

БЕЛОРУССКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ
Факультет радиофизики и компьютерных технологий
Кафедра интеллектуальных систем

Аннотация к дипломной работе

**«Методы обработки пользовательских данных для
прогнозирования бизнес-процессов»**

Дрозд Павел Сергеевич

Научный руководитель: старший преподаватель Адуцкевич Иван
Анатольевич

2018

РЕФЕРАТ

Дипломная работа: 74 страницы, 37 рисунков, 12 использованных источников, 3 приложения.

СЕГМЕНТАЦИЯ РЫНКА, КЛАСТЕРНЫЙ АНАЛИЗ, НЕЙРОННЫЕ СЕТИ, ЛОКАЛЬНО-ЧУВСТВИТЕЛЬНОЕ ХЕШИРОВАНИЕ, BIG DATA.

Объект исследования – транзакционные наборы данных о покупках, совершённых клиентами.

Цель работы – создание системы для решения задачи сегментации рынка с помощью разработанных высокоэффективных методов кластерного анализа

В ходе работы были предложены два эффективных метода кластеризации транзакционных данных. Первый метод основан на самоорганизующихся картах Кохонена, второй – на алгоритме локально-чувствительного хеширования.

Для анализа точности и устойчивости разработанных методов была реализована система сегментации рынка, поддерживающая технологии BigData, а также одновременное использование нескольких алгоритмов машинного обучения. Система была протестирована на реальных данных, был найден потенциальный заказчик.

РЭФЕРАТ

Дыпломная праца: 74 старонкі, 37 малюнкаў, 12 выкарыстаных крыніц, 3 дадаткі.

СЕГМЕНТАЦЫЯ РЫНКУ, КЛАСТАРНЫ АНАЛІЗ, НЕЙРОНАВЫЯ СЕТКИ, ЛАКАЛЬНАЕ АДЧУВАЛЬНАЕ ХЭШАВАННЕ, BIG DATA.

Аб'ект даследавання – транзакцыйных наборы дадзеных аб купляннях, учыненых кліентамі

Мэта працы – стварэнне сістэмы для вырашэння задачы сегментацыі рынку з дапамогай распрацаваных высокаэфектыўных метадаў кластэрнага аналізу

У ходзе працы былі прапанаваны два эфектыўных метаду кластарызацыі транзакцыйных дадзеных. Першы метадад заснаваны на картах Коханена, другі - на алгарытме лакальнага адчувальнага хэшавання.

Для аналізу дакладнасці і ўстойлівасці распрацаваных метадаў была рэалізавана сістэма сегментацыі рынку, якая падтрымлівае тэхналогіі BigData, а таксама адначасовае выкарыстанне некалькіх алгарытмаў машыннага навучання. Сістэма была пратэставаная на рэальных дадзеных, быў знойдзены патэнцыйны заказчык.

ABSTRACT

Thesis: 75 pages, 37 figures, 12 sources, 3 applications.

MARKET SEGMENTATION, CLUSTER ANALYSIS, NEURAL NETWORKS, LOCALITY-SENSITIVE HASHING, BIG DATA.

The object of research - customer transactional data.

Objective - create a system for solving customer segmentation problem using highly effective clustering algorithms.

In this study two effective methods for customer clustering transactional data were proposed. The first one is based on Kohonen's self-organizing maps. The second one – on the Locality-Sensitive hashing.

The customer segmentation system was implemented in order to analyse the accuracy and stability of the designed methods. The system supports Big Data technologies as well as multiple machine learning algorithms. The system was tested on real data, and a potential customer was found.