Документ подписан простой электронной подписью
Информация о владельце:
ФИО: Максимов Алексей Борисович
Должность: директор департамента по образовательной политике
Дата подписания: 24.10.2023 14:36:18
Уникальный программный ключ:
8db180d1a3f02ac9e60521a5672742735c18b1d6

МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ
РОССИЙСКОЙ ФЕДЕРАЦИИ
федеральное государственное автономное образовательное учреждение
высшего образования
«МОСКОВСКИЙ ПОЛИТЕХНИЧЕСКИЙ УНИВЕРСИТЕТ»
Факультет информационных технологий

УТВЕРЖДЕНО
Декан факультета
Информационных технологий
/ Д.Г. Демидов /
«16» 02 2023 г.

РАБОЧАЯ ПРОГРАММА ДИСЦИПЛИНЫ

«Методы работы с большими данными»

Направление подготовки
09.03.01 Информатика и вычислительная техника

Профиль
«Веб-технологии»

Квалификация
Бакалавр

Формы обучения
очная

Москва, 2023 г.

Разработчик(и):

Доцент, к.т.н. / В.Ю. Верещагин /

Согласовано:
Заведующий кафедрой «Инфокогнитивные технологии»,
к.т.н., доцент / Е.А. Пухова /

Содержание

1 Цели, задачи и планируемые результаты обучения по дисциплине 4
2 Место дисциплины в структуре образовательной программы 5
3 Структура и содержание дисциплины 5
3.1 Виды учебной работы и трудоемкость 6
3.2 Тематический план изучения дисциплины 6
3.3 Содержание дисциплины 6
4 Учебно-методическое и информационное обеспечение 7
4.1 Нормативные документы и ГОСТы 7
4.2 Основная литература 8
4.3 Дополнительная литература 8
5 Материально-техническое обеспечение 8
5.1 Требования к оборудованию и помещению для занятий 8
5.2 Требования к программному обеспечению 8
6 Методические рекомендации 8
6.1 Методические рекомендации для преподавателя по организации обучения 8
6.2 Методические указания для обучающихся по освоению дисциплины 9
7 Фонд оценочных средств 9
7.1 Методы контроля и оценивания результатов обучения 9
7.2 Шкала и критерии оценивания результатов обучения 9
7.3 Оценочные средства 14
7.3.1 Вопросы для экзамена 14
7.3.2 Типовые практические задания 14
7.3.3 Типовой экзаменационный билет 15

1 Цели, задачи и планируемые результаты обучения по дисциплине

К основным целям освоения дисциплины «Методы работы с большими данными» относится: формирование у студентов необходимой теоретической базы и практических навыков, которые позволят всесторонне и системно понимать современные проблемы прикладной математики и информатики, проблемы обработки и анализа информации, а также разрабатывать и анализировать концептуальные и теоретические модели при решении научных и прикладных задач в области информационных технологий.

К основным задачам дисциплины «Методы работы с большими данными» относятся:

- сформировать целостное представление о современных проблемах анализа и обработки больших данных;
- помочь овладеть опытом разработки и анализа концептуальных и теоретических моделей прикладных задач анализа больших данных с применением моделей Data Mining.

Обучение по дисциплине «Методы работы с большими данными» направлено на формирование у обучающихся следующих компетенций:

Код и наименование компетенций	**Индикаторы достижения компетенции**
УК-1. Способен осуществлять поиск, критический анализ и синтез информации, применять системный подход для решения поставленных задач	ИУК-1.1. Анализирует задачу, выделяя ее базовые составляющие ИУК-1.2. Осуществляет поиск, критически оценивает, обобщает, систематизирует и ранжирует информацию, требуемую для решения поставленной задачи ИУК-1.3. Рассматривает и предлагает рациональные варианты решения поставленной задачи, используя системный подход, критически оценивает их достоинства и недостатки
УК-2. Способен определять круг задач в рамках поставленной цели и выбирать оптимальные способы их решения, исходя из действующих правовых норм, имеющихся ресурсов и ограничений	ИУК-2.1.Формулирует совокупность задач в рамках поставленной цели проекта, решение которых обеспечивает ее достижение ИУК-2.2. Определяет связи между поставленными задачами, основными компонентами проекта и ожидаемыми результатами его реализации ИУК-2.3. Выбирает оптимальные способы планирования, распределения зон ответственности, решения задач, анализа результатов с учетом действующих правовых норм, имеющихся условий, ресурсов и ограничений, возможностей использования
ОПК-5. Способен инсталлировать программное и аппаратное обеспечение для информационных и автоматизированных систем	ИОПК-5.1. Знает методы информационного взаимодействия информационных и автоматизированных систем.

	ИОПК-5.3. Владеет: методами установки системного и прикладного программного обеспечения.
ПК-3. Способен разрабатывать требования и проектировать программное обеспечение	**ИПК-3.1. Знает** методологии и технологии проектирования и использования баз данных; методы и средства проектирования баз данных; методы и средства проектирования программных интерфейсов. **ИПК-3.2.** Умеет проводить анализ исполнения требований; вырабатывать варианты реализации требований; использовать существующие типовые решения и шаблоны проектирования программного обеспечения; применять методы и средства проектирования программного обеспечения, структур данных, баз данных, программных интерфейсов; осуществлять коммуникации с заинтересованными сторонами. **ИПК-3.3. Владеет** современным инструментарием и средами разработки программного кода; современным инструментарием и средами проектирования программного кода, методами тестирования ПО.

2 Место дисциплины в структуре образовательной программы

Дисциплина относится к числу учебных дисциплин формируемые участниками образовательных отношений части Дисциплины по выбору студента «Элективные дисциплины» основной профессиональной образовательной программы.

Дисциплина взаимосвязана логически и содержательно-методически со следующими дисциплинами и практиками ОПОП:

- Алгоритмическое программирование;
- Веб-разработка на стороне клиента;
- Веб-разработка;
- Серверная веб-разработка;
- Индексирование текстов и информационный поиск;
- Основы разработки виртуальной и дополнительной реальности;
- Разработка мобильных и веб-приложений;
- Основы разработки КИС.

3 Структура и содержание дисциплины

Общая трудоемкость дисциплины составляет 4 зачетных единицы, т.е. 144 академических часа (из них 72 часов – самостоятельная работа студентов и 72 часа – аудиторные занятия).

Разделы дисциплины изучаются на четвертом курсе в седьмом семестре, форма промежуточной аттестации – экзамен, курсовой проект.

3.1 Виды учебной работы и трудоемкость (по формам обучения)

3.1.1 Очная форма обучения

№ п/п	Вид учебной работы	Количество часов	Семестры	
			7	
1	**Аудиторные занятия**	**72**	72	
	В том числе:	8	8	
1.1	Лекции			
1.2	Семинарские/практические занятия			
1.3	Лабораторные занятия	64	64	
2	**Самостоятельная работа**	**72**	72	
	Курсовой проект		КП	
3	**Промежуточная аттестация**			
	Экзамен		экзамен	
	Итого:	**144**	144	

3.2 Тематический план изучения дисциплины (по формам обучения)

3.2.1 Очная форма обучения

№ п/п	Разделы/темы дисциплины	Трудоемкость, час					
		Всего	Аудиторная работа				Самостоятельная работа
			Лекции	Семинарские/ практические занятия	Лабораторные занятия	Практическая подготовка	
1	Тема 1. Обработка данных с использованием библиотеки Pandas	17	1		8		8
2	Тема 2. Визуализация данных и их анализ	17	1		8		8
3	Темам 3. Классификация данных	17	1		8		8
4	Тема 4. Кластеризация данных	17	1		8		8
5	Тема 5. Обработка текстов на естественном языке	17	1		8		8
6	Тема 6. Обработка изображений	17	1		8		8
7	Тема 7. «Плотные» нейронные сети	17	1		8		8
8	Тема 8. Глубинное обучение нейронных сетей	17	1		8		8
	Курсовой проект	8					8
Итого		**144**	**8**		**64**		**72**

3.3 Содержание дисциплины

1. Обработка данных с использованием библиотеки Pandas

Основные возможности библиотеки Pandas: загрузка и выборка данных, агрегирование данных, нормализация данных.

2. Визуализация данных и их анализ

Основные виды графиков для отображения данных: диаграммы рассеяния, размаха, гистограммы, эпюры, отображение трехмерных данных, отображение последовательностей. Элементы графика: оси, легенда, надписи. Методы снижения размерности пространства признаков: метод главных компонент, многомерное шкалирование, t-SNE, UMAP.

3. Классификация данных

Линейные методы классификации данных: линейная и логистическая регрессия, SMV. Методы, основанные на деревьях принятия решения, в том числе, метод случайного леса. Метод k ближайших соседей. Методы бустинга и стеккинга. Методы оценки результатов классификации: точность, полнота, f-мера, ROC-AUC, матрица ошибок.

4. Кластеризация данных

Метод k-средних, спектральная кластеризация. Методы, основанные на оценке плотности распределения точек в пространстве. Методы оценки точности кластеризации.

5. Обработка текстов на естественном языке

Лексический и синтаксический анализ текстов. Понятия пространства признаков для текста. Задачи обработки текстов: выделение именованных сущностей, фактов. Классификация и кластеризация текстов. Технологии Word2Vec, Glove.

6. Обработка изображений

Основные форматы хранения изображений и их отличия: BMP, PNG, JPG, GIF. Методы внесения изменений в изображения. Методы выделения областей изображения при помощи кластеризации точек. Библиотеки Python я работы с изображениями. Библиотека OpenCV и методы обработки изображений в sklearn.

7. «Плотные» нейронные сети

Бионические и искусственные нейронные сети, нейрон МакКаллока и Питтса, персептрон, сети Кохонена. Понятия порогового элемента и функции. Применение нейронных сетей для решения задач классификации и преобразования данных.

8. Глубинное обучение нейронных сетей

Понятие свертки в нейронных сетях, сверточные нейронные сети. Введение обратной связи в нейронной сети, рекуррентные сети. Виды сверточных и рекуррентных нейронных сетей. Построение архитектуры нейронной сети для решения прикладных задач.

4 Учебно-методическое и информационное обеспечение

4.1 Нормативные документы и ГОСТы

1. Федеральный закон от 29 декабря 2012 года № 273-ФЗ «Об образовании в Российской Федерации» (с изменениями и дополнениями);

2. Федеральный государственный образовательный стандарт высшего образования - бакалавриат по направлению подготовки 09.03.01 Информатика и вычислительная техника, утвержденный Приказом Министерства образования и науки РФ от 19 сентября 2017 г. N 929 "Об утверждении федерального… Редакция с изменениями N 1456 от 26.11.2020

3. Приказ Министерства образования и науки РФ от 05 апреля 2017 г. № 301 «Об утверждении Порядка организации и осуществления образовательной деятельности по образовательным программам высшего образования – программам бакалавриата, программам специалитета, программам магистратуры;

4. Порядок проведения государственной итоговой аттестации по образовательным программам высшего образования – программам бакалавриата, программам специалитета и программам магистратуры, утвержденный приказом Минобрнауки России от 29 июня 2015 г. № 636;

5. Положение о практической подготовке обучающихся, утвержденное приказом Министерства науки и высшего образования Российской Федерации и Министерства просвещения Российской Федерации от 5 августа 2020 г. № 885/390.

4.2 Основная литература

1. Чернышев, С. А. Основы программирования на Python : учебное пособие для вузов / С. А. Чернышев. — Москва : Издательство Юрайт, 2023. — 286 с. — (Высшее образование). — ISBN 978-5-534-14350-8. — Текст : электронный // Образовательная платформа Юрайт [сайт]. — URL: https://urait.ru/bcode/519949

2. Анализ данных : учебник для вузов / В. С. Мхитарян [и др.] ; под редакцией В. С. Мхитаряна. — Москва : Издательство Юрайт, 2023. — 490 с. — (Высшее образование). — ISBN 978-5-534-00616-2. — Текст : электронный // Образовательная платформа Юрайт [сайт]. — URL: https://urait.ru/bcode/511020

4.3 Дополнительная литература

1. Миркин, Б. Г. Введение в анализ данных : учебник и практикум / Б. Г. Миркин. — Москва : Издательство Юрайт, 2023. — 174 с. — (Высшее образование). — ISBN 978-5-9916-5009-0. — Текст : электронный // Образовательная платформа Юрайт [сайт]. — URL: https://urait.ru/bcode/511121

2. Черпаков, И. В. Основы программирования : учебник и практикум для вузов / И. В. Черпаков. — Москва : Издательство Юрайт, 2023. — 219 с. — (Высшее образование). — ISBN 978-5-9916-9983-9. — Текст : электронный // Образовательная платформа Юрайт [сайт]. — URL: https://urait.ru/bcode/511703

4.4 Электронные образовательные ресурсы

ЭОР не запланирован

5 Материально-техническое обеспечение

5.1 Требования к оборудованию и помещению для занятий

Лабораторные работы и самостоятельная работа студентов должны проводиться в специализированной аудитории, оснащенной современной оргтехникой и персональными компьютерами с программным обеспечением в соответствии с тематикой изучаемого материала. Число рабочих мест в аудитории должно быть достаточным для обеспечения индивидуальной работы студентов. Рабочее место преподавателя должно быть оснащены современным компьютером с подключенным к нему проектором на настенный экран, или иным аналогичным по функциональному назначению оборудованием.

5.2 Требования к программному обеспечению

Для выполнения лабораторных работ и самостоятельной работы необходимо следующее программное обеспечение:

1. Microsoft Windows.
2. Веб-браузер, Chrome.

6 Методические рекомендации

6.1 Методические рекомендации для преподавателя по организации обучения

1. При подготовке к занятиям следует предварительно проработать материал занятия, предусмотрев его подачу точно в отведенное для этого время занятия. Следует подготовить необходимые материалы – теоретические сведения, задачи и др. При проведении занятия

следует контролировать подачу материала и решение заданий с учетом учебного времени, отведенного для занятия.

2. При проверке работ и отчетов следует учитывать не только правильность выполнения заданий, но и оптимальность выбранных методов решения, правильность выполнения всех его шагов.

6.2 Методические указания для обучающихся по освоению дисциплины

Изучение дисциплины осуществляется в строгом соответствии с целевой установкой в тесной взаимосвязи с учебным планом. Основой теоретической подготовки студентов являются аудиторные занятия, лабораторные работы.

В процессе самостоятельной работы студенты закрепляют и углубляют знания, полученные во время аудиторных занятий, дорабатывают конспекты и записи, готовятся к проведению и обрабатывают результаты лабораторных работ, готовятся к промежуточной аттестации, а также самостоятельно изучают отдельные темы учебной программы.

На занятиях студентов, в том числе предполагающих практическую деятельность, осуществляется закрепление полученных, в том числе и в процессе самостоятельной работы, знаний. Особое внимание обращается на развитие умений и навыков установления связи положений теории с профессиональной деятельностью будущего специалиста.

Самостоятельная работа осуществляется индивидуально. Контроль самостоятельной работы организуется в двух формах:

- самоконтроль и самооценка студента;
- контроль со стороны преподавателей (текущий и промежуточный).

Критериями оценки результатов самостоятельной работы студента являются:

- уровень освоения студентом учебного материала;
- умения студента использовать теоретические знания при выполнении практических задач;
- сформированность компетенций;
- оформление материала в соответствии с требованиями.

7 Фонд оценочных средств

7.1 Методы контроля и оценивания результатов обучения

В процессе обучения используются следующие оценочные формы самостоятельной работы студентов, оценочные средства текущего контроля успеваемости и промежуточных аттестаций:

- лабораторные работы, экзамен.

7.2 Шкала и критерии оценивания результатов обучения

Показателем оценивания компетенций на различных этапах их формирования является достижение обучающимися планируемых результатов обучения по дисциплине.

Показатель	**Критерии оценивания**			
	2	**3**	**4**	**5**
УК-1. Способен осуществлять поиск, критический анализ и синтез информации, применять системный подход для решения поставленных задач				
ИУК-1.1. Анализирует задачу, выделяя ее	Обучающийся демонстрирует полное отсутствие	Обучающийся демонстрирует неполное соответствие	Обучающийся демонстрирует частичное соответствие	Обучающийся демонстрирует полное соответствие

базовые составляющие ИУК-1.2. Осуществляет поиск, критически оценивает, обобщает, систематизирует и ранжирует информацию, требуемую для решения поставленной задачи ИУК-1.3. Рассматривает и предлагает рациональные варианты решения поставленной задачи, используя системный подход, критически оценивает их достоинства и недостатки	или недостаточное соответствие материалу дисциплины знаний, указанных в индикаторах компетенций дисциплины.	следующих знаний, указанных в индикаторах компетенций дисциплины. Допускаются значительные ошибки, проявляется недостаточность знаний, по ряду показателей, обучающийся испытывает значительные затруднения при оперировании знаниями при их переносе на новые ситуации.	следующих знаний, указанных в индикаторах компетенций дисциплины. Но допускаются незначительные ошибки, неточности, затруднения при аналитических операциях.	следующих знаний, указанных в индикаторах компетенций дисциплины. Свободно оперирует приобретенным и знаниями.
УК-2. Способен определять круг задач в рамках поставленной цели и выбирать оптимальные способы их решения, исходя из действующих правовых норм, имеющихся ресурсов и ограничений				
ИУК-2.1. Формулирует совокупность задач в рамках поставленной цели проекта, решение которых обеспечивает ее достижение ИУК-2.2. Определяет связи между поставленными задачами, основными компонентами проекта и ожидаемыми результатами его реализации ИУК-2.3. Выбирает оптимальные	Обучающийся демонстрирует полное отсутствие или недостаточное соответствие материалу дисциплины знаний, указанных в индикаторах компетенций дисциплины.	Обучающийся демонстрирует неполное соответствие следующих знаний, указанных в индикаторах компетенций дисциплины. Допускаются значительные ошибки, проявляется недостаточность знаний, по ряду показателей, обучающийся испытывает значительные затруднения	Обучающийся демонстрирует частичное соответствие следующих знаний, указанных в индикаторах компетенций дисциплины. Но допускаются незначительные ошибки, неточности, затруднения при аналитических операциях.	Обучающийся демонстрирует полное соответствие следующих знаний, указанных в индикаторах компетенций дисциплины. Свободно оперирует приобретенным и знаниями.

способы планирования, распределения зон ответственности, решения задач, анализа результатов с учетом действующих правовых норм, имеющихся условий, ресурсов и ограничений, возможностей использования.		при оперировании знаниями при их переносе на новые ситуации.		
ОПК-5. Способен инсталлировать программное и аппаратное обеспечение для информационных и автоматизированных систем				
ИОПК-5.1. Знает методы информационного взаимодействия информационных и автоматизированных систем. ИОПК-5.3. Владеет: методами установки системного и прикладного программного обеспечения.	Обучающийся демонстрирует полное отсутствие или недостаточное соответствие материалу дисциплины знаний, указанных в индикаторах компетенций дисциплины.	Обучающийся демонстрирует неполное соответствие следующих знаний, указанных в индикаторах компетенций дисциплины. Допускаются значительные ошибки, проявляется недостаточность знаний, по ряду показателей, обучающийся испытывает значительные затруднения при оперировании знаниями при их переносе на новые ситуации.	Обучающийся демонстрирует частичное соответствие следующих знаний, указанных в индикаторах компетенций дисциплины. Но допускаются незначительные ошибки, неточности, затруднения при аналитических операциях.	Обучающийся демонстрирует полное соответствие следующих знаний, указанных в индикаторах компетенций дисциплины. Свободно оперирует приобретенным и знаниями.
ПК-3. Способен разрабатывать требования и проектировать программное обеспечение				
ИПК-3.1. Знает методологии и технологии проектирования и	Обучающийся демонстрирует полное отсутствие	Обучающийся демонстрирует неполное соответствие	Обучающийся демонстрирует частичное соответствие	Обучающийся демонстрирует полное соответствие

использования баз данных; методы и средства проектирования баз данных; методы и средства проектирования программных интерфейсов. ИПК-3.2. Умеет проводить анализ исполнения требований; вырабатывать варианты реализации требований; использовать существующие типовые решения и шаблоны проектирования программного обеспечения; применять методы и средства проектирования программного обеспечения, структур данных, баз данных, программных интерфейсов; осуществлять коммуникации с заинтересованными сторонами. ИПК-3.3. Владеет современным инструментарием и средами разработки программного кода; современным инструментарием и средами проектирования программного кода, методами тестирования ПО.	или недостаточное соответствие материалу дисциплины знаний, указанных в индикаторах компетенций дисциплины.	следующих знаний, указанных в индикаторах компетенций дисциплины. Допускаются значительные ошибки, проявляется недостаточность знаний, по ряду показателей, обучающийся испытывает значительные затруднения при оперировании знаниями при их переносе на новые ситуации.	следующих знаний, указанных в индикаторах компетенций дисциплины. Но допускаются незначительные ошибки, неточности, затруднения при аналитических операциях.	следующих знаний, указанных в индикаторах компетенций дисциплины. Свободно оперирует приобретенным и знаниями.

Шкала оценивания результатов промежуточной аттестации и её описание:
Форма промежуточной аттестации: экзамен.

Промежуточная аттестация обучающихся в форме экзамена проводится по результатам выполнения всех видов учебной работы, предусмотренных учебным планом по данной дисциплине, при этом учитываются результаты текущего контроля успеваемости в течение семестра. Оценка степени достижения обучающимися планируемых результатов обучения по дисциплине проводится преподавателем, ведущим занятия по дисциплине методом экспертной оценки. По итогам промежуточной аттестации по дисциплине выставляется оценка «отлично», «хорошо», «удовлетворительно» или «неудовлетворительно».

Шкала оценивания	**Описание**
Отлично	Выполнены все **обязательные условия подготовки студента к промежуточной аттестации**, предусмотренные программой дисциплины. Студент демонстрирует соответствие знаний, умений, навыков показателям, приведенным в таблицах, оперирует приобретенными знаниями, умениями, навыками, применяет их в ситуациях повышенной сложности. При этом могут быть допущены незначительные ошибки, неточности, затруднения при переносе знаний и умений на новые, нестандартные задачи.
Хорошо	Выполнены все **обязательные условия подготовки студента к промежуточной аттестации**, предусмотренные программой дисциплины. Студент демонстрирует соответствие знаний, умений, навыков показателям, приведенным в таблицах, оперирует приобретенными знаниями, умениями, навыками. В то же время при ответе допускает несущественные погрешности, задачи решает с недочетами, не влияющими на общий ход решения.
Удовлетворительно	Выполнены все **обязательные условия подготовки студента к промежуточной аттестации**, предусмотренные программой дисциплины. Студент демонстрирует соответствие знаний, умений, навыков показателям, приведенным в таблицах, оперирует приобретенными знаниями, умениями, навыками. Но показывает неглубокие знания, при ответе не допускает грубых ошибок или противоречий, однако в формулировании ответа отсутствует должная связь между анализом, аргументацией и выводами, в решении задач могут содержаться грубые ошибки. Для получения правильного ответа требуются уточняющие вопросы.
Неудовлетворительно	Не выполнены обязательные **условия подготовки студента к промежуточной аттестации**, предусмотренные программой дисциплины, ИЛИ студент демонстрирует неполное соответствие знаний, умений, навыков приведенным в таблицах, допускаются значительные ошибки, проявляется отсутствие знаний, умений, навыков по ряду показателей, студент испытывает значительные затруднения при оперировании знаниями и умениями.

7.3 Оценочные средства

7.3.1 Вопросы для экзамена

1. Что такое библиотека Pandas? Для чего она используется? Какие возможности предоставляет?
2. Что такое визуализация данных?
3. Перечислите основные виды графиков для отображения данных и кратко опишите каждый из них.
4. Какие элементы графика существуют? Опишите их.
5. Перечислите методы снижения размерности пространства признаков и кратко опишите каждый из них.
6. Что такое классификация данных?
7. Какие линейные методы классификации данных существуют? Опишите их.
8. Опишите методы, основанные на деревьях принятия решения, в том числе, метод случайного леса.
9. Перечислите и опишите методы оценки результатов классификации.
10. Что такое кластеризация данных?
11. Для чего используется метод k-средних? Опишите его.
12. Для чего используется спектральная кластеризация? Опишите ее.
13. Опишите методы, основанные на оценке плотности распределения точек в пространстве.
14. Опишите методы оценки точности кластеризации.
15. Что такое лексический анализ текстов?
16. Что такое синтаксический анализ текстов?
17. Опишите технологии Word2Vec и Globe.
18. Перечислите основные форматы хранения изображений и опишите их отличия.
19. Что такое «плотные» нейронные сети?
20. Что такое нейрон МакКаллока и Питтса?
21. Что такое персептрон?
22. Что такое сети Кохонена?
23. Что такое свертка в нейронных сетях?
24. Что такое рекуррентные сети?
25. Опишите виды сверточных и рекуррентых нейронных сетей.
26. Перечислите основные базы знаний Wolfram Data Platform.
27. Что представляет из себя формат CDF?
28. Что такое канонический идентификатор WDF?
29. Что такое хеширование?
30. Как осуществляется хранение и сравнение идентификационных данных пользователей компьютерных сетей?

7.3.2 Типовые практические задания

1. Оцените общий объем данных, который могло бы вместить наиболее емкое теоретически возможное хранилище данных.

2. Предложите алгоритм выявления подозрительных транзакций по банковской карте.

3. Предложите алгоритм выявления нарушения автомобилем скоростного режима на дороге на основе данных камер видеофиксации.

4. Оцените количество слов в орфографическом словаре русского языка.

5. Предложите стандарт формата цифрового следа пользователя интернета и метрику в пространстве таких следов.

7.3.3 Типовой экзаменационный билет

МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ РОССИЙСКОЙ ФЕДЕРАЦИИ
ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ АВТОНОМНОЕ ОБРАЗОВАТЕЛЬНОЕ УЧРЕЖДЕНИЕ ВЫСШЕГО ОБРАЗОВАНИЯ
«МОСКОВСКИЙ ПОЛИТЕХНИЧЕСКИЙ УНИВЕРСИТЕТ»
(МОСКОВСКИЙ ПОЛИТЕХ)

ЭКЗАМЕНАЦИОННЫЙ БИЛЕТ №1
по дисциплине
«Методы работы с большими данными»
направление подготовки 09.03.01 Информатика и вычислительная техника

ВОПРОСЫ:

1. Какие линейные методы классификации данных существуют? Опишите их.
2. Что такое персептрон?
3. Что такое канонический идентификатор WDF?

Утверждено: _____ / _____ / «__» _____ 20__г.