МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ
РОССИЙСКОЙ ФЕДЕРАЦИИ

ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ БЮДЖЕТНОЕ ОБРАЗОВАТЕЛЬНОЕ
УЧРЕЖДЕНИЕ ВЫСШЕГО ОБРАЗОВАНИЯ
«РЯЗАНСКИЙ ГОСУДАРСТВЕННЫЙ РАДИОТЕХНИЧЕСКИЙ УНИВЕРСИТЕТ ИМЕНИ В.Ф.
УТКИНА»

Кафедра «Вычислительная и прикладная математика»

МЕТОДИЧЕСКИЕ МАТЕРИАЛЫ ПО ДИСЦИПЛИНЕ
«Методы машинного обучения»

Направление подготовки – 09.04.04 Программная инженерия

ОПОП академической магистратуры

«Программно-алгоритмическое обеспечение систем искусственного интеллекта»

Квалификация (степень) выпускника – магистр

Форма обучения – очная (2 года)

Рязань

1. СПИСОК ТЕОРЕТИЧЕСКИХ ВОПРОСОВ К ЗАЧЕТУ /ЭКЗАМЕНУ

1. Определение модели и признаков.
2. Постановка задачи классификации.
3. Постановка задачи регрессии.
4. Постановка задачи кластеризации.
5. Основные методы классификации и их характеристики.
6. Основные методы регрессии и их характеристики.
7. Основные методы кластеризации.
8. Метрики качества кластеризации: RAND, индекс Жакара, F-мера.
9. Метод классификации «К ближайших соседей»: принцип реализации, достоинства и недостатки.
10. Классификация с помощью решающего дерева. Принцип построения оптимального дерева решений.
11. Метод классификации Random Forest.
12. Линейная регрессия. Метод наименьших квадратов.
13. Логистическая регрессия. Ее применение в задачах классификации.
14. Метод кластеризации «К-средних»: принцип реализации, достоинства и недостатки.
15. Метод кластеризации DBSCAN: принцип реализации, достоинства и недостатки.
16. Метод понижения размерности T-SNE, его применение.
17. Метод понижения размерности UMAP, его применение, отличия от T-SNE.
18. Байесов классификатор: принцип реализации, достоинства и недостатки.
19. Определение сильного и слабого классификатора. Определение ансамбля классификаторов.
20. Градиентный бустинг.
21. Адаптивный бустинг. Алгоритм AdaBoost.
22. Постановка задачи обучения с подкреплением. Определение среды, ее модели, агента, вознаграждения и функции ценности.
23. Определение Марковского процесса и Марковской цепи.
24. Обучение с подкреплением в постановке задачи о многоруком бандите.
25. Безмодельные методы обучения с подкреплением. Генетические алгоритмы.

2. ПЕРЕЧЕНЬ ЛАБОРАТОРНЫХ/ПРАКТИЧЕСКИХ РАБОТ ПО ДИСЦИПЛИНЕ

Лабораторная работа 1.1. Классификация. (3 час.)

Цель работы: изучение и реализация методов классификации.

Задачи работы: разработать программу для ЭВМ, реализующую метод классификации согласно варианту. Обучающая выборка предоставляется вместе с вариантом и читается программой из текстового файла. Предусмотреть сохранение параметров обученной модели в файл для повторного использования без обучения.

Вопросы по работе:

1. Перечислить основные методы классификации.
2. Продемонстрировать работоспособность разработанного классификатора и посчитанные метрики качества.

Лабораторная работа 1.2. Регрессия. (3 час.)

Цель работы: изучение и реализация методов регрессии.

Задачи работы: разработать программу для ЭВМ, реализующую метод наименьших квадратов. Обучающая выборка предоставляется вместе с вариантом и читается программой из текстового файла. Предусмотреть сохранение параметров обученной модели в файл для повторного использования без обучения.

Вопросы по работе:

1. Перечислить основные методы регрессии.
2. Продемонстрировать работоспособность разработанной программы и посчитанные метрики качества.

Лабораторная работа 1.3. Деревья решений. (3 часа)

Цель работы: изучение и реализация классификаторов на основе решающих деревьев.

Задачи работы: разработать программу для ЭВМ, реализующую метод классификации с помощью решающего дерева. Обучающая выборка предоставляется вместе с вариантом и читается программой из текстового файла. Предусмотреть сохранение параметров обученной модели в файл для повторного использования без обучения.

Вопросы по работе:

1. Продемонстрировать работоспособность разработанной программы и решение всех поставленных задач.
2. Сравнить качество классификации с результатами из первой лабораторной работы.

Лабораторная работа 1.4. Random Forest. (3 часа)

Цель работы: изучение, реализация и исследование алгоритма классификации Random Forest.

Задачи работы: разработать программу для ЭВМ, реализующую метод классификации Random Forest. Обучающая выборка предоставляется вместе с вариантом и читается программой из текстового файла. Предусмотреть сохранение параметров обученной модели в файл для повторного использования без обучения.

Вопросы по работе:

1. Продемонстрировать работоспособность разработанной программы и решение всех поставленных задач.
2. Сравнить качество классификации с результатами первой и третьей лабораторных работ.

Лабораторная работа 2.1. Статистический анализ и предобработка данных. (3 часа)

Цель работы: знакомство с основными приемами анализа исходной выборки и реализация алгоритмов предварительной обработки данных.

Задачи работы: разработать программу для ЭВМ, которая обрабатывает данные согласно варианту. Входные данные читаются из текстового файла, преобразованные данные записываются в другой текстовый файл.

Вопросы работе:

1. Продемонстрировать работоспособность разработанной программы и решение всех поставленных задач.

Лабораторная работа 2.2. Понижение размерности и визуализация. (3 часа)

Цель работы: изучение и применение алгоритмов понижения размерности пространства признаков в целях визуализации входных данных и выявления в них закономерностей.

Задачи работы: разработать программу для ЭВМ, реализующую метод понижения размерности согласно варианту. Предусмотреть графическую визуализацию выборки после снижения размерности.

Вопросы по работе:

1. Продемонстрировать работоспособность разработанной программы и решение всех поставленных задач.

Лабораторная работа 2.3. Кластеризация. (3 часа)

Цель работы: изучение, реализация и исследование алгоритмов кластеризации.

Задачи работы: разработать программу для ЭВМ, реализующую метод кластеризации согласно варианту. Обучающая выборка предоставляется вместе с вариантом и читается программой из текстового файла. Предусмотреть визуализацию результатов кластеризации. Рассчитать метрики качества кластеризации согласно варианту.

Вопросы по работе:

1. Продемонстрировать работоспособность разработанной программы и решение всех поставленных задач.

Лабораторная работа 3.1. Байесовский классификатор. (3 часа)

Цель работы: изучение, реализация и исследование байесовских классификаторов.

Задачи работы: разработать программу для ЭВМ, реализующую Байесовский классификатор. Обучающая выборка предоставляется вместе с вариантом и читается программой из текстового файла. Предусмотреть сохранение параметров обученной модели в файл для повторного использования без обучения.

Вопросы по работе:

1. Продемонстрировать работоспособность разработанной программы и решение всех поставленных задач.

Лабораторная работа 3.2. Метод опорных векторов. (3 часа)

Цель работы: изучение, реализация и исследование метода опорных векторов.

Задачи работы: разработать программу для ЭВМ, реализующую метод опорных векторов. Обучающая выборка предоставляется вместе с вариантом и читается программой из текстового файла. Предусмотреть сохранение параметров обученной модели в файл для повторного использования без обучения.

Вопросы по работе:

1. Продемонстрировать работоспособность разработанной программы и решение всех поставленных задач.

Лабораторная работа 3.3. Бустинг. (3 часа)

Цель работы: изучение методов построения и обучения ансамблей слабых классификаторов.

Задачи работы: разработать программу для ЭВМ, реализующую метод классификации согласно варианту. Обучающая выборка предоставляется вместе с вариантом и читается программой из текстового файла. Предусмотреть сохранение параметров обученной модели в файл для повторного использования без обучения.

Вопросы по работе:

1. Продемонстрировать работоспособность разработанной программы и решение всех поставленных задач.

Лабораторная работа 4.1. Марковские процессы принятия решений. (3 часа)

Цель работы: изучение и исследование формализма Марковских процессов в решении задач обучения с подкреплением.

Задачи работы: разработать программу для ЭВМ, моделирующую Марковский процесс согласно варианту. Реализовать обучающегося агента, который взаимодействует с моделью.

Вопросы по работе:

1. Продемонстрировать работоспособность разработанной программы и решение всех поставленных задач.

Лабораторная работа 4.2. Многорукие бандиты. (3 часа)

Цель работы: изучение и исследование формализма многоруких бандитов в решении задач обучения с подкреплением.

Задачи работы: разработать программу для ЭВМ, реализующую алгоритм UCB1 в задаче о многоруком бандите.

Вопросы по работе:

1. Продемонстрировать работоспособность разработанной программы и решение всех поставленных задач.

Лабораторная работа 4.3. Генетические алгоритмы. (3 часа)

Цель работы: изучение, реализация и исследование генетических алгоритмов в решении задач обучения с подкреплением.

Задачи работы: разработать программу для ЭВМ, реализующую генетический алгоритм согласно варианту.

Вопросы по работе:

1. Продемонстрировать работоспособность разработанной программы и решение всех поставленных задач.

3. МЕТОДИЧЕСКИЕ УКАЗАНИЯ ДЛЯ САМОСТОЯТЕЛЬНОЙ РАБОТЫ ОБУЧАЮЩИХСЯ

Изучение дисциплины проходит в течение одного семестра. Основные темы дисциплины осваиваются в ходе аудиторных занятий, однако важная роль отводится и самостоятельной работе студентов.

Самостоятельная работа как вид учебной работы может использоваться на лабораторных работах, а также иметь самостоятельное значение – внеаудиторная самостоятельная работа обучающихся – при подготовке к лабораторным работам, при подготовке к дифференцированному зачету.

Самостоятельная работа включает в себя следующие этапы:

- изучение теоретического материала (работа над конспектом лекции);
- самостоятельное изучение дополнительных информационных ресурсов (доработка конспекта лекции);
- выполнение заданий текущего контроля успеваемости (подготовка к лабораторным работам);
- итоговая аттестация по дисциплине (подготовка к дифференцированному зачету/ экзамену).

4. РЕКОМЕНДАЦИИ ПО ПЛАНИРОВАНИЮ И ОРГАНИЗАЦИИ ВРЕМЕНИ, НЕОБХОДИМОГО ДЛЯ ИЗУЧЕНИЯ ДИСЦИПЛИНЫ. ОПИСАНИЕ

ПОСЛЕДОВАТЕЛЬНОСТИ ДЕЙСТВИЙ СТУДЕНТА («СЦЕНАРИЙ ИЗУЧЕНИЯ ДИСЦИПЛИНЫ»)

Рекомендуется следующим образом организовать время, необходимое для изучения дисциплины.

Для освоения лекционного материала следует: изучить конспект лекции в тот же день, после лекции: 10 – 15 минут, повторно прочитать конспект лекции за день перед следующей лекцией: 10 – 15 минут. Также следует изучить теоретический лекционный материал по рекомендуемому учебнику/учебному пособию: 1 час в неделю.

Следует максимально использовать лекционное время для изучения дисциплины, понимания лекционного материала и написания конспекта лекций. В процессе лекционного занятия студент должен уметь выделять важные моменты и основные положения. При написании *конспекта лекций* следует придерживаться следующих правил и рекомендаций.

1. При ведении конспекта рекомендуется структурировать материал по разделам, главам, темам. Вести нумерацию формул. Выделять по каждой теме постановку задачи, основные положения, выводы. Кратко записывать те пояснения лектора, которые показались особенно важными. Это позволит при подготовке к сдаче зачёта не запутаться в структуре лекционного материала.

2. Лекционный материал следует записывать в конспект лишь после того, как излагаемый лектором тезис будет вами дослушан до конца и понят.

3. При конспектировании следует отмечать непонятные, на данном этапе, положения, доказательства и пр.

4. Рекомендуется по каждой теме выразить свое мнение, комментарий, вывод.

Доработка конспекта лекции с применением учебника, методической литературы, дополнительной литературы, интернет-ресурсов: этот вид самостоятельной работы студентов особенно важен в том случае, когда одну и ту же задачу можно решать различными способами, а на лекции изложен только один из них. Кроме того, рабочая программа предполагает рассмотрение некоторых относительно несложных тем только во время самостоятельных занятий, без чтения лектором.

Подготовка к лабораторным работам состоит в теоретической подготовке (изучение конспекта лекций, методических указаний к данной лабораторной работе и дополнительной литературы) и выполнении индивидуального задания. Выполнение каждой из запланированных работ заканчивается предоставлением отчета. Требования к форме и содержанию отчета приведены в методических указаниях к лабораторным работам или определяются преподавателем на первом занятии. Допускаясь к лабораторной работе, каждый студент должен представить преподавателю «заготовку» отчета, содержащую: оформленный титульный лист, цель работы, задание, проект решения, полученные результаты, выводы.

Важным этапом является защита лабораторной работы. В процессе защиты студент отвечает на вопросы преподавателя, касающиеся теоретического материала, относящегося к данной работе, и проекта, реализующего его задание, комментирует полученные в ходе работы результаты. При подготовке к защите лабораторной работы рекомендуется ознакомиться со списком вопросов по изучаемой теме и попытаться самостоятельно на них ответить, используя конспект лекций и рекомендуемую литературу.

Подготовка к сдаче дифференцированного зачета.

Зачет – форма промежуточной проверки знаний, умений, навыков, степени освоения дисциплины. Главная задача зачета состоит в том, чтобы у студента по окончанию изучения данной дисциплины сформировались определенное представление об общем содержании дисциплины, определенные теоретические знания и

практические навыки, определенный кругозор. Готовясь к зачету, студент приводит в систему знания, полученные на лекциях, на практических и лабораторных занятиях, разбирается в том, что осталось непонятным, и тогда изучаемая им дисциплина может быть воспринята в полном объеме с присущей ей строгостью и логичностью, ее практической направленностью.

Зачеты дают возможность преподавателю определить теоретические знания студента и его практические навыки при решении определенных прикладных задач. Оцениваются: понимание и степень усвоения теоретического материала; степень знакомства с основной и дополнительно литературой, а также с современными публикациями; умение применить теорию к практике, решать определенные практические задачи данной предметной области, правильно проводить расчеты и т. д.; знакомство с историей данной науки; логика, структура и стиль ответа, умение защищать выдвигаемые положения.

Значение зачета не ограничивается проверкой знаний, являясь естественным завершением обучения студента по данной дисциплине, они способствуют обобщению и закреплению знаний и умений, приведению их в стройную систему, а также устранению возникших в процессе обучения пробелов.

Подготовка к зачету – это тщательное изучение и систематизация учебного материала, осмысление и запоминание теоретических положений, формулировок, формул, установление и осмысление внутрипредметных связей между различными темами дисциплины, закрепление теоретических знаний путем решения определенных задач.

Планируйте подготовку к зачету, учитывая сразу несколько факторов: неоднородность в сложности учебного материала и степени его проработки в ходе обучения, свои индивидуальные способности. Рекомендуется делать перерывы в занятиях через каждые 50-60 минут на 10 минут. После 3-4 часов занятий следует сделать часовой перерыв. Чрезмерное утомление приведет к снижению тонуса интеллектуальной деятельности. Целесообразно разделять весь рабочий день на три рабочих периода – с утра до обеда, с обеда до ужина и с ужина до сна. Каждый рабочий период дня должен заканчиваться отдыхом не менее 1 часа. Работая в сессионном режиме, студент имеет возможность увеличить время занятий с 10 (как требовалось в семестре) до 12 часов в сутки.

Подготовку к зачету следует начинать с общего планирования своей деятельности. С определения объема материала, подлежащего проработке, необходимо внимательно сверить конспекты с программой дисциплины, чтобы убедиться, все ли разделы отражены в лекциях, отсутствующие темы изучить по учебнику. Второй этап предусматривает системное изучение материала по данному предмету с обязательной записью всех выкладок, выводов, формул. На третьем этапе – этапе закрепления – полезно чередовать углубленное повторение особенно сложных вопросов с беглым повторением всего материала.

5. РЕКОМЕНДАЦИИ ПО РАБОТЕ С ЛИТЕРАТУРОЙ

Теоретический материал курса становится более понятным, когда дополнительно к прослушиванию лекции и изучению конспекта изучаются и книги по данному предмету. Литературу по дисциплине рекомендуется читать как в бумажном, так и в электронном виде (если отсутствует бумажный аналог). Полезно использовать несколько учебников и пособий по дисциплине. Рекомендуется после изучения очередного параграфа ответить на несколько вопросов по данной теме. Кроме того, полезно мысленно задать себе следующие вопросы (и попробовать ответить на них): «о чем этот параграф?», «какие новые понятия введены, каков их смысл?», «зачем мне это нужно по специальности?».

Рекомендуется самостоятельно изучать материал, который еще не прочитан на лекции и не применялся на лабораторном или практическом занятии, тогда занятия будут гораздо понятнее. В течение недели рекомендуется выбрать время (1 час) для работы с литературой.

Оператор ЭДО ООО "Компания "Тензор"

ДОКУМЕНТ ПОДПИСАН ЭЛЕКТРОННОЙ ПОДПИСЬЮ

ПОДПИСАНО ЗАВЕДУЮЩИМ КАФЕДРЫ	**ФГБОУ ВО "РГРТУ", РГРТУ,** Овечкин Геннадий Владимирович, Заведующий кафедрой ВПМ	**12.10.26** 09:03 (MSK)	Простая подпись
ПОДПИСАНО ЗАВЕДУЮЩИМ ВЫПУСКАЮЩЕЙ КАФЕДРЫ	**ФГБОУ ВО "РГРТУ", РГРТУ,** Овечкин Геннадий Владимирович, Заведующий кафедрой ВПМ	**12.10.26** 09:03 (MSK)	Простая подпись