

УДК 004.8

## **АЛГОРИТМ ВЫБОРА ПРИЗНАКОВ НА ОСНОВЕ ДОКАЗАТЕЛЬНОЙ РЕГРЕССИИ**

**Шушарин Д.Р. (Университет ИТМО)**

**Научный руководитель – кандидат технических наук Сметанников И.Б.  
(Университет ИТМО)**

В докладе рассматривается алгоритм выбора признаков на основе доказательной регрессии, проводится анализ качества работы алгоритма на нескольких наборах данных.

В машинном обучении существует отдельный тип задач – прогнозирование целевых значений на основе наблюдений в прошлом, например предсказание цены страховки здоровья, на основе показателей человека. Такие задачи часто решаются при помощи нейронных сверточных или рекуррентных сетей, однако такие алгоритмы работают по принципу «черного ящика» и не всегда ясны причины хороших результатов работы таких алгоритмов, в связи с чем их сложно интерпретировать. Поэтому был разработан алгоритм, совмещающий в себе точность нейронных сетей и позволяющий интерпретировать полученные результаты – доказательная регрессия. Такой алгоритм в процессе обучения присваивает веса признакам в исследуемом наборе данных, что позволяет четко оценить вклад в работу каждого признака, а это как раз позволяет делать выводы о зависимостях в данных.

На основе статьи о доказательной регрессии планируется реализовать алгоритм выбора признаков на языке Python для расширения функционала библиотеки ITMO\_FS. Такой подход ранее не рассматривался в библиотеке, что значительно расширит ее инструментарий и позволит решать больше задач с ее помощью, а также может облегчить исследования в области выбора признаков для других разработчиков.

Шушарин Д.Р. (автор)

Подпись

Сметанников И.Б. (научный руководитель)

Подпись