

МИНИСТЕРСТВО ОБРАЗОВАНИЯ РЕСПУБЛИКИ БЕЛАРУСЬ
БЕЛОРУССКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ
ФАКУЛЬТЕТ ПРИКЛАДНОЙ МАТЕМАТИКИ И ИНФОРМАТИКИ
Кафедра многопроцессорных систем и сетей

Аннотация к дипломной работе
«Прогнозирование событий и Data Mining»

Литвинов Владимир Дмитриевич

Научный руководитель – зав. НИЛ ЦИРК, Довнар С.Е.

Минск, 2016

Реферат

Дипломная работа, 49 с., 15 источников, 14 рисунков, 6 таблиц.

DATA MINING, DATA SCIENCE, МАШИННОЕ ОБУЧЕНИЕ, РЕГРЕССИОННЫЙ АНАЛИЗ, БУСТИГ, ПРОГНОЗИРОВАНИЕ, BIG DATA, JAVA, MICROSOFT AZURE, БАЙЕСОВСКАЯ ЛИНЕЙНАЯ РЕГРЕССИЯ, РЕГРЕССИОННАЯ НЕЙРОННАЯ СЕТЬ, РЕГРЕССИОННОЕ ДЕРЕВО РЕШЕНИЙ, ЛИНЕЙНАЯ РЕГРЕССИЯ, РЕГРЕССИОННЫЙ ЛЕС РЕШЕНИЙ, ТУРИЗМ.

Объект исследования: методы анализа и прогнозирования событий в контексте Data Science и машинного обучения, задача о прогнозировании цен на проживание в отелях.

Цель работы: сравнение терминов и подходов Data Mining и машинного обучения, решение задачи прогнозирования.

Методы работы: теоретическое исследование, анализ данных средствами Data Mining, реализация обучающихся моделей используя алгоритмы машинного обучения и сравнение полученных результатов.

В результате работы проведен анализ и сравнение Data Mining и машинного обучения, выделены методы позволяющие решить задачу прогнозирования цен на проживания, проведены анализ и подготовка исходных данных к машинному обучению, построены и обучены различные прогнозирующие модели, проведено сравнение эффективности моделей в контексте решаемой задачи, таким образом получено окончательное решение задачи.

Область применения: туризм, решение задач прогнозирования.

Реферат

Дыпломная праца, 49 с., 15 крыніц, 14 малюнкаў, 6 табліц.

DATA MINING, DATA SCIENCE, МАШЫННАЕ НАВУЧАННЕ, РЭГРЕССІЙНЫ АНАЛІЗ, БУСЦІНГ, ПРАГНАЗАВАННЕ, BIG DATA, JAVA, MICROSOFT AZURE, БАЙЕСОВСКАЯ ЛІНЕЙНАЯ РЭГРЭСІЯ, РЭГРЭСІЙНАЯ НЕЙРОННАЯ СЕТКА, РЭГРЭСІЙНАЯ ДРЭВА РАШЭННЯЎ, ЛІНЕЙНАЯ РЭГРЭСІЯ, РЭГРЭСІЙНЫ ЛЕС РАШЭННЯЎ, ТУРЫЗМ.

Аб'ект даследавання: метады аналізу і прагназавання падзей у кантэксце Data Science і машыннага навучання, задача аб прагназаванні коштаў на пражывання ў гатэлях.

Мэта працы: параўнанне тэрмінаў і падыходаў Data Mining і машыннага навучання, рашэнне задачы прагназавання.

Метады працы: тэарэтычнае даследаванне, аналіз дадзеных сродкамі Data Mining, рэалізацыя навучаюцца мадэляў выкарыстоўваючы алгарытмы машыннага навучання і параўнанне атрыманых вынікаў.

У выніку працы праведзены аналіз і параўнанне Data Mining і машыннага навучання, вылучаны метады дазваляюць вырашыць задачу прагнілі-зирования коштаў на пражывання, праведзены аналіз і падрыхтоўка зыходных дадзеных да машынай навучання, пабудаваны і навучаны розныя Прагноз-вання мадэлі, праведзена параўнанне эфектыўнасці мадэляў у кантэксце вырашаемай задачы, такім чынам атрымана канчатковае рашэнне задачы.

Вобласць ужывання: турызм, рашэнне задач прагназавання.

Abstract

Thesis, 49 p., 15 sources, 14 pictures, 6 tables.

DATA MINING, DATA SCIENCE, MACHINE LEARNING, REGRESSION ANALYSIS, BUSTIG, FORECASTING, BIG DATA, JAVA, MICROSOFT AZURE, BAYESIAN LINEAR REGRESSION, REGRESSION NEURAL NETWORK, REGRESSION DECISION TREES, LINEAR REGRESSION, REGRESSION FOREST, TOURISM.

Object of research: methods of event analysis and forecasting in the context of the Data Science and machine learning, the problem of forecasting the price of accommodation in hotels.

Objective: To compare the terms and Data Mining and Machine Learning approach, the solution of forecasting problems.

Methods: theoretical research, data analysis Data Mining tools, implementation of training models using algorithms of Machine Learning and the results are compared.

As a result of the analysis and comparison of Data Mining and Machine Learning, highlighted methods allow to solve the problem foresight prices for accommodation, conducted the analysis and preparation of input data for machine learning, built and train various predictive models, get a comparison of model performance in the context of the problem being solved, thus obtaining a final solution to the problem.

Scope: tourism, the decision of forecasting problems.