МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ РОССИЙСКОЙ ФЕДЕРАЦИИ
ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ АВТОНОМНОЕ ОБРАЗОВАТЕЛЬНОЕ УЧРЕЖДЕНИЕ ВЫСШЕГО ОБРАЗОВАНИЯ
«Национальный исследовательский ядерный университет «МИФИ»
Обнинский институт атомной энергетики –
филиал федерального государственного автономного образовательного учреждения высшего образования
«Национальный исследовательский ядерный университет «МИФИ»
(ИАТЭ НИЯУ МИФИ)

ОТДЕЛЕНИЕ СОЦИАЛЬНО-ЭКОНОМИЧЕСКИХ НАУК (О)

Одобрено на заседании
Ученого совета
ИАТЭ НИЯУ МИФИ
Протокол №25.8 от 25.08.2025

МЕТОДИЧЕСКИЕ РЕКОМЕНДАЦИИ
для преподавателя
по дисциплине

АНАЛИЗ БОЛЬШИХ ДАННЫХ

название дисциплины

для студентов направления подготовки

38.03.05 Бизнес-информатика

код и название направления подготовки

образовательная программа

Цифровая трансформация компании

Форма обучения: очная

г. Обнинск 2025 г.

ВВЕДЕНИЕ

Методические рекомендации для преподавателей по дисциплине «Анализ больших данных» представляют собой комплекс рекомендаций и разъяснений, позволяющих преподавателю оптимальным образом организовать процесс обучения по данной дисциплине.

Дисциплина «Анализ больших данных» является одной из профильных дисциплин в обеспечении профессионального становления будущего бизнес-информатика.

Цель дисциплины – формирование у обучающихся представлений и профессиональной компетенции в области разработки и использования систем обработки и анализа больших массивов данных. Данная цель соотносится с целью образовательной программы, в частности с технологией разработки специализированных программных систем, отвечающих за обработку больших данных.

Задачи дисциплины:

- формирование у студентов представления о современных технологиях анализа больших данных;
- приобретение студентами знаний о технологиях подготовки, хранения, обработки и анализа больших данных;
- применение статистических и математических методов для анализа больших объемов информации;
- приобретение практических навыков работы с программой R-Studio.

Дисциплина реализуется в рамках части, формируемой участниками образовательных отношений, и относится к профессиональному модулю.

Дисциплина изучается на 3 курсе в 5 семестре.

Основными видами учебной работы по данной дисциплине являются лекции, практические занятия, самостоятельная работа обучающихся.

1 Лекции

Лекции являются одним из основных методов обучения по дисциплине «Анализ больших данных». Главной задачей каждой лекции является раскрытие сущности темы и анализ ее основных положений. Рекомендуется на первой лекции довести до внимания студентов структуру дисциплины и его разделы, а в дальнейшем указывать начало каждого раздела (модуля), суть и его задачи, а, закончив изложение, подводить итог по этому разделу, чтобы связать его со следующим.

Содержание лекций определяется рабочей программой дисциплины и представлено в таблице.

Неделя	**Наименование раздела /темы дисциплины**	**Содержание**
1-4	I **Введение в большие данные**	
1-2	Основные понятия и терминология, классификация анализа больших данных	Основные вызовы больших данных (4V). Определение термина «большие данные». Представление о работе аналитика. Основные определения, термины, задачи анализа больших данных. Вопросы безопасности. Понятие Data Mining. Когнитивный анализ данных. Обзор источников информации для Big Data (открытые источники информации: статистические сборники, опубликованные отчеты и результаты исследований; доступ к закрытой информации). Методики сбора данных.
3-4	Прикладные инструменты для работы с Big Data	Инструменты для обработки больших данных. Знакомство с языками и прикладными пакетами для обработки больших данных.
5-9	**II Статистические методы и программные средства для анализа больших данных**	
5-7	Аналитика больших данных. Методы анализа больших данных	Основные понятия математической статистики. данных: дескриптивная статистика, параметрические, непараметрические, номинальные методы (корреляционный, регрессионный, дисперсионный анализ, кластерный, дискриминантный, факторный анализ).
7-8	Информационное и программное обеспечение для работы по анализу больших наборов данных	Состав, виды и классификация программного обеспечения, используемого для анализа больших наборов данных. Функции, выполняемые программным обеспечением по анализу больших наборов данных. Программы, используемые в технологиях, используемых для обработки больших наборов данных.
8-9	Использование таблич-	Основные приемы работы в табличном редакторе:

Неделя	Наименование раздела /темы дисциплины	Содержание
	ного процессора MS Excel для анализа больших наборов данных	создание документа, ввод данных, редактирование, форматирование, больших наборов данных. Типы и форматы данных.
9-16	**III Использование баз данных для анализа больших наборов данных**	
9-10	Базы данных, общие сведения о РБД. Схема реляционной базы для больших наборов данных	Понятие и назначение баз данных. Система баз данных. Предметная область автоматизированных баз данных. Типы баз, используемые для работы с большими наборами данных: реляционные, иерархические, сетевые. Схема реляционной базы данных.
10-12	Система управления для работы с большими наборами данных	Понятие системы управления большими наборами данных. Виды СУБД, объекты, интерфейс. Понятие автоматизированной базы данных (АБД). Система баз больших наборов данных. Модели больших наборов данных: иерархическая, сетевая, реляционная. Виды систем управления большими наборами данных: реляционные, дедуктивные, экспертные, расширяемые, объектно-ориентированные, семантические, универсальные реляционные.
13-14	Программные инструменты для работы с большими наборами данных	Improvado - инструмент маркетинговой аналитики. Microsoft Power Query – для работы с импортируемыми большими наборами данных. Microsoft Power BI - инструмент бизнес-аналитики. Tableau Prep - известное программное обеспечение для подготовки данных. IBM SPSS Statistics для подготовки больших наборов данных. Qlik - инструмент бизнес-аналитики для БНД. DataWatch программное обеспечение для управления данными. Quest - средство для прогнозирования бизнес-данных. DataMeer - платформа на основе Saas для анализа больших наборов данных. Microstrategy - инструмент для анализа данных на рабочем столе, позволяет исследовать и анализировать данные. Paxata - самообслуживание и адаптивный инструмент для подготовки данных. Oracle - инструмент для самостоятельной подготовки данных.
15-16	OLAP и многомерные базы данных	OLAP - хранилище данных, предметно-ориентированное, привязанное ко времени и неизменяемое собрание данных для поддержки процесса принятия управляющих решений. OLAP -

Не-деля	Наименование раздела /темы дисциплины	Содержание
		инструмент анализа БНД. OLAP - совокупность средств многомерного анализа данных. Определение OLAP - Online Analytical Processing, - оперативный анализ данных. OLAP - многомерное представление табличных данных. Операции с многомерным кубом -сечения, проекции, линейные таблицы. Иерархии и уровни OLAP. Архитектура OLAP-приложений. Технические проблемы многомерного хранения данных. OLAP-серверы.

Для эффективного проведения лекционного занятия рекомендуется соблюдать последовательность ее основных этапов:

1) формулировку темы лекции;

2) указание основных изучаемых разделов или вопросов и предполагаемых затрат времени на их изложение;

3) изложение вводной части;

4) изложение основной части лекции;

5) краткие выводы по каждому из вопросов;

6) заключение;

7) рекомендации литературных источников по излагаемым вопросам.

Дадим краткую характеристику каждого из лекционных этапов.

Начальный этап каждого лекционного занятия – оглашение основной темы лекции с краткой аннотацией предлагаемых для изучения вопросов. Преподаватель должен сообщить о примерном плане проведения лекции и предполагаемом распределении бюджета времени. Если очередное занятие является продолжением предыдущей лекции, необходимо кратко сформулировать полученные ранее результаты, необходимые для понимания и усвоения изучаемых вопросов.

Во вводной части достаточно кратко характеризуется место и значение данной темы в курсе, дается обзор важнейших источников и формулируются основные вопросы или задачи, решение которых необходимо для создания стройной системы знаний в данной предметной области. В этой части лекции демонстрируются основные педагогические методы, которые будут использоваться при изложении материала и устанавливается контакт с аудиторией.

Основная часть лекции имеет своей целью раскрытие содержания основных вопросов или разделов и определяется логической структурой плана лекции. При этом используются основные педагогические способы изложения материала: описание-характеристика, повествование, объяснение и др. Преподаватель должен также умело использовать эффективные методические приемы изложения материала – анализ, обобщение, индукцию, дедукцию, противопоставления, сравнения и т.д., обеспечивающие достаточно вы-

сокий уровень качества учебного процесса.

В заключительной части лекции проводят обобщение наиболее важных и существенных вопросов, делаются выводы, формулируются задачи для самостоятельной работы слушателей и указывается рекомендуемая литература. Оставшееся время используют для ответов на вопросы, задаваемые слушателями, и для возможной дискуссии о содержании лекции.

Содержание лекционного материала должно строго соответствовать содержательной части утвержденной рабочей учебной программы дисциплины и соответствовать основным дидактическим принципам, которые обеспечивают соответствие излагаемого материала научно-методическим основам экономической деятельности. Основными из них являются целостность, научность, доступность, систематичность и наглядность.

Целостность лекции обеспечивается созданием единой ее структуры, основанной на взаимосвязи задач занятия и содержания материала, предназначенного для усвоения студентами.

Научность лекции предполагает соответствие материала основным положениям современной науки, абсолютное преобладание объективного фактора и доказательность выдвигаемых положений. Для научно обоснованной лекции характерны ясность, логичность, аргументированность, точность и сжатость.

Принцип доступности лекции предполагает, что содержание учебного материала должно быть понятным, а объем этого материала посильным для всех студентов. Это означает, что степень сложности лекционного материала должна соответствовать уровню развития и имеющемуся запасу знаний и представлений студентов.

Систематичность лекционного материала определяется взаимосвязью изучаемого материала с ранее изученным, постепенным повышением сложности рассматриваемых вопросов, взаимосвязью частей изучаемого материала, обобщением изученного материала, стройностью изложения материала по содержанию и внешней форме его подачи, рубрикацией курса, темы, вопроса и единообразием структуры построения материала.

Принцип наглядности содержания лекции требует использования при чтении лекции визуальных носителей информации в виде презентаций, поскольку основной поток информации в учебном процессе воспринимается обучаемым зрительно. Демонстрационный материал во всех случаях должен играть подчиненную роль и не подменять содержания лекции. В каждый момент лекции необходимо демонстрировать только тот наглядный материал, который иллюстрирует излагаемые положения.

При проведении лекционных занятий по дисциплине используются следующие виды лекций: информационные, проблемные, лекции-визуализации, лекции с опорным конспектированием.

Основным признаком информационной лекции является простой способ передачи готовых знаний учащимся через монологическую форму общения.

В отличие от информационной лекции, в проблемной лекции, лекции-

визуализации происходит активное освоение содержание обучения с включением механизмов творческого осмысления. В этом процессе учащиеся проявляют собственную активность в контексте диалогического взаимодействия и общения в ходе лекции.

Лекции проблемного характера отличает то, что процесс познания студентов приближается к поисковой, исследовательской деятельности. При этом обеспечивается достижение трех основных целей: усвоение студентами теоретических знаний, развитие теоретического мышления и формирование познавательного интереса к содержанию учебного предмета и профессиональной мотивации будущего специалиста. На такой лекции новое знание вводится через проблемность вопроса, задачи или ситуации. При этом процесс познания студентов в сотрудничестве и диалоге с преподавателем приближается к исследовательской деятельности. Содержание проблемы раскрывается путем организации поиска ее решения или суммирования и анализа традиционных и современных точек зрения.

Другая форма лекции – лекция-визуализация – является результатом поиска новых возможностей реализации известного в дидактике принципа наглядности, содержание которого меняется под влиянием данных психолого-педагогической науки, форм и методов активного обучения. Лекция-визуализация представляет собой визуальную форму подачи лекционного материала средствами компьютерной техники или аудио- и видеотехники (видео-лекция). Чтение такой лекции сводится к развернутому или краткому комментированию просматриваемых визуальных материалов.

Лучше использовать разные виды наглядности – натуральной, изобразительной, символической. При переходе от текста к зрительной форме или от одного вида наглядности к другому теряется некоторое количество информации. Однако это может быть преимуществом, поскольку позволяет сконцентрировать внимание на наиболее важных аспектах и особенностях содержания лекции, способствовать его пониманию и усвоению.

2 Практические занятия (семинары)

Практические занятия являются важной частью учебного процесса в вузе. Они проводятся с целью закрепления лекционного материала, овладения понятийным аппаратом предмета, методами и приёмами исследования, изучаемыми в рамках учебной дисциплины. Главной целью такого рода занятий является научиться применению теоретических знаний на практике.

Содержание практических занятий по дисциплине «Анализ больших данных» представлено в таблице.

Неделя	**Наименование раздела /темы дисциплины**	**Содержание**
1-4	I **Введение в большие данные**	
1-2	Основные понятия и терминология, классификация анализа боль-	Структура и задачи анализа больших наборов данных. Значение анализа больших наборов данных для инженерной и экономической работы. Формы

Неделя	Наименование раздела /темы дисциплины	Содержание
	ших данных	представления информации по результатам анализа больших наборов данных. Понятие большого набора данных. Роль анализа больших наборов данных в развитии современных технологий управления в технической и экономической сферах деятельности. Функции управления с использованием анализа больших наборов данных. Место и роль анализа больших наборов данных в системах обработки данных, и системах управления. Информационные процессы, опосредованные применением методов анализа больших наборов данных. Классификация информационных технологий. Технология и методы обработки информации. Информационные ресурсы общества и предприятия. Виды и формы информационного обеспечения. Классификация информационных технологий по осуществлению анализа больших наборов данных. Системы обработки больших наборов данных и автоматизированные информационные системы. Методология и технология разработки информационных технологий анализа больших наборов данных. Принципы построения автоматизированных систем управления с применением анализа больших наборов данных
3-4	Прикладные инструменты для работы с Big Data	Поиск источников информации в сети Интернет: открытые и закрытые источники данных. Портал открытых данных РФ. Сохранение данных в программе Excel. Преобразование и первичная обработка данных. Рассмотрение общей концепции и синтаксиса языка R (примеры). Базовое представление о Map Reduce и Hadoop.
5-9	**Статистические методы и программные средства для анализа больших данных**	
5-7	Аналитика больших данных. Методы анализа больших данных	Аналитика больших данных. Процесс аналитики. Стандарты жизненного цикла Big Data: CRISP-DM. Принципы и инструменты аналитики. Задачи и компетенции аналитиков Big Data. Big Data как рынок. Когнитивный анализ данных Введение в Data Mining – понятие, структура, составляющие и сопутствующие науки. • Задачи Data Mining и способы их решения.

Неделя	Наименование раздела /темы дисциплины	Содержание
		Классификация методов DM. • Области применения DM. • Классы систем DM. • Процесс накопления и анализа данных: Азбука когнитивного анализа. Аналитика больших данных. Математическая статистика Основные понятия статистики и дескриптивный анализ • Шкалы измерений. • Генеральная совокупность и выборка. • Нормальное распределение. Уровень статистической достоверности. • Свойства описательных статистик (Дескриптивный анализ) • Визуальное представление данных • Меры изменчивости Методы DATA MINING • Данные & знания • Типовые задачи Data Mining • Обучаемые и необучаемые задачи • Жизненный цикл проекта DM • Математический аппарат DM • Стандарты DM Методы анализа на графах • Случайные графы, безмасштабные графы, социальные сети – сети тесного мира. • Закономерности, методы кластеризации на графах. Прикладные инструменты анализа данных. Корреляция. • Готовые комплексные решения: Weka,RapidMiner, Knime, Orange IBM SPSS Modeler (в прошлом Clementine) • Инструменты визуализации: Tableau, • Фреймворки на JS, D3 • Корреляция. Понятие корреляции. Значимость коэффициента корреляции. Виды связи между переменными Многомерная регрессия и кластерный анализ
7-8	Информационное и программное обеспечение для работы по анализу больших наборов дан-	Представление исходных данных в программе R-Studio (векторы, массивы, матрицы, списки, таблицы). Статистическая обработка данных в программах Excel и R-Studio: подсчет описательных

Неделя	Наименование раздела /темы дисциплины	Содержание
	ных	статистик, графическое представление данных. Группировка данных, обнаружение значимых корреляций, зависимостей и тенденций в результате анализа имеющейся информации, выявления отношений между данными различного типа. Применение различных методов выделения, извлечения и группировки данных, которые позволяют выявить систематизированные структуры данных и вывести из них правила для принятия решений и прогнозирования их последствий (регрессионный, дисперсионный, кластерный, дискриминантный, факторный анализы).
8-9	Использование табличного процессора MS Excel для анализа больших наборов данных	Организация расчетов в электронных таблицах с помощью пользовательских формул и функций, выполнение расчетов с использованием условий для обработки и анализа больших наборов данных. Графическое представление - построение диаграмм, трех мерных графических систем для визуализации результатов анализа больших наборов данных в MS Excel.
9-16	**III Использование баз данных для анализа больших наборов данных**	
9-10	Базы данных, общие сведения о РБД. Схема реляционной базы для больших наборов данных	Типы баз, используемые для работы с большими наборами данных: реляционные, иерархические, сетевые. Схема реляционной базы данных.
10-12	Система управления для работы с большими наборами данных	Виды систем управления большими наборами данных: реляционные, дедуктивные, экспертные, расширяемые, объектно-ориентированные, семантические, универсальные реляционные.
13-14	Программные инструменты для работы с большими наборами данных	Возможности графического представления информации в программе R-Studio: графические функции отображения одномерных и многомерных данных, графический вывод с использованием графических параметров.
15-16	OLAP и многомерные базы данных	MOLAP (Multidimensional OLAP) - и детальные данные, и агрегаты хранятся в многомерной БД. ROLAP (Relational OLAP) - детальные данные остаются в реляционной БД. HOLAP (Hybrid OLAP) - детальные данные остаются в реляционной БД, а агрегаты хранятся в многомерной БД.

Практические занятия — метод обучения, обеспечивающий связь теории и практики, содействующий выработке у студентов умений и навыков применения знаний, полученных на лекции и в ходе самостоятельной работы.

Практические занятия представляют собой занятия по решению различных прикладных задач, образцы которых были даны на лекциях. В итоге у каждого обучающегося должен быть выработан определенный профессиональный подход к решению каждой задачи.

Практические занятия по курсу могут проводиться в различных формах. Рекомендуются активные формы занятий, такие как дискуссия, деловая игра, тренинг. Преподавателю важно давать задания в соответствии с возможностями обучающихся на данной стадии обучения, чтобы обеспечить им уверенность в своих силах.

Практическое занятие должно опираться на известный теоретический материал, который изложен или на который дана соответствующая ссылка в лекции.

Практическое занятие должно быть нацеленным на формирование определенных умений и закрепления определенных навыков, поэтому цель занятия должна быть заранее известна и понятна преподавателю и обучающимся. Лучше иметь сформулированные в письменном виде цель, задачи, содержание и последовательность занятия, ожидаемый результат.

Одно или несколько занятий желательно провести в компьютерном классе с доступом в глобальную сеть. Целью такого занятия может быть помощь в организации выполнения заданий самостоятельной работы, которые ориентированы на поиск информации в Интернет.

Обучающиеся должны всегда видеть ведущую идею курса и ее связь с практикой. Это придает учебной работе актуальность, утверждает необходимость овладения опытом профессиональной деятельности, связывает её с практикой жизни. В таких условиях задача преподавателя состоит в том, чтобы больше показывать обучающимся практическую значимость ведущих научных идей и принципиальных научных концепций и положений.

Примерные цели практических занятий:

1) помочь обучающимся систематизировать, закрепить и углубить знания теоретического характера;

2) научить студентов приемам решения практических задач, способствовать овладению навыками;

3) формировать умение учиться самостоятельно, т.е. овладевать методами, способами и приемами самообучения, саморазвития и самоконтроля.

Содержание практических занятий и методика их проведения должны обеспечивать развитие творческой активности личности. Они развивают научное мышление и речь обучающихся, позволяют проверить их знания, выступают важным средством оперативной обратной связи. Поэтому практические занятия должны выполнять не только познавательную и воспитательную функции, но и способствовать росту их креативности.

Практические занятия проводятся в двух формах: выполняются индивидуально и в групповой форме. При разработке практических занятий

должна быть учтена форма их проведения и возможности интерактивного обучения. Групповая форма предполагает обсуждение слушателями конкретной проблемы в группе по каждому этапу изучения дисциплины.

Прежде чем приступить к изучению темы, необходимо прокомментировать основные вопросы плана лекции. Такой подход преподавателя помогает студентам быстро находить нужный материал к каждому из вопросов, не задерживаясь на второстепенном.

Важно развивать у студентов умение сопоставлять источники, продумывать изучаемый материал.

Большое значение имеет совершенствование навыков конспектирования у студентов.

Преподаватель может рекомендовать студентам следующие основные формы записи: план (простой и развернутый), выписки, тезисы.

Преподаватель может предложить студентам подумать над постановкой таких вопросов по теме лекции, которые вызовут интерес своей неоднозначностью, противоречивостью, разделят участников семинара на оппонирующие группы. А это как раз то, что нужно для дискуссии, для активизации, для поиска студентами истины, которая, как известно, рождается в споре. Само собой разумеется, что и в арсенале преподавателя должны быть заготовлены вопросы для создания проблемных ситуаций, если они не будут созданы выступлениями студентов.

В процессе подготовки, прорабатывая предложенные вопросы, студент определяет для себя один-два из них (можно, конечно и больше), в которых он чувствует себя наиболее уверенно и в качестве консультанта или оппонента намерен задать тон на семинаре.

Практические занятия предполагают не просто обсуждение студентами учебного материала, а выполнение ими определенных практических заданий. Систему таких заданий часто называют практикумом.

Функции практических занятий:

1) закрепление теоретических знаний на практике;

2) усвоение умений исследовательской работы;

3) усвоение умений практической работы;

4) применение теоретических знаний для решения практических задач;

5) самопознание;

6) саморазвитие.

Соответствующие задачи ставятся преподавателем при планировании каждой работы. Те или иные функции могут выдвигаться на первый план в зависимости от того, в рамках каких образовательных программ проводятся занятия.

Практическое занятие (семинар) – один из наиболее сложных и в то же время плодотворных видов (форм) вузовского обучения и воспитания. В условиях высшей школы семинар – один из видов практических занятий, проводимых под руководством преподавателя.

Целью практических занятий (семинаров) является:

1) закрепление методов анализа;

2) проверка уровня понимания студентами вопросов, рассмотренных на лекциях и по учебной литературе, степени и качества усвоения материала студентами;

3) обучение навыкам решения поставленных задач и умение подобрать необходимый метод решения;

4) восполнение пробелов в пройденной теоретической части курса и оказание помощи в его усвоении.

При условии соблюдения требований методики их проведения семинары выполняют многогранную роль:

1) стимулируют регулярное изучение студентами первоисточников и другой литературы, а также внимательное отношение к лекционному курсу;

2) закрепляют знания, полученные студентами при прослушивании лекции и самостоятельной работе над литературой;

3) расширяют круг знаний благодаря выступлениям товарищей и преподавателя на занятии;

4)позволяют студентам проверить правильность ранее полученных знаний, вычленить в них наиболее важное, существенное;

5) способствуют превращению знаний в твердые личные убеждения, рассеивают сомнения, которые могли возникнуть на лекциях и при изучении литературы, что особенно хорошо достигается в результате столкновения мнений, дискуссии;

6) прививают навыки самостоятельного мышления, устного выступления по теоретическим вопросам, оттачивают мысль, приучают студентов свободно оперировать терминологией, экономическими понятиями и категориями;

7) предоставляют возможность преподавателю систематически контролировать уровень самостоятельной работы студентов над первоисточниками, другим учебным материалом, степень их внимательности на лекциях;

8) позволяют изучить мнения, интересы студентов, служат средством контроля преподавателя не только за работой студентов, но и за своей собственной как лектора и руководителя семинара, консультанта и т. д.

При разработке методики семинарских занятий важное место занимает вопрос о взаимосвязи между семинаром и лекцией, семинаром и самостоятельной работой студентов, о характере и способах такой взаимосвязи. Семинар не должен повторять лекцию, и, вместе с тем, его руководителю необходимо сохранить связь принципиальных положений лекции с содержанием семинарского занятия.

При подготовке к семинару студентами осуществляется весьма объемная работа по углубленному проникновению в суть вынесенной для обсуждения проблемы. В ходе семинара студент учится публично выступать, видеть реакцию слушателей, логично, ясно, четко, грамотным литературным языком излагать свои мысли, проводить доводы, формулировать аргументы в защиту своей позиции.

На семинаре каждый студент имеет возможность критически оценить свои знания, сравнить со знаниями и умениями их излагать других студентов,

сделать выводы о необходимости более углубленной и ответственной работы над обсуждаемыми проблемами.

В ходе семинара каждый студент опирается на свои конспекты, сделанные на лекции, собственные выписки из учебников, первоисточников, статей, другой специальной литературы, на словарь по данной теме. Семинар стимулирует стремление к совершенствованию конспекта, желание сделать его более информативным, качественным.

От семинара к семинару, на всех его этапах и их коррекции студент поднимается на более высокую ступеньку собственной зрелости, своего мнения более эффективно работать над проблемами, непосредственно относящимися к его будущей профессии.

Семинар – эффективная форма закрепления полученных по обсуждаемой проблеме знаний, видения этой проблемы в целом, осознания ее соотнесенности с другими темами в рамках целостной концепции.

С точки зрения методики проведения семинар представляет собой комбинированную, интегративную форму учебного занятия. Он предполагает возможность использования рефератов, фрагментов первоисточников, устных и письменных понятийных диктантов, тестов, заданий типа «закончите предложение» и др.

Для стимулирования самостоятельного мышления на занятиях используются различные активные методы обучения: проблемные ситуации, игры, педагогические задачи, тесты, интерактивный опрос.

В практике семинарских занятий используется следующий ряд форм: развернутая беседа, семинар-диспут, комментированное чтение, упражнения на самостоятельность мышления, письменная (контрольная) работа, семинар-коллоквиум и другие.

1. Развернутая беседа – наиболее распространенная форма семинарских занятий. Она предполагает подготовку всех студентов по каждому вопросу плана занятия с единым для всех перечнем рекомендуемой обязательной и дополнительной литературы; выступления студентов (по их желанию или по вызову преподавателя) и их обсуждение; вступление и заключение преподавателя. Развернутая беседа позволяет вовлечь в обсуждение изучаемой проблематики наибольшее число студентов, разумеется, при использовании всех средств их активизации: постановки хорошо продуманных, четко сформулированных дополнительных вопросов к выступающему и всей группе, умелой концентрации внимания студентов на сильных и слабых сторонах выступлений студентов, своевременном акцентировании внимания и интереса студентов на новых моментах, вскрывающихся в процессе работы и т. д.

Развернутая беседа не исключает, а предполагает и заранее запланированные выступления отдельных студентов по некоторым дополнительным вопросам. Но подобные сообщения выступают здесь в качестве не основы для обсуждения, а лишь дополнения к уже состоявшимся выступлениям.

2. Семинар-диспут имеет ряд достоинств. Кроме других задач, обычно реализуемых на семинаре, эта форма наиболее удобна для выработки у студентов навыков полемиста. Диспут может быть и самостоятельной формой

семинара, и элементом других форм практических занятий. В первом случае наиболее интересно проходят такие занятия при объединении двух или нескольких семинарских групп, когда с докладами выступают студенты одной группы, а оппонентами – другой, о чем договариваются заранее. Вопросы, выносимые на подобные семинары, должны всегда иметь теоретическую и практическую значимость.

Диспут как элемент обычного семинара может быть вызван преподавателем в ходе занятия или же заранее планируется им. Полемика возникает подчас и стихийно. В ходе полемики студенты формируют у себя находчивость, быстроту мыслительной реакции и, главное, отстаиваемое в споре мировоззрение складывается у них как глубоко личное.

3. Комментированное чтение первоисточников на семинаре преследует цель содействовать более осмысленной и тщательной работе студентов над рекомендуемой специальной литературой. Чаще всего оно составляет лишь элемент обычного семинара в виде развернутой беседы и длится всего 15-20 минут. Комментированное чтение позволяет приучать студентов лучше разбираться в специальных источниках. Комментирование может быть выделено в качестве самостоятельного пункта плана семинара.

4. Упражнения на самостоятельность мышления обычно входят в качестве одного из элементов семинарского занятия. Преподаватель подбирает задания, практические задачи, мини-кейсы, выполнение и решение которых требует от студентов самостоятельной мыслительной активности, проявление способности применять полученные знания в конкретной практико-ориентированной ситуации. Решение задач на самостоятельность мышления содействует формированию у студентов способности более глубоко вникать в профессиональные проблемы.

5. Контрольные (письменные) работы / тесты также практикуются на семинарах. На них может быть отведено от 15 минут до целой пары. Тема работы может быть сообщена студентам заранее, а иногда и без предупреждения по одному из пунктов плана текущего семинара. Такая работа носит характер фронтальной проверки знаний всех студентов по определенному разделу курса. Содержание работ анализируется преподавателем на очередном занятии, что вызывает всегда обостренный интерес студентов и активизирует их последующую подготовку к семинарским занятиям. Если на контрольную работу отводится 15-45 минут, то после ее написания работа семинара продолжается обычным порядком. В течение семинарского курса целесообразно провести несколько контрольных работ различных типов.

6. Коллоквиумы-собеседования преподавателя со студентами проводятся в конце изучаемого курса с целью выяснения знаний по обобщенным темам дисциплины, их углубленного изучения.

В целях эффективности семинарских занятий необходима обстоятельная подготовка к их проведению как со стороны преподавателей, так и обучающихся. Преподаватель в начале семестра (учебного года) должен обеспечить обучающихся методическими материалами для своевременной подготовки их к активным формам занятий, в том числе и к семинарам. Во время

лекций, связанных с темой семинарского занятия, следует обратить внимание обучающихся на то, что необходимо дополнительно изучить при подготовке к семинару (новые официальные документы, статьи в периодических журналах, вновь вышедшие монографии и т. д.).

Планы семинарских занятий, их тематика, рекомендуемая литература, цель и задачи ее изучения сообщаются преподавателем на вводных занятиях или в методических указаниях по данной дисциплине.

Начиная подготовку к семинарскому занятию, необходимо, прежде всего, указать студентам страницы в конспекте лекций, разделы учебников и учебных пособий, чтобы они получили общее представление о месте и значении темы в изучаемом курсе. Затем следует рекомендовать им поработать с дополнительной литературой, сделать записи по рекомендованным источникам.

Записи имеют первостепенное значение для самостоятельной работы студентов. Они помогают понять построение изучаемой книги, выделить основные положения, проследить их логику и тем самым проникнуть в творческую лабораторию автора.

Ведение записей способствует превращению чтения в активный процесс, мобилизует, наряду со зрительной, и моторную память. Следует помнить: у студента, систематически ведущего записи, создается свой индивидуальный фонд подсобных материалов для быстрого повторения прочитанного, для мобилизации накопленных знаний. Особенно важны и полезны записи тогда, когда в них находят отражение мысли, возникшие при самостоятельной работе.

Нередко среди начинающих преподавателей можно встретить людей, полагающих, будто записи – дело простое, требующее, в основном, усилий рук, а не головы. Это сугубо ошибочное представление. Полноценные записи отражают не только содержание прочитанного, но и результат мыслительной деятельности студента.

Важно развивать у студентов умение сопоставлять источники, продумывать изучаемый материал. Поэтому написание конспектов по рассматриваемым вопросам является обязательным элементом подготовки студентов к аудиторным занятиям.

Желательно, чтобы на занятии студент излагал материал свободно. Прикованность к конспекту объясняется обычно следующими причинами:

а) плохо продумана структура изложения, вопрос не осмыслен во всей его полноте, студент боится потерять нить мыслей, нарушить логическую последовательность высказываемых положений, скомкать выступление;

б) недостаточно развита культура устной речи, опасение говорить «коряво» и неубедительно;

в) материал списан из учебных пособий механически, без достаточного осмысливания его;

г) как исключение, материал списан у товарища или же используется чужой конспект.

Любая из перечисленных причин, за исключением второй, говорит о поверхностной или же просто недобросовестной подготовке студента к занятию.

Важно научить студентов во время выступления поддерживать постоянную – связь с аудиторией, быстро, не теряясь, реагировать на реплики, вопросы, замечания, что дается обычно не сразу, требует постоянной работы над собой. Выступающий обращается к аудитории, а не к преподавателю, как школьник на уроке. Контакт со слушателями – товарищами по группе – помогает студенту лучше выразить свою мысль, реакция аудитории позволит ему почувствовать сильные и слабые стороны своего выступления. Без «обратной связи» со слушателями выступление студента – это разговор с самим собой, обращение в пустоту; ему одиноко и неуютно за кафедрой, Поэтому на семинаре неплохо ввести в традицию анализ не только содержания выступлений, но и их формы – речи, дикции, поведения за кафедрой, характера общения с аудиторией.

Перечень образовательных технологий, используемых при осуществлении образовательного процесса по дисциплине «Анализ больших данных», представлен в таблице.

№ п/п	**Наименование темы дисциплины**	**Вид занятий (лекция, семинары, практические занятия)**	**Количество ак. час.**	**Наименование активных и интерактивных форм проведения занятий**
I	**Введение в большие данные**			
1	Основные понятия и терминология, классификация анализа больших данных	лекция	1	Проблемная лекция с обсуждением содержания основных этапов развития электронной коммерции и ее технологий
2	Прикладные инструменты для работы с Big Data	лекция / практическое занятие	3	лекция-беседа, диспут, семинар с разбором конкретных ситуаций, тесты, решение контрольных работ
II	**Статистические методы и программные средства для анализа больших данных**			
3	Аналитика больших данных. Методы анализа больших данных	лекция / практическое занятие	2	лекция-беседа, диспут, тесты, решение контрольных работ
4	Информационное и программное обеспечение для работы по анализу больших наборов данных	лекция / практическое занятие	3	Проблемная лекция и групповое обсуждение, семинар с разбором конкретных ситуаций, тесты, решение контрольных работ
5	Использование табличного процессора MS Excel для анализа	лекция / практическое занятие	2	Лекция-беседа и групповое обсуждение, семинар с разбором конкретных си-

№ п/п	Наименование темы дисциплины	Вид занятий (лекция, семинары, практические занятия)	Количество ак. час.	Наименование активных и интерактивных форм проведения занятий
	больших наборов данных			туаций, тесты, решение контрольных работ
III	**Использование баз данных для анализа больших наборов данных**			
6	Базы данных, общие сведения о РБД. Схема реляционной базы для больших наборов данных	лекция / практическое занятие	2	Практико-ориентированная лекция-дискуссия, обсуждение содержания этапов планирования электронной коммерции и проектирования электронных организаций
7	Система управления для работы с большими наборами данных	лекция / практическое занятие	2	Лекция-беседа и групповое обсуждение, семинар с разбором конкретных ситуаций, тесты, решение контрольных работ
8	Программные инструменты для работы с большими наборами данных	лекция / практическое занятие	3	Лекция-беседа и групповое обсуждение, семинар с разбором конкретных ситуаций, тесты, решение контрольных работ
9	OLAP и многомерные базы данных	лекция / практическое занятие	3	Лекция-беседа, семинар с разбором конкретных ситуаций, тесты, решение контрольных работ

В целях реализации индивидуального подхода к обучению студентов, осуществляющих учебный процесс по собственной траектории в рамках индивидуального рабочего плана, изучение данной дисциплины базируется на следующих возможностях: обеспечение внеаудиторной работы со студентами в том числе в электронной образовательной среде с использованием соответствующего программного оборудования, дистанционных форм обучения, возможностей интернет-ресурсов, индивидуальных консультаций и т.д.

3 Оценочные средства по дисциплине

Оценочные средства по дисциплине обеспечивают проверку освоения планируемых результатов обучения посредством мероприятий текущей и промежуточной аттестации.

3.1 Экзамен

а) типовые вопросы:

1. Понятие Большие данные. Роль цифровой информации в 21 веке.
2. Проблемы анализа и обработки большого объема данных.
3. Базовые принципы обработки больших данных.
4. Определение модели. Свойства модели.
5. Аналитический подход к моделированию.
6. Информационный подход к моделированию.
7. Лица, участвующие в информационном моделировании.
8. Общая схема анализа.
9. Определение тиражирования знаний. Процесс построения модели.
10. Технологии обработки больших данных: NoSQL,
11. Технологии обработки больших данных: MapReduce,
12. Технологии обработки больших данных: Hadoop, R.
13. Методика извлечения знаний Knowledge Discovery in Databases (KDD). Этапы
KDD.
14. Data Mining. Постановка основных задач.
15. Машинное обучение.
15. Бизнес-решения с помощью алгоритмов Data Mining.
16. Классификация ПО в области Data Mining и KDD.
17. Типовая схема системы на базе аналитической платформы.
18. Понятие ассоциативного правила и транзакции.
19. Определение поддержки и достоверности.
20. Определение значимости и полезности ассоциативных правил, показатели их характеризующие.
21. Формальная постановка задачи кластеризации.
22. Цели кластеризации.
23. Основные шаги алгоритма k-means. Условие остановки алгоритма k-means.
Преимущества и недостатки алгоритма k-means.
24. Кластеризация с помощью самоорганизующейся карты Кохонена
25. Этапы проведения классификации.
26. Обзор методов классификации и регрессии.
27. Задачи линейной и логистической регрессии.
28. Определение дерева решений. Структура дерева решений. Выбор атрибута разбиения в узле.

б) критерии оценивания компетенций (результатов):

Экзамен проводится устно и включает в себя ответ на три теоретических вопроса из различных разделов курса.

Освоение дисциплины оценивается по стобалльной системе, используемой в ИАТЭ НИЯУ МИФИ.

Допуск к экзамену по дисциплине осуществляется при количестве набранных в течение семестра баллов более 36.

За семестр студент может набрать от 36 до 60 баллов.

Минимальный балл за ответ на экзамене – 24, максимальный балл – 40.

Максимальная суммарная оценка за экзамен составляет 40 баллов с учетом того, что максимальная оценка работы в семестре по контрольным точкам составляет 60 баллов.

Ответ студента на экзамене оценивается по следующим критериям:

- правильность, полнота и логичность построения ответа;
- умение оперировать специальными терминами;
- использование в ответе дополнительного материала;
- умение иллюстрировать теоретические положения практическим материалом;
- владение актуальными статистическими данными по теме вопроса

Баллы	**Критерии оценки**
37-40	Студент должен: - продемонстрировать глубокое и прочное усвоение знаний программного материала; - исчерпывающе, последовательно, грамотно и логически стройно изложить теоретический материал; - правильно формулировать определения; - продемонстрировать умения самостоятельной работы с литературой; - уметь сделать выводы по излагаемому материалу
29-36	Студент должен: - продемонстрировать достаточно полное знание программного материала; - продемонстрировать знание основных теоретических понятий; - достаточно последовательно, грамотно и логически стройно излагать материал; - продемонстрировать умение ориентироваться в литературе; - уметь сделать достаточно обоснованные выводы по излагаемому материалу
25-28	Студент должен: продемонстрировать общее знание изучаемого материала; - показать общее владение понятийным аппаратом дисциплины; уметь строить ответ в соответствии со структурой излагаемого вопроса; - знать основную рекомендуемую программой учебную литературу
менее 24	Студент демонстрирует: - незнание значительной части программного материала; - не владение понятийным аппаратом дисциплины;

	- существенные ошибки при изложении учебного материала; - неумение строить ответ в соответствии со структурой излагаемого вопроса; - неумение делать выводы по излагаемому материалу

3.2 Устный опрос

Оценочные средства представлены тематикой и вопросами, разработанными для обсуждения на семинарских занятиях.

Тема 1.

Вопросы:

1. Структура и задачи анализа больших наборов данных.
2. Значение анализа больших наборов данных для инженерной и экономической работы.
3. Формы представления информации по результатам анализа больших наборов данных. Понятие большого набора данных.
4. Роль анализа больших наборов данных в развитии современных технологий управления в технической и экономической сферах деятельности.
5. Функции управления с использованием анализа больших наборов данных

Тема 2.

Вопросы:

1. Поиск источников информации в сети Интернет: открытые и закрытые источники данных. Портал открытых данных РФ.
2. Сохранение данных в программе Excel. Преобразование и первичная обработка данных.
3. Рассмотрение общей концепции и синтаксиса языка R (примеры).
4. Базовое представление о Map Reduce и Hadoop.

Тема 3.

Вопросы:

1. Аналитика больших данных. Процесс аналитики.
2. Стандарты жизненного цикла Big Data: CRISP-DM.
3. Принципы и инструменты аналитики.
4. Задачи и компетенции аналитиков Big Data.
5. Big Data как рынок.

б) критерии оценивания компетенций (результатов):

Устный опрос проходит в форме развернутой беседы – творческой дискуссии, основанной на подготовке всей группы по объявленной заранее теме при максимальном участии в обсуждении студентов группы. Как правило, один студент раскрывает один вопрос темы, давая наиболее полный ответ. Остальные делают дополнения, высказывают различные суждения и аргументацию, могут задавать вопросы друг другу и преподавателю. Преподава-

тель направляет ход дискуссии, обращая внимание на существующие научные и практические проблемы обсуждаемой темы, предлагая студентам найти собственное их решение.

в) описание шкалы оценивания:

Максимальная оценка за устное выступление и работу на семинарском занятии – 5 баллов.

5 баллов – студент дает полный ответ на поставленный вопрос, его речь свободна и грамотна, конспект не зачитывается, а используется лишь как опорный, студент делает важные дополнения по существу других вопросов, значительно проясняющие отдельные аспекты, которые не являются повторами, хорошо разбирается в обсуждаемом материале, демонстрирует знание источников, библиографии, различных точек зрения по изучаемой теме, умеет анализировать тексты, приходит к самостоятельным аргументированным выводам и отстаивает свою точку зрения, соблюдает нормы литературной речи.

2 балла – студент хорошо разбирается в обсуждаемом материале, демонстрирует умение критически анализировать источники и различные точки зрения по обсуждаемой проблеме, приходит к самостоятельным аргументированным выводам, однако не проявляет активность в работе группы на семинаре (готовится и отвечает только на один вопрос семинарского занятия).

1 балл – студент неполно владеет материалом, при изложении фактического материала допускает отдельные неточности, знает различные точки зрения по обсуждаемой проблеме, но возникают трудности с их анализом, умеет излагать собственную позицию, но не все выводы носят доказательный характер, при ответе активно пользуется конспектом вплоть до его зачитывания.

3.3 Рефераты (с последующим представлением докладов на «круглом столе»)

1. Понятие, сущность и ключевые признаки больших данных.
2. Роль и место больших данных в решении аналитических и исследовательских задач профессиональной деятельности.
3. Техника и технология Big Data
4. Техники больших данных. Консолидация данных. Визуализация.
5. Классификация. Кластеризация.
6. Регрессионный анализ.
7. Анализ ассоциативных правил.
8. Нейронные сети.
9. Технологии и инструменты больших данных.
10. Storm – система потоковой обработки.
11. Язык программирования R. Аналитика больших данных как корпоративный проект.
12. Аналитическая платформа Deductor Academic

13. Сущность и принцип работы аналитической платформы Deductor Academic. Основные функции и инструменты аналитической платформы Deductor Academic для целей анализа и исследования социально-экономических процессов и явлений в деятельности предприятий.
14. Моделирование социально-экономических процессов и явлений в деятельности предприятий с помощью платформы Deductor Academic.
15. Инструментарий прикладного компьютерного анализа и моделирования в Deductor Academic.
16. Метод факторного анализа как модель латентных переменных
17. Факторный анализ как метод понижения размерности пространства признаков
18. Экзогенные и эндогенные понятия в типологическом анализе
19. Технологии больших данных в финансах
20. Технологии больших данных в маркетинге
21. Технологии больших данных в логистике
22. Технологии больших данных в банковской сфере
23. Искусственный интеллект и его применение в управлении процессами
24. Машинное обучение.
25. Классификация клиентов организации с применением методов машинного обучения
26. Применение алгоритмов машинного обучения в задачах квантового анализа фондовой биржи
27. Применение машинного обучения для продвижения банковских продуктов
28. Применение методов машинного обучения при формировании оптимального портфеля ценных бумаг
29. Торговые роботы на основе обучения с подкреплением
30. Формирование моделей доходностей ценных бумаг на фондовой бирже на основе стратегий зависимости от новостей
31. Бизнес-аналитика на основе технологий больших данных

Показатели и критерии оценки реферата

Показатели оценки	**Критерии оценки**	**Баллы (max)**
1. Новизна реферированного текста	- актуальность проблемы и темы; - новизна и самостоятельность в постановке проблемы, в формулировании нового аспекта выбранной для анализа проблемы; - наличие авторской позиции, самостоятельность суждений	2
2. Степень раскрытия сущности проблемы	- соответствие плана теме реферата; - соответствие содержания теме и плану реферата; - полнота и глубина раскрытия основных понятий проблемы; - обоснованность способов и методов работы с мате-	2

	риалом; - умение работать с литературой, систематизировать и структурировать материал; - умение обобщать, сопоставлять различные точки зрения по рассматриваемому вопросу, аргументировать основные положения и выводы.	
3. Обоснованность выбора источников	- круг, полнота использования литературных источников по проблеме; - привлечение новейших работ по проблеме (журнальные публикации, материалы сборников научных трудов и т.д.).	2
4. Соблюдение требований к оформлению	- правильное оформление ссылок на используемую литературу; - грамотность и культура изложения; - владение терминологией и понятийным аппаратом проблемы; - соблюдение требований к объему реферата; - культура оформления: выделение абзацев. - отсутствие орфографических и синтаксических ошибок, стилистических погрешностей; - отсутствие опечаток, сокращений слов, кроме общепринятых; - литературный стиль.	2
5. Презентация на «круглом столе»	- качество и полнота презентации; - соблюдение требований к презентации по оформлению	2

Доклад – устное выступление студента, являющееся результатом его самостоятельной подготовки по заранее полученной теме и в соответствии с требованиями к самостоятельной работе студентов.

Выступление во время доклада, как правило, рассчитано на 6-7 минут, не может превышать установленное время, должно строго соответствовать объявленной теме. Приветствуются доклады с дополнительным использованием презентаций и мультимедийной техники.

Во время выступления студент может использовать свободную речь близко к тексту доклада, однако вправе зачитывать подготовленный им текст, демонстрируя владение материалом. Речь должна быть четкой, выразительной и эмоциональной. Обязательным элементом процедуры доклада является его обсуждение. Студентам группы предлагается задавать докладчику вопросы по теме доклада, вопросы может задавать и преподаватель. По окончании доклада организуется дискуссия.

Домашняя (внеаудиторная) подготовка доклада оценивается до 1 балла, выступление и ответы на вопросы на «круглом столе»- до 1 балла. Итого за выполнение данного задания студент может получить до 2 баллов.

Критерии оценки устного выступления.

2 балла (максимальная оценка) – выступление (доклад) отличается последовательностью, логикой изложения, легко воспринимается аудиторией, при ответе на вопросы выступающий демонстрирует глубину владения представленным материалом, ответы формулируются аргументированно, обосновывается собственная позиция в проблемных ситуациях.

1 балл – выступающий передает содержание проблемы, но не демонстрирует умение выделять главное, существенное, выступление воспринимается аудиторией достаточно сложно, ответы на вопросы поверхностные, либо вызывают у докладчика затруднение.

0 баллов – доклад краткий, поверхностный, несамостоятельный, докладчик не разбирается в сути вопроса, не может представить его в аудитории.

3.4 Тестовые задания

а) примеры тестов:

1. Какую часть мирового рынка Data Mining занимают услуги или консультации по эффективному внедрению этой технологии для решения актуальных бизнес-задач:
а) более 75%
б) около половины
в) менее 10% рынка

2. В основу программного продукта Cognos 4Thought положена технология:
а) множественной регрессии
б) нейронных сетей
в) деревьев решений

3. Достаточно высокая стоимость, невозможность добавлять свои функции, сложность подготовки данных, практическое отсутствие в интерфейсе терминов предметной области – это слабые стороны:
а) адаптация программного обеспечения под конкретную задачу
б) заказ готового решения у фирмы-разработчика
в) готового программного обеспечения

4. MAP-сплайны в системе STATISTICA – это:
а) непараметрическая процедура, в работе которой не используется никаких предположений об общем виде функциональных связей между зависимыми и независимыми переменными
б) процедура, опирающаяся на предположения о типе и накладывающая ограничения на класс зависимостей

в) параметрическая процедура, основанная на предположениях о виде функциональных связей между зависимыми и независимыми переменными

5. Готовые алгоритмы, полная конфиденциальность информации, техническая поддержка производителя, общение с другими пользователями пакета — это преимущества использования:
а) адаптация программного обеспечения под конкретную задачу
б) заказ готового решения у фирмы-разработчика
в) готового программного обеспечения

6. Cognos 4Thought предназначен для:
а) прогнозирования
б) моделирования
в) оба варианта верны
г) нет правильного ответа

7. Какое решение в большей мере требует наличия высококвалифицированных специалистов при внедрении и использования инструмента Data Mining:
а) адаптация программного обеспечения под конкретную задачу
б) заказ готового решения у фирмы-разработчика
в) использование готового программного обеспечения

8. Рабочее пространство STATISTICA Data Miner не включает такого элемента:
а) тестирование
б) анализ данных, моделирование
в) подготовка, преобразования и очистка данных

9. Вариант использования адаптированного программного обеспечения Data Mining:
а) имеет неоспоримые преимущества перед использованием готового программного обеспечения
б) всегда проигрывает перед использованием готового программного обеспечения
в) имеет как сильные, так и слабые стороны

10. Инструментальное средство для оперативного анализа данных и формирования отчетов по OLAP-технологии:
а) Cognos Impromptu
б) Cognos PowerPlay
в) Cognos Scenario

11. На этапе подготовки данных:
а) специалисты компании Разработчика подготавливают данные для их даль-

нейшего анализа
б) специалисты компании Разработчика и Заказчика подготавливают данные для их дальнейшего анализа
в) специалисты компании Заказчика подготавливают данные для их дальнейшего анализа

12. На каких этапах 4Thought поддерживает анализ данных:
а) сбор данных
б) преобразование данных
в) интерпретация модели
г) на всех этапах

13. Преимуществом использования адаптированного программного обеспечения Data Mining по сравнению с готовыми программными продуктами и их самостоятельным использованием является:
а) наличие терминов предметной области
б) сложность подготовки данных
в) полная конфиденциальность информации

14. Охарактеризуйте систему Cognos Scenario:
а) является средством оперативного анализа данных
б) формирует отчеты по OLAP–технологии
в) интеллектуальное инструментальное средство поиска данных

15. Преимуществом использования адаптированного программного обеспечения Data Mining по сравнению с готовыми программными продуктами и их самостоятельным использованием является:
а) адаптированность
б) не требуется дописывать программный код
в) сложность подготовки данных

16. Охарактеризуйте систему Cognos Scenario:
а) формирует отчеты по OLAP–технологии
б) позволяет руководителям выявлять скрытые тенденции и модели бизнеса
в) является средством оперативного анализа данных

17. На этапе первичного исследования данных:
а) всю работу осуществляет заказчик
б) со стороны заказчика может потребоваться максимальное участие
в) со стороны заказчика может потребоваться лишь минимальное участие

18. Модуль Oracle Data Mining доступен из таких редакций:
а) Personal Edition

б) Enterprise Edition
в) OneStandard Edition

19. Постановка бизнес-задачи – это этап, который:
а) формулирует конкретные бизнес-задачи, и они уже не могут быть изменены
б) формулирует конкретные бизнес-задачи, и они могут быть изменены в ходе прохождения именно этого цикла
в) формулирует конкретные бизнес-задачи, и они не могут быть изменены в ходе прохождения именно этого цикла

20. Архитектура хранилища типа «звезда» в Deductor называется:
а) сценарием
б) процессом
в) проектом

21. Что требуется от пользователя при работе с KXEN:
а) данные, которые необходимо проанализировать
б) выбор лучшей модели
в) тестирование модели

22. Oracle Data Mining поставляется как:
а) опция в Oracle Standard Edition
б) опция в Oracle Enterprise Edition
в) опция в любой редакции

23. Что требуется от пользователя при работе с KXEN:
а) выбор лучшей модели
б) тестирование модели
в) определение типа задачи, которую нужно решить

24. Последовательность действий, которые необходимо провести для анализа данных, называется в Deductor:
а) процессом
б) сценарием
в) этапом

25. Построение модели в KXEN можно охарактеризовать как:
а) итеративный процесс
б) набор таких шагов: подготовка модели, построение модели, тестирование модели
в) функцию предсказательного анализа в режиме on-line в формате «вопрос-ответ»

26. Один из компонентов, входящих в состав Deductor:
а) лаборатория BaseGroup Labs
б) многомерное хранилище данных Deductor Warehouse
в) лаборатория Base Labs

27. Охарактеризуйте квалификацию, которой требуется обладать пользователю для работы с KXEN:
а) пользователю не требуется обладать специальной квалификацией и знаниями в области анализа и статистики
б) пользователю требуется обладать специальной квалификацией и знаниями в области анализа, статистики и искусственного интеллекта
в) пользователю требуется обладать специальной квалификацией и знаниями в области анализа и статистики

28. Один из компонентов, входящих в состав Deductor:
а) лаборатория BaseGroup Labs
б) лаборатория Group Labs
в) аналитическое приложение Deductor Studio

29. В чем заключается основная особенность инструмента KXEN:
а) в практически полной автоматизации процесса построения моделей
б) в сложности построенных моделей
в) в возможности использования малого количества ретроспективных данных

30. Постановка задачи, построение оптимальной модели, понимание модели, применение результатов. Перечисленные выше этапы являются этапами:
а) традиционного процесса Data Mining
б) подхода KXEN
в) и того, и другого

Критерии оценивания компетенций (результатов):

Правильный ответ на каждое тестовое задание оценивается в 1 балл.

3.5. Контрольная работа

а) пример вопросов к контрольной работе

Контрольная работа «Введение в большие данные»
Вариант 1.

1. Структура и задачи анализа больших наборов данных.
2. Функции управления с использованием анализа больших наборов данных
3. Технология и методы обработки информации

4. Преобразование и первичная обработка данных
5. Принципы построения автоматизированных систем управления с применением анализа больших наборов данных

б) критерии оценивания компетенций (результатов):

Контрольные работы проводятся 1-2 раза в семестр на модульных неделях по расписанию, устанавливаемому учебно-методическим управлением. Они проводятся в форме типовых работ с вопросами, тестов или ином виде по выбору преподавателя с учетом объема изученного материала по курсу.

Контрольная работа может быть комбинированного типа: 3 вопроса, подразумевающих текстовые ответы и 5 тестовых вопросов с одним или несколькими вариантами ответа.

Оценивание студента проводится преподавателем независимо от наличия или отсутствия студента (по уважительной или неуважительной причине) на занятии. Студенту, пропустившему по уважительной причине контрольную модульную работу, предоставляется возможность отработки. Отработать занятие можно по согласованию с преподавателем в четко установленные сроки в соответствии с графиком консультаций преподавателя, который имеется в отделении социально-экономических наук.

Оценивается степень усвоения теоретических знаний по следующим критериям: правильность, полнота и логичность письменного ответа, способностью проиллюстрировать ответ примерами.

Максимальный балл за контрольную работу – 15. Каждый вопрос оценивается в 3 балла,

В случае использования вариантов заданий смешанного типа (вопросы, подразумевающие развернутый ответ, и тесты) баллы формируются следующим образом:

- 3 вопроса, максимальный балл за каждый вопрос – 4
- 2 тестовых задания, 1,5 балла за верно решенный тест, 0 баллов – если в тесте выбраны не все верные варианты, или среди выбранных вариантов ответа есть ошибочные

3.6. Индивидуальное домашнее задание

а) типовые задания (вопросы) – образец:

Выполнение индивидуального домашнего задания (ИДЗ) предполагает проведение студентами в группе из двух человек анализа ценных бумаг и производных финансовых инструментов с использованием технологий анализа больших данных и выбор оптимальных для компании или частного инвестора вариантов инвестирования с учетом максимально возможного для инвестора уровня рисков. Для проведения анализа студенты самостоятельно выбирают методы и модели. Рекомендуется использовать различные методы фундаментального и технического анализа рынка, а также методики формирования и оценки эффективности портфелей.

Вопросы к ИДЗ:

1. Проанализировать доходность акций (выбрать не менее 3 российских и 3 зарубежных компаний для анализа)
2. Проанализировать рынок российских и зарубежных корпоративных облигаций (выбрать не менее 3 российских и 3 зарубежных компаний для анализа)
3. Проанализировать доходность государственных ценных бумаг (облигации федерального займа, региональные и муниципальные облигации – не менее 2 субъектов РФ и муниципалитетов по данным за 5 лет с использованием технологий анализа больших данных)
4. Проанализировать рынок производных финансовых инструментов, выбрав не менее 3 инструментов для анализа, использовать дневные котировки за последние 5 лет (на основе использования методов анализа больших данных, в т.ч. – массивов с неполными, текстовыми данными и т.п.).
5. По каждому блоку анализа сделать выводы. Выбрать приемлемый для инвестора уровень риска и сформировать инвестиционный портфель с использованием технологий машинного обучения. Определить размер уплачиваемых инвестором налогов по операциям с ценными бумагами. Определить основные риски и возможную доходность данного портфеля.

б) критерии оценивания компетенций (результатов):

Выполнение индивидуального домашнего задания оценивается по следующим критериям:

1. правильность и полнота расчетов;
2. умение применять различные инструменты и модели анализа рынка ценных бумаг и производных финансовых инструментов для проведения анализа и обоснования выводов;
3. обоснованность выводов и предложений;
4. умение интерпретировать данные фундаментального и технического анализа более детального обоснования полученных выводов;
5. умение использовать специальную терминологию и выбирать методы анализа;
6. оригинальность работы (оценка оригинальности по системе «Антиплагиат. Вуз» - www.mephist. antiplagiat.ru должна быть не менее 70%)

в) описание шкалы оценивания:

Максимально за подготовку индивидуального домашнего задания студент может получить 10 баллов, максимальная оценка за каждый блок задания составляет 2 балла.

Оценка за выполненный блок формируется следующим образом:

2 балла - все расчеты выполнены верно, методики использованы корректно, выводы обоснованы, для обоснования выводов и предложений представлен детальный анализ дополнительных форм внутренней отчетности, терминология используется корректно и ее использование полностью соответствует содержанию блока, полученные рекомендации обоснованы анализом действующего законодательства.

1,5 балла - все расчеты выполнены верно, выводы обоснованы, для обоснования выводов и предложений представлен детальный анализ дополнительных форм отчетности, терминология используется корректно и ее использование полностью соответствует содержанию блока, полученные рекомендации обоснованы анализом действующего законодательства, но имеются незначительные неточности в выводах, или рассмотрены не все аспекты проблемы.

1 балл - расчеты выполнены в целом верно, выводы обоснованы, терминология используется корректно и ее использование полностью соответствует содержанию блока, однако для обоснования рекомендаций не достаточно полно используется дополнительная отчетность и анализ действующего законодательства.

0,5 балла - в расчетах присутствуют ошибки, выводы носят описательный характер. Рекомендации схематичны, не представлен анализ дополнительных форм отчетности и действующего законодательства для обоснования рекомендаций.

0 баллов – блок задания не выполнен, или выполнен с грубыми ошибками: ошибки в расчетах и использовании формул, отсутствие выводов или их несоответствие практике.

4 Итоговая аттестация по дисциплине

Итоговая аттестация по дисциплине является интегральным показателем качества теоретических и практических знаний и навыков обучающихся по дисциплине и складывается из оценок, полученных в ходе текущей и промежуточной аттестации.

Текущая аттестация в семестре проводится с целью обеспечения своевременной обратной связи, для коррекции обучения, активизации самостоятельной работы обучающихся.

Промежуточная аттестация предназначена для объективного подтверждения и оценивания достигнутых результатов обучения после завершения изучения дисциплины.

Текущая аттестация осуществляется два раза в семестр:

- контрольная точка № 1 (КТ № 1) – выставляется в электронную ведомость не позднее 8 недели учебного семестра. Включает в себя оценку мероприятий текущего контроля аудиторной и самостоятельной работы обучающегося по разделам/темам учебной дисциплины с 1 по 8 неделю семестра.

- контрольная точка № 2 (КТ № 2) – выставляется в электронную ведо-

мость не позднее 16 недели учебного семестра. Включает в себя оценку мероприятий текущего контроля аудиторной и самостоятельной работы обучающегося по разделам/темам учебной дисциплины с 9 по 16 неделю семестра.

Результаты текущей и промежуточной аттестации подводятся по шкале балльно-рейтинговой системы.

Этап рейтинговой системы / Оценочное средство	**Неделя**	**Балл**	
		Минимум*	Максимум
Текущая аттестация	**1-16**	**36**	**60**
Контрольная точка № 1	**7-8**	**18**	**30**
Контрольная работа	8	9	15
Опрос	1-8	3	5
Индивидуальное домашнее задание	7	6	10
Контрольная точка № 2	**15-16**	**18**	**30**
Реферат +доклад на «круглом столе»	16	6	10
Тестовые задания	9-15	9	15
Опрос	9-16	3	5
Промежуточная аттестация	**-**	**24**	**40**
Экзамен		24	40
Итого по дисциплине		**60**	**100**

* - Минимальное количество баллов за оценочное средство – это количество баллов, набранное обучающимся, при котором оценочное средство засчитывается, в противном случае обучающийся должен ликвидировать появившуюся академическую задолженность по текущей или промежуточной аттестации. Минимальное количество баллов за текущую аттестацию, в т. ч. отдельное оценочное средство в ее составе, и промежуточную аттестацию составляет 60% от соответствующих максимальных баллов.

Процедура оценивания знаний, умений, владений по дисциплине включает учет успешности по всем видам заявленных оценочных средств.

Устный опрос проводится на каждом практическом занятии и затрагивает как тематику прошедшего занятия, так и лекционный материал. Применяется групповое оценивание ответа или оценивание преподавателем.

Темы рефератов распределяются на первом занятии, готовые рефераты докладываются на занятиях в сопровождении презентаций в соответствии с установленным преподавателем графиком.

По окончании освоения дисциплины проводится промежуточная аттестация в виде экзамена, что позволяет оценить совокупность приобретенных в процессе обучения компетенций. При выставлении итоговой оценки применяется балльно-рейтинговая система оценки результатов обучения.

Экзамен предназначен для оценки работы обучающегося в течение всего срока изучения дисциплины и призван выявить уровень, прочность и систематичность полученных обучающимся теоретических знаний и умений, способности приводить примеры практического использования знаний (например, применять их в решении практических задач), приобретения навыков самостоятельной работы, развития творческого мышления.

Оценка сформированности компетенций на экзамене для тех обучающихся, которые пропускали занятия и не участвовали в проверке компетенций во время изучения дисциплины, проводится после индивидуального собеседования с преподавателем по пропущенным или не усвоенным обучающимся темам с последующей оценкой самостоятельно усвоенных знаний на экзамене.

Итоговая аттестация по дисциплине оценивается по 100-балльной шкале и представляет сумму баллов, заработанных обучающимся при выполнении заданий в рамках текущей и промежуточной аттестации

<table>
<tr><th>Сумма баллов</th><th>Оценка по 4-х балльной шкале</th><th>Оценка ECTS</th><th>Требования к уровню освоения учебной дисциплины</th></tr>
<tr><td>90-100</td><td>5- «отлично»/ «зачтено»</td><td>A</td><td>Оценка «отлично» выставляется обучающемуся, если он глубоко и прочно усвоил программный материал, исчерпывающе, последовательно, четко и логически стройно его излагает, умеет тесно увязывать теорию с практикой, использует в ответе материал монографической литературы</td></tr>
<tr><td>85-89</td><td rowspan="3">4 - «хорошо»/ «зачтено»</td><td>B</td><td rowspan="3">Оценка «хорошо» выставляется обучающемуся, если он твёрдо знает материал, грамотно и по существу излагает его, не допуская существенных неточностей в ответе на вопрос</td></tr>
<tr><td>75-84</td><td>C</td></tr>
<tr><td>70-74</td><td rowspan="2">D</td></tr>
<tr><td>65-69</td><td rowspan="2">3 - «удовлетворительно» / «зачтено»</td><td rowspan="2">Оценка «удовлетворительно» выставляется обучающемуся, если он имеет знания только основного материала, но не усвоил его деталей, допускает неточности, недостаточно правильные формулировки, нарушения логической последовательности в изложении программного материала</td></tr>
<tr><td>60-64</td><td>E</td></tr>
<tr><td>0-59</td><td>2 - «неудовлетворительно»/ «не зачтено»</td><td>F</td><td>Оценка «неудовлетворительно» выставляется обучающемуся, который не знает значительной части программного материала, допускает существенные ошибки. Как правило, оценка «неудовлетворительно» ставится обучающимся, которые не могут продолжить обучение без дополнительных занятий по соответствующей дисциплине</td></tr>
</table>