



Министерство науки  
и высшего образования РФ  
ФГБОУ ВО «НИУ «МЭИ»  
Институт дистанционного  
и дополнительного образования



**ОЦЕНОЧНЫЕ И МЕТОДИЧЕСКИЕ МАТЕРИАЛЫ ДЛЯ КОНТРОЛЯ  
ОСВОЕНИЯ РЕЗУЛЬТАТОВ ОБУЧЕНИЯ  
ДОПОЛНИТЕЛЬНОЙ ОБРАЗОВАТЕЛЬНОЙ ПРОГРАММЫ  
*повышения квалификации*  
*«Основы анализа текстовых данных»*,**

**Текущий контроль**

Текущий контроль проводится в соответствии с характеристиками контрольных заданий и представлен в табл. 1.

Таблица 1

Характеристика заданий текущего контроля

Наименование дисциплины (модуля)	Форма контроля/ наименование контрольной точки	Пример задания	Критерии оценки
<b>Основы анализа данных</b>			
Постановка задачи машинного обучения	Тестирование	<b>Что такое "машинное обучение"</b>	<i>Оценка: зачтено</i> <i>Описание характеристики выполнения знания: Тест считается выполненным на оценку «Зачтено» если даны правильные ответы не менее чем на 75% вопросов теста.</i>  <i>Оценка: не зачтено</i> <i>Описание характеристики выполнения знания: Тест считается выполненным на оценку «Не зачтено» если даны правильные ответы менее чем на 75% вопросов теста.</i>
<b>Анализ текстовых данных (Text Mining)</b>			
Особенности задачи обработки текстовых данных	Тестирование	<b>Каковы особенности задач Text Mining?</b>	<i>Оценка: зачтено</i> <i>Описание характеристики выполнения знания: Тест считается выполненным на оценку «Зачтено» если даны правильные ответы не менее чем на 75% вопросов теста.</i>  <i>Оценка: не зачтено</i> <i>Описание характеристики</i>

			<i>выполнения знания:</i> Тест считается выполненным на оценку «Не зачтено» если даны правильные ответы менее чем на 75% вопросов теста.
<b>Задача классификации текстовых документов</b>			
Методы классификации данных	Тестирование	<b>Приведены метки классов ближайших соседей к объекту X: В, А, С, С, В, С, А, В, С, В, А, С К какому классу будет отнесен объект X при параметре k = 7?</b>	<i>Оценка:</i> зачтено <i>Описание характеристики выполнения знания:</i> Тест считается выполненным на оценку «Зачтено» если даны правильные ответы не менее чем на 75% вопросов теста.  <i>Оценка:</i> не зачтено <i>Описание характеристики выполнения знания:</i> Тест считается выполненным на оценку «Не зачтено» если даны правильные ответы менее чем на 75% вопросов теста.
<b>Другие задачи, решаемые в рамках Text Mining</b>			
Кластеризация данных, выявление дубликатов текстовых документов	Тестирование	<b>Позволяет ли алгоритм автоматически определять количество кластеров?</b>	<i>Оценка:</i> зачтено <i>Описание характеристики выполнения знания:</i> Тест считается выполненным на оценку «Зачтено» если даны правильные ответы не менее чем на 75% вопросов теста.  <i>Оценка:</i> не зачтено <i>Описание характеристики выполнения знания:</i> Тест считается выполненным на оценку «Не зачтено» если даны правильные ответы менее чем на 75% вопросов теста.

### Промежуточная аттестация

Промежуточная аттестация по программе проводится в форме зачета, экзамена или отчета о стажировке в соответствии с учебным планом. Характеристика заданий представлена в табл. 2.

Таблица 2

Характеристика заданий промежуточной аттестации

Наименование дисциплины (модуля)	Пример задания	Критерии оценки
Основы анализа данных	<b>Даны два объекта, описывающиеся тремя признаками: X1 = (0, 4, 7);</b>	<i>Оценка:</i> зачтено <i>Описание характеристики выполнения знания:</i> Тест считается выполненным на

	<p><math>X_2 = (3, 1, 9)</math>.</p> <p>Чему равно расстояние городских кварталов <math>d(X_1, X_2)</math> между ними?</p>	<p>оценку «Зачтено» если даны правильные ответы не менее чем на 75% вопросов теста.</p> <p><i>Оценка:</i> не зачтено <i>Описание характеристики выполнения знания:</i> Тест считается выполненным на оценку «Не зачтено» если даны правильные ответы менее чем на 75% вопросов теста.</p>
Анализ текстовых данных (Text Mining)	<p>Дана информация о выборке:</p> <p>частота слова <math>X = 100</math> Количество документов, содержащих слово <math>X = 10</math> Количество документов, не содержащих слово <math>X = 90</math></p> <p>Посчитать tf-idf вес слова <math>X</math>. Ответ записать по правилам математического округления с точностью до 2 знаков после запятой. Использовать десятичный логарифм</p>	<p><i>Оценка:</i> зачтено <i>Описание характеристики выполнения знания:</i> Тест считается выполненным на оценку «Зачтено» если даны правильные ответы не менее чем на 75% вопросов теста.</p> <p><i>Оценка:</i> не зачтено <i>Описание характеристики выполнения знания:</i> Тест считается выполненным на оценку «Не зачтено» если даны правильные ответы менее чем на 75% вопросов теста.</p>
Задача классификации текстовых документов	<p>Даны значения признаков объектов обучающей выборки и их метки классов:</p> <p>(3, 2, 5) - 1 (4, 1, 4) - 1 (3, 3, 3) - 1 (5, 4, 5) - 0 (6, 5, 3) - 0 (4, 5, 4) - 0</p> <p>К какому классу будет отнесен объект (5, 1, 2) по методу k-ближайших соседей при <math>k = 3</math>. Использовать метрику городских кварталов</p>	<p><i>Оценка:</i> зачтено <i>Описание характеристики выполнения знания:</i> Тест считается выполненным на оценку «Зачтено» если даны правильные ответы не менее чем на 75% вопросов теста.</p> <p><i>Оценка:</i> не зачтено <i>Описание характеристики выполнения знания:</i> Тест считается выполненным на оценку «Не зачтено» если даны правильные ответы менее чем на 75% вопросов теста.</p>
Другие задачи, решаемые в рамках Text Mining	<p>Дана матрица расстояний между объектами.</p> <p>0 5 8 6 5 0 4 8 8 4 0 3 6 8 3 0</p> <p>Какие объекты будут объединены на 1 шаге при использовании принципа дальнего соседа?</p>	<p><i>Оценка:</i> зачтено <i>Описание характеристики выполнения знания:</i> Тест считается выполненным на оценку «Зачтено» если даны правильные ответы не менее чем на 75% вопросов теста.</p> <p><i>Оценка:</i> не зачтено <i>Описание характеристики выполнения знания:</i> Тест считается выполненным на оценку «Не зачтено» если даны</p>

		правильные ответы менее чем на 75% вопросов теста.
--	--	--

### Итоговая аттестация

Итоговая аттестация по программе проводится в форме *итогового зачета*. Характеристика заданий представлена в табл. 3.

Таблица 3

Характеристика заданий итоговой аттестации

Вид контроля	Краткая характеристика задания	Критерии оценки
Итоговая аттестация	<p><b>Даны два объекта: (8,0; 6,5; 4,5) (3,0; 4,5; 7,5)</b></p> <p><b>Чему равен квадрат евклидова расстояния между объектами? Ответ записать по правилам математического округления с точностью до 2 знаков после запятой</b></p>	<p><i>Оценка:</i> зачтено <i>Описание характеристики выполнения знания:</i> Тест считается выполненным на оценку «Зачтено» если даны правильные ответы не менее чем на 75% вопросов теста.</p> <p><i>Оценка:</i> не зачтено <i>Описание характеристики выполнения знания:</i> Тест считается выполненным на оценку «Не зачтено» если даны правильные ответы менее чем на 75% вопросов теста.</p>

### Независимая оценка качества обучения

Независимая оценка качества обучения предполагает внутренний аудит программ ДПО и анкетирование слушателей и/или работодателей по вопросам удовлетворенности процессом и результатами обучения.

### Учебно-методическое и информационное обеспечение

а) литература НТБ МЭИ:

1. Маннинг, К. Д. Введение в информационный поиск : пер. с англ. / К. Д. Маннинг, П. Рагхаван, Х. Шютце . – М. : Вильямс, 2011 . – 528 с. - ISBN 978-5-8459-1623-5 .;

2. Методы и модели анализа данных: OLAP и Data Mining : учебное пособие по специальности 071900 "Информационные системы и технологии" направления 654700 "Информационные системы" / А. А. Барсегян, и др. – СПб. : БХВ-Петербург, 2004 . – 336 с. + CD-ROM . - ISBN 5-941575-22-X .;

3. Мохов, А. С. Анализ и обработка текстовых данных : практикум по курсу "Интеллектуальные информационные системы" по направлению подготовки магистров 27.04.04 "Управление в технических системах" / А. С. Мохов, В. О. Толчеев, А. А. Бородкин, Нац. исслед. ун-т "МЭИ" (НИУ"МЭИ") . – Москва : Изд-во МЭИ, 2020 . – 52 с. - ISBN 978-5-7046-2284-0 .

<http://elibr.mpei.ru/elibr/view.php?id=11207>;

4. Толчеев, В. О. Современные методы обработки и анализа текстовой информации : учебное пособие по курсу "Интеллектуальные информационные системы" по специальности "Управление и информатика в технических системах" / В.

О. Толчеев, Моск. энерг. ин-т (МЭИ ТУ) . – М. : Изд-во МЭИ, 2006 . – 76 с. - ISBN 5-7046-1285-7 ..

б) литература ЭБС и БД:

1. Алексеев Д. С.,Щекочихин О. В.- "Технологии интеллектуального анализа данных", Издательство: "Лань", Санкт-Петербург, 2022 - (176 с.)

<https://e.lanbook.com/book/187559>;

2. Макшанов А. В.,Журавлев А. Е.- "Технологии интеллектуального анализа данных", (2-е изд., стер.), Издательство: "Лань", Санкт-Петербург, 2019 - (212 с.)

<https://e.lanbook.com/book/120063>.

в) используемые ЭБС:

*Не предусмотрено*

**Руководитель  
ОДПО, ЦПП УВО**  
(должность)

	<b>Подписано электронной подписью ФГБОУ ВО «НИУ «МЭИ»</b>	
	Сведения о владельце ЦЭП МЭИ	
	Владелец	Максимова А.А.
	Идентификатор	R6a033f13-VorozhtsovaAA-daecd82

**А.А.  
Максимова**  
(расшифровка  
подписи)

**Начальник ОДПО**  
(должность)

	<b>Подписано электронной подписью ФГБОУ ВО «НИУ «МЭИ»</b>	
	Сведения о владельце ЦЭП МЭИ	
	Владелец	Крохин А.Г.
	Идентификатор	R6d4610d5-KrokhinAG-aa301f84

**А.Г. Крохин**  
(расшифровка  
подписи)