

«

»

“ ”

.

31.08.2022

:

:

:

<https://www.nstu.ru/university/info/sveden/education>

РАБОЧАЯ ПРОГРАММА ДИСЦИПЛИНЫ

Анализ данных

: 22.03.01

, :

: 3, : 6

		-	,
			6
1	()		4
2			144
3	, .		60
4	, .		18
5	, .		36
6	, .		0
7	, .		6
8	, .		33
9	, .		2
10	, .		4
11	, .		84
12	(, ()/ ,)		
13			

(): 22.03.01

701 02.06.2020 ., : 10.07.2020 .

: 1,

,

(): 22.03.01

, 31.08.2022

- , 6 31.08.2022

:

,

:

. .

1.

1.1

	-2. / , , , , ,
	-2. / . 2
	-1 , ,
	-1. 1 , .

2.

,

2.1

ПК-2.В/НА. 2 Умеет проводить комплексные технологические расчеты с использованием стандартных и специальных продуктов	
	;
УК-1. 1 Знает принципы сбора, отбора и обобщения информации.	
, ,	;

3.

3.1

		„ . .	, .		
: 6					
: ,					
1. . .	2	0	0	-1.1	
2. . .	2	0	0	-1.1	
:					
3. Python	4	0	2	-1.1	
:					

4.		4	0	0	-1.1	
:						
5.		2	0	0	-1.1	
:						
6.		2	0	0	-1.1	
:						
7.		2	0	0	-1.1	

: 6						
:						
1.		3	3	0	$\frac{-2.}{2}$	
2.		4	4	0	$\frac{-2.}{2}$	
:						
3.	Python	8	6	2	$\frac{-2.}{2}$	
:						
4.		8	8	0	$\frac{-2.}{2}$	
:						
5.		9	8	1	$\frac{-2.}{2}$	
:						
6.		4	4	1	$\frac{-2.}{2}$	

: 6						
:						
1.		4	0	0	-1.1	

			()
1			:
2			:
3	Python		:
4	.		:
5			:
6			:

3.2

3.3

: 6				
1		-2. / .2, -1.1	50	0
: [] : , [2020].- : http://elibrary.nstu.ru/source?bib_id=vtls000243326.- []: - , [2020].- : http://elibrary.nstu.ru/source?bib_id=vtls000242416.-				
2		-2. / .2, -1.1	20	0
[]: - : http://elibrary.nstu.ru/source?bib_id=vtls000243326.- []: - , [2020].- : http://elibrary.nstu.ru/source?bib_id=vtls000242416.-				
3		-2. / .2, -1.1	10	4
: [] : , [2020].- : http://elibrary.nstu.ru/source?bib_id=vtls000243326.- []: - , [2020].- : http://elibrary.nstu.ru/source?bib_id=vtls000242416.-				
4		-1.1	4	0

3.3 : . . / . . ;
 [] : - / . . ;
 . . . - , [2020].- :
http://elibrary.nstu.ru/source?bib_id=vtls000243326.-
 [] :
 - / . . , . ;
 , [2020].- : http://elibrary.nstu.ru/source?bib_id=vtls000242416.-
 .

3.3

, (. 3.4).

3.4

	-
	;

4.

(), - 15- ECTS.
 . 4.1.

4.1

	.	
: 6		
<i>Самостоятельное изучение теоретического материала:</i>	15	20
<i>Лекция:</i>	5	15
<i>Практические занятия:</i>	10	25
<i>Экзамен:</i>	20	40

4.2

4.2

-2. / -2. / 2.		+
-1 -1 1. , .		+

5.

1. Современные методы и средства интеллектуального анализа данных : монография / [О. К. Альсова и др. ; под ред. Е. В. Рабиновича, А. А. Якименко, О. К. Альсовой] ; Новосиб. гос. техн. ун-т.- Новосибирск: Изд-во НГТУ, 2018.- 199 с. : ил.- Режим доступа: http://elibrary.nstu.ru/source?bib_id=vtls000239742

1. Попов А. А. Статистические методы анализа данных [Электронный ресурс] : электронный учебно-методический комплекс / А. А. Попов ; Новосиб. гос. техн. ун-т.- Новосибирск, [2020].- Режим доступа: http://elibrary.nstu.ru/source?bib_id=vtls000243326.- Загл. с экрана.

1. Github : distributed version control software : [site] / GitHub, Inc. – 2022. – URL: <http://github.com> (access data: 31.08.2022). – Text : electronic.

2. Stack Overflow : [site] / Stack Exchange Inc. – 2022. – URL: <https://stackoverflow.com/> (access data: 31.08.2022). – Text : electronic.

6.

6.1

1. Ганелина Н. Д. Теория вероятностей и математическая статистика [Электронный ресурс] : электронный учебно-методический комплекс / Н. Д. Ганелина, В. В Ландовский ; Новосиб. гос. техн. ун-т.- Новосибирск, [2020].- Режим доступа: http://elibrary.nstu.ru/source?bib_id=vtls000242416.- Загл. с экрана.

6.2

6.3

7.

1	(- , ,)	
2	(Internet)	

1. Обобщенная структура фонда оценочных средств дисциплины

Обобщенная структура фонда оценочных средств по дисциплине Анализ данных представлена в Таблице. Совокупность результатов обучения по дисциплине соотнесена с уровнями сформированности компетенций и соотнесенными с ними индикаторами. Индикаторы достижения компетенций измеряемы с помощью средств текущей и промежуточной аттестации по дисциплине Анализ данных.

Таблица

Формируемые компетенции	Индикаторы компетенций	Темы	Этапы оценки результатов обучения и соотнесенных с ними индикаторов достижения компетенций	
			Мероприятия текущего контроля (контрольная работа, курсовой проект, РГЗ(Р), реферат и др.)	Промежуточная аттестация (экзамен, зачет)
ПК-2.В/НА Способность использовать в исследованиях и расчетах знания о методах исследования, анализа, диагностики и моделирования свойств материалов, физических и химических процессах, протекающих в материалах при их получении, обработке и модификации	2. Умеет проводить комплексные технологические расчеты с использованием стандартных и специальных продуктов	Библиотеки Python для предобработки данных Классификация и кластеризация данных Линейная регрессия. Градиентный спуск Матричные преобразования Проверка статистических гипотез Стохастический градиентный спуск	Оценка устных ответов на практических занятиях: библиотеки Python для предобработки данных, градиентный спуск, проверка статистических гипотез	Экзамен, вопросы 9-18
УК-1 Способен осуществлять поиск, критический анализ и синтез информации, применять системный подход для решения поставленных задач	1. Знает принципы сбора, отбора и обобщения информации.	Классификация и кластеризация данных Линейная регрессия. Градиентный спуск. Матричные операции. Линейные преобразования. Сингулярное разложение матриц. Обучение с учителем и без учителя. Регрессионные алгоритмы. Регуляризация. Дерево решений. Кластеризация и классификация. Основные библиотеки Python для предобработки данных Основные принципы алгоритмов по обработке данных Понятие вероятности. Теорема Байеса. Вероятностные распределения. Проверка гипотез. Стохастический градиентный спуск. Снижение размерности данных	Оценка устных ответов на практических занятиях: сингулярное разложение матриц, кластеризация и классификация, снижение размерности данных.	Экзамен, вопросы 1-18

2. Методика оценки этапов формирования компетенций по дисциплине

Результаты обучения по дисциплине, соотнесенные с установленными в программе индикаторами достижения компетенций проверяются при проведении мероприятий текущей аттестации (контроля) в процессе изучения дисциплины, указанных в таблице раздела 1.

Промежуточная аттестация по **дисциплине** проводится в 6 семестре - в форме экзамена, который направлен на оценку сформированности результатов обучения по дисциплине, соотнесенных с установленными в программе индикаторами достижения компетенций ПК-2.В/НА, УК-1 и соотнесенных с ними индикаторов. (см. таблицу раздела 1).

Экзамен проводится в устной форме по билетам, содержащим два вопроса, каждый из которых требует развернутого ответа с пояснениями и обоснованием излагаемого материала. Билет формируется из приведенного в Паспорте экзамена списка вопросов, позволяющих оценить результаты обучения по дисциплине (модулю) оставить нужное, соотнесенные с установленными в программе индикаторами достижения компетенций.

Общие правила выставления оценок текущей и промежуточной аттестации по дисциплине определяются балльно-рейтинговой системой, приведенной в рабочей программе дисциплины.

На основании критериев, приведенных в п. 3, осуществляется оценка уровней достигнутых результатов обучения по дисциплине, соотнесенных с установленными в программе индикаторами достижения компетенций ПК-2.В/НА, УК-1, закрепленных за дисциплиной.

3. Общая характеристика уровней результатов обучения, соотнесенных с индикаторами достижения компетенций

Продвинутый. Теоретическое содержание курса освоено полностью. Студент демонстрирует систематическое и глубокое понимание учебного материала и способность к самостоятельному пополнению и обновлению знаний в ходе дальнейшей учебной работы и профессиональной деятельности. Сформированы необходимые навыки практической работы. Все учебные задания, предусмотренные программой обучения, выполнены качественно, без замечаний. Результаты обучения по дисциплине (модулю), соотнесенные с установленными в программе индикаторами достижения компетенций, оценены числом баллов, входящим в диапазон продвинутого уровня.

Базовый. Теоретическое содержание курса освоено в объеме, необходимом для дальнейшей учебы и профессиональной деятельности. Навыки практической работы сформированы на базовом уровне. Все предусмотренные программой обучения учебные задания выполнены с небольшими погрешностями. Результаты обучения по дисциплине (модулю), соотнесенные с установленными в программе индикаторами достижения компетенций, оценены числом баллов в пределах базового уровня.

Пороговый. Теоретическое содержание курса освоено в объеме, необходимом для дальнейшей учебы и профессиональной деятельности. Некоторые практические навыки работы сформированы с пробелами. Учебные задания, предусмотренные программой обучения, выполнялись с ошибками, исправленными под руководством преподавателя. Результаты обучения по дисциплине (модулю), соотнесенные с установленными в программе индикаторами достижения компетенций, оценены числом баллов в пределах порогового уровня.

Ниже порогового. Теоретическое содержание курса освоено фрагментарно. Необходимые навыки практической работы сформированы минимально. Большинство учебных заданий, предусмотренных программой обучения, не выполнены. Результаты обучения по дисциплине (модулю), соотнесенные с установленными в программе индикаторами достижения компетенций, оценены числом баллов, входящих в диапазон ниже порогового уровня.

Паспорт экзамена

по дисциплине «Анализ данных», 6 семестр

1. Методика оценки

Экзамен проводится в устной форме по билетам. Билет состоит из 2 вопросов и формируется по следующему правилу:

- первый вопрос выбирается из диапазона вопросов 1-9;
- второй вопрос из диапазона вопросов 10-18.

Таким образом, проверяются результаты обучения по дисциплине, соотнесенные с установленными в программе индикаторами достижения компетенций.

На экзамене преподаватель вправе задавать студенту уточняющие и дополнительные вопросы из общего перечня (п. 4).

Форма экзаменационного билета

НОВОСИБИРСКИЙ ГОСУДАРСТВЕННЫЙ ТЕХНИЧЕСКИЙ УНИВЕРСИТЕТ
Факультет МТФ

Билет № _____

к экзамену по дисциплине «Анализ данных»

1. Непрерывные случайные величины. Функция распределения и функция плотности. Нормальное распределение. Центральная предельная теорема
2. Алгоритм линейной регрессии. Метод наименьших квадратов. Градиентный спуск

Утверждаю: зав. кафедрой _____ должность, ФИО
(подпись) (дата)

2. Критерии оценки результатов обучения, соотнесенных с уровнями освоения индикаторов достижения компетенций

Ответ на экзаменационный билет засчитывается на **продвинутом** уровне, если студент проводит сравнительный комплексный анализ материала, выявляет проблемы, предлагает механизмы их решения, представляет количественные характеристики определенных процессов, приводит конкретные примеры, не допускает ошибок и способен обосновать выбор метода решения задачи. Совокупность результатов обучения по

дисциплине (модулю) и соотнесенных с ними индикаторов достижения компетенций не содержит пробелов. Установленные в программе компетенции сформированы в полном объеме. Оценка составляет *от 35 до 40 баллов*.

Ответ на экзаменационный билет засчитывается на **базовом** уровне, если студент при ответе на вопросы формулирует основные понятия, дает характеристику процессов, явлений, проводит анализ причин, условий, способен представить количественные и качественные характеристики процессов. Совокупность результатов обучения по дисциплине и соотнесенных с ними индикаторов достижения компетенций содержит несущественные пробелы. Установленные в программе компетенции сформированы на базовом уровне. Оценка составляет *от 34 до 27 баллов*.

Ответ на экзаменационный билет (тест) засчитывается на **пороговом** уровне, если студент при ответе на вопросы дает определение основных понятий, может показать причинно-следственные связи явлений. Совокупность результатов обучения по дисциплине (модулю) и соотнесенных с ними индикаторов достижения компетенций содержит пробелы. Установленные в программе компетенции сформированы на пороговом уровне. Оценка составляет *от 26 до 20 баллов*.

Ответ на экзаменационный билет (тест) считается **неудовлетворительным** (ниже порогового уровня), если студент при ответе на вопросы не дает определений основных понятий, не способен показать причинно-следственные связи явлений, допускает принципиальные ошибки. Совокупность результатов обучения по дисциплине (модулю) и соотнесенных с ними индикаторов достижения компетенций содержит существенные пробелы. Установленные в программе компетенции не сформированы. Оценка составляет *менее 20 баллов*.

3. Шкала оценки

Экзамен считается сданным, если сумма баллов по всем заданиям билета составляет от 20 до 40 баллов включительно. Сумма менее 20 баллов признается неудовлетворительным результатом промежуточной аттестации по дисциплине.

В общей оценке по дисциплине экзаменационные баллы учитываются в соответствии с правилами балльно-рейтинговой системы, установленными в НГТУ.

4. Вопросы к экзамену по дисциплине «Анализ данных»

1. Линейное пространство. Матричные операции
2. Линейные преобразования. Собственные значения.
3. Системы линейных уравнений.
4. Сингулярное разложение матриц.
5. Случайные события. Условная вероятность. Формула Байеса. Независимые испытания.
6. Дискретные случайные величины. Закон распределения вероятностей. Биномиальный закон распределения. Распределение Пуассона
7. Непрерывные случайные величины. Функция распределения и функция плотности. Нормальное распределение. Центральная предельная теорема
8. Проверка статистических гипотез. Р-значения. Доверительные интервалы
9. Взаимосвязь величин. Показатели корреляции. Корреляционный анализ. Проверка на нормальность
10. Принцип работы основных фреймворков Python.
11. Алгоритм линейной регрессии. Метод наименьших квадратов. Градиентный спуск
12. Масштабирование признаков. L₁- и L₂-регуляризация.
13. Логистическая регрессия. Log Loss.
14. Функция потерь. Информационная энтропия.

15. Алгоритм построения дерева решений.
16. Классификация методом k ближайших соседей и методом опорных векторов.
17. Стохастический градиентный спуск.
18. Способы снижения размерности данных.