

# АННОТАЦИЯ РАБОЧЕЙ ПРОГРАММЫ ДИСЦИПЛИНЫ (МОДУЛЯ)

## Б1.В.ДВ. 13.01 Анализ данных

Автор-составитель: к.т.н..

доцент кафедры

Системного анализа и

информатики Сальников

А.Ю.

Направление подготовки: 09.03.03 «Прикладная

информатика» Направленность: «Прикладная

информатика в экономике» Квалификация

выпускника: бакалавр

Форма обучения: очно-заочная

### Цели и задачи дисциплины (модуля).

Дисциплина «Анализ данных» предназначена для формирования теоретических знаний, практических умений и навыков по применению современных методов аналитической обработки электронных массивов данных в различных сферах человеческой деятельности.

Исходя из цели, в процессе изучения дисциплины решаются **следующие задачи**:

- изучение существующих технологий подготовки данных к анализу;
- изучение основных методов поиска в данных внутренних закономерностей, взаимосвязей, тенденций; □ овладение практическими умениями и навыками реализации технологий аналитической обработки данных, формирования и проверки гипотез о их природе и структуре, варьирования применяемыми моделями;
- формирование умений и навыков применения универсальных программных пакетов и аналитических платформ для анализа данных.

### План курса

№ пп	Название темы	Основные вопросы и положения, раскрывающие содержание темы
<b>Тема 1</b>	Аффинитивный анализ.	Поиск последовательных шаблонов. Введение в аффинитивный анализ (affinity analysis). Алгоритм a priori. Иерархические ассоциативные правила.
<b>Тема 2</b>	Кластерный анализ.	Введение в кластеризацию. Классификация методов кластеризации. Алгоритм кластеризации k-means. Сети Кохонена (KCN – Kohonen network). Карты Кохонена (SOM – self organizing map). Проблемы алгоритмов кластеризации.

<b>Тема 3</b>	Классификация и регрессия. Статистические методы	Введение в классификацию и регрессию. Простая линейная регрессия. Оценка соответствия простой линейной регрессии реальным данным. Простая регрессионная модель. Множественная линейная регрессия. Модель множественной линейной регрессии. Регрессия с категориальными входными переменными. Методы отбора переменных в регрессионные модели. Ограничения применимости регрессионных моделей. Основы логистической регрессии. Интерпретация модели логистической регрессии.
<b>Тема 4</b>	Классификация и регрессия. Машинное обучение.	Введение в деревья решений. Алгоритмы построения деревьев решений. Алгоритмы ID3 и C4.5. Алгоритм CART. Упрощение деревьев решений. Введение в нейронные сети. Искусственный нейрон. Принципы построения нейронных сетей. Алгоритмы обучения нейронных сетей. Алгоритм обратного распространения ошибки.
<b>Тема 5</b>	Анализ и прогнозирование временных рядов.	Временной ряд и его компоненты. Модели прогнозирования. Прогнозирование в экономике.
<b>Тема 6</b>	Ансамбли моделей	Введение в ансамбли моделей. Бэггинг. Бустинг. Альтернативные методы построения ансамблей.

### Формы текущего контроля промежуточной аттестации

По окончании изучения дисциплины «Анализ данных» студент должен:

- знать** проблемы и направления научных изысканий в области аналитической обработки данных; тенденции разработки универсальных программных средств и аналитических платформ, предназначенных для аналитической обработки данных, построения прогнозов и аналитических сценариев; основные методы консолидации, трансформации, визуализации, оценки качества, очистки и предобработки данных; принципы построения и структурную организацию хранилищ данных; алгоритмы поиска ассоциативных правил и кластерного анализа; статистические и машинные методы классификации и регрессии; методики анализа и прогнозирования временных рядов; технологию построения ансамблей и сравнения моделей; возможности отечественных и зарубежных универсальных программных средств и аналитических платформ, применяемых для анализа данных; проблемные вопросы внедрения аналитических программных продуктов и технологий в профессиональную деятельность организаций и учреждений; уметь
- уметь** практически применять методы консолидации, трансформации, визуализации, оценки качества, очистки и предобработки данных для качественной подготовки данных к анализу; применять технологии анализа электронных массивов данных для решения конкретных практических проблем; использовать возможности отечественных и зарубежных универсальных

программных средств и аналитических платформ для аналитической обработки данных, построения прогнозов и аналитических сценариев; свободно ориентироваться на современном динамичном рынке аналитических программных продуктов.

- **иметь навыки** владения технологиями аналитической обработки электронных массивов данных в целях решения практических проблем выбранной предметной области; навыками выбора и применения отечественных и зарубежных аналитических платформ, используемых для анализа табулированных массивов электронных данных.

В результате освоения программы бакалавриата у выпускника должны быть сформированы профессиональные компетенции.

Код компетенции	Содержание компетенции	Планируемые результаты обучения по дисциплине (модулю)
ОПК -6	способность анализировать и разрабатывать организационно-технические и экономические процессы с применением методов системного анализа и математического моделирования;	на уровне знаний: математические методы анализа при принятии решения
		на уровне умений: применять методы математического анализа и алгебры при решении профессиональных задач
		на уровне навыков: навыки использования методов системного анализа и математического моделирования в профессиональной деятельности
ПК-1	Способность проводить обследование организаций, выявлять информационные потребности пользователей, формировать требования к информационной системе	на уровне знаний: теории, принципов, методологий и технологий проектирования информационных систем и содержание этапов их разработки; методов анализа предметной области информационных потребностей и формирования требований к информационной системе;
		методы и средства управления проектом по разработке информационной системы;

		на уровне умений: проводить анализ предметной области, выявлять информационные потребности и разрабатывать требования к информационной системе; проводить сравнительный анализ и выбор информационно-коммуникационных технологий для решения задачи разработки информационных систем;
		на уровне навыков: владение инструментальными средствами и стандартами разработки технологической документации и управлению проектом разработки информационных систем; навыками применения современных инструментальных средств моделирования предметной области, прикладных и информационных процессов и проектирования информационных систем.

### Объем дисциплины (модуля)

#### Объем дисциплины (модуля) «Анализ данных»

Вид учебной работы		Количество часов								
		Всего по уч. плану	Семестр							
			1	2	3	4	5	6	7	8
<b>аудиторные занятия (всего):</b>		<b>32</b>							32	
<b>в том числе</b>	лекционные занятия	<b>16</b>							16	
	практические занятия	<b>16</b>							16	
<b>самостоятельная работа:</b>		<b>112</b>							112	
<b>общая</b>	<b>часы:</b>	<b>180</b>							144	

<b>трудоемкость дисциплины:</b>	<b>зачетные единицы:</b>	<b>5</b>							<b>4</b>	
Формы итогового контроля		<b>экзамен</b>							<b>36ч.</b>	

#### **Основная литература.**

1. Гуцин А.Н. Базы данных. учебник. – М.: Директ-Медиа, 2014. – 266 с.  
[с./http://biblioclub.ru/index.php?page=book&id=222149](http://biblioclub.ru/index.php?page=book&id=222149)
2. Нестеров, С. А. Интеллектуальный анализ данных средствами MS SQL Server 2008 / С. А. Нестеров. — Москва : Интернет-Университет Информационных Технологий (ИНТУИТ), 2016. — 303 с. — ISBN 2227-8397. — Текст : электронный // Электронно-библиотечная система IPR BOOKS : [сайт]. — URL: <http://www.iprbookshop.ru/62813.html>
3. Пальмов, С. В. Интеллектуальный анализ данных : учебное пособие / С. В. Пальмов. — Самара : Поволжский государственный университет телекоммуникаций и информатики, 2017. — 127 с. — ISBN 2227-8397. — Текст : электронный // Электронно-библиотечная система IPR BOOKS : [сайт]. — URL: <http://www.iprbookshop.ru/75376.html>

#### **Дополнительная литература.**

1. Информационные системы в экономике [Электронный ресурс] / Горбенко А.О. - М. : БИНОМ, 2013, - 292 с.: ил. - ISBN 978-5-9963-2268-8.  
[8.http://www.studentlibrary.ru/book/ISBN9785996322688.html](http://www.studentlibrary.ru/book/ISBN9785996322688.html)
2. Грабер М. Введение в SQL. Пер. с англ. — М.: ЛОРИ, 2008.