

Итеративное улучшение тематической модели с обратной связью от пользователя

Tuesday, 23 May 2023 13:00 (15 minutes)

В работе представлен метод тематического моделирования с использованием обратной связи от пользователя. Обратная связь заключается в определении принадлежности темы, полученной при тематическом моделировании, к одной из трёх категорий: релевантная, нерелевантная, «мусорная». Основная задача состоит в улучшении базовой модели, которое заключается в выделении новых релевантных тем при сохранении выделенных тем и уменьшении числа «мусорных» тем. В работе предлагается решение с использованием библиотек тематического моделирования и регуляризаторов сглаживания и декоррелирования. Вычислительный эксперимент проводится на текстовой коллекции, основанной на новостях сайта Lenta.ru, опубликованных в период с мая по август 2008 года.

Primary author: GORBULEV, Alex (MIPT)

Co-authors: Dr VORONTSOV, Konstantin (MIPT); ALEXEEV, Vasiliy (MIPT)

Presenter: GORBULEV, Alex (MIPT)

Session Classification: Computer & Data Science 23

Track Classification: Computer & Data Science