

БЕЛОРУССКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ

Факультет прикладной математики и информатики

Кафедра дискретной математики и алгоритмики

Аннотация к магистерской диссертации

**АНАЛИЗ СООБЩЕНИЙ ПОЛЬЗОВАТЕЛЕЙ СОЦИАЛЬНЫХ СЕТЕЙ НА ПРЕДМЕТ
ПРИНАДЛЕЖНОСТИ К СПАМОВЫМ**

Скокклеенко Даниил Артурович

Научный руководитель — кандидат. физико-математических наук,
доцент Е. П. Соболевская

2017

РЕФЕРАТ

Магистерская диссертация, 41 с., 18 рис., 11 табл., 27 источников.

СОЦИАЛЬНЫЕ СЕТИ, СПАМ, ЗАДАЧА КЛАССИФИКАЦИИ, МАШИННОЕ ОБУЧЕНИЕ

Объект исследования – применение алгоритмов машинного обучения для классификации спама в социальных сетях.

Цель работы – исследование методов определения спама в социальных сетях, сравнение подходов для решения задачи классификации спама, построение модели распознавания спама в социальной сети Twitter на основе алгоритмов машинного обучения.

Методы исследования – наивный байесовский классификатор, метод k ближайших соседей, метод опорных векторов (SVM), решающие деревья, случайные леса.

Результатом является предложенный подход для построения классификатора социального спама, не требующий наличия историчных признаков пользователя социальной сети.

Область применения — системы спамообороны.

ABSTRACT

Master thesis, 41 p., 11 tab., 18 fig., 27 ref.

SOCIAL NETWORKS, SPAM DETECTION, CLASSIFICATION PROBLEM,
MACHINE LEARNING

Object of research — Machine learning algorithms application in social spam classification problem.

Purpose — studying methods of social spam detection, comparing approaches of solving spam classification problem, building a model for Twitter spam detection based on machine learning algorithms.

Research methods — naive Bayes classifier, k -nearest neighbors algorithm, support vector machines (SVM), Decision trees, Random forest.

The results of work is the social spam classifier that does not require historical features of a social network account.

The results can be applied in antispam systems.