

**ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ БЮДЖЕТНОЕ НАУЧНОЕ
УЧРЕЖДЕНИЕ УФИМСКИЙ ФЕДЕРАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ
ЦЕНТР РОССИЙСКОЙ АКАДЕМИИ НАУК**

РЕКОМЕНДОВАНО

Директор ИНК УФИЦ РАН

 Д.Ш. Сабиров

«13» января 2026 г.

УТВЕРЖДАЮ

Руководитель УФИЦ РАН

 В.Б. Мартыненко

«13» января 2026 г.

ПРОГРАММА

кандидатского экзамена по научной специальности

1.2.1-Искусственный интеллект и машинное обучение

Программа составлена в соответствии с научной специальностью и отраслью науки, предусмотренными номенклатурой научных специальностей, по которым присуждаются учёные степени (утверждена Приказом Минобрнауки России от 24.02.2021 г. № 118).

Разработчики

 _____ д-р физ-мат. наук, проф. И.М. Губайдуллин

 _____ д-р физ.-мат. наук, доц. К.Ф. Коледина

Согласовано:

Заведующий отделом аспирантуры УФИЦ РАН

 _____ /М.Ю. Тимофеева

Ученый секретарь ИНК УФИЦ РАН

 _____ /И.Н. Павлова

**ПРОГРАММА-МИНИМУМ
КАНДИДАТСКОГО ЭКЗАМЕНА ПО СПЕЦИАЛЬНОСТИ
1.2.1-Искусственный интеллект и машинное обучение**

Введение

В основе настоящей программы лежит материал курсов: алгоритмы и структуры данных, дискретная математика, машинное обучение и анализ данных, теория вероятностей, математическая статистика, методы оптимизации.

1. Основные идеи и методы машинного обучения. Химические данные

Введение. Язык python. Анализ молекулярных данных. Основные идеи и методы машинного обучения. Химические данные. Перевод данных к машинно-читаемому виду: дескрипторы. Введение 0D-Дескрипторы, 1D-Дескрипторы, 2D-Дескрипторы Functional-Class Fingerprints (FCFPs), 3D/4D-Дескрипторы Область применимости. Метод ближайших соседей.

2. Задачи обучения по прецедентам. Байесовские методы классификации.

Метрические методы классификации

Функционал качества. Вероятностная постановка задачи обучения. Проблема переобучения и понятие обобщающей способности. Примеры прикладных задач. Задачи классификации. Задачи восстановления регрессии. Задачи ранжирования. Задачи кластеризации. Задачи поиска ассоциаций. Методология тестирования обучаемых алгоритмов. Приёмы генерации модельных данных. Вероятностная постановка задачи классификации. Функционал среднего риска. Оптимальное байесовское решающее правило. Задача восстановления плотности распределения. Непараметрическая классификация. Непараметрические оценки плотности. Метод парзеновского окна. Нормальный дискриминантный анализ. Многомерное нормальное распределение. Квадратичный дискриминант. Линейный дискриминант Фишера. Разделение смеси распределений. EM-алгоритм. Смеси многомерных нормальных распределений. Сеть радиальных базисных функций. Метрические методы классификации. Метод ближайшего соседа и его обобщения. Обобщённый метрический классификатор. Метод ближайших соседей. Метод парзеновского окна. Метод потенциальных функций. Отбор эталонных объектов. Понятие отступа объекта. Алгоритм STOLP для отбора эталонных объектов. Линейные методы классификации. Аппроксимация и регуляризация эмпирического риска.

3. Линейные методы классификации. Методы восстановления регрессии.

Искусственные нейронные сети. Кластеризация и визуализация

Линейная модель классификации. Метод стохастического градиента. Классические частные случаи. Эвристики для улучшения градиентных методов обучения. Логистическая регрессия. Обоснование логистической регрессии. Метод стохастического градиента для логистической регрессии. Скоринг и оценивание апостериорных вероятностей. Метод опорных векторов. Линейно разделимая выборка. Линейно неразделимая выборка. Ядра и спрямляющие пространства. ROC-кривая и оптимизация порога решающего правила. Метод наименьших квадратов. Линейная монотонная регрессия. Метод главных компонент. Нелинейные методы восстановления регрессии. Нелинейная модель регрессии. Метод опорных векторов в задачах регрессии.

Искусственные нейронные сети. Проблема полноты. Задача «исключающего ИЛИ». Вычислительные возможности нейронных сетей. Многослойные нейронные сети. Метод обратного распространения ошибок. Эвристики для улучшения сходимости. Оптимизация структуры сети. Кластеризация и визуализация. Алгоритмы кластеризации. Эвристические графовые алгоритмы. Функционалы качества кластеризации. Статистические алгоритмы. Иерархическая кластеризация. Сети Кохонена.

4. Вычислительная химия твердого тела. Deskрипторы кристаллической структуры. Предсказание свойств материалов

Этапы построения предсказательной модели. Классификация представлений материалов. Нейросетевые потенциалы взаимодействия. Физически интерпретируемые комбинации дескрипторов. Представления на основе химического состава. Структурные фрагменты кристаллического графа. Топологические дескрипторы.

5. Анализ данных на языке R. Введение в статистический пакет R

Команды и пакеты в R. Классы объектов, типы данных и структуры объектов в R. Среда разработки RStudio. Операторы цикла и условий. Создание собственных функций в R. Классы данных в R. Базовая графика в R.

6. Алгоритмы анализа данных и их реализация в среде RStudio

Алгоритмы линейной однофакторной и многофакторной регрессии. Задача классификации и кластеризации. Дерево решений. Обработка естественно-языковых текстов. Логистическая регрессионная модель. Реализация на языке R.

7. Введение в технологию высокопроизводительных параллельных вычислений.

Методология OpenMPI

Анализ параллельных алгоритмов. Принципы разработки параллельных методов. Параллельные алгоритмы умножения матрицы на вектор и умножения матриц. Подходы к распределению элементов матрицы между процессорами вычислительной системы. Способы распараллеливания алгоритма умножения матрицы на вектор, а также алгоритмов умножения матриц: алгоритм, основанный на ленточной схеме разделения данных, алгоритм Кэннона, быстрый алгоритм Винограда, а также умножение матриц по Штрассену. Оценки эффективности распараллеливания каждого алгоритма в зависимости от архитектуры вычислительной системы и применяемой технологии распараллеливания. Методы решения систем линейных алгебраических уравнений. Последовательные алгоритмы внутренней сортировки массива данных: сортировка пузырьком, сортировка вставками, сортировка Шелла, корневая сортировка, сортировка слиянием, быстрая сортировка. Параллельные алгоритмы сортировок. Реализация параллельных алгоритмов для различных технологий распараллеливания. Оценки эффективности распараллеливания алгоритмов сортировки данных. Директивы OpenMPI.

8. Методология MPI

Принципы построения вычислительных систем. Системы с распределенной памятью. Высокопроизводительные вычислительные системы. Классификация вычислительных систем. Программирование в MPI. Операции точка-точка. Общие

процедуры MPI. Умножение матрицы на вектор в MPI. Горизонтальное разбиение матрицы на блоки. Вертикальное разбиение матрицы на блоки.

9. Классификация задач и алгоритмов математического программирования и оптимизации

Задачи условной оптимизации и их классификация. «Сложные» задачи условной оптимизации. Пояснение использование априорной информации о целевой функции и функциях ограничениях для построения алгоритма (на примере задач линейного программирования, симплекс метод).

10. Методы машинного обучения в оптимизации

Задачи моделирования. Локальная оптимизация. Глобальная оптимизация Байесовский формализм. Эволюционные алгоритмы Swarm intelligence. Автонастройка гиперпараметров в методах оптимизации. Минусы и плюсы глобальной оптимизации.

Основная литература

1. Нейронные сети, генетические алгоритмы и нечеткие системы = Sieci neuronowe, algorytmy genetyczne i systemy rozmyte: научное издание / Д. Рутковская, М. Пилиньский, Л. Рутковский; пер. с пол. И. Д. Рудинский. М.: Горячая линия-Телеком, 2006, 452 с. Текст: непосредственный.
2. Алексеев Е.Р., Чеснокова О.В., Рудченко Е.А. Scilab. Решение инженерных и математических задач. М.: Бином, 2008, 260 с.
3. Норман Р. Дрейпер, Гарри Смит Прикладной регрессионный анализ. М.: Вильямс, 2007, 912 с.
4. Волков В.Л. Разработка алгоритмов оценивания процессов на основе Matlab. Методические указания к лабораторным работам. АПИ НГТУ, Арзамас: ОО «Ассоциация ученых», 2009, 58 с.
5. Волков В.Л. Программное обеспечение измерительных процессов. Метод. указания к лаб. работам. АПИ НГТУ, Арзамас: ОО «Ассоциация ученых», 2009, 63 с.
6. Машинное обучение: теория и практическое применение в процессах добычи нефти: учеб. пособие / А.С. Топольников, Б.М. Латыпов. Уфа: Изд-во УГНТУ, 2020, 67 с.
7. Агафонов, Е. Д. Прикладное программирование: учебное пособие / Е. Д. Агафонов, Г. В. Ващенко; Сибирский федеральный университет. Красноярск: Сибирский федеральный университет (СФУ), 2015, 112 с. Режим доступа: по подписке. URL: <https://biblioclub.ru/index.php?page=book&id=435640> (дата обращения: 07.09.2023). Текст: электронный.
8. Золотарюк, А. В. Язык и среда программирования R: учебное пособие / А.В. Золотарюк. Москва: ИНФРА-М, 2023, 162 с.
9. Зарова, Е. В. Прикладной многомерный статистический анализ: Презентации для лекций и примеры решений с использованием пакета R: Учебное пособие на английском языке / Зарова Е.В. Москва: ИНФРА-М, 2016, 370 с.
10. Федотов, И. Е. Параллельное программирование. Модели и приемы: практическое пособие / И. Е. Федотов. Москва: СОЛОН-Пресс, 2020, 390 с.
11. Афанасьев, К. Е. Основы высокопроизводительных вычислений: учебное пособие: [16+] / К. Е. Афанасьев, И. В. Григорьева, Т. С. Рейн. Кемерово: Кемеровский

государственный университет, 2012, Том 3. Параллельные вычислительные алгоритмы, 185 с.

12. Карпенко А.П. Современные алгоритмы поисковой оптимизации. Алгоритмы, вдохновленные природой: учебное пособие. М.: Изд-во МГТУ им. Н.Э. Баумана, 2014, 446с.

13. Понтрягин Л.С., Болтянский В.Г., Гамкредидзе Р.В., Мищенко Е.Ф. Математическая теория оптимальных процессов. 4-е изд. М.: Наука, Главная редакция физико-математической литературы, 1983, 392 с.

14. Беллман, Р. Динамическое программирование. М.: Иностранная литература, 1960, 400 с.

15. Самарский А. А., Михайлов А. П. Математическое моделирование. Идеи. Методы. Примеры. 2-е изд., испр. М.: Физматлит, 2001.

Дополнительная литература

1. Нейросетевые технологии: учебно-методическое пособие для выполнения лабораторных работ / УГНТУ, каф. ЦТиМ; сост. И. В. Ахметов, Р. М. Узьянбаев. Уфа: УГНТУ, 2023, 363 Кб URL: http://bibl.rusoil.net/base_docs/UGNTU/IES/Akhmetov17335.pdf (дата обращения: 10.10.2023). Текст: электронный.

2. Теофили, Т. Глубокое обучение для поисковых систем: руководство / Т. Теофили; перевод с английского Д. А. Беликова. Москва: ДМК Пресс, 2020, 318 с. ISBN 978-5-97060-776-3.

3. Коледина К.Ф., Коледин С.Н. Статистика на языке программирования R: учеб. пособие. Уфа: Изд-во УГНТУ, 2021, 107 с. ISBN 978-5-7831-2065-7.

4. Программирование на языке высокого уровня: учебно-методическое пособие для проведения лабораторных занятий со студентами направления 230100 "Информатика и вычислительная техника" / УГНТУ, каф. ВТИ; сост.: А. А. Габдуллина, Е. В. Дружинская. Уфа: Изд-во УГНТУ, 2011, 1,10 МБ. URL: http://bibl.rusoil.net/base_docs/UGNTU/VTIK/Gabdullina1.pdf. Текст: электронный.

5. Николаев, Е. И. Параллельные вычисления: учебное пособие / Е. И. Николаев; Северо-Кавказский федеральный университет. Ставрополь: Северо-Кавказский Федеральный университет (СКФУ), 2016, 185 с.

6. Абрамян, М. Э. Практикум по параллельному программированию с использованием электронного задачника Programming Taskbook for MPI: учебное пособие / М. Э. Абрамян. Ростов-на-Дону: Издательство ЮФУ, 2010.

7. Архитектура параллельных вычислительных систем: учебно-методическое пособие / УГНТУ, каф. ВТИК; сост.: А. Е. Белозеров, Г. Н. Жолобова. Уфа: УГНТУ, 2018.

8. Жалнин, Р.В. Основы параллельного программирования с использованием технологий MPI и OpenMP: учебное пособие / Р.В. Жалнин, Е.Н. Панюшкина, Е. Е. Пескова, П.А. Шаманаев. Саранск: Изд-во СВМО, 2013, С. 9-45.

9. Волков В.Л. Разработка алгоритмов оценивания процессов на основе Matlab. Методические указания к лабораторным работам. АПИ НГТУ, Арзамас: ОО «Ассоциация ученых», 2009, 58 с.

10. Семененко М. "Введение в математическое моделирование". М.: Солон-Р, 2002, 112 с.

11. Машинное обучение: теория и практическое применение в процессах добычи нефти: учеб. пособие / А.С. Топольников, Б.М. Латыпов. Уфа: Изд-во УГНТУ, 2020, 67 с.

ДОПОЛНИТЕЛЬНАЯ ПРОГРАММА КАНДИДАТСКОГО ЭКЗАМЕНА

по специальности 1.2.1-Искусственный интеллект и машинное обучение

1. Методы искусственного интеллекта и принципы создания экспертных систем

Экспертные системы и представление знаний. Подходы к построению систем искусственного интеллекта. Определение экспертной системы. Классификация экспертных систем. Структура экспертной системы. Технология разработки экспертной системы. Экспертные системы на основе нечетких знаний. Графические структуры для представления понятий и отношений. Работа с неточными данными и лингвистическими переменными.

Основные концепции нейронных сетей. Биологический нейрон и его свойства. Модель искусственного нейрона. Параметры нейрона. Характеристика функций активации. Архитектуры искусственных нейронных сетей (ИНС). Сбор данных для нейронной сети.

Генетические алгоритмы, назначение и особенности. Генетические алгоритмы, назначение и особенности. Обучение нечетких нейронных сетей на основе генетических алгоритмов. Применение генетических алгоритмов в системах искусственного интеллекта, оптимизации решения обратных задач химической кинетики.

Методы решения задач многоцелевой оптимизации. Априорные алгоритмы. Алгоритмы зондирования. Генетический алгоритм с векторной оценкой. Генетический алгоритм многокритериальной оптимизации. Генетический алгоритм недоминируемой сортировки. Подбор гиперпараметров генетического алгоритма недоминируемой сортировки.

Рекомендованная литература

Основная

1. Рутковская Д., Пилиньский М., Рутковский Л. Нейронные сети, генетические алгоритмы и нечеткие системы. Горячая Линия, Телеком, 2007.
2. Саймон Хайкин. Нейронные сети. Полный курс, Издательство: Вильямс, 2006.
3. Гладков Л.А. Генетические алгоритмы. Гриф УМО ВУЗов России, Физматлит, 2006.
4. Барский А.Б. Логические нейронные сети. М.: Интернет-ун-т инф. технол-й, Бином. Лаборатория знаний, 2007, 352 с.
5. Дьяконов В.П. MATLAB 6.5 SP1/7/7SP1/7SP2 + Simulink 5/6. Инструменты искусственного интеллекта и биоинформатики. М.: СОЛОН-Пресс, 2006, 456 с.
6. Тархов Д.А. Нейронные сети как средство математического моделирования. М.: Радиотехника, 2006, 48 с.
7. Хайкин С. Нейронные сети. Полный курс: пер. с англ., 2-е изд., испр. М.: Вильямс, 2006, 1104 с.
8. Яхьяева Г.А. Нечеткие множества и нейронные сети, 2-е изд., испр. М.: Интернет-ун-т инф. технол-й, Бином. Лаборатория знаний, 2010, 320 с.
9. Гаврилова Т.А. Базы знаний интеллектуальных систем. Учебник. СПб.: ПИТЕР, 2000.
10. Кофман А. Введение в теорию нечетких множеств: Пер. с франц. М.: Радио и связь, 1982, 432 с.

Дополнительная

1. Стюарт Рассел, Питер Норвиг, Искусственный интеллект. Современный подход, Вильямс, 2007.
2. Усков А.А., Кузьмин А.В. Интеллектуальные технологии управления. Искусственные нейронные сети и нечеткая логика, Горячая линия-Телеком, 2004.
3. Головки В.А. Нейронные сети: обучение, организация и применение. ИПРЖР, 2002.
4. Питер Джексон. Введение в экспертные системы. Издательство: Вильямс, 2001, 624 с.
5. Джордж Ф. Люгер. Искусственный интеллект: стратегии и методы решения сложных проблем. Издательство: Вильямс, 2005, 864 с.
6. Электронный учебник <http://www.neuroproject.ru/tutorial.php> (компания НейроПроект).
7. Электронный учебник <http://www.statsoft.ru/home/textbook/default.htm> (компания StatSoft Russia).
8. Шварева Е.Н., Еникеева Л.В., Губайдуллин И.М. Оптимизация роом частиц: учебное пособие. Уфа: Изд-во УГНТУ, 2021.