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

«Введение в машинное обучение»

Направление подготовки – 09.03.04 Программная инженерия

ОПОП «Программная инженерия систем искусственного интеллекта»

Квалификация (степень) выпускника – бакалавр

Форма обучения – очная (4 года)

Рязань

1. СПИСОК ТЕОРЕТИЧЕСКИХ ВОПРОСОВ К ЭКЗАМЕНУ

1. Типовые задачи машинного обучения. Поиск ассоциаций, классификация, регрессия, кластеризация, оптимизация.
2. Инструменты решения задач машинного обучения на языке Python (NumPy, Pandas, Scikit-learn, Matplotlib).
3. Постановка задачи поиска ассоциаций. Ассоциативные правила. Алгоритмы APRIORI и FPG.
4. Иерархические ассоциативные правила. Последовательные шаблоны. Применение на практике.
5. Алгоритмы линейного снижения размерности: PCA, LDA, FA.
6. Применение алгоритмов линейного снижения размерности в практических задачах.
7. Алгоритмы нелинейного снижения размерности: t-SNE, UMAP, TRIMAP, PACMAP.
8. Применение алгоритмов нелинейного снижения размерности в практических задачах.
9. Постановка задач классификации и регрессии. Переобучение и недообучение.
10. Показатели качества классификации и регрессии. k-fold валидация.
11. kNN-алгоритм. Применение в задачах классификации.
12. Логистическая регрессия (LR). Применение в задачах классификации.
13. SVM-алгоритм. Линейный SVM. Функции ядра. Трюк с ядром.
14. Мультиклассовая классификация на основе SVM. Стратегии мультиклассовой классификации.
15. Дерево решений (DT). Применение в задачах классификации и регрессии.
16. Random Forest (RF). Применение в задачах классификации и регрессии.
17. Ансамбли алгоритмов машинного обучения: бэггинг, бустинг и стекинг.
18. Современные подходы к разработке ансамблей машинного обучения.
19. Проблема балансировки классов. Алгоритмы сэмплинга: оверсэмплинг, андерсэмплинг, гибридизация.
20. Чувствительные к затратам (cost-sensitive) алгоритмы. Применение на практике.
21. История возникновения популяционных алгоритмов оптимизации. Классификация.
22. Генетический алгоритм. Применение для оптимизации.
23. Алгоритм роя частиц. Алгоритм дифференциальной эволюции.
24. Однокритериальная и многокритериальная оптимизация. Эвристики, метаэвристики, гиперэвристики.
25. Постановка задачи кластерного анализа. Иерархические и итеративные алгоритмы кластеризации.
26. K-means, K-medoids, FCM, DBSCAN, HDBSCAN, OPTICS, EM. Gaussian mixture.
27. Внешние и внутренние меры оценки качества кластеризации (Rand, Adjusted Rand, Jaccard, индекс локтя, силуэт, Данна, Davies-Bouldin).
28. Масштабируемые алгоритмы кластерного анализа: CLOPE, BIRCH, Mini-Batch K-Means, CLARANS, Affinity Propagation.
29. Понятие искусственной нейронной сети. Нейрон. Слои. Функции активации и потерь.

30. Рекуррентные нейронные сети (RNN). Проблемы исчезающего и взрывного градиента. LSTM, GRU. Dropout.

2. ПЕРЕЧЕНЬ ПРАКТИЧЕСКИХ И ЛАБОРАТОРНЫХ РАБОТ ПО ДИСЦИПЛИНЕ

№	Тема лабораторной / практической работы	Вид занятия	Кол-во часов
1	Типовые задачи машинного обучения. Инструменты Python	Лаб. / Пр.	1 / 2
2	Ассоциативные правила. Алгоритмы APRIORI и FPG	Лаб. / Пр.	1 / 2
3	Линейное снижение размерности: PCA, LDA, FA	Лаб. / Пр.	1 / 2
4	Нелинейное снижение размерности: t-SNE, UMAP, TRIMAP, PACMAP	Лаб. / Пр.	1 / 2
5	Алгоритмы knn, LR. Показатели качества. k-fold валидация	Лаб. / Пр.	1 / 2
6	SVM-алгоритм. Функции ядра. Мультиклассовая классификация	Лаб. / Пр.	1 / 2
7	DT- и RF-алгоритмы. Применение в классификации и регрессии	Лаб. / Пр.	1 / 2
8	Ансамбли: бэггинг, бустинг, стекинг	Лаб. / Пр.	1 / 2
9	Алгоритмы балансировки классов: сэмплинг, cost-sensitive	Лаб. / Пр.	1 / 2

№	Тема лабораторной / практической работы	Вид занятия	Кол-во часов
10	Чувствительные к затратам алгоритмы. Гибридные методы	Лаб. / Пр.	1 / 2
11	Генетический алгоритм оптимизации	Лаб. / Пр.	1 / 2
12	Алгоритм роя частиц. Дифференциальная эволюция	Лаб. / Пр.	1 / 2
13	Кластерный анализ: k-means, FCM, DBSCAN, иерархическая кластеризация	Лаб. / Пр.	1 / 2
14	Оценка качества кластеризации: внешние и внутренние меры	Лаб. / Пр.	1 / 2
15	Масштабируемые алгоритмы кластеризации: CLOPE, BIRCH, Mini-Batch K-Means	Лаб. / Пр.	1 / 2
16	Рекуррентные нейронные сети: RNN, LSTM, GRU	Лаб. / Пр.	1 / 2
Итого			**16 / 32**

3. МЕТОДИЧЕСКИЕ УКАЗАНИЯ ДЛЯ САМОСТОЯТЕЛЬНОЙ РАБОТЫ ОБУЧАЮЩИХСЯ

Самостоятельная работа как вид учебной работы может использоваться на практических и лабораторных занятиях, а также иметь самостоятельное значение – внеаудиторная самостоятельная работа обучающихся – при подготовке к практическим и лабораторным работам, при подготовке к экзамену.

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 практическим и лабораторным работам);
- итоговая аттестация по дисциплине (подготовка к экзамену).

При подготовке к лабораторным работам необходимо:

1. Изучить конспект лекций и методические указания к данной работе.
2. Ознакомиться с рекомендуемой основной и дополнительной литературой.
3. Выполнить индивидуальное задание.
4. Оформить отчёт в соответствии с требованиями кафедры.
5. Защитить работу, ответив на вопросы преподавателя.

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

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

Подготовка к сдаче экзамена.

Экзамен – форма промежуточной проверки знаний, умений, навыков, степени освоения дисциплины. Главная задача экзамена состоит в том, чтобы у студента по окончании изучения данной дисциплины сформировались определённое представление об общем содержании дисциплины, определённые теоретические знания и практические навыки, определённый кругозор. Готовясь к экзамену, студент приводит в систему знания, полученные на лекциях, на практических занятиях, разбирается в том, что осталось непонятным, и тогда изучаемая им дисциплина может быть воспринята в полном объёме с присущей ей строгостью и логичностью, её практической направленностью.

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

Планируйте подготовку к экзамену, 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 приведёт 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 экзамену следует начинать с общего планирования своей деятельности. С определения объёма 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 углублённое повторение особенно сложных вопросов с беглым повторением всего материала.

5. РЕКОМЕНДАЦИИ ПО РАБОТЕ С ЛИТЕРАТУРОЙ

Основная литература, литература для дополнительного чтения и методические указания к практическим (лабораторным) работам указаны в рабочей программе данной дисциплины и доступны посредством ЭБС.

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 чём этот параграф?», «какие новые понятия введены, каков их смысл?», «зачем мне это нужно по специальности?».

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

6. БИБЛИОГРАФИЧЕСКИЙ СПИСОК

6.1. Основная литература

1. Цуканова Н. И. Программирование глубоких нейронных сетей на языке Python: учеб. пособие. – Рязань: КУРС, 2023. – URL: https://elib.rsreu.ru/ebs/download/3639

2. Демидова Л. А. Разведочный анализ данных. Python. Часть 2. – Москва: РТУ МИРЭА, 2023. – 92 с. – ISBN 978-5-7339-1933-1. – URL: https://e.lanbook.com/book/382691
3. Демидова Л. А. Разведочный анализ данных. Python. Часть 1. – Москва: РТУ МИРЭА, 2022. – 107 с. – URL: https://e.lanbook.com/book/310970
4. Демидова Л. А. Кластерный анализ. Python: учебное пособие. – Москва: РТУ МИРЭА, 2022. – 103 с. – URL: https://e.lanbook.com/book/240092
5. Митяков Е. С., Шмелева А. Г., Ладынин А. И. Искусственный интеллект и машинное обучение: учебное пособие для вузов. – Санкт-Петербург: Лань, 2025. – 252 с. – ISBN 978-5-507-51465-6. – URL: https://e.lanbook.com/book/450827

6.2. Дополнительная литература

1. Баланов А. Н. Машинное обучение и искусственный интеллект: учебное пособие для спо. – Санкт-Петербург: Лань, 2024. – 80 с. – ISBN 978-5-507-49195-7. – URL: https://e.lanbook.com/book/414926
2. Орешков В. И. Интеллектуальный анализ данных: учебное пособие. – Рязань: Рязанский государственный радиотехнический университет, 2017. – 161 с. – ISSN 2227-8397. – URL: https://www.iprbookshop.ru/121842.html
3. Котельников Е. В., Котельникова А. В. Введение в машинное обучение и анализ данных: учеб. пособие. – Киров: ВятГУ, 2023. – 68 с. – URL: https://e.lanbook.com/book/390698

6.3. Методические разработки

1. Цуканова Н. И., Дмитриева Т. А. Системы искусственного интеллекта. Нейронные сети и генетические алгоритмы: метод. указ. к лаб. работам 1, 2: Методические указания. – Рязань: , 2020. – URL: https://elib.rsreu.ru/ebs/download/2842
2. Демидова Л. А. Интеллектуальный анализ данных на языке Python: учебно-методическое пособие. – Москва: РТУ МИРЭА, 2021. – 92 с. – URL: https://e.lanbook.com/book/218693

6.4. Перечень ресурсов информационно-телекоммуникационной сети «Интернет»

1. Электронная библиотека РГРТУ: https://elib.rsreu.ru/ebs/

6.5. Перечень программного обеспечения и информационных справочных систем

Лицензионное и свободно распространяемое программное обеспечение:

Наименование	Описание
PyCharm Community	Свободное ПО
Python	Свободное ПО
Anaconda3	Свободно распространяемое программное обеспечение под лицензиями

Информационные справочные системы:

1. Информационно-правовой портал ГАРАНТ.РУ: http://www.garant.ru
2. Система КонсультантПлюс: http://www.consultant.ru
3. Справочная правовая система «КонсультантПлюс» (договор об информационной поддержке №1342/455-100 от 28.10.2011 г.)