

МИНИСТЕРСТВО ОБРАЗОВАНИЯ И НАУКИ РОССИЙСКОЙ ФЕДЕРАЦИИ

Федеральное государственное бюджетное образовательное учреждение высшего образования «Московский государственный университет имени М.В.Ломоносова»

«Утверждаю»

Декан факультета ВМК МГУ
имени М.В. Ломоносова



Е.И. Моисеев

«__» _____ 2017 г.

РАБОЧАЯ ПРОГРАММА УЧЕБНОЙ ДИСЦИПЛИНЫ
« Методы машинного обучения и поиск достоверных закономерностей »

Уровень высшего образования – подготовка научно-педагогических кадров в аспирантуре

Направление подготовки – 01.06.01 «Математика и механика»

Направленность (профиль) – (01.01.09) «Дискретная математика и математическая кибернетика»

2017 г.

РАБОЧАЯ ПРОГРАММА ДИСЦИПЛИНЫ

1. НАИМЕНОВАНИЕ ДИСЦИПЛИНЫ

Методы машинного обучения и поиск достоверных закономерностей

2. УРОВЕНЬ ВЫСШЕГО ОБРАЗОВАНИЯ

Подготовка научно-педагогических кадров в аспирантуре.

3. НАПРАВЛЕНИЕ ПОДГОТОВКИ, НАПРАВЛЕННОСТЬ (ПРОФИЛЬ) ПОДГОТОВКИ

4. МЕСТО ДИСЦИПЛИНЫ В СТРУКТУРЕ ОСНОВНОЙ ОБРАЗОВАТЕЛЬНОЙ ПРОГРАММЫ

Дисциплина относится к специальным дисциплинам вариативной части образовательной программы и является обязательной для освоения.

5. ПЕРЕЧЕНЬ ПЛАНИРУЕМЫХ РЕЗУЛЬТАТОВ ОБУЧЕНИЯ

Дисциплина участвует в формировании следующих компетенций образовательной программы:

Формируемые компетенции	Планируемые результаты обучения
Способность обоснованно оценивать степень соответствия защищаемых объектов информатизации и информационных систем действующим стандартам в области информационной безопасности (ОПК-3)	З1 (ОПК-3) ЗНАТЬ принципы управления доступом в компьютерных системах, современные методы защиты информации при передаче ее по каналам связи, современные стандарты информационной безопасности У1(ОПК-3) УМЕТЬ: обосновать степень соответствия защищаемых объектов ин-

	<p>форматизации и информатизационных систем действующим стандартам в области информационной безопасности.</p>
<p>Способностью формулировать научные задачи в области обеспечения информационной безопасности, применять для их решения методологии теоретических и экспериментальных научных исследований, внедрять полученные результаты в практическую деятельность (ОПК-1);</p>	<p>З1(ОПК-1) ЗНАТЬ</p> <p>научные задачи в области обеспечения информационной безопасности</p> <p>У1(ОПК-1) УМЕТЬ:</p> <p>применять для их решения методологии теоретических и экспериментальных научных исследований, внедрять полученные результаты в практическую деятельность</p> <p>В1(ОПК-1) ВЛАДЕТЬ:</p> <p>Навыками внедрения полученных результатов в практическую деятельность</p>
<p>Владение современными методами построения и анализа математических моделей, возникающих при решении естественнонаучных задач, а также методами разработки и реализации алгоритмов их решения на основе фундаментальных знаний в области математики и информатики (ПК-1)</p>	<p>З1 (ПК-1) ЗНАТЬ:</p> <p>современные методы разработки и реализации алгоритмов организации работы вычислительных комплексов и компьютерных сетей последнего поколения</p> <p>У1 (ПК-1) УМЕТЬ:</p> <p>применять современные методы построения и анализа математических моделей, возникающих при решении естественнонаучных задач, а также современные методы разработки и</p>

	<p>реализации алгоритмов их решения</p> <p>V1 (ПК-1) ВЛАДЕТЬ:</p> <p>навыками оптимального выбора современных методов построения и анализа математических моделей, возникающих при решении естественнонаучных задач, а также современных методов разработки и реализации алгоритмов их решения</p>
<p>Способность разрабатывать и реализовывать алгоритмы организации работы современных вычислительных комплексов и компьютерных сетей (ПК-2)</p>	<p>З1 (ПК-2) ЗНАТЬ:</p> <p>современные методы разработки и реализации алгоритмов организации работы вычислительных комплексов и компьютерных сетей последнего поколения</p> <p>У1 (ПК-2) УМЕТЬ:</p> <p>применять современные методы разработки и реализации алгоритмов организации работы вычислительных комплексов и компьютерных сетей последнего поколения</p> <p>V1 (ПК-2) ВЛАДЕТЬ:</p> <p>навыками оптимального выбора современных методов разработки и реализации алгоритмов организации работы вычислительных комплексов и компьютерных сетей последнего поколения</p>

Оценочные средства для промежуточной аттестации приведены в Приложении.

6. ОБЪЕМ ДИСЦИПЛИНЫ

Объем дисциплины составляет 3 зачетных единицы, всего 108 часов.

24 часа составляет контактная работа с преподавателем – 21 часа занятий лекционного типа, 0 часов занятий семинарского типа (семинары, научно-практические занятия, лабораторные работы и т.п.), 0 часов индивидуальных консультаций, 3 часа мероприятий текущего контроля успеваемости, 0 часа групповых консультаций, 0 часов мероприятий промежуточной аттестации.

68 часов составляет самостоятельная работа аспиранта.

7. ВХОДНЫЕ ТРЕБОВАНИЯ ДЛЯ ОСВОЕНИЯ ДИСЦИПЛИНЫ

Учащиеся должны владеть знаниями по операционным системам, компьютерным сетям, базам данных, дискретной математике и основам кибернетики в объеме, соответствующем основным образовательным программам бакалавриата и магистратуры по укрупненным группам направлений и специальностей 01.00.00 «Математика и механика», 02.00.00 «Компьютерные и информационные науки».

8. ОБРАЗОВАТЕЛЬНЫЕ ТЕХНОЛОГИИ

В процессе обучения используется программное обеспечение для подготовки слайдов лекций

9. СОДЕРЖАНИЕ ДИСЦИПЛИНЫ

В курсе рассматриваются основные проблемы и задачи, связанные с теорией и применением методов машинного обучения и интеллектуального анализа данных. Рассказывается об основных направлениях и методах, технологиях машинного обучения. Даются примеры успешного решения прикладных задач.

Наименование и краткое содержание разделов и тем дисциплины (модуля),	Всего (часы)	В том числе	
		Контактная работа (работа во взаимодействии с преподавателем), часы из них	Самостоятельная работа обучающегося, часы из них

форма промежуточной аттестации по дисциплине (модулю)		Занятия лекционного типа	Занятия семинарского типа	Групповые консультации	Индивидуальные консультации	Учебные занятия, направленные на проведение текущего контроля успеваемости (коллоквиумы, практические контрольные занятия и др)*	Всего	Выполнение домашних заданий	Подготовка рефератов и т.п..	Всего
<p>Тема 1. Основные цели, понятия, проблемы машинного обучения</p> <p>Область применения методов, основанных на обучении по прецедентам (машинном обучении). Примеры применения. Понятие обучающей выборки. Способ обучения, основанный на минимизации эмпирического риска.</p> <p>Типы задач машинного обучения в зависимости от характера целевой переменной: распознавание, регрессионный анализ, анализ выживаемости</p> <p>Различные метрики для оценивания эффективности в задачах регрессион-</p>	5	2	-	-	-	-	3	2	-	2

<p>ного анализа и в задачах распознавания. ROC анализ. Основные цели метода. Способ построения ROC кривых</p> <p>Понятие обобщающей способности. Для каких алгоритмов достигается максимум обобщающей способности. Байесовский классификатор.</p> <p>Способы оценки обобщающей способности. Кросс-валидация.</p> <p>Эффект переобучения.</p> <p>1,2</p>										
<p>Тема 2. Современные методы регрессионного анализа</p> <p>Линейная регрессия. Использование метода наименьших квадратов для оценки коэффициентов. Поиск коэффициентов многомерной регрессии с помощью МНК. Явление мультиколлинеарности.</p> <p>Свойства оптимальных</p>	6	2	-	-	-	1	3	2	-	2

<p>регрессий.</p> <p>Трёхкомпонентное разложение обобщённой ошибки. Смысл шумовой составляющей, составляющей сдвига и дисперсионной составляющей. Bias-Variance дилемма.</p> <p>Методы регуляризации как средство повышения устойчивости. Гребневая регрессия, Лассо, эластичная сеть.</p> <p>[2]</p>										
<p>Тема 3. Статистические методы обучения</p> <p>Байесовские методы обучения. Аппроксимация с помощью многомерного нормального распределения.</p> <p>Линейный дискриминант Фишера. Метод k-ближайших соседей. Логистическая регрессия. Способ обучения</p> <p>[1,2]</p>	4	2	-	-	-	1	3		-	2

Тема 4. Нейросетевые методы обучения Модель искусственного нейрона. Перцептрон Розенблатта и метод его обучения, условие сходимости. Многослойный перцептрон и его структура. Аппроксимирующая способность многослойных перцептронов. Метод обратного распространения ошибки [1,2,3]	8	2	-	-	-	-	3	4	-	4
Тема 5 Комбинаторно-логические методы обучения Принцип частичной прецедентности. Понятие тупикового теста. Общая схема тестового алгоритма. Обобщение для вещественнозначной информации.	8	3	-	-	-	1	4	3	-	3

<p>Понятие тупикового представительного набора. Общая схема алгоритма распознавания, основанного на тупиковых представительных наборах. Обобщение для вещественнозначной информации.</p> <p>Модель Алгоритмов вычисления оценок. Понятия и опорного множества, функции близости, для вычисления оценок за классы. Компактные формулы для оценок в случае, когда признаки равноправны, а мощность опорных множеств фиксирована. Способы обучения для модели ABO</p> <p>[1]</p>										
<p>Тема 6 Метод опорных векторов</p> <p>Метод опорных векторов. Концепция максимального “зазора”. Сведение к задаче квадратичного программирования. Условия, налагаемые теоремой Каруша-Куна-Таккера. Двой-</p>	4	2	-	-	-	-	2	2	-	2

<p>ственная задача квадратичного программирования. Опорные вектора и их роль в формирование распознающего алгоритма.</p> <p>Обобщение исходного варианта метода опорных векторов на случай отсутствия линейной разделимости. “Смягчение” условия линейной разделимости с помощью введения дополнительных переменных. Основные отличия от исходного варианта метода.</p> <p>Обобщение метода опорных векторов, позволяющее строить нелинейные разделяющие поверхности [1,3]</p>										
<p>Тема 7 Решающие деревья и леса</p> <p>Решающие деревья. Методы обучения. Индексы неоднородности. Критерии остановки ветвления. Методы “подрезки”.</p> <p>Коллективные методы. Обоснование. Ошибка вы-</p>	4	2	-	-	-	-	2	2	-	2

<p>пуклой комбинации алгоритмов прогнозирования. Простые комитетные методы. Наивный Байесовский классификатор.</p> <p>Коллективные методы основанные на бутстрэп репликациях. Метод ы бэггинг и бустинг.</p> <p>Решающие леса [1,2,3,4]</p>										
<p>Тема 8 Голосование по системам закономерностей</p> <p>Методы, основанные на голосовании по системам логических закономерностей. Полные и частичные логические закономерности. Методы поиска. Коллективное решение.</p> <p>Метод «Статистически взвешенные синдромы». Оптимальные разбиения в рамках фиксированных моделей. Коллективное решение.</p>	4	2	-	-	-	-	1	2	-	2

[1,5]										
Тема 9 Методы анализа выживаемости Методы анализа выживаемости (надёжности). Оценки кривых выживаемости по методу Каплан-Майера. Модель Кокса. [7]	3	2	-	-	-	-	2	2	-	2
Тема 10 Методы анализа временных рядов. Стационарность временных рядов. Различные причины нестационарности. Модели прогнозирования временных рядов ARMA и ARIMA.	3									
Тема 11 Интеллектуальный анализ данных. Методы кластеризации. Цели кластерного анализа. Метод k-внутригрупповых средних. Иерархические методы кластеризации. Введение в байесовские	6	2	-	-	-	-	2	2	-	2

сети. Верификация закономерностей. Перестановочные тесты. Проблема множественного тестирования. [1, 8,9]										
	55	21				3		21	21	
11. Промежуточная аттестация – устный экзамен	33	2					31			
Итого	88	26					62			

10. УЧЕБНО-МЕТОДИЧЕСКИЕ МАТЕРИАЛЫ ДЛЯ САМОСТОЯТЕЛЬНОЙ РАБОТЫ УЧАЩИХСЯ

Самостоятельная работа учащихся состоит в изучении лекционного материала, учебно-методической литературы, подготовки к текущему контролю и промежуточной аттестации.

Литература для самостоятельной работы студентов в соответствии с тематическим планом .

Тема 1 «Основные цели, понятия, проблемы машинного обучения»

Тема 2 «Современные методы регрессионного анализа»

Тема 3 «Статистические методы обучения»

Тема 4 «Нейросетевые методы обучения»

Тема 5 «Комбинаторно-логические методы обучения »

Тема 6 «Метод опорных векторов»

Тема 7 «Решающие деревья и леса»

Тема 8 «Голосование по системам закономерностей »

Тема 9 «Методы анализа выживаемости»

Тема 10 «Интеллектуальный анализ данных»

11. РЕСУРСНОЕ ОБЕСПЕЧЕНИЕ

Основная литература

1. Журавлев Ю.И., Рязанов В.В., Сенько О.В. «Распознавание». Математические методы. Программная система. Практические применения. — М.: Фазис, 2006. 159 с.
2. T. Hastie, R. Tibshirani, J. Friedman The Elements of Statistical Learning: Data Mining, Inference, and Prediction. Springer Science & Business Media. 2013

Дополнительная литература

- 1 Мерков А.Б. Распознавание образов: Введение в методы статистического обучения. . М.: Едиториал УРСС, 2011. — 256 с.
- 2 Донской В.И. Алгоритмические модели обучения классификации: обоснование, сравнение, выбор. –Симферополь, «ДИАЙПИ», 2014,-227 с
- 3 Лбов Г.С., Старцева Н.С. Логические решающие правила и вопросы статистической устойчивости решений.Новосибирск: Изд-во Ин-та математики, 1999, 212 с.
- 4 Рязанов В.В. Логические закономерности в задачах распознавания (параметрический подход)//Ж. вычисл. Матем. И матем. Физ., 2007, [том 47, номер 10](#), страницы 1793–1808
- 5 Klein, John P., and Melvin L. Moeschberger. Survival Analysis: Techniques for Censored and Truncated Data. New York: Springer, 2003

- 6 P. GOOD, (2005), *Permutation, Parametric, and Bootstrap Tests of Hypotheses* (3rd ed.), Springer-Verlag, New York.
- 7 J.Pearl 2000 *Causality: Models, Inference and Inference*, New York, Cambridge University Press

Ресурсы информационно-телекоммуникационной сети «Интернет»

1. www.machinelearning.ru

Информационные технологии, используемые в процессе обучения

1. Программное обеспечение для подготовки слайдов лекций MS PowerPoint
2. Программное обеспечение для создания и просмотра pdf-документов Adobe Reader
3. Издательская система LaTeX.

Материально-техническая база

Для преподавания дисциплины требуется класс, оборудованный маркерной или меловой доской и проектором.

12. ЯЗЫК ПРЕПОДАВАНИЯ

Русский

13. РАЗРАБОТЧИК ПРОГРАММЫ, ПРЕПОДАВАТЕЛИ

Ведущий научный сотрудник, д.ф.-м.н Сенько Олег Валентинович

**ОЦЕНОЧНЫЕ СРЕДСТВА ДЛЯ ПРОМЕЖУТОЧНОЙ АТТЕСТАЦИИ ПО ДИСЦИПЛИНЕ
«Методы машинного обучения и поиск достоверных закономерностей»**

Средства для оценивания планируемых результатов обучения, критерии и показатели оценивания приведены ниже.

РЕЗУЛЬТАТ ОБУЧЕНИЯ по дисциплине (модулю)	КРИТЕРИИ и ПОКАЗАТЕЛИ ОЦЕНИВАНИЯ РЕЗУЛЬТАТА ОБУЧЕНИЯ по дисциплине (модулю) <i>(критерии и показатели берутся из соответствующих карт компетенций, при этом пользуются либо традиционной системой оценивания, либо БРС)</i>					ОЦЕНОЧНЫЕ СРЕДСТВА
	1	2	3	4	5	
	Неудовлетворительно	Неудовлетворительно	Удовлетворительно	Хорошо	Отлично	
ЗНАТЬ: принципы управления доступом в компьютерных системах, современные методы защиты информации при передаче ее по каналам связи, современные стандарты информационной безопасности 31 (ОПК-3)	Отсутствие знаний	Фрагментарные представления о принципах управления доступом в компьютерных системах, современных методах защиты информации при передаче ее по каналам связи, современных стандартах информационной безопасности	В целом сформированные, но неполные знания о принципах управления доступом в компьютерных системах, современных методах защиты информации при передаче ее по каналам связи, современных стандартах информационной безопасности	Сформированные, но содержащие отдельные пробелы знания о принципах управления доступом в компьютерных системах, современных методах защиты информации при передаче ее по каналам связи, современных стандартах информационной безопасности	Сформированные систематические знания о принципах управления доступом в компьютерных системах, современных методах защиты информации при передаче ее по каналам связи, современных стандартах информационной безопасности	Устный экзамен
УМЕТЬ: обосновать степень соответствия защищаемых объектов информатизации и	Отсутствие умений	Фрагментарные умения обоснования степени соответствия защищаемых объектов информатизации	В целом успешное, но не систематическое умение обоснования степени соответ-	Успешное, но содержащее отдельные пробелы умение обоснования степени соответ-	Сформированное умение обоснования степени соответствия защищаемых объектов информатизации	Контрольные работы

информатизационных систем действующим стандартам в области информационной безопасности. У1(ОПК-3)		и информатизационных систем действующим стандартам в области информационной безопасности.	ствия защищаемых объектов информатизации и информатизационных систем действующим стандартам в области информационной безопасности.	ствия защищаемых объектов информатизации и информатизационных систем действующим стандартам в области информационной безопасности.	и информатизационных систем действующим стандартам в области информационной безопасности.	
ЗНАТЬ: современные методы построения и анализа математических моделей, возникающих при решении естественнонаучных задач, а также современные методы разработки и реализации алгоритмов их решения 31 (ПК-1)	Отсутствие знаний	Фрагментарные представления о современных методах построения и анализа математических моделей, возникающих при решении естественнонаучных задач, а также современных методах разработки и реализации алгоритмов их решения	В целом сформированные, но неполные знания о современных методах построения и анализа математических моделей, возникающих при решении естественнонаучных задач, а также современных методах разработки и реализации алгоритмов их решения	Сформированные, но содержащие отдельные пробелы знания о современных методах построения и анализа математических моделей, возникающих при решении естественнонаучных задач, а также современных методах разработки и реализации алгоритмов их решения	Сформированные систематические знания о современных методах построения и анализа математических моделей, возникающих при решении естественнонаучных задач, а также современных методах разработки и реализации алгоритмов их решения	Устный экзамен
УМЕТЬ: применять современные методы построения и анализа математических моделей, возникающих при решении естественнонаучных задач, а также современные методы разработки и	Отсутствие умений	Фрагментарные умения применять современные методы построения и анализа математических моделей, возникающих при решении естественнонаучных задач, а также современные методы разработки и	В целом успешное, но не систематическое умение применять современные методы построения и анализа математических моделей, возникающих при решении ес-	Успешное, но содержащее отдельные пробелы умение применять современные методы построения и анализа математических моделей, возникающих при	Сформированное умение применять современные методы построения и анализа математических моделей, возникающих при решении естественнонаучных задач, а также современные методы разработки и	Контрольные работы

реализации алгоритмов их решения У1 (ПК-1)		реализации алгоритмов их решения	тественнонаучных задач, а также современные методы разработки и реализации алгоритмов их решения	решении естественнонаучных задач, а также современные методы разработки и реализации алгоритмов их решения	реализации алгоритмов их решения	
ВЛАДЕТЬ: навыками оптимального выбора современных методов построения и анализа математических моделей, возникающих при решении естественнонаучных задач, а также современных методов разработки и реализации алгоритмов их решения В1 (ПК-1)	Отсутствие навыков	Фрагментарное владение навыками оптимального выбора современных методов построения и анализа математических моделей, возникающих при решении естественнонаучных задач, а также современных методов разработки и реализации алгоритмов их решения	В целом успешное, но не полное владение навыками оптимального выбора современных методов построения и анализа математических моделей, возникающих при решении естественнонаучных задач, а также современных методов разработки и реализации алгоритмов их решения	Успешное, но содержащее отдельные пробелы владение навыками оптимального выбора современных методов построения и анализа математических моделей, возникающих при решении естественнонаучных задач, а также современных методов разработки и реализации алгоритмов их решения	Сформированное владение навыками оптимального выбора современных методов построения и анализа математических моделей, возникающих при решении естественнонаучных задач, а также современных методов разработки и реализации алгоритмов их решения	Контрольные работы, реферат
ЗНАТЬ: современные методы разработки и реализации алгоритмов организации работы вычислительных комплексов и компьютерных сетей последнего поколения Код 31 (ПК-2)	Отсутствие знаний	Фрагментарные представления о современных методах разработки и реализации алгоритмов организации работы вычислительных комплексов и компьютерных сетей последнего поколения	В целом сформированные, но неполные знания о современных методах разработки и реализации алгоритмов организации работы вычислительных комплексов и	Сформированные, но содержащие отдельные пробелы знания о современных методах разработки и реализации алгоритмов организации работы вычислительных	Сформированные систематические знания о современных методах разработки и реализации алгоритмов организации работы вычислительных комплексов и компьютерных сетей последнего по-	Устный экзамен

			компьютерных сетей последнего поколения	комплексов и компьютерных сетей последнего поколения	коления	
УМЕТЬ: применять современные методы разработки и реализации алгоритмов организации работы вычислительных комплексов и компьютерных сетей последнего поколения Код У1 (ПК-2)	Отсутствие умений	Фрагментарные умения применять современные методы разработки и реализации алгоритмов организации работы вычислительных комплексов и компьютерных сетей последнего поколения	В целом успешное, но не систематическое умение применять современные методы разработки и реализации алгоритмов организации работы вычислительных комплексов и компьютерных сетей последнего поколения	Успешное, но содержащее отдельные пробелы умение применять современные методы разработки и реализации алгоритмов организации работы вычислительных комплексов и компьютерных сетей последнего поколения	Сформированное умение применять современные методы разработки и реализации алгоритмов организации работы вычислительных комплексов и компьютерных сетей последнего поколения	Устный экзамен
ВЛАДЕТЬ: навыками оптимального выбора современных методов разработки и реализации алгоритмов организации работы вычислительных комплексов и компьютерных сетей последнего поколения Код В1 (ПК-2)	Отсутствие навыков	Фрагментарное владение навыками оптимального выбора современных методов разработки и реализации алгоритмов организации работы вычислительных комплексов и компьютерных сетей последнего поколения	В целом успешное, но не полное владение навыками оптимального выбора современных методов разработки и реализации алгоритмов организации работы вычислительных комплексов и компьютерных сетей последнего поколения	Успешное, но содержащее отдельные пробелы владение навыками оптимального выбора современных методов разработки и реализации алгоритмов организации работы вычислительных комплексов и компьютерных сетей последнего поколения	Сформированное владение навыками оптимального выбора современных методов разработки и реализации алгоритмов организации работы вычислительных комплексов и компьютерных сетей последнего поколения	Устный экзамен
ЗНАТЬ: научные задачи в об-	Отсутствие знаний	Фрагментарные представления о на-	В целом сформированные, но не-	Сформированные, но содержащие	Сформированные систематические	Устный экзамен

ласти обеспечения информационной безопасности З1(ОПК-1)		учных задачах в области обеспечения информационной безопасности	полные знания о научных задачах в области обеспечения информационной безопасности	отдельные пробелы о научных задачах в области обеспечения информационной безопасности	знания о научных задачах в области обеспечения информационной безопасности	
УМЕТЬ: применять для задачи в области обеспечения ИБ решения методологии теоретических и экспериментальных научных исследований, внедрять полученные результаты в практическую деятельность У1(ОПК-1)	Отсутствие умений	Фрагментарные умения применять для задачи в области обеспечения ИБ решения методологии теоретических и экспериментальных научных исследований, внедрять полученные результаты в практическую деятельность	В целом успешное, но не систематическое умение применять для задачи в области обеспечения ИБ решения методологии теоретических и экспериментальных научных исследований, внедрять полученные результаты в практическую деятельность	Успешное, но содержащее отдельные пробелы умение применять для задачи в области обеспечения ИБ решения методологии теоретических и экспериментальных научных исследований, внедрять полученные результаты в практическую деятельность	Сформированное умение применять для задачи в области обеспечения ИБ решения методологии теоретических и экспериментальных научных исследований, внедрять полученные результаты в практическую деятельность	Устный экзамен

Фонды оценочных средств, необходимые для оценки результатов обучения

Список вопросов для устного экзамена.

1	Какова область применения методов, основанных на обучении по прецедентам (машинном обучении).
2	Что такое способ обучения, основанный на минимизации эмпирического риска
3	Что такое ROC кривая ?
4	Способы оценки обобщающей способности. Кросс-валидация
5	Эффект переобучения.

6	Поиск коэффициентов многомерной регрессии с помощью МНК. Формула для регрессионных коэффициентов
7	Свойства оптимальных регрессий
8	Трёхкомпонентное разложение обобщённой ошибки. Смысл шумовой составляющей, составляющей сдвига и дисперсионной составляющей. Bias-Variance дилемма
9	Байесовские методы обучения. Что такое наивный байесовский классификатор?
10	Байесовские методы обучения. Аппроксимация с помощью многомерного нормального распределения.
11	Линейный дискриминант Фишера. Способ обучения.
12	Метод k-ближайших соседей. Способ обучения.
13	Логистическая регрессия. Способ обучения.
14	Распознавания при заданной точности распознавания одного из классов. Оптимальное решение согласно лемме Неймана-Пирсона.
15	Понятие тупикового теста. Общая схема тестового алгоритма. Обобщение для вещественнозначной информации.
16	Модель Алгоритмов вычисления оценок. Понятия и опорного множества, функции близости, для вычисления оценок за классы.
17	Модель Алгоритмов вычисления оценок. Компактные формулы для оценок в случае, когда признаки равноправны, а мощность опорных множеств фиксирована.
18	Модель искусственного нейрона. Перцептрон Розенблатта и метод его обучения, условие сходимости.
19	Многослойный перцептрон и его структура. Аппроксимирующая способность многослойных перцептронов. Метод обратного распространения ошибки.
20	Свёрточные нейронные сети.
21	Метод опорных векторов. Концепция максимального “зазора”. Сведение к задаче квадратичного программирования
21	Метод опорных векторов. Теорема Каруша-Куна-Таккера и сведение к двойственной задаче.
22	Обобщение метода опорных векторов, позволяющее строить нелинейные разделяющие поверхности.
23	Решающие деревья. Методы обучения. Индексы неоднородности. Критерии остановки ветвления
24	Коллективные методы. Обоснование. Ошибка выпуклой комбинации алгоритмов прогнозирования
25	Коллективные методы основанные на бутстрэп репликациях. Методы бэггинг и бустинг.

26	Методы бэггинг и бустинг.
27	Методы, основанные на голосовании по системам логических закономерностей
28	Метод «Статистически взвешенные синдромы»
29	Методы анализа выживаемости (надёжности). Оценки кривых выживаемости по методу Каплан-Майера.
30	Методы анализа выживаемости (надёжности). Модель Кокса.
31	Определение стационарности временного ряда в узком и широко смысле
32	Структура модели ARMA для прогнозирования временного ряда
31	Иерархические методы кластеризации
32	Метод кластеризации k-средних
33	Что такое байесовская сеть?
34	Верификация моделей с помощью перестановочных тестов.
35	В чём состоит проблема множественного тестирования?

Материалы для мероприятий текущего контроля. Примеры задач.

Задача 1

Тестирование системы распознавания для распознавания недобросовестных заёмщиков выявило связь между чувствительностью и ложной тревогой, показанную в таблице. Определить, принесёт ли эксплуатация системы к увеличению доходов банка и определить возможный прирост дохода в расчёте на одного заёмщика. Известно, что доход банка на одного заёмщика составляет 40000 денежных единиц, потери в результате отказа заёмщика от платежей составляют 90000 единиц. Доля недобросовестных заёмщиков составляет 20%.

Чувст	Лож. Тр.
0.02	0.0001
0.12	0.003
0.23	0.1

0.38	0.13
0.47	0.16
0.58	0.22
0.67	0.25
0.78	0.31
0.89	0.43
0.97	0.47
1.0	0.51

Задача 2

Наивный байесовский классификатор распознает объекты по двум признакам:

бинарный признак $X_1 \in \{a, b\}$ непрерывный признак X_2 , распределённый нормально.

Параметры распределения для классов K_1 и K_2 представлены в таблице:

μ - математическое ожидание для признака X_1 ;

σ - стандартное отклонение для признака X_2 ;

$p(a)$, $p(b)$ - вероятности значений a, b , соответственно

P_a - априорная вероятность класса

	μ	σ	$p(a)$	$p(b)$	P_a
K_1	2	1	0.3	0.7	0.7
K_2	-2	2	0.6	0.4	0.3

Выделить на числовой оси области значений показателя X с отнесением к классам K_1 и

K_2 при условии, что $X_1 = a$

Задача 3

По данным, представленным в таблице построить кусочно-непрерывную регрессию вида

$Y = \alpha_l + \beta_l X$ при $X \leq B$ и $Y = \alpha_r + \beta_r X$ при $X > B$ с точкой излома $B = 15$

. При поиске регрессионных коэффициентов использовать метод наименьших квадратов с $L2$ регуляризацией по коэффициентам β_l, β_r . Выразить регрессионные коэффициенты, а также среднеквадратичную ошибку модели на обучающей выборке, через множитель C

перед функцией штрафа .

Y	1	8	16	15	13	13
X	2	7	13	20	24	29

На экзамене достаточно решить задачи 1, 2

Или задачу 3

Методические материалы для проведения процедур оценивания результатов обучения

Особенности организации процесса обучения

Для эффективного освоения курса рекомендуется перед каждым занятием привести в порядок конспекты лекций. После каждого занятия рекомендуется найти и прочитать дополнительную литературу по теме лекции и прочитать свои конспекты.

Система контроля и оценивания

За каждую контрольную работу и реферат выставляются баллы (максимум 10 баллов за каждый вид работы). Пусть M – максимальное число баллов, которое может набрать студент. В конце семестра баллы конвертируются в оценку O_1 следующим образом:

меньше $M/2$ баллов: $O_1=2$;

больше или равно $M/2$ баллов, но меньше $2M/3$: $O_1=3$;

больше или равно $2M/3$ баллов, но меньше $5M/6$: $O_1=4$;

больше или равно $5M/6$ баллов: $O_1=5$.

На экзамене оценка O_1 является стартовой. Окончательная оценка определяется исходя из оценки устного ответа студента, при этом она не может отличаться от стартовой оценки более чем на 1 балл.

Структура и график контрольных мероприятий

Контрольная работа на 3-й, 8-й, 10-й, 14-й неделях, устный экзамен в конце семестра.