

МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ

РОССИЙСКОЙ ФЕДЕРАЦИИ

**Федеральное государственное автономное
образовательное учреждение высшего образования
«Национальный исследовательский Нижегородский государственный университет
им. Н.И. Лобачевского»**

**Институт информационных технологий,
математики и механики**

УТВЕРЖДЕНО
решением Ученого совета ННГУ
протокол от
«30» ноября 2022 г. № 13

Рабочая программа дисциплины

Инженерия больших данных

Уровень высшего образования
магистратура

Направление подготовки
090404 Программная инженерия

Профиль подготовки
Технологии цифровой трансформации

Форма обучения
Очная

Нижний Новгород
2023

1. Место и цели дисциплины в структуре ОПОП

Дисциплина «Б1.О.07, Инженерия больших данных» относится к обязательной части Блока 1 «Дисциплины (модули)» направления подготовки 09.04.04 «Программная инженерия» профиля подготовки «Технологии цифровой трансформации». Дисциплина преподается в 1 семестре. Трудоемкость дисциплины составляет 6 зачетных единиц, 216 час., экзамен.

№ варианта	Место дисциплины в учебном плане образовательной программы	Стандартный текст для автоматического заполнения в конструкторе РПД
2	Блок 1. Дисциплины (модули) Обязательная часть	Дисциплина «Б1.О.07, Инженерия больших данных» относится к обязательной части ООП направления подготовки 09.04.04 «Программная инженерия».

Цель освоения дисциплины

Цель данного курса состоит в формировании концептуальных представления об основных принципах работе с методами Data Mining. Курс включает общую концепцию проведения исследований с помощью Data Mining, предварительную обработку данных, классификацию, регрессию, временные ряды, нейронные сети, основы text mining, принципы автоматизации анализа данных. Практическая направленность курса состоит в том, чтобы научить студентов проводить самостоятельные исследования методами Data Mining и эффективно пользоваться результатами уже готовых статистических исследований.

2. Планируемые результаты обучения по дисциплине, соотнесенные с планируемыми результатами освоения образовательной программы (компетенциями выпускников)

№ п/п	Код компетенции*	Содержание компетенции	Планируемые результаты обучения	Наименование оценочного средства
1	ОПК-7	Способен применять при решении профессиональных задач методы и средства получения, хранения, переработки и трансляции информации посредством современных компьютерных технологий, в том числе, в глобальных компьютерных сетях	ОПК-7.1. Знать методы и средства получения, хранения, переработки и трансляции информации посредством современных компьютерных технологий, в том числе, в глобальных компьютерных сетях	Собеседование
			ОПК-7.2. Уметь применять методы и средства получения, хранения, переработки и трансляции информации посредством современных компьютерных технологий, в том числе, в глобальных компьютерных сетях	Собеседование Тест
			ОПК-7.3. Иметь навыки методы и средства получения, хранения, переработки и трансляции информации посредством современных компьютерных технологий, в том числе, в глобальных компьютерных	Собеседование Задание

			сетях	
--	--	--	-------	--

3. Структура и содержание дисциплины

3.1. Трудоемкость дисциплины

Объем дисциплины (модуля) составляет

6 зачетные единицы, всего **216** час., из которых

50 час. составляет **контактная** работа обучающегося с преподавателем:

32 час. занятия лекционного типа,

16 час. занятия семинарского типа,

2 час мероприятия промежуточной аттестации

166 часов составляет **самостоятельная** работа обучающегося (в т.ч. включая 36 час. подготовки к экзамену).

3.2 Содержание дисциплины

Наименование и краткое содержание разделов и тем дисциплины, форма промежуточной аттестации по дисциплине	Всего (часы)	в том числе					Самостоятельная работа студента часы
		контактная работа (работа во взаимодействии с преподавателем), часы					
		из них					
		Занятия лекционного	Занятия семинарского	Лабораторные работы	Консультации индивидуальные	Всего контактных	
Тема 1. Введение в большие данные.	19	4				8	12
Тема 2. MapReduce – теория и практика.	17		2			8	10
Тема 3. Поточковая обработка данных.	24	4	2			10	16
Тема 4. BigTable-подобные хранилища, HBase	24	4	2			10	16
Тема 5. Dynamo-подобные хранилища, Cassandra	26	4	2			11	17
Тема 6. Специализированные хранилища.	26	4	2			11	17
Тема 7. SQL over BigData	26	4	2			12	18
Тема 8. Workflow Engines & Scheduling	26	4	2			12	18
Тема 9. Архитектура систем обработки данных.	26	4	2			12	18

В т.ч. текущий контроль	2					2	
Итого:	216	32	16	0		50	166
Промежуточная аттестация - Экзамен							

Практические занятия (семинарские занятия /лабораторные работы) организуются, в том числе в форме практической подготовки, которая предусматривает участие обучающихся в выполнении отдельных элементов работ, связанных с будущей профессиональной деятельностью.

Практическая подготовка предусматривает: подготовку к вопросам для собеседования, подготовку к тестовым вопросам, выполнение практических заданий (лабораторных работ).

На проведение практических занятий (семинарских занятий /лабораторных работ) в форме практической подготовки отводится 16 часов.

Практическая подготовка направлена на формирование и развитие:

- практических навыков в соответствии с профилем ОП: создание и сопровождение архитектуры программных средств, разработка и тестирование программного обеспечения;
- компетенций – ОПК-7.

Текущий контроль успеваемости реализуется в формах опросов на занятиях семинарского типа.

Промежуточная аттестация проходит в традиционных формах (экзамен).

4. Учебно-методическое обеспечение самостоятельной работы студентов

Самостоятельная работа студентов включает:

- Изучение лекционных материалов.
- Изучение учебной литературы (см. перечень образовательных материалов).
- Выполнение заданий (лабораторных работ)
- Подготовка к экзамену

Контрольные вопросы и задания для проведения текущего контроля и промежуточной аттестации по итогам освоения дисциплины приведены в п. 5.2.

Для обеспечения самостоятельной работы обучающихся используется открытый электронный онлайн-курс (Наука о данных и аналитика больших объемов данных, <https://online.edu.ru/public/course?faces-redirect=true&cid=3792>), созданный в системе открытых онлайн-курсов российских вузов <https://online.edu.ru/>.

5. Фонд оценочных средств для промежуточной аттестации по дисциплине

5.1. Описание шкал оценивания результатов обучения по дисциплине

Уровень сформированности компетенций (индикатора достижения компетенций)	Шкала оценивания сформированности компетенций						
	плохо	неудовлетворительно	удовлетворительно	хорошо	очень хорошо	отлично	превосходно
	Не зачтено		зачтено				
<u>Знания</u>	Отсутствие знаний теоретического материала. Невозможность оценить полноту знаний вследствие отказа обучающегося от ответа	Уровень знаний ниже минимальных требований. Имели место грубые ошибки.	Минимально допустимый уровень знаний. Допущено много негрубых ошибок.	Уровень знаний в объеме, соответствующем программе подготовки. Допущено несколько негрубых ошибок	Уровень знаний в объеме, соответствующем программе подготовки. Допущено несколько незначительных ошибок	Уровень знаний в объеме, соответствующем программе подготовки, без ошибок.	Уровень знаний в объеме, превышающем программу подготовки.
<u>Умения</u>	Отсутствие минимальных умений. Невозможность оценить наличие умений вследствие отказа обучающегося от ответа	При решении стандартных задач не продемонстрированы основные умения. Имели место грубые ошибки.	Продemonстрированы основные умения. Решены типовые задачи с негрубыми ошибками. Выполнены все задания, но не в полном объеме.	Продemonстрированы все основные умения. Решены все основные задачи с негрубыми ошибками. Выполнены все задания, в полном объеме, но некоторые с недочетами.	Продemonстрированы все основные умения. Решены все основные задачи. Выполнены все задания, в полном объеме, но некоторые с недочетами.	Продemonстрированы все основные умения, решены все основные задачи с отдельными незначительными недочетами, выполнены все задания в полном объеме.	Продemonстрированы все основные умения. Решены все основные задачи. Выполнены все задания, в полном объеме без недочетов
<u>Навыки</u>	Отсутствие владения материалом. Невозможность оценить наличие навыков вследствие отказа обучающегося от ответа	При решении стандартных задач не продемонстрированы базовые навыки. Имели место грубые ошибки.	Имеется минимальный набор навыков для решения стандартных задач с некоторыми недочетами	Продemonстрированы базовые навыки при решении стандартных задач с некоторыми недочетами	Продemonстрированы базовые навыки при решении стандартных задач без ошибок и недочетов.	Продemonстрированы навыки при решении нестандартных задач без ошибок и недочетов.	Продemonстрирован творческий подход к решению нестандартных задач

Шкала оценки при промежуточной аттестации

Оценка		Уровень подготовки
зачтено	Превосходно	Все компетенции (части компетенций), на формирование которых направлена дисциплина, сформированы на уровне не ниже «превосходно»
	Отлично	Все компетенции (части компетенций), на формирование которых направлена дисциплина, сформированы на уровне не ниже «отлично», при этом хотя бы одна компетенция

		сформирована на уровне «отлично»
	Очень хорошо	Все компетенции (части компетенций), на формирование которых направлена дисциплина, сформированы на уровне не ниже «очень хорошо», при этом хотя бы одна компетенция сформирована на уровне «очень хорошо»
	Хорошо	Все компетенции (части компетенций), на формирование которых направлена дисциплина, сформированы на уровне не ниже «хорошо», при этом хотя бы одна компетенция сформирована на уровне «хорошо»
	Удовлетворительно	Все компетенции (части компетенций), на формирование которых направлена дисциплина, сформированы на уровне не ниже «удовлетворительно», при этом хотя бы одна компетенция сформирована на уровне «удовлетворительно»
не зачтено	Неудовлетворительно	Хотя бы одна компетенция сформирована на уровне «неудовлетворительно», ни одна из компетенций не сформирована на уровне «плохо»
	Плохо	Хотя бы одна компетенция сформирована на уровне «плохо»

5.2 Типовые контрольные задания или иные материалы, необходимые для оценки результатов обучения

5.2.1 Контрольные вопросы к экзамену

Вопрос	Код компетенции (согласно РПД)
1) Понятие Больших данных.	ОПК-7
2) Особенности сбора, хранения, обработки и анализа Больших данных	ОПК-7
3) Требования к распределенным информационным системам	ОПК-7
4) Средства построения распределенных информационных систем	ОПК-7
5) Технология Map-Reduce	ОПК-7
6) Система Apache Hadoop	ОПК-7
7) Базы данных NoSQL. Особенности, классификация	ОПК-7
8) Возможности NoSQL-баз данных по обеспечению целостности, доступности скорости обработки информации. CAP-теорема.	ОПК-7
9) Способы репликации и кластеризации баз данных	ОПК-7
10) Документно-ориентированные базы данных	ОПК-7
11) Возможности СУБД MongoDB	ОПК-7
12) Работа с документно-ориентированными БД на языке JSON.	ОПК-7

5.2.2. Типовые задания для оценки сформированности компетенции ОПК-7

Студентам предлагаются следующие задания (лабораторные работы) для оценивания результатов обучения в виде знаний ОПК-7.1, умений ОПК-7.2

Лабораторная работа №1

Средства построения распределенных информационных систем для BigData. Обзор возможностей. Выбор инструментария для разработки.

Лабораторная работа №2

Изучение и конфигурирование программного комплекса Apache Hadoop. Размещение набора данных по заданной тематике. Построение поисковых запросов на языке Pig

Лабораторная работа №3

Изучение основных классов NoSQL СУБД, графовых, мультиколоночных, документо-ориентированных, типа "имя=значение". Проектирование и разработка графовой базы данных в СУБД Neo4j на заданную тему. Поисковые запросы на языке Cypher

5.2.3. Типовые тестовые задания для оценки сформированности компетенции ОПК-7

Вариант 1

Понятие Больших данных.

Особенности сбора, хранения, обработки и анализа Больших данных

Требования к распределенным информационным системам

Вариант 2

Средства построения распределенных информационных систем

Технология Map-Reduce

Система Apache Hadoop

Вариант 3

Базы данных NoSQL. Особенности, классификация

Возможности NoSQL-баз данных по обеспечению целостности, доступности скорости обработки информации. CAP-теорема.

Способы репликации и кластеризации баз данных

6. Учебно-методическое и информационное обеспечение дисциплины

а. Основная литература

1. Л. А. Демидова , Е. В. Никульчев , Ю. С. Соколова. КЛАССИФИКАЦИЯ БОЛЬШИХ ДАННЫХ: ИСПОЛЬЗОВАНИЕ SVM-АНСАМБЛЕЙ И SVM-КЛАССИФИКАТОРОВ С МОДИФИЦИРОВАННЫМ РОЕВЫМ АЛГОРИТМОМ. [Электронный ресурс]. - URL: <https://e.lanbook.com/reader/journalArticle/304917/#1>
2. Адлер Ю.П., Черных Е.А. Статистическое управление процессами. «Большие данные»: учебное пособие. [Электронный ресурс]. - URL: <https://e.lanbook.com/book/93652>

б. Дополнительная литература

3. Зайцев К.С. Применение методов Data Mining для поддержки процессов управления ИТ-услугами: учебное пособие [Электронный ресурс]. - URL: <https://e.lanbook.com/book/75805>
4. О.Л. Голицына, Н.В. Максимов, И.И. Попов. Базы данных. <https://znanium.com/catalog/product/944926>

с. Программное обеспечение и Интернет-ресурсы

1. Операционные системы семейства MicrosoftWindows, лицензия по подписке MicrosoftImagine.
2. • Среда разработки семейства MicrosoftVisualStudio, лицензия по подписке MicrosoftImagine.
3. Сайт «Открытые системы» <http://www.osp.ru>.
4. Сайт «CIT Forum» <http://www.citforum.ru>

7. Материально-техническое обеспечение дисциплины

Учебная аудитория для проведения занятий лекционного типа, занятий семинарского типа, групповых и индивидуальных консультаций, текущего контроля и промежуточной аттестации, укомплектованная учебной мебелью, доской.

Лаборатория современных информационных технологий, учебная аудитория для групповых и индивидуальных консультаций, укомплектованная учебной мебелью, доской маркерной, оснащенная компьютерной техникой с возможностью подключения к сети «Интернет».

Программа составлена в соответствии с требованиями ОС ВО ННГУ с учетом рекомендаций ФГОС ВО по направлению подготовки 090404 Программная инженерия

Автор: к.ф.-м.н., доцент кафедры МОСТ, Шапошников Д.Е.

Рецензент: к.т.н., доцент кафедры ТВиАД, Борисов Н.А.

Заведующий кафедрой: д.ф.-м.н, проф. заведующий кафедрой МОСТ Стронгин Р.Г.

Программа одобрена на заседании методической комиссии института информационных технологий, математики и механики от 30 ноября 2022 года, протокол № 3.