Part					
Part	ходы и ванных	нты, подх ктуриров « данных"	струмен ки стру ванных	ине "Ин обработ стурирс	дисципли методы с и неструк всего 15/20
Part				Владиславови	
	ой <b>*</b> 0 и: 1			может быть об	формат, который г
	*1 и 1	анения	база данных	ных данных? цанных і база данных иентированная	<ul><li>неструктурирован</li><li>1. NoSQL база д</li><li>2. Реляционная</li><li>3. Объектно-орг</li></ul>
	?* 1 из	кстовых данных?	я обработки те	ec	<ul><li>1. Случайный л</li><li>2. К-средних</li><li>3. DBSCAN</li></ul>
Part	*1 us 1		структурирова i обработки.	ных данных в компьютерно обработки чел	<ul><li>неструктурирован</li><li>1. возможность</li><li>2. возможность</li><li>3. наглядность</li></ul>
	1 из	х в Интернете? <b>*</b>	ванных данных		<ul> <li>5. Какой процент н</li> <li>1. 0%</li> <li>2. 20%</li> <li>3. 80%</li> </ul>
1   1   1   1   1   1   1   1   1   1		байта) з 32 битных целых ч	и́та) мает 64 бита (8 енной) длины из	ет 32 бита (4 ба ути, double, зан ходимой (перем	<ul> <li>1. float, занимає</li> <li>2. float, но, по су</li> <li>3. массив необх</li> <li>4. два массива</li> </ul>
	* 1 из	іной регрессии? *	DВ	ции ых векторов ньших квадрат	7. Что используето  1. Метод бисеки  2. Метод опорн  3. Метод наиме
The first content of the content of	занных *1 из	ки структурирова	ся для обработ	редакторы едакторы ые системы	Данных?  1. Графические  2. Текстовые ре  3. Операционны
The factor of depotations of cores in the constraint representation of the potation of the pot	,лины <b>*</b> 1 из 1	инарных строк дл	іх различных б	тво возможнь	8?
The content of the	ании *1 из 1	я при использован	of words)? окументе; окумент; умент;	ешок слов (bag разных слов в д кдений слов в до ений слов в до	12. Какая информа представления мето раставления мето раставления мето раставления мето раставления в раставления
Section 1. Continues to continues and activation of the contin	Баллы	я Стемминг	т баллов.	гвет не приносі	
The contract of the contract o					текста на отдельные слова или словоформы
In Contractive or income in the contractive of the contractive or income					слова к его основе путём отбрасывания окончаний приведение слова к его
RECORD TO STANDARD			ываемый мето,	исание и опис	нормальной форме 11. Соотнесите оп
Source and a control of the control	Баллы	DB2CAN	Иерархическая		находит плотные
Secretaria de la companya de la contractiva de contractiva de la contractiva del cont	1 из 1				пространстве объектов и считает их кластерами, а объекты, находящиеся вне плотных регионов, - выбросами строит иерархическое
Security of the protection of a protection of					объединяя на каждом шаге два ближайших кластера разбивает данные на К кластеров, минимизируя сумму квадратов расстояний
Morran  Morran		ером: *			кластеров 13. Соотнесите тиг
In the property of the prop	Баллы		Количественні	(атегориальны	можно измерить и представить
(мателисти. В при выполнять на продоставления в продоставления произволять долго обработна продоставления в порядок применения методов продоставления долгост порядок применения порядок применения порядок продоставления применения (продоставления продоставления применения продоставления применения (продоставления применения). Удаления выбросов к баровения развитами. В порядок применения применения (продоставления применения). Удаления выбросов к баровения порядок применения методов обработам и применения применения. Удаления выбросов к баровения порядок применения методов обработам и применения методов обработам применения. В применения (применения). В применения применения (применения). В применен					(например, возраст, доход, количество товаров) могут принимать ограниченное число значений из заданного
14. Выборите правильный порядок применения методое предобработки данных.  □ 1. Обработка пропривенных эменений: Удаление выбросов; Нормализации данных (удительное выбросов; Нормализации данных признанов больботки прогуденных заменных (удительных править выбросов; Нормализации данных (удительных править заменных данных да	0 из 0				(например, цвет, пол, тип автомобиля) могут принимать только два значения (например, да/ нет, 0/1, муж/
текстовых данных.  1. Стемминг / Лемматизация; Токенизация; Синтаксический анализ; 2. Токенизация; Синтаксический анализ; 3. Токенизация; Синтаксический анализ; Стемминг / Лемматизация; 4. Синтаксический анализ; Токенизация; Стемминг / Лемматизация; 4. Синтаксический анализ; Токенизация; Стемминг / Лемматизация;  16. Какие из следующих языков программирования наиболее часто используются для обработки и анализа данных?  17. Какие из следующих методов можно использовать для обработки прогущенных значений в структурированных данных?  18. Какие из следующих методов можно использовать для обработки прогущенных значений на преднес явачение столбца  3. Удаление строк с прогущенными значениями  4. Замена пролущенных значений на среднее явачение столбца  18. Какие из следующих методов обработки текстовых данных используются для изаличении ключевых слов из текста?  1. Лемматизация  2. LDA  3. Стемминг  4. ТЕТОБ  19. Как будет высплаеть векторное представление мешка слов для набора текстов: «Это были луушие временаз и «Это было худшее время»?  1. ПП101[[1011]]  2. [11010][[1001]]  2. [11010][[1001]]  2. [11010][[1001]]  7. Какова доля правильных ответов модели, если для нее ТР = 3, FP = 4, FN = 7, FN = 7, TN = 3?  1. Тите Розітіче (ГР) Разе Розітіче (ГР) а(х) = −1 False педатіче (ГР) Тите Negative (ТР) Таблица 1. Матрица опшбок  1. 30%  2. 35%  3. 50%  4. 70%	ализация егориальных аление	е выбросов; Нормал ков ; Кодирование катего ; нных значений; Удал наков бработка пропущенн	чений; Удалению пальных признак пизация данных; енных значений; пботка пропущен приальных призн пх признаков; Об	нных.  оопущенных знование категоробросов; Нормальных; Обрирование категориальных	предобработки да  1. Обработка приданных; Кодиро  2. Удаление вы признаков; Обр  3. Нормализация выбросов; Коди
используются для обработки и анализа данных?  1. R 2. C++ 3. Java 4. Руфоп  17. Какие из следующих методов можно использовать для обработки пропущенных значений в структурированных данных?  1. Интерполяция 2. Замена пропущенных значений на медианное аначение столбца 3. Удаление строк с пропущенных значений на среднее значение столбца 13. Какие из следующих методов обработки текстовых данных используются для извлечения ключевых слов из текста?  1. Лемматизация 2. LDA 3. Стемминг  4. TF-IDF  19. Как будет выглядеть векторное представление мешка слов для набора текстов: «Это были лучшие времена» и «Это было худшее время»?  1. [11101] [11011]: 2. [110101] [11011]: 3. [1110] [1101]: 4. [01101] [1101]: 7. FN = 7, TN = 3?  20. Какова доли правильных ответов модели, если для нее TP = 3, FP = 7, FN = 7, TN = 3?  21. Ванизация [Труматизация]	1 лиз; лиз ция;	нтаксический анали нтаксический анали минг / Лемматизаци	окенизация; Сиі мматизация; Сиі ій анализ; Стемі	с. Іемматизация; ; Стемминг / Ле ; Синтаксическ	<ul><li>Текстовых данных</li><li>1. Стемминг / Л</li><li>2. Токенизация</li><li>3. Токенизация</li></ul>
пропущенных значений в структурированных данных?  ☑ 1. Интерполяция  ☑ 2. Замена пропущенных значений на медианное значение столбца  ☑ 3. Удаление строк с пропущенными значениями  ☑ 4. Замена пропущенных значений на среднее значение столбца  18. Какие из следующих методов обработки текстовых данных используются для извлечения ключевых слов из текста?  1. Лемматизация  ② 2. LDA  ③ 3. Стемминг  ☑ 4. TF-IDF  19. Как будет выглядеть векторное представление мешка слов для набора текстов: «Это были лучшие времена» и «Это было худшее время»?  ○ 1. [11101],[11011];  ② 2. [110101],[100111];  ③ 3. [1110],[11001],  20. Какова доля правильных ответов модели, если для нее TP = 3, FP = 7, FN = 7, TN = 3?  ☐ 1. 30*  ☐ 1. 30*  ☐ 2. 35%  ☐ 3. 50%  ☐ 4. 70%	iасто *1 из 1				используются для  1. R  2. C++  3. Java
используются для извлечения ключевых слов из текста?  1. Лемматизация 2. LDA 3. Стемминг 4. TF-IDF  19. Как будет выглядеть векторное представление мешка слов для набора текстов: «Это были лучшие времена» и «Это было худшее время»?  1. [11101], [11011]; 2. [1110101], [11001]; 3. [1110], [11001].  20. Какова доля правильных ответов модели, если для нее TP = 3, FP = ₹7, FN = 7, TN = 3?  10. [1101], [11001	и: 1	цанных? ое значение столбца ии	урированных д ний на медианно ными значениям	чений в струк ущенных значе рок с пропущен	пропущенных знач 1. Интерполяци 2. Замена проп 3. Удаление стр
19. Как будет выглядеть векторное представление мешка слов для набора текстов: «Это были лучшие времена» и «Это было худшее время»?  ① 1. [1 1 1 0 1], [1 1 0 1 1]; ② 2. [1 1 0 1 0 1], [1 0 0 1 1 1]; ③ 3. [1 1 1 0], [1 1 0 0 1].  20. Какова доля правильных ответов модели, если для нее TP = 3, FP = *7, FN = 7, TN = 3?  20. Какова доля правильных ответов модели, если для нее TP = 3, FP = *7, FN = 7, TN = 3?  21. 32. 32. 32. 32. 33. 43. 34. 34. 34. 34. 34. 34. 34. 34	*0 из 1			извлечения к	используются для  1. Лемматизаці  2. LDA  3. Стемминг
20. Какова доля правильных ответов модели, если для нее TP = 3, FP = *7, FN = 7, TN = 3?				Это были лучц   0 1 1];   0 0 1 1 1];   1];	19. Как будет выгл набора текстов: «З 1. [1 1 1 0 1], [1 1 2. [1 1 0 1 0 1], [1 1 3. [1 1 1 0], [1 1 0
<ul><li> 3. 50%</li><li> 4. 70%</li><li> Форма создана в домене СурГУ.</li></ul>	1	y = -1 to Positive (FP) to Negative (TN)	(TP) False (FN) True	равильных от $y=1$ True Positive alse negative	20. Какова доля проводя $a(x) = 1$ $a(x) = -1$ $a(x)$
					3. 50%