

**МИНОБРНАУКИ РОССИИ**

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

**Кафедра информационных систем и математических методов в экономике**

Авторы-составители: **Шишкин Владимир Андреевич  
Радионова Марина Владимировна**

Рабочая программа дисциплины

**MATHEMATICAL METHODS AND TECHNOLOGIES FOR PROCESSING BIG  
DATA**

Код УМК 97816

Утверждено  
Протокол №9  
от «06» июня 2022 г.

Пермь, 2022

## **1. Наименование дисциплины**

Mathematical Methods and Technologies for Processing Big Data

## **2. Место дисциплины в структуре образовательной программы**

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

Направление подготовки: **38.03.05** Бизнес-информатика

направленность Информационные системы и большие данные

### **3. Планируемые результаты обучения по дисциплине**

В результате освоения дисциплины **Mathematical Methods and Technologies for Processing Big Data** у обучающегося должны быть сформированы следующие компетенции:

**38.03.05** Бизнес-информатика (направленность : Информационные системы и большие данные)

**ОПК.9** Способен использовать информацию, методы и программные средства ее сбора, обработки и анализа для информационно-аналитической поддержки принятия управленческих решений

**Индикаторы**

**ОПК.9.2** Использует методы и программные средства сбора, обработки и анализа информации с целью информационно-аналитической поддержки принятия управленческих решений

**ПК.3** Способен к планированию и организации аналитических работ, в том числе, с использованием технологий больших данных

**Индикаторы**

**ПК.3.2** Определяет состав группы для проведения анализа данных, в том числе, больших данных

#### 4. Объем и содержание дисциплины

<b>Направление подготовки</b>	38.03.05 Бизнес-информатика (направленность: Информационные системы и большие данные)
<b>форма обучения</b>	очная
<b>№№ триместров, выделенных для изучения дисциплины</b>	10
<b>Объем дисциплины (з.е.)</b>	4
<b>Объем дисциплины (ак.час.)</b>	144
<b>Контактная работа с преподавателем (ак.час.), в том числе:</b>	56
<b>Проведение лекционных занятий</b>	28
<b>Проведение лабораторных работ, занятий по иностранному языку</b>	28
<b>Самостоятельная работа (ак.час.)</b>	88
<b>Формы текущего контроля</b>	Входное тестирование (1) Итоговое контрольное мероприятие (1) Письменное контрольное мероприятие (2)
<b>Формы промежуточной аттестации</b>	Экзамен (10 триместр)

## **5. Аннотированное описание содержания разделов и тем дисциплины**

### **Mathematical Methods and Technologies for Processing Big Data**

#### **Enter control**

#### **Introduction**

Applications of Data Science and Big Data. Data types. Data science process. Big Data Ecosystem and Data Science.

#### **1 Data processing**

Definition of goals, research and creation of a project assignment. Data collection. Cleansing, integrating and transforming data. Exploratory data analysis. Data modeling. Presentation of results.

#### **2 Machine learning**

The concept of machine learning. Modeling. Creating indicators and choosing a model. Machine learning with and without a teacher. Partially supervised machine learning.

#### **Test 1**

#### **3 Working with big data**

Problems when working with big data. General methods for processing large amounts of data. Distribution of data storage and processing in infrastructures.

#### **4 Databases**

SQL and NoSQL. Graph databases.

#### **Test 2**

#### **5 Text mining**

Methods of deep text analysis.

#### **6 Visualization of results**

Data visualization methods. Software Libraries.

#### **Final test**

## **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. Valentina Janev. Knowledge Graphs and Big Data Processing / Valentina Janev, Damien Graux, Hajira Jabeen, Emanuel Sallinger // — 2020. — 209 p. — ISBN978-3-030-53199-7. [Электронный ресурс]. <https://link.springer.com/book/10.1007/978-3-030-53199-7>
2. Vladik Kreinovich, Songsak Sriboonchitta, Nopasit Chakpitak. Predictive Econometrics and Big Data. Springer, Cham, 2018. eBook ISBN 978-3-319-70942-0. Текст электронный. <https://link.springer.com/book/10.1007/978-3-319-70942-0>

### Дополнительная:

1. Holdaway K. R. Harness oil and gas big data with analytics:optimize exploration and production with data driven models/Keith R. Holdaway.-Hoboken:John Wiley and sons,2014, ISBN 9781118779316.-364.- Библиогр. в конце глав. - Указ.: с. 351

## **9. Перечень ресурсов сети Интернет, необходимых для освоения дисциплины**

<https://ocw.mit.edu/courses/14-387-applied-econometrics-mostly-harmless-big-data-fall-2014/> Applied Econometrics: Mostly Harmless Big Data

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

Образовательный процесс по дисциплине **Mathematical Methods and Technologies for Processing Big Data** предполагает использование следующего программного обеспечения и информационных справочных систем:

- 1) presentation materials (slides on the topics of lectures);
- 2) on-line access to the Electronic library system (ELS);
- 3) access to the electronic information and educational environment of the University; 4) Internet services and electronic resources.

Free software: WPS Office - Office automation system

При освоении материала и выполнении заданий по дисциплине рекомендуется использование материалов, размещенных в Личных кабинетах обучающихся ЕТИС ПГНИУ ([student.psu.ru](http://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**

1. For lectures - a classroom with the presentation equipment (projector, screen, computer/laptop) and the suitable software, chalkboard (and) or whiteboard.
2. Practices - a computer class equipped with personal computers and appropriate software. The composition of the equipment is defined in the Passport of the computer class.
3. For self-directed study – a classroom for independent work that is equipped with computer hardware and access to the Internet and thereby to the electronic educational environment of the university. Halls of PSU Scientific Library.
4. For the current or interim knowledge assessment – a classroom with the presentation equipment (projector, screen, computer/laptop) and the suitable software, chalkboard (and) or whiteboard

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  
Mathematical Methods and Technologies for Processing Big Data**

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

**ОПК.9**

**Способен использовать информацию, методы и программные средства ее сбора, обработки и анализа для информационно-аналитической поддержки принятия управленческих решений**

<b>Компетенция (индикатор)</b>	<b>Планируемые результаты обучения</b>	<b>Критерии оценивания результатов обучения</b>
<p><b>ОПК.9.2</b> Использует методы и программные средства сбора, обработки и анализа информации с целью информационно-аналитической поддержки принятия управленческих решений</p>	<p>Knows the methods and software tools for collecting, processing and analyzing information. Able to collect, process and analyze information for the purpose of information and analytical support for managerial decision-making. Owns methods of working with information to support managerial decision-making.</p>	<p align="center"><b>Неудовлетворител</b></p> <p>Does not know the methods and software tools for collecting, processing and analyzing information. Does not know how to collect, process and analyze information for the purpose of information and analytical support for managerial decision-making. Does not own methods of working with information to support managerial decision-making.</p> <p align="center"><b>Удовлетворительн</b></p> <p>Poor knowledge of methods and software tools for collecting, processing and analyzing information. Makes gross mistakes in the process of collecting, processing or analyzing information for the purpose of information and analytical support for making managerial decisions. Owns at a satisfactory level the methods of working with information to support managerial decision-making.</p> <p align="center"><b>Хорошо</b></p> <p>Knows the methods and software tools for collecting, processing and analyzing information. Makes minor mistakes in the process of collecting, processing or analyzing information for the purpose of information and analytical support for managerial decision-making. Owns methods of working with information to support managerial decision-making.</p>

Компетенция (индикатор)	Планируемые результаты обучения	Критерии оценивания результатов обучения
		<p style="text-align: center;"><b>Отлично</b></p> <p>Knows at a high level the methods and software tools for collecting, processing and analyzing information.</p> <p>Able to collect, process and analyze information for the purpose of information and analytical support for managerial decision-making.</p> <p>He is fluent in methods of working with information to support managerial decision-making.</p>

### ПК.3

#### Способен к планированию и организации аналитических работ, в том числе, с использованием технологий больших данных

Компетенция (индикатор)	Планируемые результаты обучения	Критерии оценивания результатов обучения
<p><b>ПК.3.2</b> Определяет состав группы для проведения анализа данных, в том числе, больших данных</p>	<p>Knows methods of creating groups for data analysis.</p> <p>Knows how to determine the composition of the group for data analysis, including big data.</p> <p>Has the skills to determine the composition of the group for data analysis.</p>	<p style="text-align: center;"><b>Неудовлетворител</b></p> <p>Doesn't know how to create groups for data analysis.</p> <p>Does not know how to determine the composition of the group for data analysis, including big data.</p> <p>Does not have the skills to determine the composition of the group for data analysis.</p> <p style="text-align: center;"><b>Удовлетворительн</b></p> <p>Poor knowledge of methods for creating groups for data analysis.</p> <p>Makes gross mistakes when determining the composition of the group for data analysis, including big data.</p> <p>Has a satisfactory level of skills in determining the composition of the group for data analysis.</p> <p style="text-align: center;"><b>Хорошо</b></p> <p>Knows methods of creating groups for data analysis.</p> <p>Makes minor mistakes in determining the composition of the group for data analysis, including big data.</p> <p>Has the skills to determine the composition of the group for data analysis.</p> <p style="text-align: center;"><b>Отлично</b></p>

<b>Компетенция (индикатор)</b>	<b>Планируемые результаты обучения</b>	<b>Критерии оценивания результатов обучения</b>
		<p style="text-align: center;"><b>Отлично</b></p> <p>Knows at a high level methods of creating groups for data analysis.</p> <p>Knows how to determine the composition of the group for data analysis, including big data.</p> <p>He has excellent skills in determining the composition of the group for data analysis.</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
<b>Входной контроль</b>	Enter control <b>Входное тестирование</b>	Knowledge in the field of mathematics, logic and programming is checked.
<b>ПК.3.2</b> Определяет состав группы для проведения анализа данных, в том числе, больших данных <b>ОПК.9.2</b> Использует методы и программные средства сбора, обработки и анализа информации с целью информационно-аналитической поддержки принятия управленческих решений	Test 1 <b>Письменное контрольное мероприятие</b>	Knowledge of the basic concepts related to working with big data is tested.
<b>ПК.3.2</b> Определяет состав группы для проведения анализа данных, в том числе, больших данных <b>ОПК.9.2</b> Использует методы и программные средства сбора, обработки и анализа информации с целью информационно-аналитической поддержки принятия управленческих решений	Test 2 <b>Письменное контрольное мероприятие</b>	Knowledge and skills related to working with big data are tested. Knowledge of big data storage methods is tested.

<b>Компетенция (индикатор)</b>	<b>Мероприятие текущего контроля</b>	<b>Контролируемые элементы результатов обучения</b>
<b>ПК.3.2</b> Определяет состав группы для проведения анализа данных, в том числе, больших данных <b>ОПК.9.2</b> Использует методы и программные средства сбора, обработки и анализа информации с целью информационно-аналитической поддержки принятия управленческих решений	Final test <b>Итоговое контрольное мероприятие</b>	The level of knowledge on all studied material is checked.

### **Спецификация мероприятий текущего контроля**

#### **Enter control**

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

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

Максимальный балл, выставляемый за мероприятие промежуточной аттестации: **0**

Проходной балл: **0**

<b>Показатели оценивания</b>	<b>Баллы</b>
There are basic knowledges of linear algebra, mathematical analysis, optimization theory.	4
Have basic knowledge in programming.	3
There are basic knowledges of formal logic.	3

#### **Test 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**

<b>Показатели оценивания</b>	<b>Баллы</b>
Knows the basic concepts related to big data.	10
Knows and can use machine learning methods based on big data.	10
Knows the methods used in the process of processing big data.	10

#### **Test 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**

<b>Показатели оценивания</b>	<b>Баллы</b>
------------------------------	--------------

Knows the basic methods of processing large amounts of data.	10
Knows the types of databases suitable for working with big data.	10
Knows methods of distributed data storage and processing.	10

### **Final test**

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**

<b>Показатели оценивания</b>	<b>Баллы</b>
Knows the basic concepts related to big data.	10
Knows methods of analyzing large amounts of data and visual presentation of the results.	10
Knows how to store large amounts of data.	10
Knows and knows how to use methods for processing large amounts of data.	10