

БЕЛОРУССКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ

Факультет прикладной математики и информатики

Кафедра дискретной математики и алгоритмики

Аннотация к магистерской диссертации

**«Машинный перевод текстов медицинского и юридического тематических
доменов»**

Прокопенко Тимофей Александрович

Научный руководитель – кандидат технических наук Гецевич Ю. С.

Минск, 2021

Реферат

Магистерская диссертация, 35 страниц, 16 рисунков, 4 таблицы, 15 источников.

МАШИННЫЙ ПЕРЕВОД, НЕЙРОННЫЕ СЕТИ, КОРПУСНАЯ ЛИНГВИСТИКА, ПЕРЕВОД НА БЕЛОРУССКИЙ ЯЗЫК, ЮРИДИЧЕСКИЙ ДОМЕН, МЕДИЦИНСКИЙ ДОМЕН

Объект исследования – задачи перевода текстов медицинской и юридической тематики с русского языка на белорусский и связанные с рассматриваемой задачей создание параллельных корпусов, подготовительные этапы обработки текста и нейросетевые модели для перевода.

Цель работы – изучение современных методов построения автоматических систем перевода, реализация собственной модели, оценка полученных результатов.

Методы исследования – представляют собой анализ проблематики и существующих подходов, эксперимент, тестирование и сравнение.

Результаты –

1. Проведен обзор теоретического материала, связанного с текущим развитием машинного перевода;

2. Созданы два новых параллельных корпуса медицинского юридического тематических доменов;

3. Разработан сервис с двумя различными интерфейсами, позволяющий переводить тексты различной тематики, а также проводить последующую адаптацию при появлении новых данных.

Область применения разработанной модели – прикладные задачи, связанные с переводом юридических и медицинских текстов с русского языка на белорусский.

Abstract

Master thesis, 35 pages, 16 figures, 4 tables, 15 sources.

MACHINE TRANSLATION, NEURAL NETWORKS, CORPUS LINGUISTICS, BELARUSIAN TRANSLATION, LEGAL DOMAIN, MEDICAL DOMAIN

The object of research of the master's thesis is the task of translating medical and legal texts from Russian into Belarusian and the creation of parallel corpora related to the task under consideration, preparatory stages of text processing and neural network models for translation.

The aim of the work is to study modern methods of building automatic translation systems, to implement the own model, to evaluate the obtained results.

Research methods are analysis of problems and existing approaches, experiment, testing and comparison.

Key results of the study:

1. A review of the theoretical material related to the current development of the field of machine translation was conducted;
2. Two new parallel corpuses of medical and legal thematic domains were created;
3. A service with two different interfaces developed, which allows you to translate texts of various topics, as well as carry out subsequent adaptation when new data appears.

Field of application of the developed model is applied problems related to the translation of legal and medical texts from Russian into Belarusian.