

МИНОБРНАУКИ РОССИИ

**Федеральное государственное автономное образовательное
учреждение высшего образования "Пермский
государственный национальный исследовательский
университет"**

Кафедра информационных систем и математических методов в экономике

**Авторы-составители: Бердышев Олег Вячеславович
Шишкин Владимир Андреевич**

Рабочая программа дисциплины
АНАЛИЗ БОЛЬШИХ ДАННЫХ (BIG DATA)
Код УМК 98276

Утверждено
Протокол №10
от «06» июня 2022 г.

Пермь, 2022

1. Наименование дисциплины

Анализ больших данных (Big Data)

2. Место дисциплины в структуре образовательной программы

Дисциплина входит в базовую часть Блока « М.1 » образовательной программы по направлениям подготовки (специальностям):

Направление: **38.04.02** Менеджмент

направленность Международный бизнес в цифровой экономике

3. Планируемые результаты обучения по дисциплине

В результате освоения дисциплины **Анализ больших данных (Big Data)** у обучающегося должны быть сформированы следующие компетенции:

38.04.02 Менеджмент (направленность : Международный бизнес в цифровой экономике)

ОПК.2 Способен применять современные техники и методики сбора данных, продвинутые методы их обработки и анализа, в том числе использовать интеллектуальные информационно-аналитические системы, при решении управленческих и исследовательских задач

Индикаторы

ОПК.2.1 Собирает информацию, выбирает инструментарий и осуществляет расчет социально-экономических показателей предприятия, отрасли, региона и экономики в целом

ОПК.2.2 Обосновывает и производит выбор управленческих решений на основе результатов инструментальных методов экономического анализа

4. Объем и содержание дисциплины

Направления подготовки	38.04.02 Менеджмент (направленность: Международный бизнес в цифровой экономике)
форма обучения	очная
№№ триместров, выделенных для изучения дисциплины	2
Объем дисциплины (з.е.)	4
Объем дисциплины (ак.час.)	144
Контактная работа с преподавателем (ак.час.), в том числе:	48
Проведение лекционных занятий	12
Проведение практических занятий, семинаров	36
Самостоятельная работа (ак.час.)	96
Формы текущего контроля	Итоговое контрольное мероприятие (1) Письменное контрольное мероприятие (2)
Формы промежуточной аттестации	Экзамен (2 триместр)

5. Аннотированное описание содержания разделов и тем дисциплины

Анализ больших данных (Big Data)

В курсе дается характеристика понятий «данные», «информация», «знания»; выделяется проблема «больших данных»; дается обзор источников информации; основные определения, термины и задачи анализа больших данных; характеристика больших данных – 5V; предпосылки формирования тренда; драйверы рынка больших данных; обзор технологий хранения и обработки больших данных; современные программные средства анализа больших данных; процесс аналитики (стандарт CRISP-DM); понятие машинного обучения и его компоненты; основные определения и термины машинного обучения; виды задач в машинном обучении и процесс их решения; обучение без учителя: кластерный анализ; обучение с учителем: логистическая регрессия, деревья решений, регрессионный анализ.

Интеллектуальный анализ данных

Введение в datamining. Данные и метаданные. Методы и стадии Data Mining. Задачи Data Mining. Извлечение, данных, информации и знаний. Сферы применения Data Mining.

Сбор, хранение, обработка, извлечение и анализ больших потоков данных

Основные характеристики и типы больших данных. Что такое статистическое мышление? Обработка данных для контроля качества процессов и технологии six sigma. Статистический подход к анализу данных. Точечное и интервальное оценивание.

Методы классификации и прогнозирования

Анализ последовательностей. Отбор признаков и снижение размерности. Работа с выбросами и пропущенными значениями. Классификация и кластеризация. Поиск признаковых зависимостей и частых множеств. Мультимодальная кластеризация. Предиктивные алгоритмы. Прогнозирование и визуализация данных

Введение в машинное обучение и разработку данных

Основные аналитические методы обработки данных. Машинное обучение и майнинг больших данных (Big Data). Нейронные сети как реализация алгоритмов машинного обучения.

Способы визуализации, интерпретации, проектная организация и бизнес рекомендации

Применение технологий больших данных для задач управления в банковской, страховой, финансовой индустриях, в масс медиа, FinTech как область применения бизнес анализа. Способы визуального представления данных. Методы визуализации. Когнитивная информатика, экономика знаний, требования к специалистам в области интеллектуальной обработки данных для бизнеса

КТ 1

Контрольная работа включает проверку знаний по темам Интеллектуальный анализ данных, Сбор, хранение, обработка, извлечение и анализ больших потоков данных

КТ 2

Контрольная работа по темам Методы классификации и прогнозирования, Введение в машинное обучение и разработку данных

Итоговое контрольное мероприятие

Итоговое контрольное мероприятие по всем темам дисциплины

6. Методические указания для обучающихся по освоению дисциплины

Освоение дисциплины требует систематического изучения всех тем в той последовательности, в какой они указаны в рабочей программе.

Основными видами учебной работы являются аудиторские занятия. Их цель - расширить базовые знания обучающихся по осваиваемой дисциплине и систему теоретических ориентиров для последующего более глубокого освоения программного материала в ходе самостоятельной работы. Обучающемуся важно помнить, что контактная работа с преподавателем эффективно помогает ему овладеть программным материалом благодаря расстановке необходимых акцентов и удержанию внимания интонационными модуляциями голоса, а также подключением аудио-визуального механизма восприятия информации.

Самостоятельная работа преследует следующие цели:

- закрепление и совершенствование теоретических знаний, полученных на лекционных занятиях;
- формирование навыков подготовки текстовой составляющей информации учебного и научного назначения для размещения в различных информационных системах;
- совершенствование навыков поиска научных публикаций и образовательных ресурсов, размещенных в сети Интернет;
- самоконтроль освоения программного материала.

Обучающемуся необходимо помнить, что результаты самостоятельной работы контролируются преподавателем во время проведения мероприятий текущего контроля и учитываются при промежуточной аттестации.

Обучающимся с ОВЗ и инвалидов предоставляется возможность выбора форм проведения мероприятий текущего контроля, альтернативных формам, предусмотренным рабочей программой дисциплины. Предусматривается возможность увеличения в пределах 1 академического часа времени, отводимого на выполнение контрольных мероприятий.

Процедура оценивания результатов обучения инвалидов и лиц с ограниченными возможностями здоровья по дисциплине предусматривает предоставление информации в формах, адаптированных к ограничениям их здоровья и восприятия информации.

При проведении текущего контроля применяются оценочные средства, обеспечивающие передачу информации, от обучающегося к преподавателю, с учетом психофизиологических особенностей здоровья обучающихся.

7. Перечень учебно-методического обеспечения для самостоятельной работы обучающихся по дисциплине

При самостоятельной работе обучающимся следует использовать:

- конспекты лекций;
- литературу из перечня основной и дополнительной учебной литературы, необходимой для освоения дисциплины (модуля);
- текст лекций на электронных носителях;
- ресурсы информационно-телекоммуникационной сети "Интернет", необходимые для освоения дисциплины;
- лицензионное и свободно распространяемое программное обеспечение из перечня информационных технологий, используемых при осуществлении образовательного процесса по дисциплине;
- методические указания для обучающихся по освоению дисциплины.

8. Перечень основной и дополнительной учебной литературы

Основная:

1. Билл, Фрэнк Революция в аналитике: Как в эпоху Big Data улучшить ваш бизнес с помощью операционной аналитики / Фрэнк Билл ; перевод И. Евстигнеева ; под редакцией В. Мылова. — Москва : Альпина Паблишер, 2020. — 320 с. — ISBN 978-5-9614-5302-7. — Текст : электронный // Электронно-библиотечная система IPR BOOKS : [сайт]. <http://www.iprbookshop.ru/93032.html>
2. Воронова, Л. И. Big Data. Методы и средства анализа : учебное пособие / Л. И. Воронова, В. И. Воронов. — Москва : Московский технический университет связи и информатики, 2016. — 33 с. — Текст : электронный // Цифровой образовательный ресурс IPR SMART : [сайт]. <http://www.iprbookshop.ru/61463.html>

Дополнительная:

1. Федин, Ф. О. Анализ данных. Часть 2. Инструменты Data Mining : учебное пособие / Ф. О. Федин, Ф. Ф. Федин. — Москва : Московский городской педагогический университет, 2012. — 308 с. — ISBN 2227-8397. — Текст : электронный // Электронно-библиотечная система IPR BOOKS : [сайт]. <http://www.iprbookshop.ru/26445>

9. Перечень ресурсов сети Интернет, необходимых для освоения дисциплины

<https://www.ibm.com/analytics/big-data-analytics> Онлайн-ресурс

<https://www.heavy.ai/learn/big-data-analytics> Онлайн-ресурс

10. Перечень информационных технологий, используемых при осуществлении образовательного процесса по дисциплине

Образовательный процесс по дисциплине **Анализ больших данных (Big Data)** предполагает использование следующего программного обеспечения и информационных справочных систем:

- доступ в режиме on-line в Электронную библиотечную систему (ЭБС);
- доступ в электронную информационно-образовательную среду университета
- Интернет-сервисы и электронные ресурсы

Необходимое лицензионное и свободно распространяемое программное обеспечение:

- офисный пакет приложений Microsoft Office Professional 31.05.13.

При освоении материала и выполнения заданий по дисциплине рекомендуется использование материалов, размещенных в Личных кабинетах обучающихся ЕТИС ПГНИУ (student.psu.ru).

При организации дистанционной работы и проведении занятий в режиме онлайн могут использоваться:

система видеоконференцсвязи на основе платформы BigBlueButton (<https://bigbluebutton.org/>).

система LMS Moodle (<http://e-learn.psu.ru/>), которая поддерживает возможность использования текстовых материалов и презентаций, аудио- и видеоконтент, а так же тесты, проверяемые задания, задания для совместной работы.

система тестирования Indigo (<https://indigotech.ru/>).

11. Описание материально-технической базы, необходимой для осуществления образовательного процесса по дисциплине

Материально-техническая база обеспечивается наличием:

1. Лекционные занятия - аудитория, оснащенная презентационной техникой (проектор, экран, компьютер/ноутбук) с соответствующим программным обеспечением, меловой (и) или маркерной доской.
2. Занятий семинарского типа (практические занятия) - аудитория, оснащенная презентационной техникой (проектор, экран, компьютер/ноутбук) с соответствующим программным обеспечением, меловой (и) или маркерной доской.
3. Лабораторные занятия - Компьютерный класс, оснащенный персональными ЭВМ и соответствующим программным обеспечением. Состав оборудования определен в Паспорте компьютерного класса. Аудитория №109/12 (105э) (Комп. класс) 12 корпус
4. Самостоятельная работа - аудитория для самостоятельной работы, оснащенная компьютерной техникой с возможностью подключения к сети «Интернет», обеспеченная доступом в электронную информационно-образовательную среду университета. Помещения Научной библиотеки ПГНИУ.
5. Текущий контроль и промежуточная аттестация - аудитория, оснащенная меловой (и) или маркерной доской.
6. Индивидуальные и групповые консультации - аудитория, оснащенная презентационной техникой (проектор, экран, компьютер/ноутбук) с соответствующим программным обеспечением, меловой (и) или маркерной доской или аудитория, оснащенная меловой (и) или маркерной доской.

Помещения научной библиотеки ПГНИУ для обеспечения самостоятельной работы обучающихся:

1. Научно-библиографический отдел, корп.1, ауд. 142. Оборудован 3 персональными компьютера с

доступом к локальной и глобальной компьютерным сетям.

2. Читальный зал гуманитарной литературы, корп. 2, ауд. 418. Оборудован 7 персональными компьютерами с доступом к локальной и глобальной компьютерным сетям.

3. Читальный зал естественной литературы, корп.6, ауд. 107а. Оборудован 5 персональными компьютерами с доступом к локальной и глобальной компьютерным сетям.

4. Отдел иностранной литературы, корп.2 ауд. 207. Оборудован 1 персональным компьютером с доступом к локальной и глобальной компьютерным сетям.

5. Библиотека юридического факультета, корп.9, ауд. 4. Оборудована 11 персональными компьютерами с доступом к локальной и глобальной компьютерным сетям.

6. Читальный зал географического факультета, корп.8, ауд. 419. Оборудован 6 персональными компьютерами с доступом к локальной и глобальной компьютерным сетям.

Все компьютеры, установленные в помещениях научной библиотеки, оснащены следующим программным обеспечением:

Операционная система ALT Linux;

Офисный пакет Libreoffice.

Справочно-правовая система «КонсультантПлюс»

**Фонды оценочных средств для аттестации по дисциплине
Анализ больших данных (Big Data)**

**Планируемые результаты обучения по дисциплине для формирования компетенции.
Индикаторы и критерии их оценивания**

ОПК.2

Способен применять современные техники и методики сбора данных, продвинутые методы их обработки и анализа, в том числе использовать интеллектуальные информационно-аналитические системы, при решении управленческих и исследовательских задач

Индикатор	Планируемые результаты обучения	Критерии оценивания результатов обучения
<p>ОПК.2.2 Обосновывает и производит выбор управленческих решений на основе результатов инструментальных методов экономического анализа</p>	<p>Знает основные методы анализа больших данных для анализа и интерпретации социально-экономической информации. Умеет применять основные методы анализа больших данных для обоснования управленческих решений. Владеет навыками применения инструментальных методов анализа больших данных для обоснования и выбора управленческих решений.</p>	<p align="center">Неудовлетворител</p> <p>Не знает основные методы анализа больших данных для анализа и интерпретации социально-экономической информации. Не умеет применять основные методы анализа больших данных для обоснования управленческих решений. Не владеет навыками применения инструментальных методов анализа больших данных для обоснования и выбора управленческих решений.</p> <p align="center">Удовлетворительн</p> <p>Знает на удовлетворительном уровне основные методы анализа больших данных для анализа и интерпретации социально-экономической информации. Умеет на удовлетворительном уровне применять основные методы анализа больших данных для обоснования управленческих решений. Владеет удовлетворительно навыками применения инструментальных методов анализа больших данных для обоснования и выбора управленческих решений.</p> <p align="center">Хорошо</p> <p>Хорошее знание основных методов анализа больших данных для анализа и интерпретации социально-экономической информации. В целом сформировано умение применять основные методы анализа больших данных для обоснования управленческих решений. на хорошем уровне владеет навыками применения инструментальных методов</p>

Индикатор	Планируемые результаты обучения	Критерии оценивания результатов обучения
		<p style="text-align: center;">Хорошо</p> <p>анализа больших данных для обоснования и выбора управленческих решений.</p> <p style="text-align: center;">Отлично</p> <p>Отличное знание основных методов анализа больших данных для анализа и интерпретации социально-экономической информации.</p> <p>Сформировано уверенное умение применять основные методы анализа больших данных для обоснования управленческих решений. на высоком уровне владеет навыками применения инструментальных методов анализа больших данных для обоснования и выбора управленческих решений.</p>
<p>ОПК.2.1 Собирает информацию, выбирает инструментарий и осуществляет расчет социально-экономических показателей предприятия, отрасли, региона и экономики в целом</p>	<p>Знает основные методы и инструменты для анализа больших данных социально-экономических показателей предприятия, отрасли, региона и экономики в целом.</p> <p>Умеет применять основные методы сбора, анализа и расчета социально-экономических показателей предприятия, отрасли, региона и экономики в целом на основе больших данных.</p> <p>Владеет навыками сбора информации, выбора инструментария и осуществления расчетов социально-экономических показателей предприятия, отрасли, региона и экономики в целом на основе больших данных.</p>	<p style="text-align: center;">Неудовлетворител</p> <p>Не знает основные методы и инструменты для анализа больших данных социально-экономических показателей предприятия, отрасли, региона и экономики в целом.</p> <p>Не умеет применять основные методы сбора, анализа и расчета социально-экономических показателей предприятия, отрасли, региона и экономики в целом на основе больших данных.</p> <p>Не владеет навыками сбора информации, выбора инструментария и осуществления расчетов социально-экономических показателей предприятия, отрасли, региона и экономики в целом на основе больших данных.</p> <p style="text-align: center;">Удовлетворительн</p> <p>Знает на удовлетворительном уровне основные методы и инструменты для анализа больших данных социально-экономических показателей предприятия, отрасли, региона и экономики в целом.</p> <p>Умеет на удовлетворительном уровне применять основные методы сбора, анализа и расчета социально-экономических показателей предприятия, отрасли, региона и экономики в целом на основе больших данных.</p> <p>Владеет на удовлетворительном уровне</p>

Индикатор	Планируемые результаты обучения	Критерии оценивания результатов обучения
		<p style="text-align: center;">Удовлетворительн</p> <p>навыками сбора информации, выбора инструментария и осуществления расчетов социально-экономических показателей предприятия, отрасли, региона и экономики в целом на основе больших данных.</p> <p style="text-align: center;">Хорошо</p> <p>Хорошее знание основные методы и инструменты для анализа больших данных социально-экономических показателей предприятия, отрасли, региона и экономики в целом.</p> <p>В целом сформировано умение применять основные методы сбора, анализа и расчета социально-экономических показателей предприятия, отрасли, региона и экономики в целом на основе больших данных.</p> <p>На хорошем уровне владеет навыками сбора информации, выбора инструментария и осуществления расчетов социально-экономических показателей предприятия, отрасли, региона и экономики в целом на основе больших данных.</p> <p style="text-align: center;">Отлично</p> <p>Отличное знание основные методы и инструменты для анализа больших данных социально-экономических показателей предприятия, отрасли, региона и экономики в целом.</p> <p>Сформировано уверенное умение применять основные методы сбора, анализа и расчета социально-экономических показателей предприятия, отрасли, региона и экономики в целом на основе больших данных.</p> <p>На высоком уровне владеет навыками сбора информации, выбора инструментария и осуществления расчетов социально-экономических показателей предприятия, отрасли, региона и экономики в целом на основе больших данных.</p>

Оценочные средства текущего контроля и промежуточной аттестации

Схема доставки : Базовая

Вид мероприятия промежуточной аттестации : Экзамен

Способ проведения мероприятия промежуточной аттестации : Оценка по дисциплине в рамках промежуточной аттестации определяется на основе баллов, набранных обучающимся на контрольных мероприятиях, проводимых в течение учебного периода.

Максимальное количество баллов : 100

Конвертация баллов в отметки

«отлично» - от 81 до 100

«хорошо» - от 61 до 80

«удовлетворительно» - от 43 до 60

«неудовлетворительно» / «незачтено» менее 43 балла

Компетенция (индикатор)	Мероприятие текущего контроля	Контролируемые элементы результатов обучения
ОПК.2.2 Обосновывает и производит выбор управленческих решений на основе результатов инструментальных методов экономического анализа ОПК.2.1 Собирает информацию, выбирает инструментарий и осуществляет расчет социально-экономических показателей предприятия, отрасли, региона и экономики в целом	КТ 1 Письменное контрольное мероприятие	Знает и умеет применять основные понятия «данные», «информация», «знания»; проблема «больших данных»; источники информации; основные определения, термины и задачи анализа больших данных; характеристика больших данных– 5V; предпосылки формирования тренда; драйверы рынка больших данных; обзор технологий хранения и обработки больших данных; современные программные средства анализа больших данных.
ОПК.2.2 Обосновывает и производит выбор управленческих решений на основе результатов инструментальных методов экономического анализа ОПК.2.1 Собирает информацию, выбирает инструментарий и осуществляет расчет социально-экономических показателей предприятия, отрасли, региона и экономики в целом	КТ 2 Письменное контрольное мероприятие	Знает и умеет выбирать для последующего анализа применять методы кластеризации больших данных, поиска закономерностей, основные аналитические методы обработки данных, машинное обучение и майнинг больших данных (Big Data), нейронные сети.

Компетенция (индикатор)	Мероприятие текущего контроля	Контролируемые элементы результатов обучения
ОПК.2.2 Обосновывает и производит выбор управленческих решений на основе результатов инструментальных методов экономического анализа ОПК.2.1 Собирает информацию, выбирает инструментарий и осуществляет расчет социально-экономических показателей предприятия, отрасли, региона и экономики в целом	Итоговое контрольное мероприятие Итоговое контрольное мероприятие	Знает и умеет применять основные понятия анализа больших данных: «данные», «информация», «знания», понятие машинного обучения и его компоненты; основные определения и термины машинного обучения; виды задач в машинном обучении и процесс их решения; обучение без учителя: кластерный анализ, современные программные средства анализа больших данных.

Спецификация мероприятий текущего контроля

КТ 1

Продолжительность проведения мероприятия промежуточной аттестации: **2 часа**

Условия проведения мероприятия: **в часы аудиторной работы**

Максимальный балл, выставляемый за мероприятие промежуточной аттестации: **30**

Проходной балл: **13**

Показатели оценивания	Баллы
Верно решенное задание	5

КТ 2

Продолжительность проведения мероприятия промежуточной аттестации: **2 часа**

Условия проведения мероприятия: **в часы аудиторной работы**

Максимальный балл, выставляемый за мероприятие промежуточной аттестации: **30**

Проходной балл: **13**

Показатели оценивания	Баллы
Верно решенное задание	5

Итоговое контрольное мероприятие

Продолжительность проведения мероприятия промежуточной аттестации: **2 часа**

Условия проведения мероприятия: **в часы аудиторной работы**

Максимальный балл, выставляемый за мероприятие промежуточной аттестации: **40**

Проходной балл: **17**

Показатели оценивания	Баллы
Верно решенное задание	5