

ГОСУДАРСТВЕННОЕ ОБРАЗОВАТЕЛЬНОЕ УЧРЕЖДЕНИЕ  
ВЫСШЕГО ПРОФЕССИОНАЛЬНОГО ОБРАЗОВАНИЯ  
«ДОНЕЦКИЙ НАЦИОНАЛЬНЫЙ ТЕХНИЧЕСКИЙ УНИВЕРСИТЕТ»



**УТВЕРЖДАЮ:**

Проректор

по научно-педагогической работе

А.Б. Бирюков

(подпись)

20 июня 20 20 года

**РАБОЧАЯ ПРОГРАММА ДИСЦИПЛИНЫ**

Б1.В2 Интеллектуальный анализ данных

(код и наименование дисциплины согласно учебному плану)

Направление подготовки:

09.04.04 Программная инженерия

(код и наименование направления / специальности)

Магистерская программа:

Методы и средства разработки программного обеспечения

(наименование профиля / магистерской программы / специализации)

Уровень образования:

Магистратура

(бакалавриат, магистратура, специалитет)

Форма обучения:

очная, заочная

(очная, заочная, очно-заочная)

Формы обучения:	Очная	Заочная
Семестр(ы)	3	3
Общая трудоемкость в ЗЕТ/часах	6 / 216	6 / 216
Контактная работа (час.), в том числе:	72	26
лекции (час.)	34	10
лабораторные работы (час.)	34	10
практические (семинарские) занятия (час.)	-	-
Самостоятельная работа (час.), в том числе:	112	160
курсовой проект (работа) (семестр/час)	-	-
индивидуальное задание (кол./час)	-	1/9
Контроль (экзамен, час./зачёт)	экзамен, 36	экзамен, 36

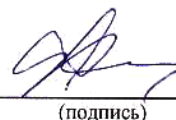
Донецк, 2020 г.

Рабочая программа дисциплины Б1.В2 Интеллектуальный анализ данных составлена в соответствии с учебным планом подготовки магистров по направлению 09.04.04 Программная инженерия, магистерская программа Методы и средства разработки программного обеспечения для 2020 года приёма.

**Составитель:**

зав. кафедрой

программной инженерии, к.т.н., доцент



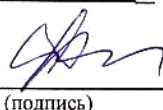
(подпись)

Федяев О.И.

Рабочая программа **рассмотрена и принята** на заседании кафедры программной инженерии.

Протокол от «10» марта 2020 года № 9.

Заведующий кафедрой



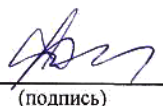
(подпись)

Федяев О.И.

Рабочая программа **одобрена учебно-методической комиссией** ГОУВПО «ДОННТУ» по направлению подготовки 09.03.04 Программная инженерия

Протокол от «20» мая 2020 года № 10.

Председатель



(подпись)

Федяев О.И.

Рабочая программа **продлена** для 20\_\_ года приёма на заседании кафедры программной инженерии.

Протокол от «  » \_\_\_\_\_ 20\_\_ года №   .

Заведующий кафедрой

(подпись)

(Ф.И.О.)

Рабочая программа **продлена** для 20\_\_ года приёма на заседании кафедры программной инженерии.

Протокол от «  » \_\_\_\_\_ 20\_\_ года №   .

Заведующий кафедрой

(подпись)

(Ф.И.О.)

Рабочая программа **продлена** для 20\_\_ года приёма на заседании кафедры программной инженерии.

Протокол от «  » \_\_\_\_\_ 20\_\_ года №   .

Заведующий кафедрой

(подпись)

(Ф.И.О.)

## 1 ОБЪЕКТ, ЦЕЛЬ И ЗАДАЧИ ОСВОЕНИЯ ДИСЦИПЛИНЫ

Дисциплина ориентирована на продолжение изучения современной технологии обработки данных – интеллектуальному анализу данных (Data Mining). Отдельные разделы этого направления были рассмотрены ранее в дисциплинах «Системы искусственного интеллекта» и «Нейросетевые и нечёткие системы». В данной дисциплине сделан акцент на систематическое изложение традиционных методов математической статистики, которые по-прежнему остаются основными инструментами анализа данных. Они математически обоснованы и полезны для проверки заранее сформулированных гипотез, для «грубого» разведочного анализа, составляющего основу оперативной аналитической обработки данных. Современные версии статистических пакетов реализуют мощные классические методики: корреляционный, регрессионный, факторный и временной анализ. Поэтому от пользователя такими пакетами требуется серьёзная математическая подготовка в этих вопросах.

В результате изучения данной дисциплины студент получит знания о современных моделях статического прогнозирования (как наиболее строго обоснованных математически), а также практические навыки анализа, моделирования и прогнозирования элементов выборки различной природы с использованием современных инструментальных пакетов программ.

Положительной особенностью составленной рабочей программы является то, что учебный материал составлен из двух взаимосвязанных и взаимодополняющих друг друга частей: теоретической, в которой прорабатываются основные подходы и методы теории статистического анализа и прогнозирования, и практической, в которой закрепляются приёмы анализа и прогнозирования в инструментальной системе STATISTICA. Совмещение теории и практики, как показывает опыт преподавания, способствует тому, что студенты смогут получить, во-первых, качественные профессиональные знания в области статистического анализа данных, а во-вторых, работа с пакетами улучшит понимание назначения используемых методов и их свойства. Пакет STATISTICA сейчас занимает лидирующее место на рынке статистического программного обеспечения.

Таким образом, предметом курса являются современные статистические методы и программные средства автоматизации анализа данных.

**Цель дисциплины** состоит в изучении основных методов и программных средств обнаружения в данных ранее неизвестных, нетривиальных, практически полезных и доступных интерпретации знаний, необходимых для принятия решений в различных сферах человеческой деятельности.

**В задачи дисциплины** входит: сфера применения и рынок инструментов Data Mining; архитектура современных систем для интеллектуального анализа данных (ИАД); усвоение классических статистических методов решения задач: классификации, регрессии, прогнозирования, кластеризации, определения взаимосвязей (задачи поиска ассоциативных правил), анализа последовательностей, факторного анализа; изучение методов приобретения знаний из баз данных и текста; изучение методов и инструментальных средств Data Mining.

В результате освоения дисциплины студент должен:

**знать:**

- методы анализа данных «с учителем», использующие обучающую информацию;
- методы исследования многомерных данных без использования обучающей информации, направленные на выяснение структуры взаимоотношений объектов и признаков;
- методы выявления логических закономерностей с помощью деревьев решений;
- концептуальные основы построения программных инструментальных систем Data Mining и методики их применения для ИАД;
- возможности современных и перспективных средств разработки программных продуктов ИАД;

**уметь:**

- формулировать задачи анализа данных;
- выбирать адекватные алгоритмы их решения;
- оценивать качество получаемых решений;
- подготавливать данные для анализа;
- разрабатывать инструментальные программные средства анализа данных;
- применять знания данной дисциплины к решению практических задач в сфере профессиональной деятельности, проектировании и разработке информационно-аналитических систем, систем поддержки принятия решений;

**владеть:**

- основными методами и приёмами исследовательской и практической работы в области ИАД;
- навыками работы с современными методами и инструментами для ИАД;
- технологиями разработки алгоритмов и программных систем анализа данных;
- комплексным подходом к разработке и внедрению аналитических подсистем в системах подготовки принятия решений.

Перечисленные результаты обучения являются основой для формирования следующих компетенций у студентов:

- способен осуществлять критический анализ проблемных ситуаций на основе системного подхода, вырабатывать стратегию действий (УК-1);
- знание методов организации и управления информационными процессами (ПК-1);
- владение навыками создания программного обеспечения для анализа, распознавания и обработки информации, систем цифровой обработки сигналов (ПК-3);
- способен выполнить постановку новых задач анализа и синтеза новых проектных решений (ПК-5).

## 2 МЕСТО ДИСЦИПЛИНЫ В ОСНОВНОЙ ОБРАЗОВАТЕЛЬНОЙ ПРОГРАММЕ

Дисциплина относится к части, формируемой участниками образовательных отношений, блока 1 «Дисциплины (модули) Б1» учебного плана.

Базируется на знаниях, умениях и навыках, которые студент приобрел при освоении предшествующих дисциплин в соответствии с учебным планом по направлению 09.04.04 Программная инженерия (магистерская программа «Методы и средства разработки программного обеспечения»):

- «Информационные и телекоммуникационные технологии»;
- «Цифровая обработка сигналов и распознавание речи»;
- «Нейросетевые и нечёткие системы»;
- «Распознавание образов»;
- «Технологии облачных вычислений».

Знания, умения и навыки, приобретённые при освоении данной дисциплины, реализуются студентом при освоении следующих учебных дисциплин:

- «Производственная практика: преддипломная»;
- «Выполнение выпускной квалификационной работы».

## 3. СТРУКТУРА И СОДЕРЖАНИЕ ДИСЦИПЛИНЫ

### 3.1 Распределение учебных часов по темам дисциплины и видам занятий

№ темы	Наименование темы (содержательных модулей)	Количество часов (очная/заочная)				
		Всего	В том числе			
			Лекции	Лабор.	Практ.	СР
1	Цель, предмет и задачи курса. Источники знаний для интеллектуальных систем. Общее представление о технологиях Data Mining и Big Data	8/9	2/1			6/8
2	Современные методы статистического анализа данных: пакет Statistica. Методы анализа временных рядов. Основные компоненты ряда	12/11	2/1	2/0		8/10
3	Выделение сезонной компоненты: аддитивная и мультипликативная модели. Разностные операторы. Преобразование шкалы	14/15	2/1	4/2		8/12
4	Модели случайной компоненты. AR-модели. Свойства моделей и их идентификация	20/17	4/1	4/0		12/16
5	МА-модели. Свойства моделей и их идентификация	20/15	4/1	4/0		12/14
6	Модели ARMA и ARIMA. Идентификация параметров моделей	14/15	2/1	4/2		8/12
7	Методы, использующие обучающую	16/14	2/0	2/0		12/14

	информацию: множественная регрессия, методы сравнения с образцом					
8	Дискриминантный анализ (ДА). Характеристика методов ДА (метод дискриминации и метод классификации)	20/19	4/1	4/2		12/16
9	Методы автоматического группирования (факторный анализ, кластерный анализ)	21/20	5/1	4/2		12/17
10	Обнаружение логических закономерностей в данных с помощью деревьев решений. Алгоритмы построения деревьев решений	23/21	5/1	6/2		12/18
11	Применение инструментов Data Mining, Text Analysis и Web Mining к интеллектуальному анализу данных	12/15	2/1			10/14
Индивидуальное задание		0/9				0/9
Курсовой проект						
Итого по видам занятий		180/180	34/10	34/10		112/160
Контроль		36/36				
<b>ИТОГО:</b>		<b>216/216</b>				

### Формирование компетенций в результате освоения тем дисциплины

Компетенции	Темы дисциплины, нацеленные на формирование компетенции
УК-1	Темы 1, 2, 6, 7, 8, 9, 10, 11
ПК-1	Темы 1, 6, 11
ПК-3	Темы 2, 3, 4, 5, 8, 9, 10
ПК-5	Темы 1, 2, 7, 8, 9, 10, 11

### 3.2 Лекции

Тема 1. Цель, предмет и задачи курса. Источники знаний для интеллектуальных систем. Общее представление о технологиях Data Mining и Big Data.

#### Содержание темы 1:

Современные методы анализа данных. Стадии и методы анализа. Общие представления о технологии Data Mining. Основные подходы к извлечению знаний. Источники знаний.

Литература к теме 1: [2, 4].

Тема 2. Современные методы статистического анализа данных: пакет Statistica. Методы анализа временных рядов. Основные компоненты ряда.

#### Содержание темы 2:

Статистические методы анализа данных. Цели и способы анализа временных рядов. Тренд, сезонная и циклическая компоненты ряда. Модели тренда.

Литература к теме 2: [1, 2, 3].

Тема 3. Выделение сезонной компоненты: аддитивная и мультипликативная модели. Разностные операторы. Преобразование шкалы.

Содержание темы 3:

Аддитивная модель, мультипликативная модель. Прогнозирование с использованием сезонной компоненты. Простые и сезонные разностные операторы. Способы преобразования шкалы.

Литература к теме 3: [1, 2, 3].

Тема 4. Модели случайной компоненты. AR-модели. Свойства моделей и их идентификация.

Содержание темы 4:

Авторегрессионные модели AR. Случайный ряд типа «белый шум». Стационарность случайного процесса. Выборочная автокорреляционная функция ряда. Коррелограмма. Интерпретация графика коррелограммы. Числовые характеристики процессов AR(1) и AR(2). Оценивание параметров AR-моделей.

Литература к теме 4: [1, 2, 3].

Тема 5. МА-модели. Свойства моделей и их идентификация.

Содержание темы 5:

Частная автокорреляционная функция, её свойства. Свойства модели MA(q). Нахождение параметров модели MA(q).

Литература к теме 5: [1, 2, 3].

Тема 6. Модели ARMA и ARIMA. Идентификация параметров моделей.

Содержание темы 6:

Смешанная модель авторегрессии – скользящего среднего ARMA(p,q). Оценивание параметров ARMA(p,q). Идентификация стационарных моделей AR, MA, ARMA. Модель ARMA в операторной форме. Модель авторегрессии – проинтегрированного скользящего среднего ARIMA(p,d,q). Идентификация модели ARIMA(p,d,q). Проверка нормальности распределения.

Литература к теме 6: [1, 2, 3].

Тема 7. Методы, использующие обучающую информацию: множественная регрессия, методы сравнения с образцом.

Содержание темы 7:

Методы анализа данных «с учителем». Множественный регрессионный анализ. Применение множественной регрессии в различных областях.

Литература к теме 7: [2, 3, 5].

Тема 8. Дискриминантный анализ (ДА). Характеристика методов ДА (метод дискриминации и метод классификации).

Содержание темы 8:

Области применения дискриминантного анализа. Предмет и задачи дискриминантного анализа. Метод дискриминации, метод классификации.

Литература к теме 8: [1 - 5].

Тема 9. Методы автоматического группирования (факторный анализ, кластерный анализ).

Содержание темы 9:

Статистические методы анализа структур данных. Типы задач и методы их решения. Метод главных компонент. Факторный анализ. Задача кластерного анализа. Методы кластерного анализа (иерархические, неиерархические). Метод К-средних.

Литература к теме 9: [1 - 4].

Тема 10. Обнаружение логических закономерностей в данных с помощью деревьев решений. Алгоритмы построения деревьев решений.

Содержание темы 10:

Язык логических высказываний. Деревья решений – как метод выявления логических закономерностей. Этапы конструирования деревьев. Критерий расщепления (выбор атрибута разбиения). Алгоритмы построения деревьев решений: ID3, C4.5, CART.

Литература к теме 10: [2, 4, 5].

Тема 11. Применение инструментов Data Mining, Text Analysis и Web Mining к интеллектуальному анализу данных.

Содержание темы 11:

Обзор инструментальных средств интеллектуального анализа данных. Технологические особенности их применения.

Литература к теме 11: [2, 4, 5].

### **3.3 Практические (семинарские) занятия**

Практические (семинарские) занятия по дисциплине учебным планом не предусмотрены.

### **3.4 Лабораторные работы**

№ п/п	Тема занятия	Объём, час. очн./заоч.	Литера- тура
1	Построение детерминированной модели для анализа и прогнозирования динамики процесса	6/2	[1, 2, 3, 6]
2	Модель прогнозирования динамики процесса с учётом случайной компоненты	12/2	[1, 2, 3, 6]
3	Выявление закономерностей в данных с помощью дискриминантного анализа	6/2	[2, 3, 5, 6]
4	Факторный анализ – метод исследования структуры данных	4/2	[2, 4, 5, 6]
5	Обнаружение логических закономерностей в данных с помощью деревьев решений	6/2	[2, 4, 5, 6]
<b>ИТОГО:</b>		<b>34/10</b>	



### 3.5 Самостоятельная работа студента

№ п/п	Виды самостоятельной работы студента	Объём, час. очн./заоч.
1	Изучение лекционного материала	50/60
2	Подготовка к лабораторным занятиям	62/91
3	Выполнение индивидуального задания	0/9
<b>ИТОГО:</b>		<b>112/160</b>

### 3.6 Курсовой проект (работа), индивидуальное задание

Учебным планом в рамках освоения дисциплины не предусмотрено выполнение студентами курсового проекта.

В 3 семестре студентом заочной формы обучения выполняется индивидуальное задание. Индивидуальное задание посвящено метрикам качества классификации.

Цель – приобретение студентами опыта в применении различных мер качества классификации (долю правильных ответов, точность, полноту, AUC-ROC) для сравнения алгоритмов классификации при наличии ограничений на точность или полноту.

Отчёт о работе состоит из текстовой части на листах формата А4. Рекомендуемый объём отчёта по индивидуальному заданию – не более 12 страниц.

## 4 ФОНД ОЦЕНОЧНЫХ СРЕДСТВ

### 4.1 Критерии и шкалы для интегрированной оценки уровня сформированности компетенций

#### *Составляющая компетенции – полнота знаний*

- нулевой уровень: неверные, не аргументированные, с множеством грубых ошибок ответы на вопросы; уровень знаний ниже минимальных требований;
- минимальный уровень: даны не полные, неточные и неаргументированные ответы на вопросы, допущено много грубых ошибок, уровень знаний ниже минимальных требований;
- пороговый уровень: даны недостаточно полные, точные и аргументированные ответы на вопросы, плохо знает термины, определения и понятия, основные закономерности, соотношения, принципы; допущено много негрубых ошибок;
- средний уровень: даны достаточно полные, точные и аргументированные ответы на вопросы; в целом знает термины, определения и понятия, основные закономерности, соотношения, принципы; допущено несколько негрубых ошибок;
- продвинутый уровень: даны полные, точные и аргументированные ответы на вопросы, знает термины, определения и понятия, основные закономерности, соотношения, принципы; допущено несколько негрубых ошибок;
- высокий уровень: даны полные, точные и аргументированные ответы на вопросы, знает термины, определения и понятия, основные закономерности, соотношения, принципы; допущено несколько неточностей.

#### *Составляющая компетенции – умения*

- нулевой уровень: полное отсутствие понимания сути методики решения задачи, допущено множество грубейших ошибок, задания не выполнены вообще;
- минимальный уровень: слабое понимание сути методики решения задачи, допущены грубые ошибки, решения не обоснованы, не умеет использовать нормативно-техническую литературу;
- пороговый уровень: достаточное понимание сути методики решения задачи, допущены ошибки, решения не всегда обоснованы, умеет использовать нормативно-техническую литературу, слабо ориентируется в специальной научной литературе;
- средний уровень: в целом понимает суть методики решения задачи, допущены ошибки, решения не всегда обоснованы, умеет использовать нормативно-техническую и специальную научную литературу;
- продвинутый уровень: в целом понимает суть методики решения задачи, допущены неточности, способен обосновать решения, умеет использовать нормативно-техническую и специальную научную литературу;
- высокий уровень: понимает суть методики решения задачи, способен обосновать решения, умеет использовать нормативно-техническую и специальную научную литературу, передовой производственный опыт.

### *Составляющая компетенции – владение навыками*

- нулевой уровень: не демонстрирует владение навыками выполнения профессиональных задач, не может выполнить задания;
- минимальный уровень: не демонстрирует владение навыками выполнения профессиональных задач, испытывает существенные трудности при выполнении отдельных заданий;
- пороговый уровень: владеет навыками выполнения профессиональных задач на пороговом уровне, задания выполняет медленно и некачественно;
- средний уровень: владеет навыками выполнения профессиональных задач, задания выполняет на среднем уровне по скорости и качеству;
- продвинутый уровень: владеет уверенными навыками выполнения профессиональных задач, быстро и качественно выполняет задания, иногда допуская незначительные погрешности;
- высокий уровень: владеет уверенными навыками выполнения профессиональных задач, быстро и качественно выполняет задания, при необходимости демонстрируя творческий подход.

### *Обобщенная оценка сформированности компетенций*

- нулевой уровень: на нулевом уровне сформированы: все составляющие; одна или две из трёх, остальные – на более высоком уровне;
- минимальный уровень: на минимальном уровне сформированы: все составляющие; одна или две из трёх, остальные – на более высоком уровне;
- пороговый уровень: на пороговом уровне сформированы: все составляющие; одна или две из трёх, остальные – на более высоком уровне;
- средний уровень: на среднем уровне сформированы: все составляющие; одна или две из трёх, остальные – на более высоком уровне;
- продвинутый уровень: на продвинутом уровне сформированы: все составляющие; одна или две из трёх, остальные – на высоком уровне;
- высокий уровень: на высоком уровне сформированы все составляющие компетенций.

## **4.2 Вопросы к экзамену и пример экзаменационного билета**

### **Вопросы к экзамену:**

1. Основные подходы к извлечению знаний. Применение и рынок инструментов Data Mining и Big Data. Источники знаний для интеллектуальных систем. Общее представление о технологиях Data Mining и Big Data.
2. Типы закономерностей, выявляемые методами Data Mining. Их характеристика.
3. Статистические методы анализа данных. Прогнозирование – как способ изучения путей развития на основе выявленных знаний о динамике процесса. Постановка задачи анализа временного ряда, цели анализа.

4. Пакет Statistica. Методы анализа временных рядов. Основные компоненты временного ряда.
5. Детерминированная и случайная составляющие временного ряда, аддитивная и мультипликативная модели.
6. Модели тренда, сезонная и циклическая компоненты временного ряда, их краткая характеристика.
7. Стационарность случайного процесса. Методы сведения к стационарности. Понятие разностного оператора 1-го порядка, его назначение. Пример применения разностного оператора.
8. Выделение сезонных эффектов, способы оценки сезонной компоненты, сезонный индекс, удаление сезонной компоненты.
9. Прогнозирование с использованием сезонной компоненты. Пример прогнозирования.
10. Учёт случайной составляющей (остатков) временного ряда. Выборочная автокорреляционная функция временного ряда. Оценка коэффициентов корреляции. Коррелограмма. Интерпретация графика коррелограммы.
11. Линейная модель для случайной компоненты временного ряда. Модель авторегрессии первого порядка  $AR(1)$ . Оценивание параметров модели  $AR(1)$ .
12. Модель авторегрессии второго порядка  $AR(2)$ . Оценивание параметров  $AR(2)$ .
13. Частная автокорреляционная функция (ЧАКФ), её свойства. Применение этой функции.
14. Модель скользящего среднего  $MA(q)$ . Нахождение параметров модели  $MA(q)$ .
15. Смешанные модели авторегрессии–скользящего среднего  $ARMA(p,q)$  и  $ARIMA(p,d,q)$ . Идентификация моделей.
16. Статистические методы анализа структур данных. Постановка задачи, типы задач, методы решения, области применения. Идея метода главных компонент и факторного анализа.
17. Иллюстративный пример применения факторного анализа к распознаванию образов (арабских цифр).
18. Методы статистического анализа данных, использующих обучающую информацию (многомерный анализ данных «с учителем»).
19. Множественная регрессия, методы сравнения с образцом.
20. Дискриминантный анализ (ДА), области применения ДА, типы задач ДА, характеристика методов ДА (метод дискриминации и метод классификации). Общая постановка задачи ДА, программные пакеты реализации ДА.
21. Кластерный анализ данных, Задача кластерного анализа, пути её решения. Метрики в кластерном анализе.
22. Методы и алгоритмы кластерного анализа, их идея.
23. Обнаружение логических закономерностей в данных с помощью деревьев решений. Алгоритмы построения деревьев решений.
24. Основные принципы нейросетевого представления неизвестных знаний и закономерностей. Возможности инструментов Data Mining, Text Analysis и Web Mining для интеллектуального анализа данных.

## Пример экзаменационного билета:

ГОУВПО «ДОНЕЦКИЙ НАЦИОНАЛЬНЫЙ ТЕХНИЧЕСКИЙ УНИВЕРСИТЕТ»

Дисциплина: Интеллектуальный анализ данных

Направление подготовки: Программная инженерия

Магистерская программа: Методы и средства разработки программного обеспечения

Шифр направления: 09.04.04

Группа: ПИМ-19

### Экзаменационный билет № 2

1. Методы статистического анализа данных, использующие обучающую информацию (многомерный анализ данных «с учителем»). Дискриминантный анализ (ДА), области применения ДА, типы задач ДА, общая постановка задачи ДА, программные пакеты реализации ДА.

2. Задача.

Осуществить прогноз динамики заданного процесса (поквартальную прибыль компании, тыс. руб.) с помощью построенной мультипликативно - аддитивной модели.

t	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
X <sub>t</sub>	72	100	90	64	70	92	80	58	62	80	68	48	52	60	50	30

3. Задача.

С помощью факторного анализа минимизировать число признаков и определить какие именно взять признаки (элементы шаблона), по которым можно распознавать символы: **0, 5, 6, 7, 8, 9**. Символы высвечиваются на экране калькулятора своими отрезками.



Утверждено на заседании кафедры программной инженерии

Протокол №\_\_\_\_\_от \_\_\_\_\_ 20\_\_\_\_г.

Зав. кафедрой

О.И. Федяев

Экзаменатор

О.И. Федяев

### КРИТЕРИИ

#### оценивания экзаменационной работы

по дисциплине «Интеллектуальный анализ данных»

для обучающихся по направлению 09.04.04 Программная инженерия

(магистерская программа Методы и средства разработки программного обеспечения)

Экзамен проводится письменно по билетам. Билет содержит 3 вопроса: один на знание теории по тематике дисциплины, а второй и третий на умение решить конкретную задачу. Каждый вопрос требует конкретного ответа. При необходимости отвечающий должен сопроводить написанное поясняющей информацией о структуре исходных данных для анализа, о построенной модели, об алгоритме интеллектуального анализа данных, о динамике процесса в виде графиков, о дереве решений, о выполненных расчётах и т.п.

Вопросы охватывают теоретическую часть курса, а также требуют демонстрации практических навыков, полученных студентом в ходе выполнения лабораторных работ.

Правильный ответ на каждый вопрос оценивается в 15 баллов. Если ответ на вопрос не полный, то он оценивается в диапазоне от 1 до 15 баллов. При отсутствии правильного ответа на поставленный вопрос обучающийся получает ноль баллов. Баллы, полученные за ответы на вопросы билета, суммируются и с учётом результатов текущего контроля работы студента выводится итоговая оценка по 100-балльной шкале.

Полученная оценка по 100-балльной шкале определяет оценку по государственной шкале и шкале ECTS.

Утверждено на заседании кафедры программной инженерии,  
протокол № 9 от 10.03.2020 г.

Заведующий кафедрой \_\_\_\_\_ Федяев О.И..

### 4.3 Критерии оценивания

Оценивание уровня усвоения студентом учебного материала дисциплины «Интеллектуальный анализ данных» производится в ходе текущего контроля и промежуточной аттестации (семестрового контроля).

**Текущий контроль** знаний студента производится по результатам выполнения лабораторных работ, а для студента заочной формы обучения ещё и по результатам выполнения индивидуального задания.

Распределение баллов текущего контроля работы студента на протяжении семестра приведено в табл. 1.

Таблица 1 – Распределение баллов текущего контроля

Форма контроля	Возможное количество баллов	Примечание
Для студентов очной формы обучения		
Отчёт о выполнении лабораторной работы	11	Оценивается в баллах каждое задание. Индивидуальное задание по лабораторной работе выполнено правильно, проектные решения обоснованы, приведен анализ полученного результата
	6	Минимальное количество баллов за одно задание. Задание выполнено в целом правильно, проектные решения не всегда обоснованы, возникли трудности в объяснении полученных результатов
<b>Итого по лабораторным работам (максимально возможное)</b>	<b>55</b>	Из расчёта выполнения студентом 5 индивидуальных лабораторных работ.
<b>ИТОГО:</b>	<b>55</b>	Максимально возможное
Для студентов заочной формы обучения		
Отчёт о выполнении лабораторной работы	10	Оценивается в баллах каждое задание. Индивидуальное задание по лабораторной работе выполнено правильно, проектные решения обоснованы, приведен анализ полученного результата

Форма контроля	Возможное количество баллов	Примечание
	6	Минимальное количество баллов за одно задание. Задание выполнено в целом правильно, проектные решения не всегда обоснованы, возникли трудности в объяснении полученных результатов, имеются замечания по оформлению
<b>Итого по лабораторным работам (максимально возможное)</b>	<b>50</b>	Из расчёта выполнения студентом 5 индивидуальных лабораторных работ.
Выполнение индивидуального задания	5	При выполнении задания приняты правильные проектные решения, изложение материала аргументированное, последовательное, работа оформлена грамотно
	3	Задание выполнено в целом правильно, проектные решения не всегда обоснованы, имеются замечания по оформлению работы
<b>ИТОГО:</b>	<b>55</b>	Максимально возможное

**Промежуточная аттестация** по результатам освоения дисциплины в семестре проводится в форме семестрового экзамена. Форма проведения экзамена – письменная. Экзаменационный билет включает в себя 3 вопроса: один теоретический и два практических (задачи). При оценивании студента на экзамене преподаватель руководствуется критериями, приведенными в табл. 2.

Максимальное количество баллов за ответ на вопрос экзаменационного билета засчитывается студенту в случае, если ответ подтверждает владение студентом знаниями в полном объёме учебной программы, материал изложен в логической последовательности с выделением главного, содержит точные формулировки, сопровождается иллюстрирующими схемами и рисунками (при необходимости).

В случае, если ответ на вопрос не в полной мере отвечает приведенным требованиям, студенту начисляется количество баллов в диапазоне от 1 до 15 баллов. При отсутствии правильного ответа на поставленный вопрос студент получает 0 баллов.

Таблица 2 – Распределение баллов по семестровому экзамену

Форма контроля		Максимально возможное количество баллов
Ответ на вопросы экзаменационного билета	Вопрос 1	15
	Вопрос 2 (задача)	15
	Вопрос 3 (задача)	15
<b>ИТОГО:</b>		<b>45</b>

**Итоговая оценка** определяется путём суммирования количества баллов по результатам текущего контроля и количества баллов по результатам семестрового экзамена. **Максимально возможное количество баллов – 100.**

Полученная оценка по 100-балльной шкале определяет оценку по государственной шкале и шкале ECTS:

Сумма баллов по 100-балльной шкале	Оценка по шкале ECTS	Оценка по государственной шкале
90-100	A	Отлично
80-89	B	Хорошо
75-79	C	
70-74	D	Удовлетворительно
60-69	E	
35-59	FX	Неудовлетворительно
0-34	F*	

\* – с обязательным повторным изучением дисциплины.

#### 4.4 Пример текущего опроса на лабораторных занятиях

На примере темы «Модель прогнозирования динамики процесса с учётом случайной компоненты»:

1. Какой случайный процесс называют процессом авторегрессии первого порядка?
2. Какими свойствами обладает модель  $AR(1)$ ?
3. В каких случаях целесообразно использовать модель  $AR(2)$ ?
4. Какой случайный процесс называют процессом скользящего среднего первого порядка?
5. Какими свойствами обладает модель  $MA(1)$ ?
6. В каких случаях целесообразно использовать модель  $MA(2)$ ?
7. Как находятся параметры модели  $MA(1)$ ?
8. Как находятся параметры модели  $MA(2)$ ?
9. Каковы особенности смешанной модели авторегрессии-скользящего среднего  $ARMA(p,q)$ ?
10. Как выполняется оценивание параметров модели  $ARMA(p,q)$ ?
11. Что понимается под идентификацией стационарных моделей  $AR$ ,  $MA$  и  $ARMA$ ?
12. Каким способом выполняется идентификация моделей  $AR$ ,  $MA$  и  $ARMA$ ?
13. Каковы особенности процесса авторегрессии-проинтегрированного скользящего среднего  $ARIMA(p,d,q)$ ?
14. Как выполняется идентификация модели  $ARIMA(p,d,q)$ ?

Ответы на вопросы входного контроля учитываются преподавателем в результатах текущего контроля работы студента.



## **5 РЕКОМЕНДУЕМАЯ ЛИТЕРАТУРА**

### **I Основная литература**

1. Сизова Т.М. Статистика [Электронный ресурс]: учебное пособие для вузов / Т.М. Сизова. - 3 Мб. - Санкт-Петербург: ИТМО, 2013. - 1 файл. - Систем. требования: Acrobat Reader. <http://ed.donntu.org/books/cd5533.pdf>

2. Сирота А.А. Методы и алгоритмы анализа данных и их моделирование в MATLAB [Электронный ресурс]: учебное пособие для вузов / А.А. Сирота. - 1 Мб. - Санкт-Петербург: БХВ-Петербург, 2016. - 1 файл. - Систем. требования: Acrobat Reader. <http://ed.donntu.org/books/17/cd8002.pdf>

3. Фадеева Л.Н. Теория вероятностей и математическая статистика [Электронный ресурс]: учебное пособие для вузов / Л.Н. Фадеева, А.В. Лебедев; под ред. Л.Н. Фадеевой. - 10 Мб. - Москва: Рид Групп, 2011. - 1 файл. - Систем. требования: Acrobat Reader. <http://ed.donntu.org/books/cd5943.pdf>

### **II Дополнительная литература**

4. Силен Д. Основы Data Science и Big Data. Python и наука о данных [Электронный ресурс] / Д. Силен, А. Мейсман, М. Али; Д. Силен, А. Мейсман, М. Али; пер. с англ. Е. Матвеев. - 113 Мб. - Санкт-Петербург: Питер, 2017. - 1 файл. - (Библиотека программиста). - Перевод изд.: Introducing Data Science: big Data, Machine learning, and more, using Python tools / D. Cielen, A.D.B. Meysman, M. Ali. - Систем. требования: Acrobat Reader. <http://ed.donntu.org/books/17/cd6499.pdf>

5. Нечаев Ю.И. Интеллектуальные системы: концепция и приложения [Электронный ресурс]: учебное пособие для вузов / Ю.И. Нечаев, А.Б. Дегтярев; С.-Пб. гос. ун-т. - 7 Мб. - Санкт-Петербург: Изд-во СПбГПУ, 2011. - 1 файл. - Систем. требования: Просмотрщик djvu-файлов. <http://ed.donntu.org/books/17/cd8012.djvu>

## **6 УЧЕБНО-МЕТОДИЧЕСКОЕ ОБЕСПЕЧЕНИЕ ДИСЦИПЛИНЫ**

### **Учебно-методические издания, разработанные в ДонНТУ:**

6. Методические указания к лабораторным занятиям по дисциплине «Интеллектуальный анализ данных» [Электронный ресурс]: для студентов уровня профессионального образования «магистр» направления подготовки 09.04.04 «Программная инженерия» магистерской программы «Методы и средства разработки программного обеспечения» всех форм обучения / ГОУВПО «ДОННТУ», каф. программной инженерии; сост. О.И.Федяев, О.В.Рычка, А.В.Московченко. – Электрон. дан. (1 файл: 3,6 Мб). – Донецк: ДОННТУ, 2020. – Систем. требования: Acrobat Reader.

### **Электронно-информационные ресурсы**

ЭБС ДОННТУ – <http://donntu.org/library>

## **7 МАТЕРИАЛЬНО-ТЕХНИЧЕСКОЕ ОБЕСПЕЧЕНИЕ ДИСЦИПЛИНЫ**

### **7.1 Лекционные занятия:**

Учебная аудитория № 8.704 (учебный корпус 8) для проведения занятий лекционного типа, групповых и индивидуальных консультаций, текущего контроля и промежуточной аттестации, (мультимедийное оборудование: компьютер, операционная система Windows 7 Professional x86/64 (академическая подписка DreamSparkPremium), LibreOffice 4.3.2.2, Google Slides (бесплатная версия)), мультимедийный проектор, экран; специализированная мебель: доска аудиторная, парты.

### **7.2 Лабораторные занятия:**

Компьютерная аудитория № 5.429 (учебный корпус 5) для проведения лабораторных занятий, групповых и индивидуальных консультаций, текущего контроля и промежуточной аттестации, (мультимедийное оборудование: компьютер IntelCore 2Duo E8200 2.66MHz/4 Gb ОЗУ/160 Gb HDD, операционная система Windows XP Professional x86/64 (академическая подписка DreamSparkPremium), FeatureCAMDemo (бесплатная лицензия), Гемма 3D (коробочная версия 2008 года), WPSOffice (бесплатная лицензия), Open Office 2.0.3 – общественная лицензия, Google Slides (бесплатная версия), X3d-player (условно-бесплатная), FreeWRL (бесплатная лицензия), OpenVRML (бесплатная лицензия), 3dExperience for Education (учебная лицензия), Visual Studio Community (freeware), Netbeans (freeware), fxSolver (бесплатная лицензия), GeoGebra (бесплатная лицензия), SolidWorks for students (студ. лицензия), SIMULIA Research & Teaching Suites (студ. лицензия), Rockwell Arena (студ. лицензия), Fusion 360 (студенческая лицензия), GNU Octave (свободная система), Sage (GNU General Public License), Scilab (полусвободная), R (programming language) (GNU GPL), Sage (GNU GPL), Maxima (GNU GPL), Visual Prolog (студ. лицензия), Малая экспертная система 2.0 (freeware), Simintech (проприетарная), 3D Max (студ. лицензия), Eclipse (freeware), BlueJ (freeware), Elmer (freeware), CP2K (freeware).

### **7.3 Самостоятельная работа:**

Помещения для самостоятельной работы с возможностью подключения к сети «Интернет» и обеспечением доступа в электронную информационно-образовательную среду организации: читальные залы, учебные корпуса 2, 3 (Компьютерная техника с возможностью подключения к сети и обеспечением доступа в электронную информационно-образовательную среду (ЭИОС ДОННТУ) и электронно-библиотечную систему (ЭБС IPRbooks), а также возможностью индивидуального неограниченного доступа обучающихся в ЭБС и ЭИОС посредством Wi-Fi с персональных мобильных устройств. ОС- Microsoft Windows 7, Open Office 2.0.3 - общественная лицензия MPL 2.0/ Grub loader for ALT Linux - лицензия GNU LGPL v3/ Mozilla Firefox - лицензия MPL2.0, Moodle (Modular Object-Oriented Dynamic Learning Environment) - лицензия GNU GPL.