


МИНОБРНАУКИ РОССИЙСКОЙ ФЕДЕРАЦИИ
НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ
ТОМСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ

Институт прикладной математики и компьютерных наук

УТВЕРЖДАЮ

Директор института прикладной
математики и компьютерных наук

« 11 »  2021 г.



**Технологии высокопроизводительной обработки больших
данных**

рабочая программа дисциплины

Закреплена за кафедрой Учебный план	<i>теоретических основ информатики 01.03.02 Прикладная математика и информатика, профиль «Математические методы в экономике»</i>
Форма обучения	<i>очная</i>
Общая трудоёмкость	<i>3 з.е.</i>
Часов по учебному плану	<i>108</i>
в том числе:	
аудиторная контактная работа	<i>54.7</i>
самостоятельная работа	<i>53.3</i>
Вид(ы) контроля в семестрах <i>экзамен/зачет/зачет с оценкой</i>	<i>Семестр 8 – экзамен</i>

Программу составил:
канд. техн. наук,
доцент кафедры теоретических
основ информатики

Д.В. Дружинин

Рецензент:
д-р техн. наук, профессор,
профессор кафедры теоретических основ информатики

Ю.Л. Костюк

Рабочая программа дисциплины «Технологии высокопроизводительной обработки больших данных» разработана в соответствии с образовательным стандартом высшего образования – бакалавриат, самостоятельно устанавливаемым федеральным государственным автономным образовательным учреждением высшего образования «Национальный исследовательский Томский государственный университет» по направлению подготовки 01.03.02 Прикладная математика и информатика (утвержден Ученым советом НИ ТГУ, протокол от 27.10.2021 г. № 08).

Рабочая программа одобрена на заседании кафедры теоретических основ информатики

Протокол от 04 июня 2021 г. № 05

Заведующий кафедрой теоретических основ информатики,
д-р техн. наук, профессор

А.В. Замятин

Рабочая программа одобрена на заседании учебно-методической комиссии института прикладной математики и компьютерных наук (УМК ИПМКН)

Протокол от 17 июня 2021 г. № 05

Председатель УМК ИПМКН,
д-р техн. наук, профессор

С.П. Сущенко

Цель освоения дисциплины

Цель – обучить студентов различным методам организации высокопроизводительных вычислений, сформировать навыки использования различных программных инструментов для организации параллельных вычислений.

1. Место дисциплины в структуре ОПОП

Дисциплина «Технологии высокопроизводительной обработки больших данных» относится к факультативным дисциплинам, входит в факультативный модуль по выбору «Введение в искусственный интеллект».

Для успешного освоения дисциплины необходимо знание основ информатики, опыт разработки простых программ на одном из языков программирования.

Пререквизиты дисциплины: «Введение в интеллектуальный анализ данных».

Постреквизиты дисциплины: нет.

2. Компетенции и результаты обучения, формируемые в результате освоения дисциплины

Таблица 1.

Компетенция	Индикатор компетенции	Код и наименование результатов обучения (планируемые результаты обучения по дисциплине, характеризующие этапы формирования компетенций)
ОПК-3. Способен применять и модифицировать математические модели для решения задач в области профессиональной деятельности	ИОПК-3.1. Демонстрирует навыки применения современного математического аппарата для построения адекватных математических моделей реальных процессов, объектов и систем в своей предметной области. ИОПК-3.2. Демонстрирует умение собирать и обрабатывать статистические, экспериментальные, теоретические и т.п. данные для построения математических моделей, расчетов и конкретных практических выводов.	ОР-3.1.1. обучающийся сможет: - знать базовую терминологию, иметь представление об архитектуре высокопроизводительных систем; - знать методы, математические модели и алгоритмы высокопроизводительной обработки данных; - владеть навыками использования углубленных теоретических и практических знаний в области технологий и математических моделей высокопроизводительной обработки данных для решения различных задач. ОР-3.2.1. обучающийся сможет: - уметь применять математические модели, системное и прикладное программное обеспечение для решения научных и практических задач технологической деятельности; - уметь выбирать методы, математические модели и средства высокопроизводительной обработки данных для решения научно-исследовательских и прикладных задач; - уметь разрабатывать и применять математические методы для решения научно-исследовательских задач.

3. Структура и содержание дисциплины

3.1. Структура и трудоемкость видов учебной работы по дисциплине

Общая трудоемкость дисциплины составляет 3 зачетные единицы, 108 часов.

Таблица 2.

Вид учебной работы	Трудоемкость в академических часах	
	8 семестр	всего
Общая трудоемкость	108	108
Контактная работа:	54,7	54,7
Лекции (Л):	16	16
Практики (ПЗ)	32	32
Лабораторные работы (ЛР)		
Семинары (СЗ)		
Групповые консультации	2	2
Индивидуальные консультации	2,4	2,4
Промежуточная аттестация	2,3	2,3
Самостоятельная работа обучающегося:	53,3	53,3
- изучение учебного материала, публикаций	7,6	7,6
- подготовка к практическим занятиям	14	14
- подготовка к рубежному контролю по теме/разделу	31,7	31,7
Вид промежуточной аттестации (зачет, зачет с оценкой, экзамен)	Экзамен	Экзамен

3.2. Содержание и трудоемкость разделов дисциплины

Таблица 3.

Код занятия	Наименование разделов и тем и их содержание	Вид учебной работы, занятий, контроля	С е м е с т р	Часы в электронной форме	Всего (час.)	Литература	Код (ы) результата(ов) обучения
	Раздел 1. Введение		8		3	Основная литература: [4], [5] Дополнительная литература: [2]	ОР-3.1.1
1.1.	Актуальность, базовая терминология и тенденции развития.	Лекции	8		1		
1.2.	Параллелизм компьютерных вычислений.	Лекции	8		1		
1.3.	Изучение учебного материала, публикаций; подготовка к практическим занятиям	СРС	8		1		
	Текущий контроль успеваемости	Контрольная работа	8				
	Раздел 2. Архитектура вычислительных систем		8		5	Основная литература: [4], [5] Дополнительная литература: [2], [5]	ОР-3.2.1
2.1	Архитектура высокопроизводительных вычислительных систем	Лекции	8		1		
2.2	Классификация вычислительных систем	Лекции	8		1		
2.3	Изучение учебного материала, публикаций; подготовка к практическим занятиям	СРС	8		3		
	Раздел 3. Облачные технологии, их свойства и типы		8		15	Основная литература: [1], [2], [3] Дополнительная литература: [1], [2], [3] Интернет-ресурсы: [1], [2], [3]	ОР-3.1.1; ОР-3.2.1
3.1	Облачные технологии, их свойства и типы	Лекции	8		3		
		Практики	8		8		
3.2	Изучение учебного материала, публикаций; подготовка к практическим занятиям	СРС	8		4		

	Раздел 4. Технология вычислений MapReduce		8		14,6	Основная литература: [1], [3] Интернет-ресурсы: [4]	ОП-3.1.1; ОП-3.2.1
4.1	Технология вычислений MapReduce	Лекции	8		3		
		Практики	8		8		
4.2	Изучение учебного материала, публикаций; подготовка к практическим занятиям	СРС	8		3,6		
	Раздел 5. Распределённые файловые системы		8		16	Основная литература: [1], [3] Дополнительная литература: [2] Интернет-ресурсы: [5]	ОП-3.1.1; ОП-3.2.1
5.1	Распределённые файловые системы	Лекции	8		3		
		Практики	8		8		
5.2	Изучение учебного материала, публикаций; подготовка к практическим занятиям	СРС	8		5		
	Раздел 6. Программирование для высокопроизводительных вычислений		8		16	Основная литература: [4] Дополнительная литература: [4]	ОП-3.1.1; ОП-3.2.1
6.1.	Методы программирования для высокопроизводительных вычислений	Лекции	8		1		
		Практики	8		4		
6.2.	Методология проектирования параллельных алгоритмов	Лекции	8		2		
		Практики	8		4		
6.3	Изучение учебного материала, публикаций; подготовка к практическим занятиям	СРС	8		5		
	Консультации в период теоретического обучения	К	8		4,4		
	Подготовка к промежуточной аттестации в форме экзамена	СРС	8		31,7		
	Прохождение промежуточной аттестации в форме экзамена	Э	8		2,3		

4. Образовательные технологии, учебно-методическое и информационное обеспечение для освоения дисциплины

Текущий контроль освоения дисциплины осуществляется в виде проведения регулярных опросов по пройденному материалу и проверки знаний при допуске к выполнению практических работ.

Также текущий контроль успеваемости осуществляется в виде сдачи практических работ.

Промежуточная аттестация по курсу проводится в форме экзамена.

Самостоятельная работа студента включает:

- изучение разделов курса при подготовке к лекционным и практическим работам;
- самостоятельный поиск необходимых источников информации, включая научно-образовательные ресурсы сети Интернет;
- подготовка ответов на контрольные вопросы;
- подготовка к экзамену.

Типовые контрольные задания или иные материалы, необходимые для оценки результатов обучения, характеризующих этапы формирования компетенций, и методические материалы, определяющие процедуры оценивания результатов обучения, приведены в Приложении 1 к рабочей программе «Фонд оценочных средств».

4.1. Рекомендуемая литература и учебно-методическое обеспечение

№ п/п	Авторы / составители	Заглавие	Издательство	Год издания, количество страниц
Основная литература				
1.	S. Srinivasan	Cloud Computing Basics electronic resource	New York, NY: Springer New York: Imprint: Springer,	2014 г., 162 с.
2.	A Ohri	R for Cloud Computing electronic resource: An Approach for Data Scientists	New York, NY: Springer New York: Imprint: Springer	2014 г., 284 с.
3.	Zaigham Mahmood	Cloud Computing: Methods and Practical Approaches	London: Springer London: Imprint: Springer XVII, 347 p. 97 illus.: online resource.	2013 г., 364 с.
4.	К. Ю. Богачев	Основы параллельного программирования	Москва: БИНОМ. Лаборатория знаний	2015 г., 342 с.
5.	В. П. Гергель	Теория и практика параллельных вычислений: учебное пособие	Москва: Интернет- Университет Информационных Технологий	2016 г., 421 с.
Дополнительная литература				
1.	Christoph Fehling, Frank Leymann, Ralph Retter, Walter Schupeck, Peter	Cloud Computing Patterns electronic resource: Fundamentals to Design, Build, and Manage Cloud	Vienna: Springer Vienna: Imprint: Springer	2014 г., 393 с.

	Arbitter	Applications		
2.	Xiaolin Li, Judy Qiu	Cloud Computing for Data-Intensive Applications electronic resource	New York, NY: Springer New York: Imprint: Springer	2014 г., 427 с.
3.	Zaigham Mahmood	Cloud Computing electronic resource: Challenges, Limitations and R&D Solutions	New York, NY: Springer New York: Imprint: Springer	2014 г., 376 с.
4.	А. В. Линеv, Д. К. Боголепов, С. И. Баcтраков	Технологии параллельного программирования для процессоров новых архитектур	Нижегородский гос. ун-т им. Н. И. Лобачевского – Москва: Изд-во Московского университета	2010 г., 160 с.

4.2. Базы данных и информационно-справочные системы, в том числе зарубежные

1. Облачный сервис – Режим доступа: <http://www.ncloudtech.ru>
2. Статья «Бизнес в облаках. Чем полезны облачные технологии для предпринимателя» – Режим доступа: <https://kontur.ru/articles/225>
3. Батура Т.В., Мурзин Ф.А., Семич Д.Ф. Облачные технологии: основные понятия, задачи и тенденции развития – Режим доступа: <http://swsys-web.ru/cloud-computing-basic-concepts-problems.html>
4. MapReduce Tutorial – Режим доступа: https://hadoop.apache.org/docs/r1.2.1/mapred_tutorial.html
5. HDFS Architecture Guide – Режим доступа: https://hadoop.apache.org/docs/r1.2.1/hdfs_design.html

4.3. Перечень лицензионного и программного обеспечения

Microsoft Visual Studio Community Edition 2013 (C++), библиотека классов, реализующая технологию MapReduce (например, <https://github.com/cdmh/mapreduce>). Все используемые программные продукты являются свободно распространяемыми.

4.4. Оборудование и технические средства обучения

Для материально-технического обеспечения дисциплины требуется наличие компьютерной техники с установленным соответствующим программным обеспечением и другого оборудования, поддерживающего проведение презентаций.

Технические характеристики: Компьютер INTANT i6250_T (i5-6500, 3.2GHz\GA-N110M-S2V, mATX\8ГБ DDR3, 1600МГц\1000ГБ 7200об/мин., 32МБ SATAIII\Microtower 400Вт\DVD±RW\2 x 23,6" АOC I2476VW \кл-ра+мышь). Монитор LCD 23" Philips 234E5QDAB – 13 шт.

Интерактивная доска SMART Board SB480, ультракороткофокусный проектор SMART UF70 с настенным креплением

5. Методические указания обучающимся по освоению дисциплины

Основой обучения является курс лекций, читаемый преподавателем. Важным аспектом овладения перечисленными компетенциями является выполнение практических работ. Для самостоятельной работы и дополнительного расширения круга знаний желательно использовать литературу, приведенную в разделе 4.1.

6. Преподавательский состав, реализующий дисциплину

Дружинин Денис Вячеславович, канд. техн. наук, доцент кафедры теоретических основ информатики ТГУ.

7. Язык преподавания – русский язык.