

БЕЛОРУССКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ

Факультет прикладной математики и информатики

Кафедра теории вероятностей и математической статистики

Аннотация к дипломной работе

**ИССЛЕДОВАНИЕ ЭФФЕКТИВНОСТИ
ПОСЛЕДОВАТЕЛЬНЫХ ТЕСТОВ С ДИСКРЕТИЗАЦИЕЙ
ПРОСТРАНСТВА СОСТОЯНИЙ**

Шатерник Артём Александрович

Научный руководитель — заведующий кафедрой теории вероятностей и математической статистики ФПМИ, доктор физико-математических наук, профессор А. Ю. Харин

Минск 2025

РЕФЕРАТ

Дипломная работа, 47 страниц, 16 рисунков, 6 таблиц, 7 источников.

Ключевые слова: ПРОВЕРКА ГИПОТЕЗ, ПОСЛЕДОВАТЕЛЬНЫЙ АНАЛИЗ, ЦЕПИ МАРКОВА, ДИСКРЕТИЗАЦИЯ ПОСЛЕДОВАТЕЛЬНОГО ТЕСТА, КРИТЕРИЙ ВАЛЬДА, МАШИННОЕ ОБУЧЕНИЕ, ПОДБОР ГИПЕРПАРАМЕТРОВ.

Объект исследования — последовательный критерий отношения вероятностей Вальда с дискретизацией пространства состояний.

Цель работы — изучить влияние дискретизации пространства состояний статистики последовательного теста на его характеристики, такие как среднее число итераций и вероятности ошибок, а также представить вариант аппроксимации теста с помощью цепей Маркова, и продемонстрировать применимости предложенных подходов на примере задачи подбора гиперпараметров моделей машинного обучения.

Методы исследования — исследование литературы, методы математической статистики, теория цепей Маркова, компьютерное моделирование последовательных тестов, методы машинного обучения.

Полученные результаты — построена модель последовательного критерия Вальда с помощью языка программирования Python, на её основе проведён анализ характеристик теста, рассмотрен процесс дискретизации пространства состояний статистики теста, приведён и смоделирован вариант аппроксимации теста с помощью цепи Маркова, а также предложен способ применения изученных методов к актуальной задаче подбора гиперпараметров для моделей машинного обучения, позволяющий ускорить этот процесс.

РЭФЕРАТ

Дыпломная праца, 47 старонак, 16 малюнкаў, 6 табліц, 7 крыніц.

Ключавыя словы: ТЭСТАВАННЕ ГІПОТЭЗ, ПАСЛЯДОЎНЫ АНАЛІЗ, ЛАНЦУГІ МАРКАВА ДЫСКРЭТЫЗАЦЫЯ ПАСЛЯДОЎНАГА ТЭСТУ, КРЫТЭРЫЙ ВАЛЬДА, МАШЫННАЕ НАВУЧАННЕ, ПАДБОР ГІПЕРПАРАМЕТРАЎ.

Аб’ект даследавання — паслядоўны крытэрыі адносінаў імавернасцяў Вальда з дыскрэтызацыяй прасторы станаў.

Мэта даследавання — вывучэнне ўплыву дыскретизацыі прасторы статыстыкі паслядоўнага тэсту на яго характарыстыкі, такія як сярэдняя колькасць ітэрацый і верагоднасці памылак, прапанаванне варыянту апраксімацыі тэсту з дапамогай ланцугоў Маркава, і дэманстрацыя прымянімасці прапанаваных падыходаў на прыкладзе задачы падбору гіперпараметраў для мадэляў машыннага навучання.

Метады даследавання — аналіз літаратуры, метады матэматычнай статыстыкі, тэорыя ланцугоў Маркава, камп’ютарнае мадэліраванне паслядоўных тэстаў, метады машыннага навучання.

Атрыманыя вынікі — пабудавана мадэль паслядоўнага крытэрыя Вальда на мове праграмавання Python, на аснове гэтай мадэлі праведзен аналіз характарыстык тэсту, разгледжан працэс дыскрэтызацыі прасторы станаў статыстыкі тэсту, праванаван і змадэляван варыянт апраксімацыі тэсту з дапамогай ланцугоў Маркава, таксама праванаван падыход да прымянення атрыманых вынікаў у актуальнай задачы падбору гіперпараметраў для мадэлей машыннага навучання, які дазваляе паскорыць гэты працэс.

ANNOTATION

Degree paper, 47 pages, 16 illustrations, 6 tables, 7 sources.

Key words: HYPOTHESIS TESTING, SEQUENTIAL ANALYSIS, MARKOV CHAINS, DISCRETIZATION OF SEQUENTIAL TEST, WALD'S LIKELIHOOD RATIO TEST, MACHINE LEARNING, HYPERPARAMETER TUNING.

Object of research — Wald's sequential likelihood ratio test with discretization of the state space.

Purpose of the work — To study the effect of discretizing the state space of the sequential test statistic on its characteristics such as average number of iterations and error probabilities, present an approximation method of the test using Markov chains, and to demonstrate the applicability of the proposed approaches on the example of hyperparameter tuning for machine learning models.

Research methods — Literature review, mathematical statistics, Markov chain theory, computer simulation of sequential tests, machine learning methods.

Obtained results — A computational model of Wald's sequential likelihood ratio test was implemented in Python, with an analysis of its key characteristics. The effect of state space discretization was investigated, and a Markov chain-based approximation method was proposed and evaluated. Furthermore, an application of the developed approaches to the problem of hyperparameter tuning in machine learning was demonstrated.