

БЕЛОРУССКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ

Факультет прикладной математики и информатики

Кафедра дискретной математики и алгоритмики

Аннотация к дипломной работе

**«Алгоритмы стохастического отбора значимых признаков в
задаче полногеномного поиска ассоциаций»**

Климук Геннадий Владимирович

Научный руководитель - доктор физико-математических наук, член
корреспондент НАН Беларуси, профессор Тузиков А.В.

2016

РЕФЕРАТ

Дипломна работа, 58 стр., 20 рис., 6 табл., 11 источников, 2 приложения.

ГЕНОМ, ТОЧЕЧНАЯ МУТАЦИЯ, МАШИННОЕ ОБУЧЕНИЕ, ЗАДАЧА КЛАССИФИКАЦИИ, ОТБОР ЗНАЧИМЫХ ПРИЗНАКОВ, ПОИСК ЗАВИСИМОСТЕЙ, МОДЕЛЬ, АЛГОРИТМ.

Объектом исследования являются методы отбора значимых признаков в задаче полногеномного поиска ассоциаций.

Целью работы является изучение методов стохастического отбора признаков, применение изученных методов для анализа геномных последовательностей, сравнение методов с общими подходами к решению задачи отбора значимых признаков, а также разработка программы, позволяющей решать данную и аналогичные задачи.

В результате работы была разработана программа для автоматизированного анализа геномных последовательностей, поиска ассоциаций и определения лекарственной устойчивости препаратов.

Методы исследования: изучение литературы, поиск ассоциаций, отбор признаков, машинное обучение.

Область применения: назначение наилучших лекарственных препаратов пациентам, у которых выявлена устойчивость к основным лекарствам.

ABSTRACT

Graduation work, 58 p., 20 pictures, 6 tables, 11 sources, 2 appendices.

GENOME, POINT MUTATION, MACHINE LEARNING,
CLASSIFICATION PROBLEM, IMPORTANT FEATURES SELECTION,
DEPENDENCIES SEARCH, MODEL, ALGORITHM.

Object of research: methods of important features selection in the problem of genome-wide associations search.

Goal of research: study of methods of stochastic important features selection, application of the studied methods for the analysis of genomic sequences, comparison of methods with general approaches to solving the problem of important features selection, as well as program development, for solving this and similar tasks.

The result of research: a program has been developed for automated analysis of genomic sequences and definition of medicament resistance of preparations.

Research methods: analysis of topic relevant literature, associations search, features selection, machine learning.

Applications: prescribing the most suitable medicines to patients, who are known to have resistance to general medicines.