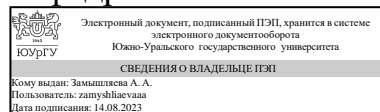


УТВЕРЖДАЮ:
Заведующий выпускающей
кафедрой



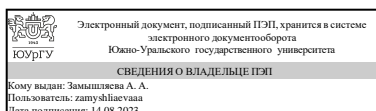
А. А. Замышляева

РАБОЧАЯ ПРОГРАММА

дисциплины 1.Ф.П0.15 Высокопроизводительные параллельные вычисления на кластерных системах
для направления 01.03.02 Прикладная математика и информатика
уровень Бакалавриат
профиль подготовки Прикладная математика и искусственный интеллект
форма обучения очная
кафедра-разработчик Прикладная математика и программирование

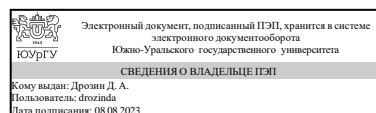
Рабочая программа составлена в соответствии с ФГОС ВО по направлению подготовки 01.03.02 Прикладная математика и информатика, утверждённым приказом Минобрнауки от 10.01.2018 № 9

Зав.кафедрой разработчика,
д.физ.-мат.н., проф.



А. А. Замышляева

Разработчик программы,
к.экон.н., доцент



Д. А. Дрозин

1. Цели и задачи дисциплины

Целью преподавания дисциплины является приобретение студентами знаний и умений в области параллельного программирования задач машинного обучения. В результате изучения дисциплины студенты должны: 1) знать архитектурную организацию обычных компьютеров и суперкомпьютеров; 2) знать и владеть методологией проведения массово параллельных вычислений для ускорения машинного обучения; 3) знать и владеть методами и технологиями обработки и анализа больших данных; 4) знать и уметь разрабатывать компьютерные программы анализа больших данных.

Краткое содержание дисциплины

Изучение архитектурных особенностей обычных компьютеров и суперкомпьютеров
Изучение методов ускорения машинного обучения за счет использования систем массово параллельных вычислений
Изучение методов и технологий обработки и анализа больших данных
Изучение методов разработки компьютерных программ анализа больших данных

2. Компетенции обучающегося, формируемые в результате освоения дисциплины

Планируемые результаты освоения ОП ВО (компетенции)	Планируемые результаты обучения по дисциплине
ПК-8 (ПК-5 модели) Способен использовать инструментальные средства для решения задач машинного обучения	Знает: [ПК-5.3. 3-2.] методологию проведения массово параллельных вычислений для ускорения машинного обучения (с использованием GPU); [ПК-5.3. 3-3.] принципы работы распределённых кластерных систем Умеет: [ПК-5.3. У-2.] работать с распределённой кластерной системой при создании, поддержке и использовании систем искусственного интеллекта
ПК-10 (ПК-7 модели) Способен осуществлять сбор и подготовку данных для систем искусственного интеллекта	Умеет: [ПК-7.2. У-5.] использовать методы и технологии массово параллельной обработки и анализа данных
ПК-11 (ПК-8 модели) Способен разрабатывать системы анализа больших данных	Умеет: [ПК-8.1. У-4.] использовать шины данных (Apache Kafka)

3. Место дисциплины в структуре ОП ВО

Перечень предшествующих дисциплин, видов работ учебного плана	Перечень последующих дисциплин, видов работ
Сбор, анализ и предобработка данных, Администрирование и проектирование хранилищ больших данных, Алгоритмы машинного обучения, Современные архитектуры глубоких искусственных нейронных сетей, Программирование для анализа данных	Не предусмотрены

Требования к «входным» знаниям, умениям, навыкам студента, необходимым при освоении данной дисциплины и приобретенным в результате освоения предшествующих дисциплин:

Дисциплина	Требования
Современные архитектуры глубоких искусственных нейронных сетей	<p>Знает: [ПК-6.1. 3-2.] функциональность современных инструментальных средств и систем программирования в области создания моделей искусственных нейронных сетей; [ПК-6.2. 3-1.] принципы построения систем искусственного интеллекта на основе искусственных нейронных сетей, методы и подходы к планированию и реализации проектов по созданию систем искусственного интеллекта в том числе в условиях малого количества данных, [ПК-8.2. 3-4.] методы и технологии машинного обучения на больших данных, [ПК-5.2. 3-1.] функциональные возможности современных инструментальных средств и систем программирования в области создания моделей и методов машинного обучения; [ПК-5.2. 3-2.] принципы проведения машинного эксперимента, проблемы переобучения и недообучения модели, требования к обучающей, тестовой и валидационной выборкам для решения задач анализа данных и машинного обучения; [ПК-5.3. 3-1.] принципы построения систем искусственного интеллекта, методы и подходы к планированию и реализации проектов по созданию систем искусственного интеллекта с применением машинного обучения, [ПК-7.2. 3-2.] методы планирования вычислительного эксперимента, формирования обучающей и контрольной выборок</p> <p>Умеет: [ПК-6.1. У-1.] проводить оценку и выбор моделей искусственных нейронных сетей и инструментальных средств для решения задачи машинного обучения; [ПК-6.2. У-1.] решать задачи по выполнению коллективной проектной деятельности для создания, поддержки и использования системы искусственного интеллекта на основе искусственных нейронных сетей, [ПК-5.3. У-1.] решать задачи по выполнению коллективной проектной деятельности для создания, поддержки и использования системы искусственного интеллекта с применением машинного обучения и массово параллельных вычислений для ускорения машинного обучения, [ПК-7.2. У-2.] выделять входные и выходные переменные с целью использования предиктивных моделей</p> <p>Имеет практический опыт: машинного обучения на больших данных</p>
Сбор, анализ и предобработка данных	Знает: [ПК-7.2. 3-1.] методы редукции размерности элементов набора данных и их

	<p>предварительной статистической обработки, разметки структурированных и неструктурированных данных; Умеет: [ПК-1.3. У-1.] осуществлять сбор исходной информации с использованием платформ данных (облачных и внутрикорпоративных), [ПК-7.2. У-1.] выявлять и исключать из массива данных ошибочные данные и выбросы; [ПК-7.2. У-3.] осуществлять разметку структурированных и неструктурированных данных Имеет практический опыт: сбора информации с использованием платформ данных (облачных и внутрикорпоративных)</p>
<p>Программирование для анализа данных</p>	<p>Знает: [ПК-5.1. 3-1.] возможности современных инструментальных средств и систем программирования для решения задач анализа данных и машинного обучения Умеет: [ПК-2.2. У-1.] разрабатывать программные приложения систем искусственного интеллекта, с использованием современных языков программирования, библиотек и программных платформ функционального, логического, объектно-ориентированного программирования (Python, R, C++, C#) Имеет практический опыт: участия в разработке программных приложений систем искусственного интеллекта, с использованием современных языков программирования</p>
<p>Администрирование и проектирование хранилищ больших данных</p>	<p>Знает: [ПК-7.1. 3-1.] виды представления данных, методы поиска и парсинга данных; [ПК-7.1. 3-2.] уровни представления данных (ODS, DDL, семантический слой, модель данных); [ПК-7.1. 3-3.] основные инструменты и технологии Data Science, [ПК-8.1. 3-1.] общедоступные репозитории и специализированные библиотеки, содержащие наборы больших данных; [ПК-8.1. 3-2.] принципы работы экосистемы Hadoop. фреймворка SPARK; [ПК-8.1. 3-3.] устройство интерфейсов между реляционными SQL-хранилищами данных и нереляционными NoSQL-хранилищами данных; [ПК-8.1. 3-4.] предметно-ориентированные языки; [ПК-8.2. 3-2.] устройство и принципы работы систем обработки и анализа больших массивов данных (SQL, NoSQL, Hadoop, ETL) Умеет: [ПК-7.1. У-1.] отделять достоверные источники данных от сомнительных, осуществлять критический отбор данных, проверять их на целостность и непротиворечивость, [ПК-8.1. У-1.] настраивать и оптимизировать конфигурацию программного и аппаратного обеспечения с целью интеграции больших данных; [ПК-8.1. У-2.] разрабатывать программное обеспечение для очистки и валидации наборов больших данных; [ПК-8.1. У-3.] выполнять потоковую обработку данных (data streaming, event processing); [ПК-8.1. У-5.]</p>

	использовать языки запросов, в том числе нереляционных, для поддержки различных типов данных (например, XML, RDF, JSON, мультимедиа) и операций с большими данными (например, матричные операции);[ПК-8.2. У-3.] использовать системы обработки и анализа больших массивов данных (SQL, NoSQL, Hadoop, ETL, процессы и инструменты) Имеет практический опыт:
Алгоритмы машинного обучения	Знает: [ПК-4.1. 3-1.] принципы и методы машинного обучения, типы и классы задач машинного обучения, методологию ML Ops;[ПК-4.1. 3-2.] статистические методы анализа данных;[ПК-4.3. 3-1.] классические методы и алгоритмы машинного обучения: предиктивные - обучение с учителем, дескриптивные - обучение без учителя Умеет: [ПК-5.1. У-1.] проводить сравнительный анализ и осуществлять выбор инструментальных средств для решения задач машинного обучения, [ПК-4.1. У-1.] сопоставить задачам предметной области классы задач машинного обучения;[ПК-4.1. У-2.] использовать статистические методы анализа данных при решении задач машинного обучения;[ПК-4.3. У-1.] проводить сравнительный анализ и осуществлять выбор, настройку при необходимости, разработку методов и алгоритмов для решения задач машинного обучения Имеет практический опыт: участия в разработке алгоритмов для решения задач машинного обучения

4. Объём и виды учебной работы

Общая трудоемкость дисциплины составляет 2 з.е., 72 ч., 40,25 ч. контактной работы

Вид учебной работы	Всего часов	Распределение по семестрам в часах
		Номер семестра
		8
Общая трудоёмкость дисциплины	72	72
<i>Аудиторные занятия:</i>	36	36
Лекции (Л)	24	24
Практические занятия, семинары и (или) другие виды аудиторных занятий (ПЗ)	0	0
Лабораторные работы (ЛР)	12	12
<i>Самостоятельная работа (СРС)</i>	31,75	31,75
Подготовка к промежуточной аттестации	1,75	1,75
Подготовка к лекциям	30	30
Консультации и промежуточная аттестация	4,25	4,25
Вид контроля (зачет, диф.зачет, экзамен)	-	зачет

5. Содержание дисциплины

№ раздела	Наименование разделов дисциплины	Объем аудиторных занятий по видам в часах			
		Всего	Л	ПЗ	ЛР
1	Архитектурные особенности обычных компьютеров и суперкомпьютеров	10	6	0	4
2	Методы ускорения машинного обучения за счет использования систем массово параллельных вычислений	2	2	0	0
3	Методы и технологии обработки и анализа больших данных	14	10	0	4
4	Методы разработки компьютерных программ анализа больших данных	10	6	0	4

5.1. Лекции

№ лекции	№ раздела	Наименование или краткое содержание лекционного занятия	Кол-во часов
1	1	Устройство обычного компьютера. Операции с числами. Иерархия памяти. Языки программирования и программы. Узкие места. Принципы работы распределённых кластерных систем.	2
2	1	Усложнение и наращивание аппаратных средств. Повышение интеллектуальности управления компьютером. Система функциональных устройств.	2
3	1	Классификация параллельных компьютеров и систем. Векторно-конвейерные компьютеры. Параллельные компьютеры с общей памятью. Работа с распределённой кластерной системой при создании, поддержке и использовании систем искусственного интеллекта	2
4	2	Подключение, настройка и использование GPU. Подключение, настройка и использование систем массово параллельных вычислений. Методология проведения массово параллельных вычислений для ускорения машинного обучения (с использованием GPU).	2
5-7	3	Первичный анализ рядов. Исследование рядов данных. Очистка рядов. Методы и технологии массово параллельной обработки и анализа данных.	5
7-9	3	Корреляционный анализ. Автокорреляционный анализ данных. Кластеризация.	5
10-12	4	Разработка компьютерных программ. UML-диаграммы, диаграммы действий. Кодирование. Интерфейс пользователя. Использование шины данных (Apache Kafka)	6

5.2. Практические занятия, семинары

Не предусмотрены

5.3. Лабораторные работы

№ занятия	№ раздела	Наименование или краткое содержание лабораторной работы	Кол-во часов
1,2	1	Изучение параметров обычного компьютера: быстродействие, память, точность и проведение анализа ведущих суперкомпьютеров мира из TOP-500	4
3,4	3	Обработка и анализ больших данных. Изучение методов и технологии массово параллельной обработки и анализа данных	4

5,6	4	Разработка компьютерной программы анализа больших данных	4
-----	---	--	---

5.4. Самостоятельная работа студента

Выполнение СРС			
Подвид СРС	Список литературы (с указанием разделов, глав, страниц) / ссылка на ресурс	Семестр	Кол-во часов
Подготовка к промежуточной аттестации	Воеводин, В. В. Параллельные вычисления Учеб. пособие для вузов по направлению 510200 "Прикладная математика и информатика" В. В. Воеводин, Вл. В. Воеводин. - СПб.: БХВ-Петербург, 2004. - 599 с. Антонио, Д. Библиотека Keras – инструмент глубокого обучения. Реализация нейронных сетей с помощью библиотек Theano и TensorFlow / Д. Антонио, П. Суджит ; перевод с английского А. А. Слинкин. — Москва : ДМК Пресс, 2018. — 294 с. Шолле, Ф. Глубокое обучение с R и Keras / Ф. Шолле ; перевод с английского В. С. Яценкова. — Москва : ДМК Пресс, 2023. — 646 с.	8	1,75
Подготовка к лекциям	Воеводин, В. В. Параллельные вычисления Учеб. пособие для вузов по направлению 510200 "Прикладная математика и информатика" В. В. Воеводин, Вл. В. Воеводин. - СПб.: БХВ-Петербург, 2004. - 599 с. Антонио, Д. Библиотека Keras – инструмент глубокого обучения. Реализация нейронных сетей с помощью библиотек Theano и TensorFlow / Д. Антонио, П. Суджит ; перевод с английского А. А. Слинкин. — Москва : ДМК Пресс, 2018. — 294 с. Шолле, Ф. Глубокое обучение с R и Keras / Ф. Шолле ; перевод с английского В. С. Яценкова. — Москва : ДМК Пресс, 2023. — 646 с.	8	30

6. Фонд оценочных средств для проведения текущего контроля успеваемости, промежуточной аттестации

Контроль качества освоения образовательной программы осуществляется в соответствии с Положением о балльно-рейтинговой системе оценивания результатов учебной деятельности обучающихся.

6.1. Контрольные мероприятия (КМ)

№ КМ	Се-местр	Вид контроля	Название контрольного мероприятия	Вес	Макс. балл	Порядок начисления баллов	Учи-тыва-ется в ПА
1	8	Текущий	Анализ ведущих	20	1	Если задание выполнено без ошибок,	зачет

		контроль	суперкомпьютеров мира из TOP-500			написан отчет в соответствии с требованиями, то выставляется 1 балл. Иначе 0 баллов.	
2	8	Текущий контроль	Обработка и анализ больших данных. Изучение методов и технологии массово параллельной обработки и анализа данных	20	1	Если задание выполнено без ошибок, написан отчет в соответствии с требованиями, то выставляется 1 балл. Иначе 0 баллов.	зачет
3	8	Текущий контроль	Разработка компьютерной программы анализа больших данных	20	1	Если задание выполнено без ошибок, написан отчет в соответствии с требованиями, то выставляется 1 балл. Иначе 0 баллов.	зачет
4	8	Промежуточная аттестация	Ответ по билету	-	3	Если вопрос раскрыт полностью - 1 балл. Если вопрос раскрыт, но не полностью - 0.5 балла. Если вопрос не раскрыт - 0 баллов	зачет

6.2. Процедура проведения, критерии оценивания

Вид промежуточной аттестации	Процедура проведения	Критерии оценивания
зачет	На промежуточной аттестации происходит оценивание учебной деятельности обучающихся по дисциплине на основе полученных оценок за контрольно-рейтинговые мероприятия текущего контроля. Студент может улучшить свой рейтинг, пройдя контрольное мероприятие промежуточной аттестации, которое не является обязательным. Контрольное мероприятие промежуточной аттестации проводится в смешанной форме - письменно-устной. Студенту выдается билет, содержащий 3 вопроса. На подготовку выделяется 1 час, после чего студент сдает работу в письменном виде. Затем проводится собеседование.	В соответствии с пп. 2.5, 2.6 Положения

6.3. Паспорт фонда оценочных средств

Компетенции	Результаты обучения	№ КМ			
		1	2	3	4
ПК-8	Знает: [ПК-5.3. З-2.] методологию проведения массово параллельных вычислений для ускорения машинного обучения (с использованием GPU); [ПК-5.3. З-3.] принципы работы распределённых кластерных систем	+	+		+
ПК-8	Умеет: [ПК-5.3. У-2.] работать с распределённой кластерной системой при создании, поддержке и использовании систем искусственного интеллекта			+	+
ПК-10	Умеет: [ПК-7.2. У-5.] использовать методы и технологии массово параллельной обработки и анализа данных			+	+
ПК-11	Умеет: [ПК-8.1. У-4.] использовать шины данных (Apache Kafka)				+

Типовые контрольные задания по каждому мероприятию находятся в приложениях.

7. Учебно-методическое и информационное обеспечение дисциплины

Печатная учебно-методическая документация

а) основная литература:

1. Воеводин, В. В. Параллельные вычисления Учеб. пособие для вузов по направлению 510200 "Прикладная математика и информатика" В. В. Воеводин, Вл. В. Воеводин. - СПб.: БХВ-Петербург, 2004. - 599 с.
2. Антонов, А. С. Технологии параллельного программирования MPI и OpenMP [Текст] учеб. пособие для вузов по направлениям ВПО 010400 "Прикладная математика и информатика" и 010300 "Фундамент. информатика и информационные технологии" А. С. Антонов ; Моск. гос. ун-т им. М. В. Ломоносова. - М.: Издательство Московского университета, 2012. - 339 с.

б) дополнительная литература:

Не предусмотрена

в) отечественные и зарубежные журналы по дисциплине, имеющиеся в библиотеке:

Не предусмотрены

г) методические указания для студентов по освоению дисциплины:

1. Методические указания к лабораторным работам (приложенный файл)

из них: учебно-методическое обеспечение самостоятельной работы студента:

1. Методические указания к лабораторным работам (приложенный файл)

Электронная учебно-методическая документация

№	Вид литературы	Наименование ресурса в электронной форме	Библиографическое описание
1	Основная литература	Электронно-библиотечная система издательства Лань	Антонио, Д. Библиотека Keras – инструмент глубокого обучения. Реализация нейронных сетей с помощью библиотек Theano и TensorFlow / Д. Антонио, П. Суджит ; перевод с английского А. А. Слинкин. — Москва : ДМК Пресс, 2018. — 294 с. https://e.lanbook.com/book/111438
2	Дополнительная литература	Электронно-библиотечная система издательства Лань	Шолле, Ф. Глубокое обучение с R и Keras / Ф. Шолле ; перевод с английского В. С. Яценкова. — Москва : ДМК Пресс, 2023. — 646 с. https://e.lanbook.com/book/315488

Перечень используемого программного обеспечения:

Нет

Перечень используемых профессиональных баз данных и информационных справочных систем:

Нет

8. Материально-техническое обеспечение дисциплины

Вид занятий	№ ауд.	Основное оборудование, стенды, макеты, компьютерная техника, предустановленное программное обеспечение, используемое для различных видов занятий
Практические занятия и семинары	332 (36)	Компьютеры
Лабораторные занятия	340 (36)	Компьютеры
Лекции	332 (36)	Меловая доска, проектор