

ЯСОВІ 4: АЛГОРИТМЫ МНОГОМЕРНОГО АНАЛИЗА, СПЕЦИАЛЬНЫЕ БИОИНФОРМАТИЧЕСКИЕ МЕТОДЫ

Д. А. Полунин

Институт цитологии и генетики СО РАН, г. Новосибирск
Новосибирский государственный университет

При анализе многомерных данных часто возникает необходимость в выполнении однопипных преобразований данных, кроме того важной задачей является оценка достоверности полученного результата. Наилучшим универсальным средством для этого является бутстреп-анализ, который требует многократного (от тысяч до миллионов раз) просчета тех же данных с намеренно вносимым случайным искажением, что невозможно без автоматизации обработки данных.

ЯСОВІ 4 состоит из программы-диспетчера, интерпретирующей пользовательский скрипт, и набора директив, реализующих преобразования данных.

В рамках работы над проектом реализовано более сорока вспомогательных директив для подготовки данных, набор директив для вычисления матриц сходства-различия, а также методы статистического анализа, среди которых можно выделить директивы для понижения размерности данных с минимальными потерями значимой информации: метод главных компонент, метод главных координат, неметрическое многомерное шкалирование; директивы для анализа взаимосвязи признаков: дискриминантный анализ, множественная линейная регрессия, нейронные сети с обратным распространением ошибки; директивы для ПЛС анализа [1]: 2В-PLS анализ, PLS регрессия; директивы для кластеризации: алгоритм объединения, алгоритм ближайшего соседа.

1. Агбаш И. А., Полунин Д. А., Сулопаров Д. С., Штайгер И. А. «Программный комплекс для многомерного анализа микрочиповых данных. Многомерный PLS анализ». // Материалы 50-й Международной научной студенческой конференции «Студент и научно-технический прогресс»: Информационные технологии. 13-19 апреля 2012 г., НГУ, Новосибирск. – Новосибирск, 2012. – С. 123.

Научный руководитель – д-р биол. наук, доцент В. М. Ефимов