Аннотация к рабочей программе дисциплины

Обработка естественного языка

**Направление подготовки:** 09.04.01 Информатика и вычислительная техника

**Направленность (профиль*):*** Инженерия искусственного интеллекта

**Квалификация выпускника:** магистр

**Целью освоения дисциплины является:** изучение математического аппарата, используемого в основе методов обработки естественных языков, программных инструментов для обработки естественных языков и приобретение практических навыков в профессиональной деятельности.

**Объем дисциплины:** 3 зачетные единицы – 108 часов

**Семестр:** 3

**Краткое содержание основных разделов дисциплины:**

|  |  |  |
| --- | --- | --- |
| № п/п  раздела | Основные разделы  дисциплины | Краткое содержание разделов дисциплины |
| 1 | **Введение в обработку естественного языка** | Введение в обработку естественного языка;  Определение авторства текста:  N-грамм модели |
| 2 | **Классификация и кластеризация текстов** | Классификация и кластеризация текстов;  Word2Vec, кластеризация текста |
| 3 | **Информационный поиск** | Информационный поиск;  Поиск коллокаций |
| 4 | **Введение в машинный перевод** | Введение в машинный перевод;  GPT2 для генерации текста |
| 5 | **Введение в извлечение информации** | Введение в извлечение информации  BERT для анализа настроения |
| 6 | **Методы машинного обучения в задаче извлечения информации** | Методы машинного обучения в задаче извлечения информации;  Выделение именованных сущностей |
| 7 | **Извлечение мнений** | Извлечение мнений  Грамматический разбор в тексте |

**Форма промежуточной аттестации:** зачет