

УДК 004.85

## ОТБОР ПРИЗНАКОВ НА ОСНОВАНИИ КРИТЕРИЯ ВЗАИМНОЙ ИНФОРМАЦИИ

Нериновский А.К. (Университет ИТМО)

Научный руководитель – к.ф.-м.н., Сметанников И.Б.

(Университет ИТМО)

В данной работе будет предложен алгоритм отбора признаков основанный на вычислении максимальной статистической зависимости.

**Введение.** Выбор признаков является важной проблемой для систем основанных на методах машинного обучения. В данной работе мы рассмотрим методы отбора признаков в соответствии с критерием максимальной статистической зависимости, основанным на мере взаимной информации.

**Основная часть.** Из-за трудности непосредственного вычисления максимальной статистической зависимости, мы сначала выводим эквивалентную форму, называемую критерий минимальной избыточности-максимальной релевантности (mRMR).

В данной работе описывается двухэтапный алгоритм выбора признаков, получающийся в результате комбинирования mRMR и других сложных функций отбора признаков. Это позволяет выбрать компактный набор признаков при очень низких затратах.

**Выводы.** В данной работе описывается обширное экспериментальное сравнение предлагаемого алгоритма и других методов с использованием трех различных классификаторов (наивный байесовский, методом опорных векторов и линейный дискриминационный анализ) и четырех различных наборах данных (рукописные цифры, аритмия, раковые клеточные линии NCI, и ткани лимфомы). Результаты подтверждают, что mRMR приводит к многообещающему улучшению выбора признаков и точности классификации.

Нериновский А.К. (автор)

Подпись

Сметанников И.Б. (научный руководитель)

Подпись