

**АЛГОРИТМ ВЫБОРА ПРИЗНАКОВ
ОСНОВАННЫЙ НА ЛАПЛАССОВОЙ РЕГУЛЯРИЗАЦИИ ДЛЯ
ЧАСТИЧНО-РАЗМЕЧЕННЫХ ДАННЫХ С ПРИЗНАКАМИ РАЗНОГО РОДА**

Терещук М.А., Университет ИТМО

Научный руководитель – Сметанников И.Б., к.т.н, доцент факультета информационных технологий и программирования, Университет ИТМО

Аннотация

В настоящее время задача понижения размерности становится крайне актуальной, поскольку в силу возрастающих вычислительных возможностей становится возможным исследовать высокоразмерные данные. В этом докладе рассматривается один из таких алгоритмов – алгоритм выбора признаков на основе Лапласовской регуляризации.

Введение

В ИТМО активно разрабатывается библиотека выбора признаков ИТМО_FS, призванная облегчить работу программистам, реализуя наиболее популярные подходы для выбора признаков и покрывая как можно больше способов работы и типов данных.

Основная часть

Одним из пока недостаточно мощно покрытых типов данных являются алгоритмы, работающие с частично-размеченными данными. Такой тип данных особенно важен в связи с активным ростом числа картинок и фотографий, появляющихся в Сети, и соответственно увеличением числа возможных задач. Целью этой научно-исследовательской работы была реализация алгоритма выбора признаков для частично-размеченных данных с признаками разного рода, а именно алгоритма выбора признаков на основе Лапласовской регуляризации.

Выводы

В результате проведенной работы был имплементирован алгоритм выбора признаков, умеющий работать с несколькими типам частично-размеченных данных, такой алгоритм позволит улучшить качество работы различных алгоритмов распознавания, поскольку сможет учитывать несколько источников данных.

Терещук М.А. (автор)

Подпись

Сметанников И.Б. (научный руководитель)

Подпись