

МИНИСТЕРСТВО ОБРАЗОВАНИЯ РЕСПУБЛИКИ БЕЛАРУСЬ
БЕЛОРУССКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ
ФАКУЛЬТЕТ ПРИКЛАДНОЙ МАТЕМАТИКИ И ИНФОРМАТИКИ
Кафедра информационных систем управления

Аннотация к дипломной работе

ПОЛЬЗОВАТЕЛЬСКИ ОРИЕНТИРОВАННОЕ
АВТОМАТИЧЕСКОЕ РЕФЕРИРОВАНИЕ ТЕКСТОВЫХ ДОКУМЕНТОВ

Головач Андрей Юрьевич

Научный руководитель:
доктор технических наук,
профессор Совпель И.В.

Минск, 2017

Реферат

Дипломная работа, 42 страницы, 15 иллюстраций, 11 источников.

Ключевые слова: АВТОМАТИЧЕСКОЕ РЕФЕРИРОВАНИЕ, АВТОМАТИЧЕСКАЯ АННОТАЦИЯ, ОБРАБОТКА ЕСТЕСТВЕННОГО ЯЗЫКА, TF-IDF, СТЕММИНГ, НЕЙРОННЫЕ СЕТИ.

Объект исследования – алгоритмы и пользовательски ориентированные системы автоматического реферирования и аннотирования текстовых документов.

Цель работы – исследовать и улучшить известные алгоритмы автоматического реферирования и аннотирования текстовых документов, разработать пользовательски ориентированную систему автоматического реферирования.

Методы исследования – обработка естественного языка, статистический и семантический анализ текста, алгоритмы поиска информации и анализа данных.

В результате разработана пользовательски ориентированная система автоматического реферирования и аннотирования текстовых документов.

Областью применения являются системы информационного поиска, классификации и анализа текстовых документов.

Рэферат

Дыпломная праца, 42 старонкі, 15 малюнкаў, 11 крыніц.

Ключевые слова: АЎТАМАТЫЧНАЕ РЭФЕРЫРАВАННЕ, АЎТАМАТЫЧНАЯ АНАТАЦЫЯ, АПРАЦОЎКА НАТУРАЛЬнай МОВЫ, TF-IDF, СТЭМІНГ НЕЙРОНАВЫЯ СЕТКІ.

Аб'ект даследвання – алгарытмы і карыстальніка арыентаваныя сістэмы аўтаматычнага рэферыравання і анатыравання тэкставых дакументаў.

Мэта працы – даследаваць і палепшыць вядомыя алгарытмы аўтаматычнага рэферыравання і анатыравання тэкставых дакументаў, распрацаваць карыстальніка-арыентаваную сістэму аўтаматычнага рэферыравання.

Метады даследвання – апрацоўка натуральнай мовы, статыстычны і семантычны аналіз тэкста, алгарытмы пошуку іфармацыі і аналізу дадзеных.

У выніку распрацавана разработана карыстальніка арыентаваная сістэма аўтаматычнага рэферыравання і анатыравання тэкставых дакументаў.

Вобласцю ўжывання з'яўляюцца сістэмы інфармацыйнага пошуку, класіфікацыі і аналізу тэкставых дакументаў.

Abstract

Diploma thesis, 42 pages, 15 figures, 11 Sources.

Keywords: AUTOMATIC SUMMARIZATION, ABSTRACTING, EXTRACTING, NATURAL LANGUAGE PROCESSING, TF-IDF, NEURAL NETS.

Object of study – algorithms and user friendly systems for automatic summarization and abstracting of text documents.

Purpose – research and improve known algorithms of automatic summarization and abstracting of text documents, develop user friendly system for automatic summarization.

Research methods – natural language processing, statistic and semantic text analysis, information retrieval and data mining algorithms.

As a result user friendly automatic summarization and text document abstracting system is developed.

Area of applications is information retrieval systems, text documents classification and analysis.