

МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ
РОССИЙСКОЙ ФЕДЕРАЦИИ
ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ БЮДЖЕТНОЕ ОБРАЗОВА-
ТЕЛЬНОЕ УЧРЕЖДЕНИЕ ВЫСШЕГО ОБРАЗОВАНИЯ
«ЛУГАНСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ
ИМЕНИ ВЛАДИМИРА ДАЛЯ»

Факультет компьютерных систем и информационных технологий
Кафедра информатики и программной инженерии

УТВЕРЖДАЮ



Декан факультета компьютерных си-
стем и информационных технологий

Кочевский А. А.

» 04 2023 г.

ФОНД ОЦЕНОЧНЫХ СРЕДСТВ
по учебной дисциплине

«Методы и технологии интеллектуального анализа данных»

09.04.03 Прикладная информатика
«Прикладная информатика в экономике»

Разработчик
к.э.н, доцент

Степанова Е.М.

ФОС рассмотрен и одобрен на заседании кафедры информатики и программ-
ной инженерии
от «18» апреля 2023 г., протокол № 17

Заведующий кафедрой информатики
и программной инженерии

Кочевский А.А.

(подпись)

Луганск 2023 г.

**Паспорт
фонда оценочных средств по учебной дисциплине
«Методы и технологии интеллектуального анализа данных»**

**Перечень компетенций (элементов компетенций),
формируемых в результате освоения учебной дисциплины**

№ п/п	Код контроли- руемой компе- тенции	Формулировка контроли- руемой компетенции	Контролируемые темы учеб- ной дисциплины, практики	Этапы форми- рования (се- местр изу- чения)
1	ПК-5	Способен использовать и развивать методы науч- ных исследований и ин- струментария в области проектирования и управ- ления информационными системами в прикладных областях.	Тема 1. Возникновение, про- блемы и перспективы техно- логий интеллектуального ана- лиза данных (ИАД). Тема 2. Методы и стадии ИАД. Тема 4. Информация и знания. Тема 3. Задачи ИАД. Тема 5. Инструменты ИАД. Тема 6. Методы классифика- ции и прогнозирования. Тема 7. Методы кластерного анализа. Тема 8. Методы поиска ассо- циативных правил. Тема 9. Способы визуального представления данных. Тема 10 Сферы применения ИАД. Тема 11. Использование мето- дов ИАД в перспективных направлениях развития ин- формационных технологий. Тема 12. Применение ИАД в самостоятельном исследова- нии и проектировании соб- ственных программных про- дуктов.	1

Показатели и критерии оценивания компетенций, описание шкал оценивания

№ п/п	Код контролируемой компетенции	Показатель оценивания (знания, умения, навыки)	Контролируемые темы учебной дисциплины	Наименование оценочного средства
1.	ПК-5	знать современные методы научных исследований в области проектирования информационных систем в экономике; уметь анализировать и выбирать инструментарий проектирования и управления информационными системами в экономике; владеть навыками использования в практике проектирования информационных систем современного программного и методического инструментария.	Тема 1. Тема 2. Тема 3. Тема 4. Тема 5. Тема 6. Тема 7. Тема 8. Тема 9. Тема 10. Тема 11. Тема 12.	Фронтальные и индивидуальные опросы; Курсовая работа; промежуточная аттестация (зачет)

Фонды оценочных средств по дисциплине «Методы и технологии интеллектуального анализа данных»

Вопросы для фронтальных и индивидуальных опросов:

Тема 1. Возникновение, проблемы и перспективы технологий интеллектуального анализа данных (ИАД).

1. Назначение, задачи и общая характеристика курса, общие понятия и определения.
2. Основные классы интеллектуальных информационных систем.
3. Основные цели интеллектуализации информационных систем для бизнеса.
4. Концепция ИИС.
5. Классификация проблем, возникающих в экономике
5. Идентификация проблем и ситуаций.
7. Генерация, хранение и необходимость оперативной обработки огромных массивов данных современными организациями.

Тема 2. Методы и стадии ИАД.

1. Основные методы ИАД.
2. Добыча данных - data mining: классификация задач.
3. Классификация, регрессия, поиск ассоциативных правил, кластеризация.
4. Области практического применения data mining.

5. Методы data mining: базовые, нечеткая логика, генетические алгоритмы, нейронные сети.
6. Процесс обнаружения знаний.
7. Стадии ИАД.
8. Выявление закономерностей (свободный поиск).
9. Использование выявленных закономерностей для предсказания неизвестных значений (прогностическое моделирование).
10. Анализ исключений, предназначенный для выявления и толкования аномалий в найденных закономерностях.
11. Промежуточная стадия проверки достоверности найденных закономерностей между их нахождением и использованием (стадия валидации).

Тема 3. Задачи ИАД.

1. Обзор основных задач и методов ИАД.
2. Классификация и регрессия.
3. Постановка задач и представление результатов.
4. Правила классификации, деревья решений, математические функции и методы их представления.
5. Карта Кохонена.
6. Поиск ассоциативных правил.
7. Постановка и формализация задачи.
8. Сиквенциальный анализ.
9. Разновидности задачи и алгоритмы их решения.
10. Формы визуализации и представления результатов.

Тема 4. Информация и знания.

1. Проблемы представления и моделирования знаний.
2. Продукционные модели.
3. Логические модели.
4. Исчисление высказываний и исчисление предикатов.
5. Семантические сети.
6. Интенциональные и экстенциональные модели семантических сетей.
7. Фреймы: структура, динамика.
8. Представление и формализация нечетких знаний.
9. Обработка знаний и вывод решений в интеллектуальных системах.

Тема 5. Инструменты ИАД.

1. Аналитический и информационный подходы к моделированию.
2. Формы представления, типы и виды анализируемых данных.
3. Источники данных для анализа.
4. Технология KDD (Knowledge Discovery in Databases).
5. Введение в технологию Data Mining.
6. Программный инструментарий анализа данных.

Тема 6. Методы классификации и прогнозирования.

1. Статистические методы.
2. Введение в классификацию и регрессию.
3. Простая линейная регрессия.
4. Оценка соответствия простой линейной регрессии реальным данным.
5. Простая регрессионная модель.
6. Множественная линейная регрессия.
7. Модель множественной линейной регрессии.
8. Регрессия с категориальными входными переменными.
9. Методы отбора переменных в регрессионные модели.
10. Ограничения применимости регрессионных моделей.
11. Основы логистической регрессии.
12. Интерпретация модели логистической регрессии. Множественная логистическая регрессия.
12. Анализ и прогнозирование временных рядов.
12. Временной ряд и его компоненты.
13. Модели прогнозирования.
14. Прогнозирование в торговле и логистике.

Тема 7. Методы кластерного анализа.

1. Суть метода кластерного анализа.
2. Меры близости, основанные на расстояниях.
3. Формы представления результатов. Базовые алгоритмы кластеризации.
4. Иерархические алгоритмы: агломеративные, дивизимные.
5. Неиерархические и нечеткие алгоритмы.

Тема 8. Методы поиска ассоциативных правил.

1. Задача поиска ассоциативных правил.
2. Постановка задачи.
3. Алгоритмы решения задачи.
4. Представление результатов.

Тема 9. Способы визуального представления данных.

1. Введение в визуализацию данных.
2. Визуализаторы общего назначения.
3. OLAP-анализ.
4. Визуализаторы, применяемые для оценки качества моделей.
5. Визуализаторы, применяемые для интерпретации результатов анализа.

Тема 10. Сферы применения ИАД.

1. Практическое применение Data Mining:
интернет-технологии;
- торговля;

- телекоммуникации;
- промышленное производство;
- медицина;
- банковское дело;
- страховой бизнес.

2. Применение ИАД в информационной безопасности.

3. Примеры применения ИАД по работе с клиентами, по выявлению мошенничества, по прямому маркетингу, по кредитному скорингу.

Тема 11. Использование методов ИАД в перспективных направлениях развития информационных технологий.

1. Обзор наиболее успешных продуктов, библиотек, toolkit, framework в выбранном направлении из области анализа данных.

2. История развития, степень развития, назначение, функции.

3. Конкретные примеры использования методов ИАД.

4. Применяемые средства анализа данных.

5. Совместимые технологии.

Тема 12. Применение ИАД в самостоятельном исследовании и проектировании собственных программных продуктов.

1. Государственные стандарты оформления технического и программной документации.

2. Проблемы разработки интеллектуальных систем (точность, надежность, вычислительная эффективность и пр.).

Критерии и шкала оценивания по оценочному средству «фронтальный и индивидуальный опрос»

Шкала оценивания	Критерий оценивания
отлично (5)	Студент глубоко и в полном объеме владеет программным материалом. Грамотно, исчерпывающе и логично его излагает в устной или письменной форме. При этом знает рекомендованную литературу, проявляет творческий подход в ответах на вопросы и правильно обосновывает принятые решения, хорошо владеет умениями и навыками при выполнении практических задач.
хорошо (4)	Студент знает программный материал, грамотно и по сути излагает его в устной или письменной форме, допуская незначительные неточности в утверждениях, трактовках, определениях и категориях или незначительное количество ошибок. При этом владеет необходимыми умениями и навыками при выполнении практических задач.
удовлетворительно (3)	Студент знает только основной программный материал, допускает неточности, недостаточно четкие формулировки, непоследовательность в ответах, излагаемых в устной или письменной форме. При этом недостаточно владеет умениями и навыками при выполнении практических задач. Допускает до 30% ошибок в излагаемых ответах.

неудовлетворительно (2)	Студент не знает значительной части программного материала. При этом допускает принципиальные ошибки в доказательствах, в трактовке понятий и категорий, проявляет низкую культуру знаний, не владеет основными умениями и навыками при выполнении практических задач. Студент отказывается от ответов на дополнительные вопросы
----------------------------	--

КУРСОВАЯ РАБОТА

Целями написания курсовой работы по дисциплине «Методы и технологии интеллектуального анализа данных» являются:

- углубление, систематизация и интеграция теоретических знаний и практических навыков, полученных при изучении данного курса;
- применение полученных теоретических знаний и практических навыков при решении прикладных задач;
- стимулирование навыков самостоятельной аналитической работы
- овладение современными методами научного исследования;
- выработка навыков презентации публичной дискуссии и защиты научных идей, предложений и рекомендаций.

Курсовая работа - исследовательская работа, выполняемая студентом самостоятельно под руководством научного руководителя.

Курсовая работа выполняется студентом по материалам, собранным им лично.

Рекомендуемый объем курсовой работы – от 20 до 30 страниц печатного текста с приложениями.

Примерное содержание курсовой работы:

1. Описание предметной области исследования Постановка проблемы.
2. Обоснование актуальности исследования.
3. Цели, задачи и методологическая база исследования.
4. Проведение исследования с использованием:
 - описательных статистик;
 - разведочного анализа данных,
 - выявления связей и зависимостей,
 - сравнения групп,
 - регрессионного анализа,
 - методов классификации:
 - кластерного анализа,
 - логистической регрессии,
 - деревьев решений,
 - нейронных сетей;
 - факторного анализа;
 - анализа и прогнозирования временных рядов и др
5. Построение модели.
6. Интерпретация полученных результатов.

7. Выводы.
8. Список использованных литературных и интернет источников (со ссылками на них в тексте).

**Примерная тематика курсовых работ по дисциплине
«Методы и технологии интеллектуального анализа данных»**

1. Сравнительный анализ финансового состояния предприятий различной отраслевой принадлежности в ряде регионов России.
2. Сравнительный анализ кредитоспособности предприятий различной отраслевой принадлежности в ряде регионов России.
3. Сравнительный анализ платежеспособности предприятий различной отраслевой принадлежности в ряде регионов России.
4. Сравнительный анализ финансовой устойчивости предприятий различной отраслевой принадлежности в ряде регионов России.
5. Анализ и прогнозирование вероятности банкротства предприятий различной отраслевой принадлежности в ряде регионов России.
6. Анализ транзакций по вкладам физических лиц для выработки оптимальных программ работы с частными вкладчиками.
7. Анализ и прогнозирование инвестиций и кредитных рисков.
8. Классификация кредиторов на основании их финансово-экономических и организационных характеристик.
9. Прогнозирование финансовых показателей.
10. Анализ и прогнозирование эффективности рекламной кампании.
11. Анализ и прогнозирование ценообразования.
12. Анализ и прогнозирование спроса на продукцию.
13. Сегментация рынка с использованием статистических методов.
14. Выявление, анализ и прогнозирование предпочтений клиентов.
15. Анализ и прогнозирование конкурентоспособности производимой продукции.
16. Сравнительный анализ поставщиков
17. Сегментация рынка услуг по степени удовлетворенности.
18. Анализ текущей социально-экономической ситуации и прогнозирование воздействия различных факторов на общество.
19. Анализ и прогнозирование доходов населения.
20. Оптимизация налоговых льгот для осуществления инвестиций.
21. Оценка и прогнозирование технического состояния транспортных средств города.
22. Анализ и прогнозирование рынка труда.
23. Анализ и прогнозирование безработицы с учетом сезонных колебаний.
24. Анализ и прогнозирование экспортно-импортной политики региона.
25. Анализ и прогнозирование повышения эффективности производства и сокращения затрат.
26. Анализ и прогнозирование стабильности показателей качества продукции.

27. Анализ и прогнозирование перевозок.
28. Анализ продаж билетов и прогнозирование динамики изменения пассажирских потоков.
29. Оценка доходности пассажирских перевозок от различных экономических показателей
30. Оценка и прогнозирование нагрузки телекоммуникационной сети
31. Сегментация и классификация абонентов телекоммуникационной сети.
32. Анализ и прогнозирование эффективности промо-акций в телекоммуникационной сфере.
33. Анализ и прогнозирование капитальных затрат на строительство объектов.
34. Оптимизация финансовых потоков, материальных и людских ресурсов в строительстве.
35. Прогнозирование стоимости проектов.
36. Прогнозирование стоимости недвижимости в зависимости от различных факторов.
37. Прогнозирование стоимости прокладки дорог.
38. Анализ и прогнозирование страховых выплат.
39. Анализ рисков и таблиц жизни.
40. Прогноз денежных поступлений от продаж полисов и их количества.

Критерии и шкала оценивания по оценочному средству курсовая работа

Шкала оценивания (интервал баллов) ²	Критерий оценивания
5	Курсовая работа выполнена на высоком уровне (правильные ответы даны на 90-100% вопросов/задач)
4	Курсовая работа выполнена на среднем уровне (правильные ответы даны на 75-89% вопросов/задач)
3	Курсовая работа выполнена на низком уровне (правильные ответы даны на 50-74% вопросов/задач)
2	Курсовая работа выполнена на неудовлетворительном уровне (правильные ответы даны менее чем на 50%)

Оценочные средства для промежуточной аттестации (зачет)

1. Пирамида знаний.
2. Задача анализа данных (Постановка, история возникновения, онтология методов)
3. Методика анализа данных.
4. Подготовка исходного набора данных.
5. Моделирование неопределённости. Вероятностные рассуждения.
6. Условные Байесовские сети.
7. Визуальный анализ данных.
8. Задача анализа временных рядов (определение временного ряда, классификация рядов динамики, разложение по компонентам).
9. Аномальные уровни временных рядов причины их возникновения и методы их устранения.
10. Тренд-сезонные экономические процессы и их анализ.
11. Задача дисперсионного анализа (постановка и методы решения).
12. Задача корреляционного анализа (постановка и методы решения).
13. Алгоритм Naïve Bayes.
14. Алгоритм C4.5.
15. Алгоритм Apriori.
16. Алгоритм C-means.
17. Основные статистические показатели временных рядов.
18. Метод Ирвина.
19. Метод проверки разности средних уровней.
20. Метод Фостера – Стюарта.
21. Методы сглаживания временных рядов.
22. Метод Четверикова.
23. Основные статистические показатели временных рядов.
24. Интеллектуальный анализ данных (определение и классификация задач).
25. Knowledge Discovery in Databases (KDD).
26. Что такое методы интеллектуального анализа данных?
27. Основные задачи data mining
28. Метод скользящего контроля
29. Проверка статистических гипотез о распределении. Оптимальное байесовское решение
30. Задача классификации (постановка и методы решения).
31. Наивный байесовский классификатор
32. Групповая классификация выборок объектов. Связь с принципом минимума информационного рассогласования Кульбака-Лейблера.
33. Критерии проверки однородности выборок
34. Описание алгоритмов классификации и обучения перцептрона Розенблатта.
35. Задача поиска ассоциативных правил (постановка и методы решения).

36. Задача секвенциального анализа (постановка и методы решения).
37. Задача кластеризации (постановка и методы решения).
38. Оперативная аналитическая обработка (OLAP) и многомерная модель данных.
39. Метод обратного распространения ошибки
40. Применение искусственной нейронной сети прямого распространения в задаче классификации с числом классов большим 2.
41. Обучение машин опорных векторов
42. Вероятностные нейронные сети. Реализация для дискретных величин
43. Приближенные методы ближайшего соседа
44. Методы построения коллективов решающих правил
45. Задачи обработки изображения. Признаки (дескрипторы) изображений
46. Распознавание речи. Скрытые марковские модели и MFCC.
47. Понятия экспертных систем
48. Экспертные системы (определение, назначение, модели представления знаний, достоинства).

**Критерии и шкала оценивания по оценочному средству
промежуточный контроль (зачет)**

шкала оценивания (интервал баллов)	Критерий оценивания
отлично (5)/ зачтено	Студент глубоко и в полном объеме владеет программным материалом. Грамотно, исчерпывающе и логично его излагает в устной или письменной форме. При этом знает рекомендованную литературу, проявляет творческий подход в ответах на вопросы и правильно обосновывает принятые решения, хорошо владеет умениями и навыками при выполнении практических задач.
хорошо (4)/ зачтено	Студент знает программный материал, грамотно и по сути излагает его в устной или письменной форме, допуская незначительные неточности в утверждениях, трактовках, определениях и категориях или незначительное количество ошибок. При этом владеет необходимыми умениями и навыками при выполнении практических задач.
удовлетворительно (3)/ зачтено	Студент знает только основной программный материал, допускает неточности, недостаточно четкие формулировки, непоследовательность в ответах, излагаемых в устной или письменной форме. При этом недостаточно владеет умениями и навыками при выполнении практических задач. Допускает до 30% ошибок в излагаемых ответах.
неудовлетворительно (2)/ не зачтено	Студент не знает значительной части программного материала. При этом допускает принципиальные ошибки в доказательствах, в трактовке понятий и категорий, проявляет низкую культуру знаний, не владеет основными умениями и навыками при выполнении практических задач. Студент отказывается от ответов на дополнительные вопросы

Лист изменений и дополнений

№ п/п	Виды дополнений и изменений	Дата и номер протокола заседания кафедры (кафедр), на котором были рассмотрены и одобрены изменения и дополнения	Подпись (с расшифровкой) заведующего кафедрой (заведующих кафедрами)

Экспертное заключение

Представленный фонд оценочных средств (далее - ФОС) *по дисциплине «Методы и технологии интеллектуального анализа данных»* соответствует требованиям ФГОС ВО.

Предлагаемые формы и средства текущего и промежуточного контроля адекватны целям и задачам реализации основной профессиональной образовательной программы по направлению подготовки 09.04.03 Прикладная информатика.

Оценочные средства для текущего контроля успеваемости, промежуточной аттестации по итогам освоения дисциплины представлены в полном объеме.

Виды оценочных средств, включенные в представленный фонд, отвечают основным принципам формирования ФОС.

Разработанный и представленный для экспертизы фонд оценочных средств рекомендуется к использованию в процессе подготовки обучающихся по указанному направлению.

Председатель учебно-методической комиссии факультета компьютерных систем и информационных технологий



Ветрова Н. Н.