

МИНОБРНАУКИ РОССИИ
Федеральное государственное бюджетное образовательное учреждение
высшего образования
«Тверской государственный технический университет»
(ТвГТУ)

УТВЕРЖДАЮ
Проректор по учебной
работе

Смирнов М.А.
«_____» 20__ г.

РАБОЧАЯ ПРОГРАММА

Дисциплины, части формируемой участниками образовательных отношений
Блока 1 «Дисциплины (модули)»

«Алгоритмы для работы с большими данными»

Направления подготовки бакалавров 09.03.04 Программная инженерия

Направленность (профиль) – Разработка систем искусственного интеллекта.

Типы задач профессиональной деятельности – производственно-технологический.

Форма обучения - очная

Факультет информационных технологий
Кафедра «Программное обеспечение»

Тверь 20__

Рабочая программа дисциплины соответствует ОХОП подготовки бакалавров в части требований к результатам обучения по дисциплине и учебному плану.

Разработчик программы

А.Л. Калабин

Программа рассмотрена и одобрена на заседании кафедры ПО
«___» 20___ г., протокол № ____.

Заведующий кафедрой

А.Л. Калабин

Согласовано:

Начальник учебно-методического
отдела УМУ

Е.Э. Наумова

Начальник отдела
комплектования
зональной научной библиотеки

О.Ф. Жмыхова

1. Цели и задачи дисциплины.

Цель изучения дисциплины «Алгоритмы для работы с большими данными» формирование у обучающихся системного понимания принципов, методов и технологий разработки и оптимизации алгоритмов для обработки больших объёмов данных, включая структурирование данных, эффективные методы поиска и сортировки, параллельные и распределённые вычисления, использование алгоритмов машинного обучения и анализа данных. Дисциплина направлена на формирование компетенций, необходимых для профессиональной деятельности в области проектирования, разработки и внедрения информационных систем, работающих с большими данными.

Задачи дисциплины:

- формирование знаний о структурах данных, алгоритмах и их применении для обработки больших объёмов данных;
- освоение методов анализа и оптимизации алгоритмов с учётом сложности и производительности;
- изучение принципов параллельной и распределённой обработки данных;
- развитие навыков применения алгоритмов машинного обучения и интеллектуального анализа данных в контексте больших данных;
- освоение инструментальных средств разработки, тестирования и оптимизации алгоритмов;
- формирование умений разрабатывать алгоритмические решения, проектную документацию и спецификации для систем работы с большими данными;
- стимулирование самостоятельной работы обучающихся, направленной на исследование, проектирование и анализ алгоритмов обработки больших объёмов данных.

2. Место дисциплины в структуре ОП.

Дисциплина относится к части, формируемой участниками образовательных отношений Блока 1 ОП ВО. Для освоения дисциплины «Алгоритмы для работы с большими данными» студенты используют знания, умения и навыки, сформированные в процессе изучения базовых дисциплин программной инженерии и математического цикла, таких как «Информатика и программирование», «Алгоритмизация и программирование», «Структуры и алгоритмы обработки данных», «Объектно-ориентированное программирование», «Базы данных», «Теоретическая информатика», «Теория алгоритмов», «Методы и алгоритмы машинного обучения», «Основы работы с большими данными», «Введение в искусственный интеллект и основные методы машинного обучения», «Математический анализ», «Линейная алгебра», «Дифференциальные уравнения», «Теория вероятностей и математическая статистика для анализа данных».

Освоение дисциплины является необходимой основой для последующего изучения дисциплин, сопряжённых с профессиональными стандартами и

профильной подготовкой в сфере искусственного интеллекта, а также при выполнении проектной и преддипломной практики и выпускной квалификационной работы.

3. Планируемые результаты обучения по дисциплине.

3.1. Перечень компетенций, закреплённых за дисциплиной в ОХОП.

Компетенция, закрепленная за дисциплиной ОХОП:

ПК-5. Способен разрабатывать, адаптировать, применять в профессиональной деятельности алгоритмы, программные средства, системы и комплексы обработки данных, методы и алгоритмы машинного обучения, программно-технические платформы, электронные библиотеки, программные оболочки приложений, сетевые технологии для решения задач в сфере искусственного интеллекта и смежных областях.

Индикаторы компетенции, закреплённые за дисциплиной в ОХОП:

ИПК-5.4. Осуществляет сбор данных по запросам при помощи специализированных программ, форматирование, создание набора данных для машинного обучения, интеграция технологий машинного обучения в оптимизацию приложений для улучшения производительности и функциональности.

ИПК-5.9. Использует вычислительные методы и алгоритмы для анализа и извлечения значимой информации из больших объемов различного контента с применением различных методов обработки, алгоритмов машинного обучения, интеллектуального анализа данных.

Показатели оценивания индикаторов достижения компетенций

Знать:

31. Классы алгоритмов обработки больших данных, методы распределённой и потоковой обработки, принципы выбора структур данных и оптимизации алгоритмов для высоконагруженных систем.

Уметь:

У1. Выбирать, применять и настраивать алгоритмы обработки больших данных для решения задач предметной области;

Иметь опыт практической подготовки:

ПП1. Анализировать сложность алгоритмов, оценивать их эффективность и обосновывать выбор используемых методов.

Компетенция, закрепленная за дисциплиной ОХОП:

ПК-6. Способен выбирать, применять и проводить экспериментальную проверку работоспособности программных компонентов систем, включающих модули по созданию искусственного интеллекта.

Индикаторы компетенции, закреплённые за дисциплиной в ОХОП:

ИПК-6.1. Применяет средства проектирования с различными формами представления знаний и выбора режима работы для конкретной предметной области.

ИПК-6.2. Самостоятельно создает на основе стандартных оболочек с привлечением искусственного интеллекта программное обеспечение для решения необходимых функций, профессиональных задач предприятий или организаций.

Показатели оценивания индикаторов достижения компетенций

Знать:

31. Инструментальные средства и программные платформы для обработки больших данных;

32. Технологии распределённых вычислений, средства выполнения и тестирования алгоритмов анализа данных.

Уметь:

У1. Выбирать подходящие программные инструменты и компоненты, интегрировать их в архитектуру разрабатываемой системы;

Иметь опыт практической подготовки:

ПП1. Реализовывать, настраивать и экспериментально проверять алгоритмы обработки больших данных, оценивать корректность, производительность и устойчивость работы системы.

3.2. Технологии, обеспечивающие формирование компетенций

Проведение лекционных и лабораторных занятий, самостоятельная работа под руководством преподавателя.

4. Трудоемкость дисциплины и виды учебной работы.

Таблица 1. Распределение трудоемкости дисциплины по видам учебной работы

Вид учебной работы	Зачетные единицы	Академические часы
Общая трудоемкость дисциплины	3	108
Аудиторные занятия (всего)		75
В том числе:		
Лекции		30
Практические занятия (ПЗ)		не предусмотрены
Лабораторные работы (ЛР)		45
Самостоятельная работа обучающихся (всего)		33
В том числе:		
Курсовая работа (КР)		не предусмотрена
Курсовой проект (КП)		не предусмотрен
Расчетно-графические работы		не предусмотрены
Другие виды самостоятельной работы: - Подготовка к лабораторным работам и тестам		25
Текущий контроль успеваемости и промежуточная аттестация (зачет)		8
Текущий контроль успеваемости и промежуточная аттестация (экзамен)		не предусмотрен
Практическая подготовка при реализации дисциплины (всего)		45

Вид учебной работы	Зачетные единицы	Академические часы
В том числе:		
Практические занятия (ПЗ)		не предусмотрены
Лабораторные работы (ЛР)		45
Курсовая работа (КР)		не предусмотрена
Курсовой проект (КП)		не предусмотрен

5. Структура и содержание дисциплины.

Структура и содержание дисциплины построены по модульно-блочному принципу. Под модулем дисциплины понимается укрупненная логико-понятийная тема, характеризующаяся общностью использованного понятийно-терминологического аппарата.

5.1. Структура дисциплины.

Таблица 2. Модули дисциплины, трудоемкость в часах и виды учебной работы.

№	Наименование модуля	Труд-ть часов	Лекции	Практич. занятия	Лаб. практикум	Сам. работа
1	Основы больших данных и архитектуры обработки	22	8	-	10	4
2	Алгоритмы и структуры данных для больших объёмов данных	32	10	-	14	8
3	Потоковые алгоритмы и алгоритмы приближенных вычислений	30	7	-	16	7
4	Прикладные алгоритмы анализа больших данных и машинного обучения	24	5	-	5	14
Всего на дисциплину		108	30	-	45	33

5.2. Содержание дисциплины.

МОДУЛЬ 1. «Основы больших данных и архитектуры обработки»

Понятие больших данных. Источники и характеристики Big Data. Архитектуры хранения и обработки данных. Файловые распределённые системы. Модели вычислений при работе с большими данными. Основы Hadoop/Spark-подходов. Принципы отказоустойчивости, масштабируемости и устойчивости систем. Механизмы организации потоков данных. Базовые паттерны обработки больших массивов данных.

МОДУЛЬ 2. «Алгоритмы и структуры данных для больших объёмов данных»

Эффективные структуры данных: деревья, индексы, хеш-структуры. Алгоритмы поиска, сортировки и агрегации для больших данных. Алгоритмы фильтрации и сжатия. Алгоритмы оценки частотности и значений. Методы

оптимизации алгоритмов под высокую нагрузку. Компромисс точность–скорость–память. Оценка сложности алгоритмов на больших данных. Выбор стратегий реализации.

МОДУЛЬ 3. «Потоковые алгоритмы и алгоритмы приближенных вычислений»

Основы потоковой обработки данных. Однопроходные алгоритмы и ограничения модели *streaming*. Приближенные структуры данных: Bloom-фильтры, Counting Bloom, HyperLogLog, Count-Min Sketch. Методы детекции аномалий в потоках. Алгоритмы оценки уникальности, частотности, трендов и распределений. Приближенные алгоритмы поиска и кластеризации. Алгоритмы скользящих окон. Ограничения память и обработка высоких скоростей поступления данных. Применение *streaming*-подходов в аналитике реального времени.

МОДУЛЬ 4. «Прикладные алгоритмы анализа больших данных и машинного обучения»

Методы анализа больших данных. Подходы машинного обучения, адаптированные под большие объёмы. Обработка признаков и построение конвейеров обработки данных. Алгоритмы классификации, регрессии и кластеризации, масштабируемые для Big Data. Алгоритмы для анализа логов, событий, взаимодействия пользователей. Методы оценки качества алгоритмов. Средства экспериментальной проверки и сравнения моделей.

5.3. Лабораторные работы

Таблица 3. Лабораторные работы и их трудоемкость

Модуль	Цели лабораторных работ	Наименование лабораторных работ	Трудоёмкость (ч)
Модуль 1	Освоить базовые операции работы с большими данными и архитектурами хранения	Работа с распределённой файловой системой. Загрузка и первичная обработка данных	5
Модуль 2	Освоить методы реализации эффективных алгоритмов и структур данных для больших массивов	Реализация потоковых структур: Bloom-фильтр, Count-Min Sketch	7
		Реализация алгоритмов поиска, сортировки и фильтрации на больших наборах данных	7
Модуль 3	Изучить потоковые алгоритмы и методы приближенных вычислений для Big Data	Реализация алгоритмов HyperLogLog, оценка уникальности и частотности	6
		Реализация алгоритмов обработки потоков данных со скользящими окнами	6
Модуль 4	Применить алгоритмы анализа больших данных и машинного обучения	Построение и обучение масштабируемой модели анализа данных	7
		Экспериментальная проверка, сравнение и оценка качества алгоритмов	7

5.4. Практические занятия.

Учебным планом практические занятия не предусмотрены.

6. Самостоятельная работа обучающихся и текущий контроль успеваемости.

6.1. Цели самостоятельной работы

Формирование способностей к самостоятельному познанию и обучению, поиску литературы, обобщению, оформлению и представлению полученных результатов, их критическому анализу, поиску новых и неординарных решений, аргументированному отстаиванию своих предложений, умений подготовки выступлений и ведения дискуссий.

6.2. Организация и содержание самостоятельной работы.

Самостоятельная работа заключается в изучении отдельных тем курса по заданию преподавателя по рекомендуемой им учебной литературе, в подготовке к лабораторным работам, к текущему контролю успеваемости, экзамену.

В рамках дисциплины выполняется 5 лабораторных работ, которые защищаются устным опросом. Максимальная оценка за каждую выполненную лабораторную работу – 5 баллов, минимальная – 3 балла.

Выполнение всех лабораторных работ обязательно.

7. Учебно-методическое и информационное обеспечение дисциплины.

7.1. Основная литература

1. Миркин, Б. Г. Введение в анализ данных : учебник и практикум / Б. Г. Миркин. — Москва : Юрайт, 2022. — (Высшее образование). — Образовательная платформа Юрайт. — Текст : электронный. — ISBN 978-5-9916-5009-0. — URL: <https://urait.ru/book/vvedenie-v-analiz-dannih-469306> . — (ID=134346-0)
2. Нестеров, С. А. Интеллектуальный анализ данных с использованием SQL Server : учебник / С. А. Нестеров. — Санкт-Петербург [и др.] : Лань, 2023. — ЭБС Лань. — Текст : электронный. — ISBN 978-5-507-45535-5. — URL: <https://e.lanbook.com/book/311861> . — (ID=157185-0)
3. Кобзаренко, Д. Н. Учебное пособие дисциплины «Анализ больших данных»... : курс лекций / Д. Н. Кобзаренко, А. Г. Кобзаренко. — Махачкала : ДГУНХ, 2019. — ЭБС Лань. — Текст : электронный. — URL: <https://e.lanbook.com/book/246542> . — (ID=150643-0)
4. Конкина, В. В. Введение в большие данные и анализ информации : учебное пособие / В. В. Конкина, А. Б. Борисенко, И. Л. Коробова. — Тамбов : ТГТУ, ЭБС АСВ, 2024. — 81 с. — ISBN 978-5-8265-2749-8. — Текст : электронный. — URL: <https://www.iprbookshop.ru/145326.html> . — (ID=165560-0)

7.2. Дополнительная литература

1. Белов, В. С., Захаров, Ю. Н. Информационно-аналитические системы. Основы проектирования и применения : учебное пособие. — Москва : Евразийский открытый институт, 2010. — ЦОР IPR SMART. — URL: <https://www.iprbookshop.ru/10678.html> . — (ID=147644-0)
2. Зайцев, К. С. Применение методов Data Mining для поддержки процессов управления ИТ-услугами : учебное пособие. — Москва : НИЯУ МИФИ, 2009. — ЭБС Лань. — URL: <https://e.lanbook.com/book/75805> . — (ID=145909-0)
3. Мхитарян, В. С. и др. Анализ данных : учебник для вузов. — Москва : Юрайт, 2025. — 448 с. — ISBN 978-5-534-19964-2. — URL: <https://urait.ru/bcode/560311> . — (ID=100456-0)
4. Синева, И. С. Анализ данных в среде R : учебное пособие. Ч. 1. — Москва : МТУСИ, 2018. — ЦОР IPR SMART. — URL: <https://www.iprbookshop.ru/92422.html> . — (ID=146948-0)
5. Вольфсон, М. Б. Анализ данных : учебное пособие. — Санкт-Петербург : СПбГУТ им. Бонч-Бруевича, 2015. — ЭБС Лань. — URL: <https://e.lanbook.com/book/180254> . — (ID=145915-0)
6. Вольфсон, М. Б. Анализ данных : методические указания к лабораторным работам. — Санкт-Петербург : СПбГУТ, 2013. — ЭБС Лань. — URL: <https://e.lanbook.com/book/181497> . — (ID=145916-0)
7. Федин, Ф. О., Федин, Ф. Ф. Анализ данных : учебное пособие. Ч. 2 : Инструменты Data Mining. — Москва : МГПУ, 2012. — ЦОР IPR SMART. — URL: <https://www.iprbookshop.ru/26445.html> . — (ID=146080-0)
8. Парфенов, Ю. П. Постреляционные хранилища данных : учебное пособие для вузов. — Москва : Юрайт, 2025. — 97 с. — ISBN 978-5-534-21173-3. — URL: <https://urait.ru/bcode/559502> . — (ID=144706-0)
9. Кадырова, Н. О., Павлова, Л. В. Статистический анализ больших данных: подход на основе машин опорных векторов. — Санкт-Петербург : СПбПУ, 2022. — 60 с. — ISBN 978-5-7422-7813-9. — URL: <https://www.iprbookshop.ru/128651.html> . — (ID=165561-0)
10. Миркин, Б. Г. Базовые методы анализа данных : учебник и практикум. — Москва : Юрайт, 2025. — ISBN 978-5-534-19709-9. — URL: <https://urait.ru/bcode/560414> . — (ID=165559-0)
11. Воронов, В. И., Воронова, Л. И., Усачев, В. А. Data Mining — технологии обработки больших данных : учебное пособие. — Москва : МТУСИ, 2018. — ЦОР IPR SMART. — URL: <https://www.iprbookshop.ru/81324.html> . — (ID=145723-0)

7.3. Методические материалы

1. Учебно-методический комплекс дисциплины обязательной части Блока 1 "Алгоритмы для работы с большими данными". Направление подготовки 09.03.04 Программная инженерия. Направленность

(профиль) - Разработка систем искусственного интеллекта : ФГОС 3++ / Каф. Программное обеспечение ; сост. А.Л. Калабин. - 2025. - (УМК). - Текст : электронный. - URL:
<https://elib.tstu.tver.ru/MegaPro/GetDoc/Megapro/189513> . - (ID=189513-0)

7.4. Программное обеспечение по дисциплине

1. Операционная система Microsoft Windows: лицензии № ICM-176609 и № ICM-176613 (Azure Dev Tools for Teaching).
2. Microsoft Office 2019 Russian Academic: OPEN No Level: лицензия № 41902814.

7.5. Специализированные базы данных, справочные системы, электронно-библиотечные системы, профессиональные порталы в Интернет.

ЭБС и лицензионные ресурсы ТвГТУ размещены:

1. Ресурсы: <https://lib.tstu.tver.ru/header/obr-res>
2. ЭБ ТвГТУ: <https://elib.tstu.tver.ru/MegaPro/Web>
3. ЭБС "Лань": <https://e.lanbook.com/>
4. ЭБС "Университетская библиотека онлайн": <https://biblioclub.ru/>
5. Национальная электронная библиотека: <https://rusneb.ru>
6. ЦОР IPRSmart: <https://www.iprbookshop.ru/>
7. Электронная образовательная платформа "Юрайт": <https://urait.ru/>
8. Научная электронная библиотека eLIBRARY: <https://elibrary.ru/>
9. Информационная система "ТЕХНОМАТИВ". Конфигурация "МАКСИМУМ" : сетевая версия (годовое обновление) : [нормативно-технические, нормативно-правовые и руководящие документы (ГОСТы, РД, СНиПы и др.). Диск 1, 2, 3, 4. - М. :Технорматив, 2014. - (Документация для профессионалов). - CD. - Текст : электронный. - 119600 р. - (105501-1)
10. База данных учебно-методических комплексов: <https://lib.tstu.tver.ru/header/umk.html>

УМК размещен: <https://elib.tstu.tver.ru/MegaPro/GetDoc/Megapro/189513>

8. Материально-техническое обеспечение.

При изучении дисциплины «Алгоритмы для работы с большими данными» используются современные образовательные и программные ресурсы: мультимедийный проектор и экран для демонстрации материалов; интерактивные схемы архитектур Big Data и моделей обработки данных; доступ к компьютерным классам с высокоскоростным Интернетом; программное обеспечение для разработки и анализа алгоритмов больших данных (Python 3.x с библиотеками NumPy, Pandas, PySpark, Dask, Scikit-learn; среди разработки Visual Studio Code, PyCharm; инструменты работы с распределёнными системами и хранилищами данных); средства визуализации данных; доступ к электронно-библиотечным системам и справочным ресурсам. Перечень основного оборудования: компьютерный класс с современными рабочими местами на Windows 10/11 или

Linux; программные платформы для обработки больших данных; мультимедийные средства визуализации и презентации.

9. Оценочные средства для проведения промежуточной аттестации

9.1. Оценочные средства для проведения промежуточной аттестации в форме экзамена

Учебным планом экзамен по дисциплине не предусмотрен.

9.2. Оценочные средства для проведения промежуточной аттестации в форме зачёта.

1. Шкала оценивания промежуточной аттестации в форме зачёта – «зачтено», «не зачтено».

2. Вид промежуточной аттестации в форме зачёта.

Вид промежуточной аттестации устанавливается преподавателем: по результатам текущего контроля знаний обучающегося без дополнительных контрольных испытаний или с выполнением дополнительного итогового контрольного испытания при наличии у студентов задолженностей в текущем контроле.

3. Для дополнительного итогового контрольного испытания студенту в обязательном порядке предоставляется: база заданий, предназначенных для предъявления обучающемуся на дополнительном итоговом контрольном испытании (типовой образец задания приводится в Приложении), задание выполняется письменно.

База заданий, предъявляемая обучающимся на дополнительном итоговом контролльном испытании:

1. Понятие больших данных и их основные характеристики.
2. Архитектуры хранения и обработки больших данных.
3. Модели вычислений при работе с большими данными.
4. Принципы организации распределённых файловых систем.
5. Концепции устойчивости и масштабируемости Big Data-систем.
6. Основы потоковой обработки данных.
7. Однопроходные алгоритмы и ограничения streaming-модели.
8. Приближенные алгоритмы и их применение.
9. Структуры данных Bloom Filter, Count-Min Sketch, HyperLogLog.
10. Алгоритмы оценки частотности и уникальности.
11. Методы фильтрации и агрегирования больших наборов данных.
12. Алгоритмы поиска и сортировки для больших объёмов данных.
13. Алгоритмы обработки данных в режиме реального времени.
14. Принципы организации скользящих окон.
15. Методы анализа потоков событий.

- 16.Масштабируемые алгоритмы машинного обучения.
- 17.Понятие распределённого обучения моделей.
- 18.Методы оценки качества алгоритмов для больших данных.
- 19.Основы инструментов Apache Spark, Hadoop, Dask (при наличии).
- 20.Основы оптимизации алгоритмов при ограничениях по памяти и времени.

Методические материалы определяют процедуру проведения дополнительного итогового испытания и присвоения зачёта.

Критерии выполнения контрольного испытания и условия присвоения зачёта: критерии оценки для категории «знать» (количественный критерий):

Ниже базового – 0 баллов,

Базовый уровень – 1 балл.

Критерии оценки для категории «уметь» (бинарный критерий):

Отсутствие умения – 0 баллов,

Наличие умения – 1 балл.

Критерии итоговой оценки за зачёт:

«зачтено» – при сумме баллов 2 или 3;

«не зачтено» – при сумме баллов 0 или 1.

Число заданий для дополнительного итогового контрольного испытания – 20. Число вопросов – 3. Продолжительность – 60 минут.

4. При промежуточной аттестации без выполнения дополнительного итогового контрольного испытания студенту в обязательном порядке описываются критерии присвоения зачёта: «зачтено» – выставляется обучающемуся при условии выполнения им всех контрольных мероприятий.

9.3. Оценочные средства для проведения промежуточной аттестации в форме курсового проекта или курсовой работы.

Учебным планом курсовой проект или курсовая работы не предусмотрены.

10. Методические рекомендации по организации изучения дисциплины.

Студенты перед началом изучения дисциплины должны быть ознакомлены с системами кредитных единиц и балльно-рейтинговой оценки, которые должны быть опубликованы и размещены на сайте вуза или кафедры.

В учебном процессе рекомендуется внедрение субъект-субъектной педагогической технологии, при которой в расписании каждого преподавателя определяется время консультаций студентов по закрепленному за ним модулю дисциплины.

Рекомендуется обеспечить студентов, изучающих дисциплину, электронными учебниками, учебно-методическим комплексом по дисциплине, включая методические указания к выполнению лабораторных работ, а также всех видов самостоятельной работы.

11. Внесение изменений и дополнений в рабочую программу дисциплины

Кафедра ежегодно обновляет содержание рабочих программ дисциплин, которые оформляются протоколами заседаний дисциплин, форма которых утверждена Положением о рабочих программах дисциплин, соответствующих ФГОС ВО.

Приложение

МИНОБРНАУКИ РОССИИ

Федеральное государственное бюджетное образовательное учреждение
высшего образования
«Тверской государственный технический университет»

Направление подготовки бакалавров 09.03.04 Программная инженерия
Профиль – Разработка систем искусственного интеллекта.
Кафедра «Программное обеспечение»
Дисциплина «Алгоритмы для работы с большими данными»

ЗАДАНИЕ ДЛЯ ДОПОЛНИТЕЛЬНОГО ИТОГОВОГО КОНТРОЛЬНОГО ИСПЫТАНИЯ №_1_

1. Вопрос для проверки уровня «ЗНАТЬ» – 0 или 1 балл:

Понятие больших данных, их основные характеристики и свойства.

2. Задание для проверки уровня «ЗНАТЬ» - 0 или 1 балл:

Назначение и принципы работы приблизительных структур данных (например, Bloom Filter или HyperLogLog).

3. Задание для проверки уровня «УМЕТЬ» – 0 или 1 балл:

Построить и применить алгоритм подсчёта частотности элементов в потоке данных с использованием Count-Min Sketch.

Критерии итоговой оценки за зачет:
«зачтено» - при сумме баллов 2 или 3;
«не зачтено» - при сумме баллов 0 или 1.

Составитель: д.ф.-м.н., профессор_____ Калабин А.Л.

Заведующий кафедрой ПО, д.ф.-м.н., профессор _____ Калабин А.Л.