

1. Воронцов К. В. (2020) [Вероятностное тематическое моделирование: теория, модели и проект BigARTM](#)

Описывает работа алгоритма тематического моделирования: формирование данных для модели, EM-алгоритм, добавление регуляризации к модели. В главе 7 описываются мультимодельные языковые модели, приводится оптимизационная формула и упоминается про веса модальностей.

2. Янина А.О. (2016) [Мультимодальные тематические модели статей коллективных блогов для разведочного поиска](#)

Описывается решение задачи выдачи рекомендаций и разведочного информационного поиска с помощью мультимодальной модели. Показывается, что мультимодальная модель превосходит по качеству ранжирования модели без учета модальностей (метрики Precision@k, recall@k, MAP@k). Озвучено, что веса модальностей подбираются вручную

3. Timothy N. Rubin · America Chambers · Padhraic Smyth · Mark Steyvers (2010) [Statistical topic models for multi-label document classification](#)

4. Никитин Ф. А. (2018) [Применение мультимодальных тематических моделей к анализу транзакционных данных банков](#)

Описывается задача выявления профиля потребностей клиента по транзакционным данным: идентификатор клиента, совершившего транзакцию, сумма транзакции и МСС-код. Предлагается способ подбора весовых коэффициентов для модальностей.

5. Ефимова И. В. (2017) [Иерархическая мультимодальная тематическая модель коллекции научно-популярных текстов](#)

Строится иерархическая тематическая модель на 2 уровня. На первом уровне названия тем заданы экспертами, на втором автоматически определяются моделью. Для первого уровня в качестве модальностей выступают теги документов. Вес модальности задается вручную “Обычно число тегов $|G|$ много меньше числа слов. Поэтому модальность тегов в уравнении (1) предлагается учитывать с большим весом ηg (≈ 102 , при условии, что вес модальности слов равен 1). Иначе модальность тегов не будет вносить вклад в модель”

6. К. В. Воронцов [Тематическое моделирование](#) Видео с описанием алгоритмов тематического моделирования