвержденным в 2016 году.

2. Цели освоения дисциплины

Целями освоения дисциплины «Математические основания компьютерной лингвистики» являются

- Знакомство студентов со следующими разделами: линейная алгебра, теория вероятностей, математический анализ, необходимых для изучения прикладных дисциплин и дальнейшего практического применения. Формирование у слушателей ясного представления о базисных понятиях и методах данных разделов.
- развитие логического мышления и умения оперировать абстрактными объектами, привитие навыков корректного употребления математических понятий и символов для выражения различных количественных и качественных отношений;
- развитие навыка строгих математических рассуждений и доказательств.

3. Компетенции обучающегося, формируемые в результате освоения дисциплины

В результате освоения дисциплины студент должен:

- Знать: базовые математические понятия и определения, необходимые для дальнейшего изучения прикладных математических дисциплин;
- Уметь: применять необходимый математический инструментарий при решении задач;
- Иметь навыки: математической формализации задач, использования основных принципов и методов анализа, линейной алгебры и теории вероятностей в последующей профессиональной деятельности;

<u>В результате освоения дисциплины студент осваивает следующие компетенции:</u>

Компетенция	Код по ФГОС/ НИУ	Дескрипторы – основные признаки освоения (показатели достижения результата)	Формы и методы обучения, способствующие формированию и развитию компетенции
умение формулировать результат	ПК-3	Правильно воспроизводит чужие результаты Правильно формулирует собственные результаты	Компетенция формируется в любом сегменте учебного процесса Формируется в процессе активных занятий (участие в семинарах).
умение строго доказать утверждение	ПК-4	Воспроизводит доказательства стандартных результатов, услышанных на лекциях Оценивает строгость и корректность научных текстов анализа многих переменных	Изучение базового курса За счет повышения общей математической культуры в процессе обучения
умение грамотно пользоваться языком предметной области	ПК-7	Владеет профессиональной лексикой в области анализа многих переменных Распознает и воспроизводит названия основных математических структур, возникающих при изучении данной дисциплины, умеет корректно формулировать утверждения и их доказательства	Продумывание и повторение услышанного на семинарах и лекциях. Беседы с преподавателями во время консультаций. Компетенция достигается в процессе накопления опыта работы по данной теме и общения с преподавателями.
понимание корректности постановок задач	ПК-10	Понимает постановки проблем Адекватно оценивает корректность использования тех или иных математических методов, применяемых при формулировке и решении задач	Продумывание базовых понятий курса Вырабатывается в процессе решения задач, самостоятельного чтения, работы над курсовыми заданиями
выделение главных смысловых аспектов в доказательствах	ПК-16	Понимает и воспроизводит ключевые идеи, методы и геометрические конструкции пройденного математического материала. Обосновывает и оценивает мотивировки и логические ходы доказательств основных результатов анализа многих переменных.	Продумывание ключевых моментов лекций Вырабатывается путем активного решения задач, самообразования, общения с преподавателем

4. Место дисциплины в структуре образовательной программы;

Настоящая дисциплина относится к математическому и естественнонаучному циклу дисциплин направления 45.04.03 «Фундаментальная и прикладная лингвистика» подготовки магистра. В рабочем учебном плане 1 курса магистратуры дисциплина не является обязательной.

Изучение данной дисциплины базируется на знаниях математики в объеме базового курса «Математика» (адаптационного).

Основные положения дисциплины должны быть использованы в дальнейшем при изучении следующих дисциплин:

- Компьютерная лингвистика
- Математические основания компьютерной лингвистики (2 курс)

- Анализ лингвистический данных: квантитативные методы и визуализация (1 курс)
- Машинное обучение (1 курс)
- Социальные сети

5. Тематический план учебной дисциплины

№	Наименование разделов	Аудиторные часы		Всего аудиторных часов	Самостоятельная работа
		Лекции	Семинары		
1	Теория вероятностей	4	8	12	20
2	Линейная алгебра	4	16	20	60
3	Математический анализ	4	12	16	24
	ИТОГО	**12**	**36**	**48**	**104**

6. Формы контроля знаний студентов

Тип контроля	Форма контроля	1 год				Параметры
		1	2	3	4	
Текущий	Домашнее задание 1		*			Письменная работа
	Домашнее задание 2			*		Письменная работа
	Домашнее задание 3			*		Письменная работа
Итоговый	Итоговый экзамен				*	Письменная работа

6.1. Критерии оценки знаний, навыков

По всем формам отчетности оценки выставляются по 10-балльной шкале, при этом оценки могут быть дробными, округление проводится до первой цифры после запятой. При выставлении итоговой (окончательной) оценки по дисциплине производится округление до целого числа по обычным арифметическим правилам (если дробная часть оценки равна 0.5 или более, то округление производится в большую сторону). При выставлении накопленной оценки и оценки за экзаменационную контрольную работу в ведомость также производится их округление до целого числа по обычным арифметическим правилам. Перевод в 5-балльную шкалу из 10-балльной шкалы осуществляется согласно следующему правилу: $0 \le X < 4$ неудовлетворительно, $4 \le X < 6$ удовлетворительно, $6 \le X < 8$ хорошо, $8 \le X \le 10$ отлично. Аналогичная шкала используется и для итоговой оценки.

Оценки по всем формам текущего контроля выставляются по 10-ти балльной шкале. Основная форма текущего контроля – решение задач из домашних заданий (5-7 задач по каждой теме). Задачи подбираются так, чтобы их решение потребовало от студента свободного владения основными понятиями и умения пользоваться техническими (вычислительными) приемами, которые изучаются в соответствующем разделе курса. Обсуждение подходов к решению этих задач происходит на семинарах. Кроме того в рамках курса предусмотрена одна письменная контрольная работа.

Экзамен включает в себя письменное решение задач, решение которых требует от студента владения как понятийным, так и техническим аппаратом по изучавшимся в течение модуля темам

6.2. Порядок формирования оценок по дисциплине

Промежуточная оценка за курс = средний балл за домашние задания*4/7+оценка за промежуточную

контрольную*3/7.

Окончательная оценка = 0,3* экзамен + 0,7*накопленная оценка

Способ округления накопленной оценки текущего контроля: в пользу студента.

7. Содержание дисциплины

Тема 1. Теория вероятностей.

Основные законы распределения случайных величин. Дискретные и непрерывные случайные величины. Основные характеристики. Биномиальное и геометрическое распределения. Распределение Пуассона. Ковариация и коэффициент корреляции. Неравенство Чебышева. Центральная предельная теорема. Нормальное распределение на прямой.

Тема 2. Линейная алгебра.

Понятие линейного пространства и базиса. Матрицы и операции над матрицами. Определитель и ранг матрицы. Решение систем линейных уравнений, метод Гаусса. Линейные операторы. Преобразование матрицы линейного оператора при замене базиса. Собственные векторы и собственные значения линейного оператора. Приведение матрицы линейного оператора к диагональному виду. Характеристический многочлен линейного оператора. Элементы аналитической геометрии.

Тема 3. Математический анализ.

Пределы числовых последовательностей и рядов. Пределы функций. Непрерывность и дифференцируемость функции, производная функции. Касательная к графику функции. Исследование функции. Неопределенный и определенный интеграл. О-большое и о-малое. Ряд Тейлора. Геометрический смысл интеграла. Примеры функций нескольких переменных, графики и исследование функций нескольких переменных.

Изучение дисциплины проводится в режиме лекций и семинаров. Для текущего контроля используются четыре домашних задания.

9. Примеры заданий итогового контроля.

Задача 1. Вероятность встретить слово из шести и более букв в русских, английских, немецких и французских текстах художественных литературных произведений равна, соответственно, 31% , 16% ,23% и 20% . Наугад берется слово из каждого из четырех текстов на данных языках. Какова вероятность, что среди полученных таким образом 4 слов длину шесть и более букв имеют:

a) 0 слов; b) 1 слово; c) 2 слова; d) 3 слова; e) все 4 слова.

Задача 2. Пусть есть монета со смещенным центром тяжести.
a) Воспользуйтесь неравенством Чебышева, чтобы подсчитать сколько нужно совершить бросков монеты, чтобы с надежностью 95% найти вероятность выпадения орла с точностью до 0.01.
b) Теперь допустим, что нам не нужна такая большая надежность нашего вывода. Допустим, что нас устраивает и надежность 90% . Как это повлияло на ответ? Насколько меньше бросков нужно сделать по сравнению с предыдущим пунктом?
c) Пусть теперь надежность 95% , а точность требуется небольшая 0.1. Насколько уменьшилось необходимое число бросков?

Задача 3. Линейный оператор f растягивает векторы e1 = (1, 2, 3), e2 = (1, −1, 2) и e3 = (2, 1, 1) в 2, 3 и 4 раза соответственно. Найдите матрицу оператора в стандартном базисе.

Задача 4. Найдите сообщенные векторы и собственные значения данного линейного оператора. Возведите в степень матрицу линейного оператора.

Задача 5. Исследуйте функцию на максимумы, минимумы, промежутки монотонности. Найдите

возможные вертикальные/горизонтальные асимптоты. Нарисуйте график функции.. Примеры функций: 2x+e^{-2x}, x/(x^2+9).

Задача 6. Найдите площадь, ограниченную: графиком синуса и отрезком [0, π] оси Ox.

10. Учебно-методическое и информационное обеспечение дисциплины

1. Тюрин Ю.Н., Макаров А.А., Симонова Г.И. Теория вероятностей: учебник для экономических и гуманитарных специальностей. – М.: МЦНМО, 2009.
2. Шведов А.С. Теория вероятностей и математическая статистика: учебное пособие для вузов. – М.: Изд. дом ГУ ВШЭ. – 254 с.
3. Н.Ш. Кремер. Теория вероятностей и математическая статистика. - М.: ЮНИТИ, 2004.
4. S. Ross. A first course in probability. Prentice hall, Upper saddle river, New Jersey, 1997.
5. Гельфанд И.М. Лекции по линейной алгебре, изд. Наука, 1971.
6. Винберг Э.Б. Курс алгебры, изд. Факториал пресс, 2002.
7. J. Stewart., Calculus – Early Transcendentals 6e 2008