

ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ АВТОНОМНОЕ ОБРАЗОВАТЕЛЬНОЕ
УЧРЕЖДЕНИЕ ВЫСШЕГО ОБРАЗОВАНИЯ
«НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ УНИВЕРСИТЕТ
«ВЫСШАЯ ШКОЛА ЭКОНОМИКИ»

МОСКОВСКИЙ ИНСТИТУТ ЭЛЕКТРОНИКИ И МАТЕМАТИКИ
им. А.Н. ТИХОНОВА

«УТВЕРЖДАЮ»

Академический руководитель образовательной
программы
«Информатика и вычислительная техника»

_____ Ю.И. Гудков

« ____ » декабря 2017 г.

ЗАДАНИЕ
на выпускную квалификационную работу бакалавра

студенту группы БИВ-144 Вдовкин Василий Алексеевич

1. Тема работы
Методы сокращения текста и извлечения ключевой информации из русскоязычных новостных статей
2. Требования к работе
Получение моделей, способных извлекать факты из текста и производить его сокращение с относительно хорошей точностью.
Практическое применение данных моделей на реальных данных в виде web-сервиса, анализирующего русскоязычные новости.
3. Содержание работы
Обоснование значимости работы, описание предметной области.
Исследование алгоритмов нормализации текста, векторизации, кластеризации (TF-IDF, K-means и т.д.), методов извлечения информации и сокращения текста.
Сбор и подготовка датасета русскоязычных новостных статей.
Выбор метрик, определяющих точность моделей.
Использование рассмотренных алгоритмов на собранных данных для получения моделей извлечения фактов и сокращения текста.
Разработка web-сервиса, демонстрирующего результаты работы полученных моделей, в режиме реального времени.
Выводы о проделанной работе.
4. Сроки выполнения этапов работы

Проект ВКР представляется студентом в срок до	«9» февраля 2018 г.
Первый вариант ВКР представляется студентом в срок до	«16» апреля 2018 г.
Итоговый вариант ВКР представляется студентом руководителю до загрузки работы в систему «Антиплагиат» в срок до	«15» мая 2018 г.

Задание выдано

«15» декабря 2017 г.


подпись руководителя

Ф.В. Строк

Задание принято к
исполнению

«15» декабря 2017 г.


подпись студента

В.А. Вдовкин

ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ АВТОНОМНОЕ ОБРАЗОВАТЕЛЬНОЕ
УЧРЕЖДЕНИЕ ВЫСШЕГО ОБРАЗОВАНИЯ
«НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ УНИВЕРСИТЕТ
«ВЫСШАЯ ШКОЛА ЭКОНОМИКИ»

МОСКОВСКИЙ ИНСТИТУТ ЭЛЕКТРОНИКИ И МАТЕМАТИКИ
им. А.Н. ТИХОНОВА

**График сдачи этапов
выпускной квалификационной работы бакалавра**

студента группы БИВ-144 Вдовкин Василий Алексеевич

Тема работы

Методы сокращения текста и извлечения ключевой информации из
русскоязычных новостных статей

Дата представления проекта
ВКР

«9» февраля 2018 г.


подпись руководителя

Ф.В. Строк

Дата представления первого
варианта ВКР

«16» апреля 2018 г.


подпись руководителя

Ф.В. Строк

Дата представления итогового
варианта ВКР

«30» апреля 2018 г.


подпись руководителя

Ф.В. Строк