

Федеральное государственное образовательное бюджетное учреждение
высшего профессионального образования
«Финансовый университет при Правительстве Российской Федерации»

Департамент анализа данных, принятия решения и финансовых
технологий

Курсовая работа

по дисциплине "Технологии анализа данных и машинное обучение" на
тему:

"Классификация текстов методами машинного обучения"

Вид исследуемых данных: Корпус новостей с сайта Lenta.Ru

Выполнила:
студентка группы ПМ17-1
Баданина Н. Д.
Научный руководитель:
доктор техн. наук, доцент
Судаков Владимир Анатольевич

Москва 2020

Содержание

Введение	3
0.1 Сложности при обработке текстов на естественном языке	4
0.2 Перспективы развития технологии	5
1 Теоретическая часть	6
1.1 Задача классификации текстов	7
1.2 Модели классификации	8
2 Практическая часть	9
2.1 Работа с данными	9
Заключение	10
Список использованных источников	11
Приложение	12

Введение

0.1 Сложности при обработке текстов на естественном языке

0.2 Перспективы развития технологии

1 Теоретическая часть

1.1 Задача классификации текстов

1.2 Модели классификации

2 Практическая часть

2.1 Работа с данными

Заключение

Список использованных источников

- [1] Applied Text Analysis with Python by Benjamin Bengfort, Rebecca Bilbro, and Tony Ojeda (O'Reilly). 9781491963043

Приложение

Исходный код можно найти по [ссылке](#).