

МИНИСТЕРСТВО ОБРАЗОВАНИЯ РЕСПУБЛИКИ БЕЛАРУСЬ
БЕЛОРУССКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ
ФАКУЛЬТЕТ ПРИКЛАДНОЙ МАТЕМАТИКИ И ИНФОРМАТИКИ
Кафедра математического моделирования и анализа данных

Аннотация к дипломной работе

**Статистическая классификация дискретных многомерных
наблюдений и дискретных временных рядов**

Колячко Мария Николаевна

Научный руководитель: доктор физико-математических наук, академик
НАН Беларуси Юрий Семенович Харин

Минск, 2022

Аннотация

Колячко М.Н. Статистическая классификация дискретных многомерных наблюдений и дискретных временных рядов / Минск: БГУ, 2022.

В данной работе рассмотрено понятие байесовского решающего правила на примере основных дискретных многомерных распределений и временных рядов, построены подстановочные байесовские решающие правила, а также разработано программное обеспечение для классификации и оценивания параметров рассмотренных моделей.

Анатацыя

Калячка М.М. Статыстычная класіфікацыя дыскрэтных шматмерных назіранняў і дыскрэтных часовых радоў / Мінск: БДУ, 2022.

У дадзенай працы разгледжана паняцце байесаўскага вырашальнага правіла на прыкладзе асноўных шматмерных дыскрэтных размеркаванняў і дыскрэтных часавых радоў, пабудаваны падстановачныя байесаўскія вырашальныя правілы, а таксама распрацавана праграмае забяспечанне для класіфікацыі і ацэньвання параметраў разгледжаных мадэляў.

Annotation

Kolyachko M.N. Statistical classification of discrete multivariate observations and discrete time series / Minsk: BSU, 2022.

In this work, the concept of a Bayesian decision rule is considered using the example of basic discrete multivariate distributions and time series, substitutional Bayesian decision rules are constructed, and software is developed for classifying and estimating the parameters of the considered models.

РЕФЕРАТ

Дипломная работа, 77 страниц, 20 рисунков, 13 источников, 10 приложений.

СТАТИСТИЧЕСКАЯ КЛАССИФИКАЦИЯ ДИСКРЕТНЫХ МНОГОМЕРНЫХ НАБЛЮДЕНИЙ И ДИСКРЕТНЫХ ВРЕМЕННЫХ РЯДОВ

Ключевые слова: ДИСКРЕТНЫЕ МНОГОМЕРНЫЕ НАБЛЮДЕНИЯ, ДИСКРЕТНЫЕ ВРЕМЕННЫЕ РЯДЫ, КЛАССИФИКАЦИЯ, БАЙЕСОВСКОЕ РЕШАЮЩЕЕ ПРАВИЛО, СТАТИСТИЧЕСКОЕ ОЦЕНИВАНИЕ ПАРАМЕТРОВ.

Объект исследования — методы классификации, оценивания параметров.

Цель работы — разработка оптимальных решающих правил для статистической классификации дискретных многомерных наблюдений и дискретных временных рядов.

Результаты работы — построенные байесовские решающие правила и подстановочные байесовские решающие правила, разработанное программное обеспечение.

Область применения — медицина, экономика, финансы и т. п.

РЭФЕРАТ

Дыпломная праца, 77 старонак, 20 малюнкаў, 13 крыніц, 10 дадаткаў.

СТАТЫСТЫЧНАЯ КЛАСІФІКАЦЫЯ ДЫСКРЭТНЫХ ШМАТМЕРНЫХ НАЗІРАННЯЎ І ДЫСКРЭТНЫХ ЧАСАВЫХ РАДОЎ

Ключавыя словы: ДЫСКРЭТНЫЯ ШМАТМЕРНЫЯ НАЗІРАННІ, ДЫСКРЭТНЫЯ ЧАСАВЫЯ РАДЫ, КЛАСІФІКАЦЫЯ, БАЕЎСКАЕ ВЫРАШАЛЬНАЕ ПРАВІЛА, СТАТЫСТЫЧНАЕ АЦЭНЬВАННЕ ПАРАМЕТРАЎ.

Аб’ект даследавання — метады класіфікацыі, ацэньванні параметраў.

Мэты працы — распрацоўка аптымальнага вырашальнага правіла для статыстычнай класіфікацыі дыскрэтных шматмерных назіранняў і дыскрэтных часавых радоў.

Вынікі працы — пабудаваныя баесаўскія вырашальныя правілы і падстаноўныя баесаўскія вырашальныя правілы, распрацаванае праграмае забеспечанне.

Галіна выкарыстання — медыцына, эканоміка, фінансы і г.д.

ABSTRACT

Diploma thesis, 77 pages, 20 figures, 13 sources, 10 appendices.

STATISTICAL CLASSIFICATION OF DISCRETE MULTIVARIATE OBSERVATIONS AND DISCRETE TIME SERIES

Keywords: DISCRETE MULTIVARIATE OBSERVATIONS, DISCRETE TIME SERIES, CLASSIFICATION, BAYESIAN DECISION RULE, STATISTICAL PARAMETER ESTIMATION.

Object of study — methods of classification, estimation of parameters.

Objectives of study — development of an optimal decision rule for the statistical classification of discrete multivariate observations and time series.

Results — constructed Bayesian decision rules and substitutional Bayesian decision rules, software developed.

Field of application — medicine, economics, finance and etc.

ВВЕДЕНИЕ

Тема дипломной работы посвящена классификации дискретных многомерных наблюдений и дискретных временных рядов. Многомерные данные и временные ряды — это актуальные инструменты, применимые во множестве задач, от предсказания цен на акции, прогнозов погоды, планирования бизнеса, до распределения ресурсов.

Многомерные данные содержат данные о двух или более признаках для каждого изучаемого объекта. В дополнение к той информации, которую можно извлечь из одномерных наборов, многомерные данные можно использовать для получения информации о том, существует ли простая зависимость между этими признаками, насколько они взаимосвязаны.

Характерной чертой временного ряда является то, что наблюдения за выбранным объектом являются последовательными во времени. Временные ряды используются для аналитики и прогнозирования, когда важно определить, что будет происходить с показателями на протяжении заданного промежутка времени.

Многие прикладные задачи прогнозирования сводятся к задачам прогнозирования многомерных наблюдений и временных рядов. В таких задачах часто прогнозируемая величина оказывается дискретной. Математической моделью для решения подобных прикладных задач является модель статистической классификации.

Под задачей оптимальной статистической классификации понимается классификация с минимизацией вероятности ошибки. Методы решения этой задачи зависят от количества известной априорной информации о вероятностных характеристиках исследуемых классов. В случае, когда все характеристики априорно известны, решение задачи заключается в построении байесовского решающего правила. В случае априорно неизвестных вероятностных характеристик, как правило, следуют распространенному в математической статистике "подстановочному" принципу, состоящему из двух этапов: первый этап — это построение математической модели на основании имеющихся статистических данных и известной априорной информации, второй этап — это вычисление оптимальной классификации на основе модели, полученной на первом этапе.

Цель данной работы заключается в построении оптимальной классификации для моделей основных дискретных распределений и дискретных временных рядов, а также оценивании вероятности ошибки построенной классификации.