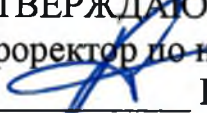


МИНОБРНАУКИ РОССИИ
Федеральное государственное бюджетное образовательное учреждение высшего
образования «Челябинский государственный университет»
(ФГБОУ ВО «ЧелГУ»)

УТВЕРЖДАЮ
Проректор по научной работе

И.В. Бычков
«31» _____ 2022 г.



**ПРОГРАММА ВСТУПИТЕЛЬНОГО ИСПЫТАНИЯ
ПО СПЕЦДИСЦИПЛИНЕ**

**Группа научных специальностей – 2.3. Информационные технологии
и телекоммуникации**

Научная специальность

2.3.1. Системный анализ, управление и обработка информации

Уровень образования

Высшее образование - подготовка кадров высшей квалификации

Форма обучения - очная

Челябинск, 2022

1. ОБЩИЕ ПОЛОЖЕНИЯ

Программа вступительного испытания по научной специальности - 2.3.1. Системный анализ, управление и обработка информации, относящихся к группе научных специальностей – 2.3. Информационные технологии и телекоммуникации, составлена на основе федеральных государственных образовательных стандартов высшего образования соответствующих уровней образования (специалитет, магистратура).

Вступительное испытание нацелено на оценку знаний поступающих лиц, полученных ими в ходе освоения программ высшего образования и на отбор среди поступающих лиц наиболее способных и подготовленных к освоению программ подготовки научных и научно-педагогических кадров в аспирантуре.

Вступительное испытание проводится в рамках нескольких конкурсов и сдается однократно.

Вступительное испытание принимает экзаменационная комиссия.

Вступительное испытание проводится на русском языке.

Вступительное испытание проводится очно или с использованием дистанционных технологий в случаях, предусмотренных Правилами приема.

2. СОДЕРЖАНИЕ ПРОГРАММЫ

1. Основные понятия и задачи системного анализа

- 1) Понятие о системном подходе, системном анализе. Выделение системы из среды, определение системы. Системы и закономерности их функционирования и развития.
- 2) Управляемость, достижимость, устойчивость. Модели систем: статические, динамические, концептуальные, топологические, формализованные (процедуры формализации моделей систем), информационные, логико-лингвистические, семантические, теоретико-множественные и др.
- 3) Классификация систем. Естественные, концептуальные и искусственные, простые и сложные, целенаправленные, целеполагающие, активные и пассивные, стабильные и развивающиеся системы.
- 4) Основные методологические принципы анализа систем. Задачи системного анализа. Роль человека в решении задач системного анализа.

2. Методы оптимизации

- 1) Математическое программирование. Типы экстремумов функций многих переменных, условия локального экстремума, метод множителей Лагранжа. Основные понятия выпуклого программирования. Седловые точки. Функция Лагранжа. Теорема Куна – Таккера и ее геометрическая интерпретация. Современные методы градиентной оптимизации.

Формулировка задачи линейного программирования (ЛП), экономическая интерпретация.

- 2) Понятия опорного плана и базиса, вырожденность и невырожденность задач ЛП, основные принципы симплекс-метода. Основные теоремы ЛП.
- 3) Потоки в сетях. Теорема Форда – Фалкерсона. Транспортная задача.
- 4) Динамическое программирование. Примеры задач, решаемых методом динамического программирования. Задача оптимального управления и принцип максимума Понтрягина.

3. Алгебра

- 1) Линейное пространство. Линейная зависимость. Базис. Системы линейных уравнений (СЛУ). Критерий совместности СЛУ. Обратная и псевдообратная матрицы.
- 2) Линейные операторы. Собственные векторы и собственные значения квадратной матрицы и симметричной квадратной матрицы. Диагонализация матрицы линейного оператора. Сингулярные числа прямоугольных матриц и их связь с собственными числами ассоциированных матриц. Матричные разложения. Скалярное произведение. Ортогональность. Булева алгебра.

4. Основы теории вероятностей и математической статистики

- 1) Случайные величины. Распределение дискретных случайных величин. Характеристики распределений. Основные законы распределения непрерывных случайных величин. Функции плотности распределения, свойства и квантили одномерной, двумерной и n -мерной нормальной случайной величины. Распределения хи-квадрат, Стьюдента, Снедекора – Фишера, логнормальное и равномерное.
- 2) Закон больших чисел (в форме Чебышева) как выражение свойства статистической устойчивости среднего значения. Центральная предельная теорема.
- 3) Генеральная совокупность, выборка и ее основные характеристики (среднее значение, дисперсия, асимметрия, квантили, функции распределения и плотности). Понятие статистической гипотезы и статистического критерия. Основные понятия теории статистических оценок и свойства оценок (несмещенность, состоятельность, асимптотическая нормальность, эффективность).

5. Интеллектуальные информационные системы. Машинное обучение

- 1) Классификация интеллектуальных информационных систем. Обеспечение работы интеллектуальных информационных систем. Этапы разработки интеллектуальных информационных систем.

- 2) Экспертные системы. Характеристики экспертных систем. Базовые функции экспертных систем. Классификация систем, основанных на знаниях.
- 3) Понятия информация, данные, знания. Отличия знаний от данных. Инженерия знаний. Процедурные и декларативные знания.
- 4) Модели представления знаний в интеллектуальных системах. Сетевые модели (семантические сети). Фреймовые модели. Логические модели представления знаний. Системы продукций. Нечеткие модели. Нейросетевые модели. Байесовские модели (сети доверия). Графы знаний (Knowledge Graph).
- 5) Методы приобретения знаний. Построение баз знаний для экспертных систем. Проблемы обучения интеллектуальных систем. Методы и средства интеллектуального анализа данных. Методы и проблемы автоматического извлечения знаний.
- 6) Машинное обучение. Задачи обучения с учителем и без учителя. Алгоритмы классификации. Деревья решений. Случайный лес. Линейная регрессия. Логистическая регрессия. Метод опорных векторов. Стохастический градиентный спуск. Ансамбли алгоритмов. Регуляризация. Проблемы переобучения. Алгоритмы кластеризации. Методы снижения размерности, метод главных компонент. Проблемы обработки больших данных. Методы online-обучения. Метрики оценки качества моделей машинного обучения, способы валидации. Проблема несбалансированных обучающих выборок и методы ее решения. Инженерия признаков, методы нормализации и предобработки данных.
- 7) Нейронные сети. Элементы и архитектура нейронных сетей. Модель нейрона. Функции активации. Методы обучения нейронных сетей. Оптимизация обучения нейронных сетей. Архитектуры глубоких нейронных сетей. Сверточные и рекуррентные сети. Механизмы внимания (attention). Перенос обучения (transfer learning). Self-attention и архитектуры трансформеров (Transformer).

6. Информационный поиск и автоматическая обработка текстов на естественном языке

- 1) Цели и задачи информационного поиска. Вопросно-ответный поиск.
- 2) Типы индексов в информационном поиске. Подготовка данных для индексирования. Словари и нечеткий поиск.
- 3) Методы ранжирования, взвешивания терминов и модели векторного пространства. Ранжирование на основе машинного обучения.
- 4) Оценка качества информационного поиска. Метрики оценка точности поиска и качества ранжирования.
- 5) Задачи обработки естественного языка. Основные этапы обработки естественного языка. Предобработка текста.
- 6) Семантическая близость и семантическая связанность. Меры семантической близости.

- 7) Способы векторных представлений слов и документов. Подходы «bag-of-words» и подход дистрибутивной семантики. Нейросетевые модели языка.
- 8) Методы вероятностного тематического моделирования. Вероятностный латентно-семантический анализ и латентное размещение Дирихле.
- 9) Методы классификации текстовых документов. Применение нейронных сетей в задачах классификации текстов. Архитектуры сверточных и рекуррентных нейронных сетей в задачах обработки естественного языка. Трансформеры в обработке естественного языка.

7. Алгоритмы и вычислительная сложность

- 1) Алгоритмы и структуры данных. Асимптотический анализ сложности алгоритмов. Подходы к проектированию алгоритмов: «разделяй и властвуй», динамическое программирование, жадная стратегия. Алгоритмы сортировки, двоичного поиска. Алгоритмы на графах: обход графа, поиск кратчайших путей, построение минимального остовного дерева. Двоичные деревья поиска, кучи, хеш-таблицы.
- 2) Машина Тьюринга. Тезис Черча. Неразрешимость проблемы останова машины Тьюринга.
- 3) Классы задач P и NP, примеры. Сводимость задач по Карпу и Тьюрингу. NP-полнота. Теорема Кука.

8. Дискретный анализ, теория графов

- 1) Комбинаторные методы дискретного анализа. Классические задачи комбинаторного анализа. Разбиения и размещения. Основные комбинаторные тождества. Задачи о кодировании информации.
- 2) Бинарные отношения и графы. Основы теории графов: определение графа, цепи, циклы, пути, контуры. Способы представления графов. Пути в графе. Связность. Связные и сильно связные графы. Теорема о связности двух вершин, имеющих нечетную локальную степень. Максимальное число ребер в графе с n вершинами и k связными компонентами. Достаточное условие связности графа с n вершинами. Матрица смежности графа. Матрица инцидентий дуг и ребер графов.
- 3) Деревья. Связанность любых двух вершин дерева единственным простым путем. Проблема визуализации деревьев. Кратчайшие пути и контуры.

3. ПРОЦЕДУРА ВСТУПИТЕЛЬНОГО ЭКЗАМЕНА

Вступительное испытание проводится в устной форме по билетам (приложение 1). Каждый билет содержит 2 вопроса. На подготовку вопроса отводится 30 минут. Записи при подготовке к ответу поступающие делают на учтенном комиссией листе, где указывается фамилия, номер билета и время его получения.

Во время вступительного испытания комиссией могут быть заданы дополнительные или уточняющие вопросы. После ответа черновые записи и билет сдаются председателю комиссии. Записи должны быть подписаны с указанием даты вступительного экзамена. При подготовке к ответу разрешается пользоваться программой вступительного испытания, выдаваемой комиссией.

Программа вступительного экзамена содержит 20 вопросов. Экзамен проводится по билетам, каждый из которых содержит один теоретический вопрос (часть I) и один практический (часть II).

Вступительное испытание поступающий сдаёт один раз. Передача вступительного испытания не допускается, за исключением случаев удовлетворения апелляции о нарушении процедуры вступительного испытания.

Во время испытания не разрешается пользоваться справочными материалами на бумажных или электронных носителях.

4. ВОПРОСЫ К ВСТУПИТЕЛЬНОМУ ИСПЫТАНИЮ

- 1) Понятие системы и основные свойства систем. Проиллюстрировать на примерах. Описать закономерности развития для интеллектуальных систем.
- 2) Задачи системного анализа. Применение системного анализа при построении интеллектуальных систем. Основные сложности и нерешенные научные проблемы.
- 3) Модели систем. Процедуры формализации моделей систем. Сопоставить модели систем с моделями представления знаний. Сложности автоматического извлечения знаний. Нерешенные научные проблемы извлечения и представления знаний.
- 4) Изложить основные понятия теории графов. Методы описания графов. Аспекты применения графов в интеллектуальных системах: графы знаний и графовые нейронные сети.
- 5) Отличие понятий: данные, информация, знания. Модели представления знаний в интеллектуальных системах. Продемонстрировать особенности различных моделей на примерах анализа некоторой предметной области.
- 6) Стохастический градиентный спуск. Изложить суть метода. Связь метода с обучением нейронных сетей. Сложности оптимизации обучения нейронных сетей.

- 7) Изложить основы глубокого обучения нейронных сетей. Сверточные и рекуррентные нейронные сети. Преимущества и недостатки их применения в задачах обработки естественного языка.
- 8) Методы ранжирования, взвешивания терминов и модели векторного пространства в информационном поиске. Привести различные примеры применения методов машинного обучения в информационном поиске. Актуальные научные проблемы информационного поиска с применением методов машинного обучения.
- 9) Изложить основные понятия построения искусственных нейронных сетей. Способы оптимизации обучения нейронных сетей. Механизмы внимания при обучении нейронных сетей и их практическое применение в задачах обработки естественного языка.
- 10) Архитектуры трансформеров и механизм Self-attention в нейронных сетях. Какие преимущества и недостатки в сравнении с сетями прямого распространения, сверточными и рекуррентными сетями. Привести примеры применения трансформер-архитектур в задачах обработки естественного языка. Сложности и нерешенные проблемы обучения трансформер-архитектур.
- 11) Понятие алгоритма. Анализ сложности алгоритмов. В качестве примера провести оценку сложности не менее четырех различных алгоритмов машинного обучения на примерах в программном коде (в среде Jupyter Notebook).
- 12) Вероятностный латентно-семантический анализ текстов, вероятностное тематическое моделирование. Проиллюстрировать на примерах в программном коде (в среде Jupyter Notebook), изложив основные понятия теории вероятностей, включая основные законы распределения непрерывных случайных величин.
- 13) Пояснить отличия реализации методов линейной регрессии и логистической регрессии, изложив необходимые понятия линейной алгебры. Проиллюстрировать работу методов на примерах в программном коде (в среде Jupyter Notebook).
- 14) Методы построения ансамблей алгоритмов машинного обучения. Пояснить связь с понятиями теории вероятностей и математической статистики. Проиллюстрировать работу методов на примерах в программном коде (в среде Jupyter Notebook).
- 15) Проблемы переобучения. Способы борьбы с переобучением. Виды регуляризации. Пояснить на примерах в программном коде (в среде Jupyter Notebook) с изложением понятий линейной алгебры.
- 16) Семантическая близость: понятие и постановка задачи. Меры семантической близости. Описать различные способы расчета семантической близости: статистические методы, машинное обучение, нейросетевые алгоритмы. Как оценивается качество методов. Проиллюстрировать работу методов на примерах в программном коде (в среде Jupyter Notebook).

- 17) Постановка задачи кластеризации. Методы обучения без учителя. Способы кластеризации текстовых документов. Применение нейронных сетей и методов тематического моделирования в задачах кластеризации текстов. Особенности кластеризации коротких текстов. Проиллюстрировать на примерах в программном коде (в среде Jupyter Notebook).
- 18) Пояснить работу метода опорных векторов, изложив необходимый математический аппарат, уделив внимание принципам математического программирования. Теорема Куна – Таккера и ее геометрическая интерпретация. Проиллюстрировать работу метода на примерах в программном коде (в среде Jupyter Notebook).
- 19) Методы снижения размерности, метод главных компонент. Смысл методов и практика их применения. Проиллюстрировать на примерах в программном коде (в среде Jupyter Notebook), изложив необходимые понятия линейной алгебры, уделив особое внимание сингулярному разложению.
- 20) Методы машинного обучения: деревья решений и случайный лес. Проиллюстрировать на примерах работу методов в программном коде (в среде Jupyter Notebook), изложив необходимые понятия теории графов и теории вероятностей. Преимущества и недостатки данных методов в сравнении с другими методами машинного обучения.

5. КРИТЕРИИ ОЦЕНКИ РЕЗУЛЬТАТОВ ВСТУПИТЕЛЬНОГО ИСПЫТАНИЯ

Максимальное количество баллов за вступительное испытание – 100 баллов.

Минимальное количество баллов за успешное прохождение вступительного испытания, независимо от условия поступления, соответствует минимальным баллам, утвержденным Правилами на текущий год.

«Отлично» (от 91 до 100) – поступающий обнаружил всестороннее, систематическое и глубокое знание учебно-программного материала, исчерпывающе, последовательно, грамотно и логически стройно его излагает, не затрудняется с ответом при видоизменении задания, свободно справляется с поставленными задачами, показывает знания монографического материала, правильно обосновывает принятые решения, владеет разносторонними навыками и приемами выполнения практических работ, обнаруживает умение самостоятельно обобщать и излагать материал, не допуская ошибок, уяснил взаимосвязь основных понятий дисциплины и их значение для приобретения профессии.

«Хорошо» (от 76 до 90) – поступающий твердо знает учебно-программный материал, грамотно и по существу излагает его, не допускает существенных неточностей в ответе на вопрос, может правильно применить теоретические

положения и владеет необходимыми навыками при выполнении практических задач.

«Удовлетворительно» (от 40 до 75) – поступающий усвоил только основной материал, но не знает отдельных деталей, допускает неточности, недостаточно правильные формулировки, нарушает последовательность в изложении программного материала и испытывает затруднения в выполнении практических заданий.

«Неудовлетворительно» (от 0 до 39) – поступающий не знает значительной части программного материала, допускает существенные ошибки, с большим затруднением выполняет практические работы.

6. СПИСОК РЕКОМЕНДУЕМЫХ ИСТОЧНИКОВ ДЛЯ ПОДГОТОВКИ К ВСТУПИТЕЛЬНОМУ ИСПЫТАНИЮ

Источники, отмеченные знаком «*», имеются в научной библиотеке ЧелГУ в печатном или электронном виде в ЭБС «Университетская библиотека онлайн» и «ЛАНЬ», к которым имеется подписка по договорам с правообладателями на текущий учебный год.

Основные:

1. * Коэльо, Л. П. Построение систем машинного обучения на языке Python / Л. П. Коэльо, В. Ричарт ; перевод с английского А. А. Слинкин. — 2-е изд. — Москва : ДМК Пресс, 2016. — 302 с. — ISBN 978-5-97060-330-7. — Текст : электронный // Лань : электронно-библиотечная система. — URL: <https://e.lanbook.com/book/82818> (дата обращения: 29.03.2022). — Режим доступа: для авториз. пользователей.

2. *Шалев-Шварц, Ш. Идеи машинного обучения : учебное пособие / Ш. Шалев-Шварц, Бен-ДавидШ. ; перевод с английского А. А. Слинкина. — Москва : ДМК Пресс, 2019. — 436 с. — ISBN 978-5-97060-673-5. — Текст : электронный // Лань : электронно-библиотечная система. — URL: <https://e.lanbook.com/book/131686> (дата обращения: 29.03.2022). — Режим доступа: для авториз. пользователей.

Дополнительные:

1.* Волкова, В. Н. Системный анализ информационных комплексов : учебное пособие / В. Н. Волкова. — 3-е изд., стер. — Санкт-Петербург : Лань, 2020. — 336 с. — ISBN 978-5-8114-5601-7. — Текст : электронный // Лань : электронно-библиотечная система. — URL: <https://e.lanbook.com/book/143131> (дата обращения: 29.03.2022). — Режим доступа: для авториз. пользователей.

2.*Флах, П. Машинное обучение. Наука и искусство построения алгоритмов, которые извлекают знания из данных / П. Флах. — Москва : ДМК Пресс, 2015. — 400 с. — ISBN 978-5-97060-273-7. — Текст : электронный // Лань : электронно-библиотечная система. — URL: <https://e.lanbook.com/book/69955> (дата обращения: 29.03.2022). — Режим доступа: для авториз. пользователей.

**Рекомендуемые ресурсы информационно-коммуникационной сети
«Интернет»:**

1. <http://www.machinelearning.ru/>
2. <https://academy.yandex.ru/dataschool/online>
3. <http://systems-analysis.ru/>
4. <https://nlpub.ru/>

МИНОБРНАУКИ РОССИИ
Федеральное государственное бюджетное образовательное учреждение
высшего образования «Челябинский государственный университет»
(ФГБОУ ВО «ЧелГУ»)

Уровень образования

Высшее образование - подготовка кадров высшей квалификации

ВСТУПИТЕЛЬНОЕ ИСПЫТАНИЕ ПО СПЕЦДИСЦИПЛИНЕ

**Группа научных специальностей – 2.3. Информационные технологии
и телекоммуникации**

Научная специальность

2.3.1. Системный анализ, управление и обработка информации

БИЛЕТ № 1

1. Задачи системного анализа. Применение системного анализа при построении интеллектуальных систем. Основные сложности и нерешенные научные проблемы.
2. Методы машинного обучения: деревья решений и случайный лес. Проиллюстрировать на примерах работу методов в программном коде (в среде Jupyter Notebook), изложив необходимые понятия теории графов и теории вероятностей. Преимущества и недостатки данных методов в сравнении с другими методами машинного обучения.

Председатель предметной комиссии

ФИО