

МИНОБРНАУКИ РОССИЙСКОЙ ФЕДЕРАЦИИ  
НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ  
ТОМСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ

Институт прикладной математики и компьютерных наук

УТВЕРЖДАЮ  
Директор института прикладной  
математики и компьютерных наук  
А.В. Замятин  
« 16 » \_\_\_\_\_ 2022 г.



Оценочные материалы текущего контроля и промежуточной аттестации по дисциплине  
(Оценочные средства по дисциплине)

**Обработка естественного языка**

по направлению подготовки

**02.03.02 Фундаментальная информатика и информационные технологии**

Направленность (профиль) подготовки:

**Искусственный интеллект и разработка программных продуктов**

ОС составил(и):

канд. техн. наук, доцент,  
доцент кафедры теоретических основ информатики

М.С. Пожидаев

Рецензент:

д-р. техн. наук., профессор,  
профессор кафедры теоретических основ информатики

Ю.Л. Костюк

Оценочные средства одобрены на заседании учебно-методической комиссии  
института прикладной математики и компьютерных наук (УМК ИПМКН)

Протокол от 16.05, 2022 г. № 4

Председатель УМК ИПМКН,  
д-р техн. наук, профессор

С.П. Сущенко

**Оценочные средства (ОС)** являются элементом системы оценивания сформированности компетенций у обучающихся в целом или на определенном этапе их формирования.

ОС разрабатываются в соответствии с рабочей программой (РП) дисциплины.

### 1. Компетенции и результаты обучения, формируемые в результате освоения дисциплины

Компетенция	Индикатор компетенции	Код и наименование результатов обучения (планируемые результаты обучения, характеризующие этапы формирования компетенций)	Критерии оценивания результатов обучения	
			Зачтено	Незачтено
ОПК-2. Способен применять компьютерные/суперкомпьютерные методы, современное программное обеспечение, в том числе отечественного происхождения, для решения задач профессиональной деятельности	ИОПК-2.1. Обладает необходимыми знаниями основных концепций современных вычислительных систем деятельности.	ОР-2.1.1. Знать процедуры выявления, формирования и согласования требований к результатам аналитических работ с применением технологий обработки естественного языка;	Сформированные систематические знания	Отсутствие знаний
	ИОПК-2.2. Использует методы высокопроизводительных вычислительных технологий, современного программного обеспечения, в том числе отечественного происхождения	ОР-2.2.1. Знать принципы планирования и организации аналитических работ с использованием технологий обработки естественного языка;	Сформированные систематические знания	Отсутствие знаний
	ИОПК-2.3. Использует инструментальные средства высокопроизводительных вычислений в научной и практической деятельности	ОР-2.3.1. Уметь подготавливать данные для проведения аналитических работ по исследованию текстовых массивов данных методами обработки естественного языка;	Сформированные систематические умения	Отсутствие умений

ПК-1. Способен осуществлять программирование, тестирование и опытную эксплуатацию ИС с использованием технологических и функциональных стандартов, современных моделей и методов оценки качества и надежности программных средств	ИПК-1.3. Кодирует на языках программирования и проводит модульное тестирование ИС	ОП-1.3.1. Уметь проводить аналитическое исследование и разрабатывать приложения для решения задач обработки естественного языка соответствии с требованиями заказчика	Сформированные систематические умения	Отсутствие умений
---	---	---	---------------------------------------	-------------------

## 2. Этапы формирования компетенций и виды оценочных средств

№	Этапы формирования компетенций (разделы дисциплины)	Код и наименование результатов обучения	Вид оценочного средства (тесты, задания, кейсы, вопросы и др.)
1.	Раздел 1. Введение, история развития дисциплины, решаемые задачи, подходы, методы и инструменты	ОР-2.1.1, ОР-2.2.1, ОР-2.3.1, ОР-1.3.1	Выполнение практической работы №1
2.	Раздел 2. Предварительная обработка текстовых данных	ОР-2.1.1, ОР-2.2.1, ОР-2.3.1, ОР-1.3.1	Выполнение практической работы №2
3.	Раздел 3. Вероятностные алгоритмы	ОР-2.1.1, ОР-2.2.1, ОР-2.3.1, ОР-1.3.1	Выполнение практической работы №3
4.	Раздел 4. Формальные грамматики	ОР-2.1.1, ОР-2.2.1, ОР-2.3.1, ОР-1.3.1	Выполнение практической работы №4
5.	Раздел 5. Векторное представление слов	ОР-2.1.1, ОР-2.2.1, ОР-2.3.1, ОР-1.3.1	Выполнение практической работы №5
6.	Раздел 6. Модель Seq2seq	ОР-2.1.1, ОР-2.2.1, ОР-2.3.1, ОР-1.3.1	Выполнение практической работы №6
7.	Раздел 7. Self-attention и Трансформер	ОР-2.1.1, ОР-2.2.1, ОР-2.3.1, ОР-1.3.1	Выполнение практической работы №7
8.	Раздел 8. BERT и GPT-3	ОР-2.1.1, ОР-2.2.1, ОР-2.3.1, ОР-1.3.1	Выполнение практической работы №8

## 3. Типовые контрольные задания или иные материалы, необходимые для оценки образовательных результатов обучения

3.1. Типовые задания для проведения текущего контроля успеваемости по дисциплине

Перечень практических работ:

Практическая работа № 1. Парсинг сайтов / использование `api` для получения текстовых данных.

Практическая работа № 2. Реализация стеммера Портера.

Практическая работа № 3. Использование библиотек для морфологического анализа, решение задачи частеречной разметки.

Практическая работа № 4. Векторное представление текста, `word2vec`, модели `skip-gram` и `CBOW`.

Практическая работа № 5. Тематическое моделирование с использованием библиотеки `gensim`.

Практическая работа № 6. Анализ тональности текстовых данных. Развертывание обученной модели в вебе.

Практическая работа № 7. Построение языковой модели, порождение текста.

Практическая работа № 8. Генерация подписи к изображению.

#### **4. Методические материалы, определяющие процедуры оценивания образовательных результатов обучения**

4.1. Методические материалы для оценки текущего контроля успеваемости по дисциплине.

Текущий контроль успеваемости проводится во время сдачи практических работ. Каждая работа оценивается по следующим параметрам, согласно таблице раздела 1:

- полнота реализации практической работы,
- ответы на вопросы по практической работе,
- ответы на вопросы по теории из соответствующего раздела курса,
- умение исправлять ошибки и оперативно вносить изменения в практическую работу.

4.2. Методические материалы для проведения промежуточной аттестации по дисциплине.

«Зачтено» – студент выполнил все практические работы, ответил на все вопросы по практической работе;

«Не зачтено» – студент не сдал какие-либо практические работы, не ответил на вопросы по практической работе.

Во время зачета студент может повысить свою оценку, сдав заново соответствующую практическую работу.