

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)

Одобрено на заседании УМС
ИАТЭ НИЯУ МИФИ Протокол
от 30.08.2022 №2-8/2022

**ФОНД
ОЦЕНОЧНЫХ СРЕДСТВ**

ПО УЧЕБНОЙ ДИСЦИПЛИНЕ

«Большие данные»

Направление подготовки:	09.04.01 «Информатика и вычислительная техника»
Профиль:	«Большие данные и машинное обучение для атомной энергетики»
Квалификация (степень) выпускника:	магистр
Форма обучения:	очная

2022 г.

Фонд оценочных средств составлен в соответствии с образовательным стандартом высшего образования НИЯУ МИФИ по направлению подготовки 09.04.01 Информатика и вычислительная техника.

Фонд оценочных средств составил:

_____ С.В. Грицюк, доцент, к.т.н.

Фонд оценочных средств рассмотрен на заседании отделения интеллектуальных кибернетических систем (О)

(протокол № _____ от « _____ » _____ 2022 г.)

Начальник отделения интеллектуальных кибернетических систем

_____ С.О. Старков

« _____ » _____ 2022 г.

Область применения

Фонд оценочных средств (ФОС) – является неотъемлемой частью учебно-методического комплекса учебной дисциплины «Большие данные» и предназначен для контроля и оценки образовательных достижений обучающихся, освоивших программу данной дисциплины.

Цели и задачи фонда оценочных средств

Целью Фонда оценочных средств является установление соответствия уровня подготовки обучающихся требованиям федерального государственного образовательного стандарта.

Для достижения поставленной цели Фондом оценочных средств по дисциплине «Большие данные» решаются следующие задачи:

- контроль и управление процессом приобретения обучающимися знаний, умений и навыков, предусмотренных в рамках данного курса;
- контроль и оценка степени освоения компетенций, предусмотренных в рамках данного курса;
- обеспечение соответствия результатов обучения задачам будущей профессиональной деятельности через совершенствование традиционных и внедрение инновационных методов обучения в образовательный процесс в рамках данного курса.

1. Перечень компетенций с указанием этапов их формирования в процессе освоения образовательной программы

1.1. Перечень планируемых результатов обучения по дисциплине, соотнесенных с планируемыми результатами освоения образовательной программы

В результате освоения ОПОП магистратуры обучающийся должен овладеть следующими результатами обучения по дисциплине:

Коды компетенций	Результаты освоения ООП Содержание компетенций	Перечень планируемых результатов обучения по дисциплине
ПК-1	способен применять научно обоснованные перспективные методы исследования и решать задачи на основе знания мировых тенденций развития вычислительной техники и информационных технологий с внедрением результатов исследований в реальный сектор экономики	<p>Знать: существующие подходы и технологии работы с большими объемами данных, их особенности, предпосылки возникновения, сильные и слабые стороны.</p> <p>Уметь: составлять список требований к разрабатываемой информационной системе и технологиям работы с большими данными.</p> <p>Владеть: навыками создания новых алгоритмов и систем обработки больших объемов данных.</p>
СПК-1	способен использовать и развивать методы научных исследований и инструментарий в области интеллектуального анализа данных	<p>Знать: подходы к обработке данных, их особенности.</p> <p>Уметь: использовать современные информационные технологии поиска, анализа и обработки данных для последующего использования в рамках информационной системы.</p> <p>Владеть: навыками практического применения методов и современных компьютерных технологий поиска и анализа информации.</p>
		<p>Знать: теоретические основы больших данных; возможности современных технологий и их использование для решения прикладных задач в различных областях профессиональной деятельности.</p> <p>Уметь: использовать технологии и системы анализа больших данных</p> <p>Владеть: технологиями анализа больших данных</p>
		<p>Знать: актуальные подходы и технологии в области обработки больших объемов данных</p> <p>Уметь: применять современные подходы и технологии больших данных в рамках реальных создаваемых систем</p> <p>Владеть: навыками проектирования и создания систем на базе современных технологий обработки больших данных</p>

		<p>Знать: современные разработки в области интеллектуального анализа данных</p> <p>Уметь: создавать новые подходы/инструменты в области интеллектуального анализа данных</p> <p>Владеть: навыками поиска и интеллектуального анализа данных</p>
--	--	--

1.2. Этапы формирования компетенций в процессе освоения ООП магистратуры

Компоненты компетенций, как правило, формируются при изучении нескольких дисциплин, а также в немалой степени в процессе прохождения практик, НИР и во время самостоятельной работы обучающегося. Выполнение и защита ВКР являются видом учебной деятельности, который завершает процесс формирования компетенций.

Место дисциплины и соответствующий этап формирования компетенций в целостном процессе подготовки по образовательной программе можно определить по матрице компетенций, которая приводится в Приложении.

Этапы формирования компетенции в процессе освоения дисциплины:

- **начальный** этап – на этом этапе формируются знаниевые и инструментальные основы компетенции, осваиваются основные категории, формируются базовые умения. Студент воспроизводит термины, факты, методы, понятия, принципы и правила; решает учебные задачи по образцу;
- **основной** этап – знания, умения, навыки, обеспечивающие формирование компетенции, значительно возрастают, но еще не достигают итоговых значений. На этом этапе студент осваивает аналитические действия с предметными знаниями по дисциплине, способен самостоятельно решать учебные задачи, внося коррективы в алгоритм действий, осуществляя коррекцию в ходе работы, переносит знания и умения на новые условия;
- **завершающий** этап – на этом этапе студент достигает итоговых показателей по заявленной компетенции, то есть осваивает весь необходимый объем знаний, овладевает всеми умениями и навыками в сфере заявленной компетенции. Он способен использовать эти знания, умения, навыки при решении задач повышенной сложности и в нестандартных условиях.

Этапы формирования компетенций в ходе освоения дисциплины отражаются в тематическом плане (см.п. 4 рабочей программы дисциплины).

1.3. Паспорт фонда оценочных средств по дисциплине

№ п/п	Контролируемые разделы (темы) дисциплины (результаты по разделам)	Код контролируемой компетенции (или её части) / и ее формулировка	Наименование оценочного средства
1-8.	1. Введение в область больших данных 2. Введение в программирование на языке Scala 3. MapReduce, Hadoop и Apache Spark 4. Системы хранения 5. Экосистема Hadoop 6. Анализ естественного языка 7. Машинное обучение 8. Практические примеры из области больших данных	ПК-1 (знать, уметь, владеть) СПК-1 (знать, уметь, владеть)	Лабораторная работа №1 (демонстрация на компьютере выполненного проекта и защита работы в форме собеседования с преподавателем); Контрольная работа №1 (в форме письменных ответов и устного собеседования на теоретические вопросы); Экзамен (в форме письменных ответов и устного

			собеседования на теоретические вопросы)
9.	Анализ естественного языка на примере литературного произведения	ПК-1 (знать, уметь, владеть) СПК-1 (знать, уметь, владеть)	Лабораторная работа №2 (демонстрация на компьютере выполненного проекта и защита работы в форме собеседования с преподавателем)
10.	Spark SQL для анализа данных	ПК-1 (знать, уметь, владеть) СПК-1 (знать, уметь, владеть)	Лабораторная работа №3 (демонстрация на компьютере выполненного проекта и защита работы в форме собеседования с преподавателем)
11.	Основы машинного обучения на Spark ML	ПК-1 (знать, уметь, владеть) СПК-1 (знать, уметь, владеть)	Лабораторная работа №4 (демонстрация на компьютере выполненного проекта и защита работы в форме собеседования с преподавателем)

2. Описание показателей и критериев оценивания компетенций на различных этапах их формирования, описание шкал оценивания

Конечными результатами освоения программы дисциплины являются сформированные когнитивные дескрипторы «знать», «уметь», «владеть», расписанные по отдельным компетенциям, которые приведены в п.1.1. Формирование этих дескрипторов происходит в процессе изучения дисциплины по этапам в рамках различного вида учебных занятий и самостоятельной работы.

Выделяются три уровня сформированности компетенций на каждом этапе: пороговый, продвинутый и высокий.

Уровни	Содержательное описание уровня	Основные признаки выделения уровня	БРС, % освоения	ECTS/Пятибалльная шкала для оценки экзамена/зачета
Высокий <i>Все виды компетенций сформированы на высоком уровне в соответствии с целями и задачами дисциплины</i>	Творческая деятельность	<i>Включает нижестоящий уровень.</i> Студент демонстрирует свободное обладание компетенциями, способен применить их в нестандартных ситуациях: показывает умение самостоятельно принимать решение, решать проблему/задачу теоретического или прикладного характера на основе изученных методов, приемов, технологий	90-100	A/ Отлично/ Зачтено
Продвинутый <i>Все виды компетенций сформированы на продвинутом уровне в соответствии с целями и задачами дисциплины</i>	Применение знаний и умений в более широких контекстах учебной и профессиональной деятельности, нежели по образцу, большей долей самостоятельности и инициативы	<i>Включает нижестоящий уровень.</i> Студент может доказать владение компетенциями: демонстрирует способность собирать, систематизировать, анализировать и грамотно использовать информацию из самостоятельно найденных теоретических источников и иллюстрировать ими теоретические положения или обосновывать практику применения.	85-89	B/ Очень хорошо/ Зачтено
			75-84	C/ Хорошо/ Зачтено
Пороговый <i>Все виды компетенций сформированы на пороговом уровне</i>	Репродуктивная деятельность	Студент демонстрирует владение компетенциями в стандартных ситуациях: излагает в пределах задач курса теоретически и практически контролируемый материал.	65-74	D/Удовлетворительно/ Зачтено
			60-64	E/Посредственно /Зачтено
Ниже порогового	Отсутствие признаков порогового уровня: компетенции не сформированы. Студент не в состоянии продемонстрировать обладание компетенциями в стандартных ситуациях.		0-59	Неудовлетворительно/ Зачтено

Оценивание результатов обучения студентов по дисциплине осуществляется по регламенту текущего контроля и промежуточной аттестации.

Критерии оценивания компетенций на каждом этапе изучения дисциплины для каждого вида оценочного средства и приводятся в п. 4 ФОС. Итоговый уровень сформированности компетенции при изучении дисциплины определяется по таблице. При этом следует понимать, что граница между уровнями для конкретных результатов освоения образовательной программы может смещаться.

Уровень сформированности компетенции	Текущий контроль	Промежуточная аттестация
высокий	высокий	высокий
	<i>продвинутой</i>	<i>высокий</i>
	<i>высокий</i>	<i>продвинутой</i>
продвинутой	<i>пороговый</i>	<i>высокий</i>
	<i>высокий</i>	<i>пороговый</i>
	продвинутой	продвинутой
	<i>продвинутой</i>	<i>пороговый</i>
	<i>пороговый</i>	<i>продвинутой</i>
пороговый	пороговый	пороговый
ниже порогового	пороговый	ниже порогового
	ниже порогового	-

3. Методические материалы, определяющие процедуры оценивания знаний, умений, навыков или опыта деятельности, характеризующих этапы формирования компетенций.

Рейтинговая оценка знаний является интегральным показателем качества теоретических и практических знаний и навыков студентов по дисциплине и складывается из оценок, полученных в ходе текущего контроля и промежуточной аттестации.

Текущий контроль в семестре проводится с целью обеспечения своевременной обратной связи, для коррекции обучения, активизации самостоятельной работы студентов.

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

Текущий контроль осуществляется два раза в семестр: контрольная точка № 1 (*КТ № 1*) и контрольная точка № 2 (*КТ № 2*).

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

Вид контроля	Этап рейтинговой системы Оценочное средство	Балл	
		Минимум	Максимум
Текущий	Контрольная точка № 1	18	30
	Лабораторная работа №1	6	10
	Лабораторная работа №2	6	10
	Контрольная работа №1 (2 вопроса – 5 и 5 баллов)	6	10
	Контрольная точка № 2	18	30
	Лабораторная работа №3	9	15
	Лабораторная работа №4	9	15
Промежуточный	Экзамен	24	40
ИТОГО по дисциплине		60	100

За несвоевременную сдачу любого из указанных в таблице оценочных средств оценка может быть снижена от 1 до 2 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. Ответ оценивается преподавателем.

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 и систематичность полученных обучающимся теоретических знаний, приобретенных навыков самостоятельной работы.

Оценка сформированных 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.

4. Типовые контрольные задания или иные материалы, необходимые для оценки знаний, умений, навыков

4.1. *Комплект экзаменационных билетов по дисциплине*

МИНИСТЕРСТВО ОБРАЗОВАНИЯ И НАУКИ РОССИЙСКОЙ ФЕДЕРАЦИИ
ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ АВТОНОМНОЕ ОБРАЗОВАТЕЛЬНОЕ УЧРЕЖДЕНИЕ
ВЫСШЕГО ПРОФЕССИОНАЛЬНОГО ОБРАЗОВАНИЯ

«Национальный исследовательский ядерный университет «МИФИ»

Обнинский институт атомной энергетики –

филиал федерального государственного автономного образовательного учреждения высшего профессионального образования «Национальный исследовательский ядерный университет «МИФИ»

Отделение Интеллектуальных кибернетических систем

Направление/ Специальность	090401 «Информатика и вычислительная техника»
Профиль/ Специализация	«Большие данные и машинное обучение для атомной энергетики»
Дисциплина	«Большие данные»

ЭКЗАМЕНАЦИОННЫЙ БИЛЕТ №1

1. Apache Spark. Работа в кластере.
2. Масштабируемость. Виды.

Составитель _____ С.В. Грицюк
(подпись)

Руководитель направления _____ С.О. Старков
(подпись)

« ____ » _____ 2021 г.

МИНИСТЕРСТВО ОБРАЗОВАНИЯ И НАУКИ РОССИЙСКОЙ ФЕДЕРАЦИИ
ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ АВТОНОМНОЕ ОБРАЗОВАТЕЛЬНОЕ УЧРЕЖДЕНИЕ
ВЫСШЕГО ПРОФЕССИОНАЛЬНОГО ОБРАЗОВАНИЯ

«Национальный исследовательский ядерный университет «МИФИ»

Обнинский институт атомной энергетики –

филиал федерального государственного автономного образовательного учреждения высшего профессионального образования «Национальный исследовательский ядерный университет «МИФИ»

Отделение Интеллектуальных кибернетических систем

Направление/ Специальность	090401 «Информатика и вычислительная техника»
Профиль/ Специализация	«Большие данные и машинное обучение для атомной энергетики»
Дисциплина	«Большие данные»

ЭКЗАМЕНАЦИОННЫЙ БИЛЕТ №2

1. NoSQL. Колоночные БД.
2. CAP теорема.

Составитель _____ С.В. Грицюк
(подпись)

Руководитель направления _____ С.О. Старков
(подпись)

« ____ » _____ 2021 г.

МИНИСТЕРСТВО ОБРАЗОВАНИЯ И НАУКИ РОССИЙСКОЙ ФЕДЕРАЦИИ
ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ АВТОНОМНОЕ ОБРАЗОВАТЕЛЬНОЕ УЧРЕЖДЕНИЕ
ВЫСШЕГО ПРОФЕССИОНАЛЬНОГО ОБРАЗОВАНИЯ

«Национальный исследовательский ядерный университет «МИФИ»

Обнинский институт атомной энергетики –

филиал федерального государственного автономного образовательного учреждения высшего профессионального образования «Национальный исследовательский ядерный университет «МИФИ»

Отделение Интеллектуальных кибернетических систем

Направление/ Специальность	090401 «Информатика и вычислительная техника»
Профиль/ Специализация	«Большие данные и машинное обучение для атомной энергетики»
Дисциплина	«Большие данные»

ЭКЗАМЕНАЦИОННЫЙ БИЛЕТ №3

1. Apache Hadoop. Особенности. Плюсы и минусы.
2. Примеры применения машинного обучения (2-3).

Составитель _____ С.В. Грицюк
(подпись)

Руководитель направления _____ С.О. Старков
(подпись)

« ____ » _____ 2021 г.

МИНИСТЕРСТВО ОБРАЗОВАНИЯ И НАУКИ РОССИЙСКОЙ ФЕДЕРАЦИИ
ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ АВТОНОМНОЕ ОБРАЗОВАТЕЛЬНОЕ УЧРЕЖДЕНИЕ
ВЫСШЕГО ПРОФЕССИОНАЛЬНОГО ОБРАЗОВАНИЯ

«Национальный исследовательский ядерный университет «МИФИ»

Обнинский институт атомной энергетики –

филиал федерального государственного автономного образовательного учреждения высшего профессионального образования «Национальный исследовательский ядерный университет «МИФИ»

Отделение Интеллектуальных кибернетических систем

Направление/ Специальность	090401 «Информатика и вычислительная техника»
Профиль/ Специализация	«Большие данные и машинное обучение для атомной энергетики»
Дисциплина	«Большие данные»

ЭКЗАМЕНАЦИОННЫЙ БИЛЕТ №4

1. Apache Hadoop. HDFS. MapReduce.
2. Классификация и различия NoSQL хранилищ.

Составитель _____ С.В. Грицюк
(подпись)

Руководитель направления _____ С.О. Старков
(подпись)

« ____ » _____ 2021 г.

МИНИСТЕРСТВО ОБРАЗОВАНИЯ И НАУКИ РОССИЙСКОЙ ФЕДЕРАЦИИ
ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ АВТОНОМНОЕ ОБРАЗОВАТЕЛЬНОЕ УЧРЕЖДЕНИЕ
ВЫСШЕГО ПРОФЕССИОНАЛЬНОГО ОБРАЗОВАНИЯ

«Национальный исследовательский ядерный университет «МИФИ»

Обнинский институт атомной энергетики –

филиал федерального государственного автономного образовательного учреждения высшего профессионального образования «Национальный исследовательский ядерный университет «МИФИ»

Отделение Интеллектуальных кибернетических систем

Направление/ Специальность	090401 «Информатика и вычислительная техника»
Профиль/ Специализация	«Большие данные и машинное обучение для атомной энергетики»
Дисциплина	«Большие данные»

ЭКЗАМЕНАЦИОННЫЙ БИЛЕТ №5

1. NoSQL. Графовые БД.
2. Основы языка Scala. Особенности. Плюсы и минусы.

Составитель _____ С.В. Грицюк
(подпись)

Руководитель направления _____ С.О. Старков
(подпись)

« ____ » _____ 2021 г.

МИНИСТЕРСТВО ОБРАЗОВАНИЯ И НАУКИ РОССИЙСКОЙ ФЕДЕРАЦИИ
ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ АВТОНОМНОЕ ОБРАЗОВАТЕЛЬНОЕ УЧРЕЖДЕНИЕ
ВЫСШЕГО ПРОФЕССИОНАЛЬНОГО ОБРАЗОВАНИЯ

«Национальный исследовательский ядерный университет «МИФИ»

Обнинский институт атомной энергетики –

филиал федерального государственного автономного образовательного учреждения высшего профессионального образования «Национальный исследовательский ядерный университет «МИФИ»

Отделение Интеллектуальных кибернетических систем

Направление/ Специальность	090401 «Информатика и вычислительная техника»
Профиль/ Специализация	«Большие данные и машинное обучение для атомной энергетики»
Дисциплина	«Большие данные»

ЭКЗАМЕНАЦИОННЫЙ БИЛЕТ №6

1. Анализ естественного языка. Мотивация. Направления.
2. Классификация и различия NoSQL хранилищ.

Составитель _____ С.В. Грицюк
(подпись)

Руководитель направления _____ С.О. Старков
(подпись)

« ____ » _____ 2021 г.

МИНИСТЕРСТВО ОБРАЗОВАНИЯ И НАУКИ РОССИЙСКОЙ ФЕДЕРАЦИИ
ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ АВТОНОМНОЕ ОБРАЗОВАТЕЛЬНОЕ УЧРЕЖДЕНИЕ
ВЫСШЕГО ПРОФЕССИОНАЛЬНОГО ОБРАЗОВАНИЯ

«Национальный исследовательский ядерный университет «МИФИ»

Обнинский институт атомной энергетики –

филиал федерального государственного автономного образовательного учреждения высшего профессионального образования «Национальный исследовательский ядерный университет «МИФИ»

Отделение Интеллектуальных кибернетических систем

Направление/ Специальность	090401 «Информатика и вычислительная техника»
Профиль/ Специализация	«Большие данные и машинное обучение для атомной энергетики»
Дисциплина	«Большие данные»

ЭКЗАМЕНАЦИОННЫЙ БИЛЕТ №7

1. Машинное обучение. Deep Learning.
2. NoSQL и транзакции.

Составитель _____ С.В. Грицюк
(подпись)

Руководитель направления _____ С.О. Старков
(подпись)

« ____ » _____ 2021 г.

МИНИСТЕРСТВО ОБРАЗОВАНИЯ И НАУКИ РОССИЙСКОЙ ФЕДЕРАЦИИ
ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ АВТОНОМНОЕ ОБРАЗОВАТЕЛЬНОЕ УЧРЕЖДЕНИЕ
ВЫСШЕГО ПРОФЕССИОНАЛЬНОГО ОБРАЗОВАНИЯ

«Национальный исследовательский ядерный университет «МИФИ»

Обнинский институт атомной энергетики –

филиал федерального государственного автономного образовательного учреждения высшего профессионального образования «Национальный исследовательский ядерный университет «МИФИ»

Отделение Интеллектуальных кибернетических систем

Направление/ Специальность	090401 «Информатика и вычислительная техника»
Профиль/ Специализация	«Большие данные и машинное обучение для атомной энергетики»
Дисциплина	«Большие данные»

ЭКЗАМЕНАЦИОННЫЙ БИЛЕТ №8

1. Машинное обучение. Техники/задачи.
2. Основы языка Scala. Методы на коллекциях.

Составитель _____ С.В. Грицюк
(подпись)

Руководитель направления _____ С.О. Старков
(подпись)

« ____ » _____ 2021 г.

МИНИСТЕРСТВО ОБРАЗОВАНИЯ И НАУКИ РОССИЙСКОЙ ФЕДЕРАЦИИ
ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ АВТОНОМНОЕ ОБРАЗОВАТЕЛЬНОЕ УЧРЕЖДЕНИЕ
ВЫСШЕГО ПРОФЕССИОНАЛЬНОГО ОБРАЗОВАНИЯ

«Национальный исследовательский ядерный университет «МИФИ»

Обнинский институт атомной энергетики –

филиал федерального государственного автономного образовательного учреждения высшего профессионального образования «Национальный исследовательский ядерный университет «МИФИ»

Отделение Интеллектуальных кибернетических систем

Направление/ Специальность	090401 «Информатика и вычислительная техника»
Профиль/ Специализация	«Большие данные и машинное обучение для атомной энергетики»
Дисциплина	«Большие данные»

ЭКЗАМЕНАЦИОННЫЙ БИЛЕТ №9

1. Анализ естественного языка. Мотивация. Направления.
2. Шардинг.

Составитель _____ С.В. Грицюк
(подпись)

Руководитель направления _____ С.О. Старков
(подпись)

« ____ » _____ 2021 г.

МИНИСТЕРСТВО ОБРАЗОВАНИЯ И НАУКИ РОССИЙСКОЙ ФЕДЕРАЦИИ
ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ АВТОНОМНОЕ ОБРАЗОВАТЕЛЬНОЕ УЧРЕЖДЕНИЕ
ВЫСШЕГО ПРОФЕССИОНАЛЬНОГО ОБРАЗОВАНИЯ

«Национальный исследовательский ядерный университет «МИФИ»

Обнинский институт атомной энергетики –

филиал федерального государственного автономного образовательного учреждения высшего профессионального образования «Национальный исследовательский ядерный университет «МИФИ»

Отделение Интеллектуальных кибернетических систем

Направление/ Специальность	090401 «Информатика и вычислительная техника»
Профиль/ Специализация	«Большие данные и машинное обучение для атомной энергетики»
Дисциплина	«Большие данные»

ЭКЗАМЕНАЦИОННЫЙ БИЛЕТ №10

1. Apache Spark. Стэк технологий.
2. Масштабируемость. Виды.

Составитель _____ С.В. Грицюк
(подпись)

Руководитель направления _____ С.О. Старков
(подпись)

« ____ » _____ 2021 г.

МИНИСТЕРСТВО ОБРАЗОВАНИЯ И НАУКИ РОССИЙСКОЙ ФЕДЕРАЦИИ
ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ АВТОНОМНОЕ ОБРАЗОВАТЕЛЬНОЕ УЧРЕЖДЕНИЕ
ВЫСШЕГО ПРОФЕССИОНАЛЬНОГО ОБРАЗОВАНИЯ

«Национальный исследовательский ядерный университет «МИФИ»

Обнинский институт атомной энергетики –

филиал федерального государственного автономного образовательного учреждения высшего профессионального образования «Национальный исследовательский ядерный университет «МИФИ»

Отделение Интеллектуальных кибернетических систем

Направление/ Специальность	090401 «Информатика и вычислительная техника»
Профиль/ Специализация	«Большие данные и машинное обучение для атомной энергетики»
Дисциплина	«Большие данные»

ЭКЗАМЕНАЦИОННЫЙ БИЛЕТ №11

1. Машинное обучение. Категории.
2. Репликация.

Составитель _____ С.В. Грицюк
(подпись)

Руководитель направления _____ С.О. Старков
(подпись)

« ____ » _____ 2021 г.

МИНИСТЕРСТВО ОБРАЗОВАНИЯ И НАУКИ РОССИЙСКОЙ ФЕДЕРАЦИИ
ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ АВТОНОМНОЕ ОБРАЗОВАТЕЛЬНОЕ УЧРЕЖДЕНИЕ
ВЫСШЕГО ПРОФЕССИОНАЛЬНОГО ОБРАЗОВАНИЯ

«Национальный исследовательский ядерный университет «МИФИ»

Обнинский институт атомной энергетики –

филиал федерального государственного автономного образовательного учреждения высшего профессионального образования «Национальный исследовательский ядерный университет «МИФИ»

Отделение Интеллектуальных кибернетических систем

Направление/ Специальность	090401 «Информатика и вычислительная техника»
Профиль/ Специализация	«Большие данные и машинное обучение для атомной энергетики»
Дисциплина	«Большие данные»

ЭКЗАМЕНАЦИОННЫЙ БИЛЕТ №12

1. NoSQL. Документные БД.
2. Популярные NoSQL-дистрибутивы и их различия.

Составитель _____ С.В. Грицюк
(подпись)

Руководитель направления _____ С.О. Старков
(подпись)

« ____ » _____ 2021 г.

МИНИСТЕРСТВО ОБРАЗОВАНИЯ И НАУКИ РОССИЙСКОЙ ФЕДЕРАЦИИ
ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ АВТОНОМНОЕ ОБРАЗОВАТЕЛЬНОЕ УЧРЕЖДЕНИЕ
ВЫСШЕГО ПРОФЕССИОНАЛЬНОГО ОБРАЗОВАНИЯ

«Национальный исследовательский ядерный университет «МИФИ»

Обнинский институт атомной энергетики –

филиал федерального государственного автономного образовательного учреждения высшего профессионального образования «Национальный исследовательский ядерный университет «МИФИ»

Отделение Интеллектуальных кибернетических систем

Направление/ Специальность	090401 «Информатика и вычислительная техника»
Профиль/ Специализация	«Большие данные и машинное обучение для атомной энергетики»
Дисциплина	«Большие данные»

ЭКЗАМЕНАЦИОННЫЙ БИЛЕТ №13

1. Парадигма MapReduce.
2. Популярные Hadoop-дистрибутивы и их различия.

Составитель _____ С.В. Грицюк
(подпись)

Руководитель направления _____ С.О. Старков
(подпись)

« ____ » _____ 2021 г.

МИНИСТЕРСТВО ОБРАЗОВАНИЯ И НАУКИ РОССИЙСКОЙ ФЕДЕРАЦИИ
ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ АВТОНОМНОЕ ОБРАЗОВАТЕЛЬНОЕ УЧРЕЖДЕНИЕ
ВЫСШЕГО ПРОФЕССИОНАЛЬНОГО ОБРАЗОВАНИЯ

«Национальный исследовательский ядерный университет «МИФИ»

Обнинский институт атомной энергетики –

филиал федерального государственного автономного образовательного учреждения высшего профессионального образования «Национальный исследовательский ядерный университет «МИФИ»

Отделение Интеллектуальных кибернетических систем

Направление/ Специальность	090401 «Информатика и вычислительная техника»
Профиль/ Специализация	«Большие данные и машинное обучение для атомной энергетики»
Дисциплина	«Большие данные»

ЭКЗАМЕНАЦИОННЫЙ БИЛЕТ №14

1. Apache Spark. Работа в кластере.
2. Примеры работы с Большими данными (2-3).

Составитель _____ С.В. Грицюк
(подпись)

Руководитель направления _____ С.О. Старков
(подпись)

« ____ » _____ 2021 г.

МИНИСТЕРСТВО ОБРАЗОВАНИЯ И НАУКИ РОССИЙСКОЙ ФЕДЕРАЦИИ
ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ АВТОНОМНОЕ ОБРАЗОВАТЕЛЬНОЕ УЧРЕЖДЕНИЕ
ВЫСШЕГО ПРОФЕССИОНАЛЬНОГО ОБРАЗОВАНИЯ

«Национальный исследовательский ядерный университет «МИФИ»

Обнинский институт атомной энергетики –

филиал федерального государственного автономного образовательного учреждения высшего профессионального образования «Национальный исследовательский ядерный университет «МИФИ»

Отделение Интеллектуальных кибернетических систем

Направление/ Специальность	090401 «Информатика и вычислительная техника»
Профиль/ Специализация	«Большие данные и машинное обучение для атомной энергетики»
Дисциплина	«Большие данные»

ЭКЗАМЕНАЦИОННЫЙ БИЛЕТ №15

1. Apache Hadoop. HDFS. MapReduce.
2. Репликация.

Составитель _____ С.В. Грицюк
(подпись)

Руководитель направления _____ С.О. Старков
(подпись)

« ____ » _____ 2021 г.

МИНИСТЕРСТВО ОБРАЗОВАНИЯ И НАУКИ РОССИЙСКОЙ ФЕДЕРАЦИИ
ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ АВТОНОМНОЕ ОБРАЗОВАТЕЛЬНОЕ УЧРЕЖДЕНИЕ
ВЫСШЕГО ПРОФЕССИОНАЛЬНОГО ОБРАЗОВАНИЯ

«Национальный исследовательский ядерный университет «МИФИ»

Обнинский институт атомной энергетики –

филиал федерального государственного автономного образовательного учреждения высшего профессионального образования «Национальный исследовательский ядерный университет «МИФИ»

Отделение Интеллектуальных кибернетических систем

Направление/ Специальность	090401 «Информатика и вычислительная техника»
Профиль/ Специализация	«Большие данные и машинное обучение для атомной энергетики»
Дисциплина	«Большие данные»

ЭКЗАМЕНАЦИОННЫЙ БИЛЕТ №16

1. NoSQL. Ключ-значение.
2. Дистрибутив Cloudera. Сервисы в составе дистрибутива.

Составитель _____ С.В. Грицюк
(подпись)

Руководитель направления _____ С.О. Старков
(подпись)

« ____ » _____ 2021 г.

МИНИСТЕРСТВО ОБРАЗОВАНИЯ И НАУКИ РОССИЙСКОЙ ФЕДЕРАЦИИ
ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ АВТОНОМНОЕ ОБРАЗОВАТЕЛЬНОЕ УЧРЕЖДЕНИЕ
ВЫСШЕГО ПРОФЕССИОНАЛЬНОГО ОБРАЗОВАНИЯ

«Национальный исследовательский ядерный университет «МИФИ»

Обнинский институт атомной энергетики –

филиал федерального государственного автономного образовательного учреждения высшего профессионального образования «Национальный исследовательский ядерный университет «МИФИ»

Отделение Интеллектуальных кибернетических систем

Направление/ Специальность	090401 «Информатика и вычислительная техника»
Профиль/ Специализация	«Большие данные и машинное обучение для атомной энергетики»
Дисциплина	«Большие данные»

ЭКЗАМЕНАЦИОННЫЙ БИЛЕТ №17

1. Apache Spark. Особенности. Плюсы и минусы.
2. Примеры применения машинного обучения (2-3).

Составитель _____ С.В. Грицюк
(подпись)

Руководитель направления _____ С.О. Старков
(подпись)

« ____ » _____ 2021 г.

МИНИСТЕРСТВО ОБРАЗОВАНИЯ И НАУКИ РОССИЙСКОЙ ФЕДЕРАЦИИ
ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ АВТОНОМНОЕ ОБРАЗОВАТЕЛЬНОЕ УЧРЕЖДЕНИЕ
ВЫСШЕГО ПРОФЕССИОНАЛЬНОГО ОБРАЗОВАНИЯ

«Национальный исследовательский ядерный университет «МИФИ»

Обнинский институт атомной энергетики –

филиал федерального государственного автономного образовательного учреждения высшего профессионального образования «Национальный исследовательский ядерный университет «МИФИ»

Отделение Интеллектуальных кибернетических систем

Направление/ Специальность	090401 «Информатика и вычислительная техника»
Профиль/ Специализация	«Большие данные и машинное обучение для атомной энергетики»
Дисциплина	«Большие данные»

ЭКЗАМЕНАЦИОННЫЙ БИЛЕТ №18

1. Парадигма MapReduce.
2. Области применения технологий Больших данных.

Составитель _____ С.В. Грицюк
(подпись)

Руководитель направления _____ С.О. Старков
(подпись)

« ____ » _____ 2021 г.

МИНИСТЕРСТВО ОБРАЗОВАНИЯ И НАУКИ РОССИЙСКОЙ ФЕДЕРАЦИИ
ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ АВТОНОМНОЕ ОБРАЗОВАТЕЛЬНОЕ УЧРЕЖДЕНИЕ
ВЫСШЕГО ПРОФЕССИОНАЛЬНОГО ОБРАЗОВАНИЯ

«Национальный исследовательский ядерный университет «МИФИ»

Обнинский институт атомной энергетики –

филиал федерального государственного автономного образовательного учреждения высшего профессионального образования «Национальный исследовательский ядерный университет «МИФИ»

Отделение Интеллектуальных кибернетических систем

Направление/ Специальность	090401 «Информатика и вычислительная техника»
Профиль/ Специализация	«Большие данные и машинное обучение для атомной энергетики»
Дисциплина	«Большие данные»

ЭКЗАМЕНАЦИОННЫЙ БИЛЕТ №19

1. NoSQL. Мотивация. Классификация.
2. Что такое Большие данные? Характеристика.

Составитель _____ С.В. Грицюк
(подпись)

Руководитель направления _____ С.О. Старков
(подпись)

« ____ » _____ 2021 г.

МИНИСТЕРСТВО ОБРАЗОВАНИЯ И НАУКИ РОССИЙСКОЙ ФЕДЕРАЦИИ
ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ АВТОНОМНОЕ ОБРАЗОВАТЕЛЬНОЕ УЧРЕЖДЕНИЕ
ВЫСШЕГО ПРОФЕССИОНАЛЬНОГО ОБРАЗОВАНИЯ

«Национальный исследовательский ядерный университет «МИФИ»

Обнинский институт атомной энергетики –

филиал федерального государственного автономного образовательного учреждения высшего профессионального образования «Национальный исследовательский ядерный университет «МИФИ»

Отделение Интеллектуальных кибернетических систем

Направление/ Специальность	090401 «Информатика и вычислительная техника»
Профиль/ Специализация	«Большие данные и машинное обучение для атомной энергетики»
Дисциплина	«Большие данные»

ЭКЗАМЕНАЦИОННЫЙ БИЛЕТ №20

1. Машинное обучение. Основы.
2. Дистрибутив Cloudera. Сервисы в составе дистрибутива.

Составитель _____ С.В. Грицюк
(подпись)

Руководитель направления _____ С.О. Старков
(подпись)

« ____ » _____ 2021 г.

Критерии и шкала оценивания

Оценка	Критерии оценки
Отлично 36-40	Студент должен: <ul style="list-style-type: none">- 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.
Хорошо 30-35	Студент должен: <ul style="list-style-type: none">- 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.
Удовлетворительно 24-29	Студент должен: <ul style="list-style-type: none">- 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.
Неудовлетворительно 23 и меньше	Студент демонстрирует: <ul style="list-style-type: none">- 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.

4.2. Список экзаменационных вопросов

МИНИСТЕРСТВО ОБРАЗОВАНИЯ И НАУКИ РОССИЙСКОЙ ФЕДЕРАЦИИ
ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ АВТОНОМНОЕ ОБРАЗОВАТЕЛЬНОЕ УЧРЕЖДЕНИЕ
ВЫСШЕГО ПРОФЕССИОНАЛЬНОГО ОБРАЗОВАНИЯ

«Национальный исследовательский ядерный университет «МИФИ»

Обнинский институт атомной энергетики –

филиал федерального государственного автономного образовательного учреждения высшего профессионального образования «Национальный исследовательский ядерный университет «МИФИ»

Отделение Интеллектуальных кибернетических систем

Направление/ Специальность	090401 «Информатика и вычислительная техника»
Профиль/ Специализация	«Большие данные и машинное обучение для атомной энергетики»
Дисциплина	«Большие данные»

ВОПРОСЫ К ЭКЗАМЕНУ

1. Большие данные. Что это такое? Определение. История.
2. Распределенные системы
3. Нереляционные базы данных – NoSQL
4. Свойства нереляционных баз данных
5. Типы данных
6. Масштабируемость систем
7. Целостность данных
8. ACID свойства
9. ACID свойства и NoSQL
10. CAP теорема
11. Язык Scala. Основные особенности
12. Шардинг / Sharding
13. Репликация / Replication
14. MapReduce парадигма
15. Apache Hadoop
16. WordCount
17. Apache Spark. Основные особенности
18. RDD. Основная идея
19. Spark запуск (execution)
20. Spark развертывание (deployment)
21. Стэк технологий Apache Spark
22. NLP. Основы и особенности
23. NLP. Примеры и задачи
24. NLP. Каноническая форма слова
25. NLP. Стемминг (stemming)
26. NLP. Извлечение именованных сущностей
27. ML. Основы и особенности
28. ML. Классы
29. ML. Обучение с учителем, схема
30. ML. Обучение без учителя

31. ML. Обучение с подкреплением
32. ML. Техники
33. Глубокое обучение (deep learning)
34. OLTP vs OLAP
35. NoSQL. Классы и их свойства
36. NoSQL. Базы данных ключ-значение
37. NoSQL. Документные базы данных
38. NoSQL. Колоночные базы данных
39. NoSQL. Графовые базы данных
40. NewSQL. Свойства и характеристики
41. NoSQL. Сравнение классов
42. Типы обработки данных
43. Zookeeper
44. Экосистема Hadoop. Сервисы
45. Scala. Mutability и Immutability

4.3. Комплект заданий для контрольных работ

МИНИСТЕРСТВО ОБРАЗОВАНИЯ И НАУКИ РОССИЙСКОЙ ФЕДЕРАЦИИ
ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ АВТОНОМНОЕ ОБРАЗОВАТЕЛЬНОЕ УЧРЕЖДЕНИЕ
ВЫСШЕГО ПРОФЕССИОНАЛЬНОГО ОБРАЗОВАНИЯ

«Национальный исследовательский ядерный университет «МИФИ»

Обнинский институт атомной энергетики –

филиал федерального государственного автономного образовательного учреждения высшего профессионального образования «Национальный исследовательский ядерный университет «МИФИ»

Отделение Интеллектуальных кибернетических систем

Комплект заданий для контрольной работы

по дисциплине «Большие данные»
(наименование дисциплины)

Темы: «Введение в область больших данных», «Введение в программирование на языке Scala»
«MapReduce, Hadoop и Apache Spark», «Системы хранения»

Вариант 1

1. Что такое Большие данные? Характеристика.
2. NoSQL. Ключ-значение.

Вариант 2

1. Масштабируемость. Виды.
2. Apache Spark. Особенности. Плюсы и минусы.

Вариант 3

1. Репликация и Шардинг.
2. Apache Hadoop. Особенности. Плюсы и минусы.

Вариант 4

1. NoSQL и транзакции.
2. Apache Hadoop. HDFS. MapReduce.

Вариант 5

1. CAP теорема

2. NoSQL. Документные БД.

Вариант 6

1. Что такое Большие данные? Характеристика.
2. Apache Spark. Работа в кластере.

Вариант 7

1. Масштабируемость. Виды.
2. NoSQL. Графовые БД.

Вариант 8

1. Репликация и Шардинг.
2. Apache Spark. Стэк технологий.

Вариант 9

1. NoSQL и транзакции.
2. Парадигма MapReduce.

Вариант 10

1. CAP теорема
2. NoSQL. Колоночные БД.

Вариант 11

1. Что такое Большие данные? Характеристика.
2. NoSQL. Документные БД.

Вариант 12

1. Масштабируемость. Виды.
2. Apache Spark. Стэк технологий.

Вариант 13

1. Репликация и Шардинг.
2. NoSQL. Графовые БД.

Вариант 14

1. NoSQL и транзакции.
2. Apache Spark. Особенности. Плюсы и минусы.

Вариант 15

1. CAP теорема
2. Apache Spark. Работа в кластере.

Критерии и шкала оценивания

Оценка	Критерии оценки
Отлично 9-10 баллов	Студент должен: <ul style="list-style-type: none">• продемонстрировать глубокое и прочное усвоение знаний программного материала;• исчерпывающе, последовательно, грамотно и логически стройно изложить в письменной форме теоретический материал;• правильно формулировать определения.
Хорошо 8 баллов	Студент должен: <ul style="list-style-type: none">• продемонстрировать достаточно полное знание программного материала;• продемонстрировать знание основных теоретических понятий;• достаточно последовательно, грамотно и логически стройно излагать в письменной форме материал;
Удовлетворительно 6-7 баллов	Студент должен: <ul style="list-style-type: none">• продемонстрировать общее знание изучаемого материала;• показать общее владение понятийным аппаратом дисциплины;• уметь строить ответ в письменной форме в соответствии со структурой излагаемого вопроса.
Неудовлетворительно 0-5 баллов	Студент демонстрирует: <ul style="list-style-type: none">• незнание значительной части программного материала;• невладение понятийным аппаратом дисциплины;• существенные ошибки при изложении учебного материала;• неумение строить ответ в письменной форме в соответствии со структурой излагаемого вопроса.

4.4. Лабораторные работы №1, №2, №3, №4

Лабораторные работы предназначены для выработки практических навыков по материалу, полученному в рамках предмета (курс лекций), а также выявления качества усвоения знаний по дисциплине.

По завершению каждой из лабораторных работ студент должен продемонстрировать ее результат на компьютере и защитить в форме собеседования с преподавателем. На собеседование выносятся вопросы, касающиеся теоретических аспектов выполняемой работы, последовательности используемых для решения задачи шагов/процедур, а также анализа полученных результатов.

Лабораторная работа №1 включает установку, запуск, изучение интерфейса и встроенных средств среды разработки Scala IDE, а также основы программирования на языке Scala. Студент получает практические навыки создания и конфигурирования Maven-проекта в среде Scala IDE. По завершению лабораторной работы №1 в ходе устного опроса у компьютера студент показывает реализацию Scala программы в соответствии со своим вариантом.

Лабораторная работа №2 включает подключение и запуск библиотек Apache Spark в standalone режиме, а также реализацию программы для анализа текста своего варианта литературного произведения (очистка текста, реализация WordCount, стемминг, ранжирование встречаемости слов). По завершению лабораторной работы №2 в ходе устного опроса у компьютера студент демонстрирует код программы на языке Scala, объясняет основные проблемы, с которыми пришлось столкнуться и методы их решения, показывает результаты.

Лабораторная работа №3 включает освоение основ библиотеки Spark SQL, а также анализ и описание набора данных в соответствии с вариантом. Студент выполняет запросы к данным через Spark SQL, проводит визуализацию данных. По завершению лабораторной работы №3 в ходе устного опроса у компьютера студент демонстрирует код программы на языке Scala, объясняет полученные результаты.

Лабораторная работа №4 включает освоение основ библиотеки Spark ML. Студент формулирует простейшую задачу машинного обучения (регрессия, классификация) по набору данных для своего варианта. Далее, с помощью библиотеки Spark ML студент решает поставленную задачу. По завершению лабораторной работы №4 в ходе устного опроса у компьютера студент демонстрирует код программы на языке Scala, объясняет полученные результаты.

Критерии и шкала оценивания

Критерий оценки – полнота, качество, своевременность выполненной работы и успешная ее защита. Лабораторные работы №1 и №2 оцениваются в баллах от 0 до 10, а лабораторные работы №3 и №4 от 0 до 15. Каждая лабораторная работа считается сданной при получении оценки не ниже 60% от максимального балла.

	Балл	
	Минимум	Максимум
Лабораторная работа №1	6	10
Лабораторная работа №2	6	10
Лабораторная работа №3	9	15
Лабораторная работа №4	9	15