

МИНОБРНАУКИ РОССИИ

Федеральное государственное автономное образовательное
учреждение высшего образования «Южный федеральный университет»
(ЮЖНЫЙ ФЕДЕРАЛЬНЫЙ УНИВЕРСИТЕТ)

Программа утверждена на заседании
Учёного совета Института
компьютерных технологий и
информационной безопасности

УТВЕРЖДАЮ

Директор Института компьютерных
технологий и информационной
безопасности



Т. Е. Веселов

Программа вступительного испытания по магистерской программе
Машинное обучение и технологии больших данных

Направление подготовки
09.04.03 Прикладная информатика

Уровень высшего образования
магистратура

Форма обучения
очная

Ростов-на-Дону – Таганрог

I. ОБЩИЕ ПОЛОЖЕНИЯ

Вступительное испытание по магистерской программе «Машинное обучение и технологии больших данных» направления подготовки 09.04.03 Прикладная информатика проводится в форме устного экзамена (собеседование).

Вступительное испытание проводится в соответствии с утверждённым расписанием. В начале проведения вступительного испытания поступающему выдаются вопросы в соответствии с данной программой вступительного испытания. На подготовку ответов поступающему отводится до 30 минут. При подготовке ответов поступающий имеет право конспектировать основные положения своих ответов, однако оцениванию подлежат только ответы обучающегося, даваемые им в устной форме непосредственно при проведении собеседования. Продолжительность проведения собеседования – до 15 минут на одного поступающего.

Результаты вступительного испытания оцениваются по 100-балльной шкале. Минимальные баллы, необходимые для участия в конкурсе на поступление, установлены локальными нормативными актами.

II. ВОПРОСЫ ВСТУПИТЕЛЬНОГО ИСПЫТАНИЯ

При проведении вступительного испытания каждому поступающему предлагается ответить на три вопроса:

Вопрос № 1. Поясните причины выбора данной магистерской программы Южного федерального университета и цели, которые ставите перед собой при обучении по данной магистерской программе (при пояснении можно опираться на имеющиеся результаты обучения, научной деятельности, опыт профессиональной деятельности и др.).

Вопросы № 2-3. Выбирается экзаменационной комиссией (или автоматически электронным сервисом проведения вступительного испытания) из следующего перечня вопросов, соответствующего предметной области магистерской программы:

1. Национальная стратегия развития искусственного интеллекта.
2. Принципы классификации систем искусственного интеллекта в соответствии с ГОСТ Р 59277–2020.
3. Основные критерии определения принадлежности проектов к проектам в сфере искусственного интеллекта.

4. Понятия «слабый» и «сильный» искусственный интеллект.
5. Понятия «данные», «информация» и «знания».
6. Понятия «наука о данных», «интеллектуальный анализ данных», «машинное обучение».
7. Понятия «большие данные», «малые данные».
8. Датасеты и типы данных.
9. Задача регрессии.
10. Задача классификации.
11. Задача кластеризации.
12. Задача анализа ассоциаций и последовательностей.
13. Задача выявления аномалий.
14. Обучение с учителем, без учителя, с подкреплением.
15. Отбор признаков.
16. Снижение размерности данных. Метод главных компонент.
17. Деревья решений. Основные понятия.
18. Индукция деревьев решений. Энтропия. Информационный выигрыш.
19. Редукция деревьев решений. Предредукция. Постредукция.
20. Алгоритмы ID3, CART.
21. Алгоритмы ограниченного перебора Бонгарда.
22. Метод опорных векторов. Линейная и нелинейная разделимость.
23. Байесовская классификация. Наивный байесовский классификатор.
24. Кластеризация. Типологический и таксономический подходы.
25. Визуализация кластеров. Дендрограмма. Диаграмма рассеяния.
26. Самоорганизующиеся карты Кохонена.
27. Метод k-средних. Меры расстояния.
28. Алгоритм Sobweb.
29. Статистические методы кластеризации. EM-алгоритм.
30. Графовые методы кластеризации.
31. Нечеткая кластеризация. FCM-алгоритм.
32. Задача анализа рыночных корзин.
33. Поддержка и достоверность ассоциативного правила. Алгоритм Apriori.
34. Матрица несоответствий.
35. Метрики качества: правильность, полнота, точность, F-мера, AUC.
36. Оценка моделей обучения. Метод удержания. Метод перекрестной проверки.
37. Классификация с учетом издержек (Cost-sensitive classification).
38. Графики выигрыша и роста.

39. Проблема переобучения.
40. Стратификация данных.
41. Нейронные сети и глубокое обучение.
42. Ансамбли (комитеты) моделей.
43. Бэггинг (бутстреп-агрегирование). Бэггинг с рандомизацией.
44. Бустинг (усиление) ансамбля классификаторов.
45. Алгоритм «случайный лес».
46. Метаклассификаторы. Стэкинг.
47. Понятия «социальная сеть», «социальный граф», «актор».
48. Теория «малого мира».
49. Эксцентриситет вершины, диаметр графа, радиус графа, шарнир.
50. Метрики центральности в социальной сети.

III. СТРУКТУРА И КРИТЕРИИ ОЦЕНИВАНИЯ

Результаты вступительного испытания оцениваются по 100-балльной шкале. Ответ на первый вопрос вступительного испытания до 40 баллов и ответы на вопросы 2 и 3 (выбранные экзаменационной комиссией или автоматически электронным сервисом) вступительного испытания позволяет набрать до 30 баллов за каждый из этих вопросов.

Структура и критерии оценивания ответа на вопрос №1 вступительного испытания:

- обоснование выбора Южного федерального университета и данной магистерской программы, связи предметной области магистерской программы с настоящей или будущей профессиональной деятельностью – до 20 баллов;
- обоснование целей и ожидаемых результатов обучения в магистратуре, а также результатов научной и/или проектной деятельности, планируемой к выполнению в ходе обучения – до 10 баллов;
- обоснование готовности к эффективному освоению магистерской программы с учётом имеющихся образовательных результатов, достижений в научно-исследовательской и инновационной деятельности, опыта профессиональной деятельности – до 10 баллов.

Структура и критерии оценивания ответа на вопросы №2-3 вступительного испытания:

- понимание предмета вопроса, полнота ответа на поставленный вопрос, доказывающая наличие достаточно обширных знаний о предмете вопроса – до 15 баллов;
- свободное и правильное оперировании терминами и понятиями, связанными с предметом вопроса – до 5 баллов;

– ответы на дополнительные уточняющие вопросы по ответу на основной вопрос – до 5 баллов;

– уровень общей научной культуры и аналитические способности – до 5 баллов.

IV. РЕКОМЕНДУЕМЫЕ ИСТОЧНИКИ ИНФОРМАЦИИ

1. Национальная стратегия развития искусственного интеллекта на период до 2030 года (утв. Указом Президента РФ от 10.10.2019 № 490). – URL: <https://www.garant.ru/products/ipo/prime/doc/72738946/#1000>

2. ГОСТ Р 59277–2020. Национальный стандарт Российской Федерации. Системы искусственного интеллекта. Классификация систем искусственного интеллекта (утв. и введён в действие Приказом Росстандарта от 23.12.2020 № 1372-ст). – URL: <http://protect.gost.ru/document1.aspx?control=7&id=239563>

3. Критерии определения принадлежности проектов к проектам в сфере искусственного интеллекта (утв. Приказом Министерство экономического развития Российской Федерации от 29.06.2021 № 392). – URL: <https://www.consultant.ru/cons/cgi/online.cgi?req=doc&base=LAW&n=391797>

4. Целых, А. Н. Современные методы прикладной информатики в задачах анализа данных : учебное пособие по курсу «Методы интеллектуального анализа данных» / А. Н. Целых, А. А. Целых, Э. М. Котов ; Южный федеральный университет. – Ростов-на-Дону ; Таганрог : Издательство Южного федерального университета, 2021. – 130 с. – URL: <https://hub.sfedu.ru/repository/material/801272973/>

5. Целых, А. Н. Применение временных рядов для анализа больших данных : учебное пособие по курсу «Математические методы анализа больших данных» / А. Н. Целых, В. С. Васильев, Э. М. Котов ; Южный федеральный университет. - Ростов-на-Дону ; Таганрог : Издательство Южного федерального университета, 2021. – 86 с. – URL: <https://hub.sfedu.ru/repository/material/801287277/>

6. Герасименко Е. М. Методы машинного обучения в научных исследованиях: Учебное пособие. – Ростов-на-Дону; Таганрог: Изд-во ЮФУ, 2020. – 97 с. – URL: <https://hub.sfedu.ru/repository/material/801272034/>

7. Чубукова, И. А. Data Mining : учебное пособие / И. А. Чубукова. — 3-е изд. — Москва, Саратов : Интернет-Университет Информационных Технологий (ИНТУИТ), Ай Пи Ар Медиа, 2020. — 469 с. — ISBN 978-5-4497-0289-0. — Текст : электронный // Электронно-библиотечная система IPR BOOKS : [сайт]. — Режим доступа: по подписке. — URL: <https://www.iprbookshop.ru/89404.html>

8. Келлехер, Д. Наука о данных: базовый курс : [16+] / Д. Келлехер, Б. Тирни ; науч. ред. З. Мамедьяров ; пер. с англ. М. Белоголовского. – Москва : Альпина Паблишер, 2020. – 224 с. : схем., табл. – Режим доступа: по подписке. – URL: <https://biblioclub.ru/index.php?page=book&id=598235>

9. Маккинли, Уэс. Python и анализ данных / Уэс Маккинли ; перевод А. Слинкина. — 2-е изд. — Саратов : Профобразование, 2019. — 482 с. — ISBN

978-5-4488-0046-7. — Текст : электронный // Электронно-библиотечная система IPR BOOKS : [сайт]. — Режим доступа: по подписке. — URL: <https://www.iprbookshop.ru/88752.html>

10. Алексеев Д. С. Технологии интеллектуального анализа данных [Электронный ресурс]: учебное пособие / Алексеев Д. С. – Кострома: КГУ им. Н.А. Некрасова, 2020. – 141 с. – Режим доступа: по подписке. – URL: <https://e.lanbook.com/book/160082>

11. Трофимова Е.А. Нейронные сети в прикладной экономике : учебное пособие / Трофимова Е.А., Мазуров Вл.Д., Гилёв Д.В.. — Екатеринбург : Издательство Уральского университета, 2017. — 96 с. — ISBN 978-5-7996-2018-9. — Текст : электронный // Электронно-библиотечная система IPR BOOKS : [сайт]. — Режим доступа: по подписке. — URL: <http://www.iprbookshop.ru/106462.html>

12. Железнов, М. М. Методы и технологии обработки больших данных : учебно-методическое пособие / М. М. Железнов. – Москва : МИСИ-МГСУ, ЭБС АСВ, 2020. – 46 с. – ISBN 978-5-7264-2193-3. – Текст : электронный // Электронно-библиотечная система IPR BOOKS : [сайт]. — Режим доступа: по подписке. – URL: <https://www.iprbookshop.ru/101802.html>

13. Воронов, В. И. Data Mining – технологии обработки больших данных : учебное пособие / В. И. Воронов, Л. И. Воронова, В. А. Усачев. – Москва : Московский технический университет связи и информатики, 2018. – 47 с. – ISBN 2227-8397. – Текст : электронный // Электронно-библиотечная система IPR BOOKS : [сайт]. – Режим доступа: по подписке. – URL: <https://www.iprbookshop.ru/81324.html>

14. Билл, Фрэнкс. Укрощение больших данных : как извлекать знания из массивов информации с помощью глубокой аналитики / Фрэнкс Билл ; перевод А. Баранов. – Москва : Манн, Иванов и Фербер, 2014. – 340 с. – ISBN 978-5-00057-146-0. – Текст : электронный // Электронно-библиотечная система IPR BOOKS : [сайт]. – Режим доступа: по подписке. – URL: <https://www.iprbookshop.ru/39433.html>

15. Интеллектуальный предиктивный мультимодальный анализ слабоструктурированных больших данных / Н. Г. Ярушкина, И. А. Андреев, Г. Ю. Гуськов [и др.]. – Ульяновск : Ульяновский государственный технический университет, 2020. – 221 с. – ISBN 978-5-9795-2088-9. – Текст : электронный // Электронно-библиотечная система IPR BOOKS : [сайт]. – Режим доступа: по подписке. – URL: <https://www.iprbookshop.ru/106136.html>

Разработчик программы вступительного испытания:

А. А. Целых, кандидат технических наук, доцент, доцент Института компьютерных технологий и информационной безопасности Южного федерального университета, руководитель магистерской программы «Машинное обучение и технологии больших данных» направления подготовки 09.04.03 Прикладная информатика