

УДК 930.2

EDN [SXIVNC](#)



<https://www.doi.org/10.47813/mip.5.2023.9.129-135>

Искусственный интеллект в исторических исследованиях: потенциал и пределы эффективности

А.В. Кузнецов

Институт всеобщей истории РАН, Ленинский пр., 32 а, Москва, 119334, Россия

E-mail: historyras@gmail.com

Аннотация. В статье подробно анализируются сильные и слабые стороны использования технологий искусственного интеллекта в исторических исследованиях на примере методов тематического моделирования. Рассмотрено применение популярных алгоритмов машинного обучения, таких как Latent Dirichlet Allocation (LDA), для анализа больших массивов текстовых данных. Обсуждаются ключевые преимущества тематического моделирования, включая возможности обработки больших объемов текста, выявления скрытых тематических структур, отслеживания динамики тем. В то же время подробно рассмотрены важные ограничения этого подхода, такие как предположение о статичности тем в большинстве моделей, слабая интерпретируемость результатов, их неустойчивость и сильная зависимость от настроек параметров. Это требует критического отношения к получаемым результатам и их тщательной верификации с опорой на предметные знания исследователя. На основе проведенного анализа делается вывод о необходимости сбалансированного сочетания методов искусственного интеллекта с традиционными качественными подходами гуманитарных наук.

Ключевые слова: искусственный интеллект, тематическое моделирование, исторические исследования, анализ текста, машинное обучение.

Artificial Intelligence in Historical Research: Potential and Limits of Effectiveness

A.V. Kuznetsov

Institute of World History of RAS, Leninskij pr. 32a, Moscow, 119334, Russia

E-mail: historyras@gmail.com

Abstract. The article thoroughly analyzes the strengths and weaknesses of applying artificial intelligence technologies in historical research using the example of topic modeling methods. The use of popular machine learning algorithms such as Latent Dirichlet Allocation (LDA) for analyzing large arrays of textual data is examined in detail. The key advantages of topic modeling are discussed, including the abilities to process large volumes of text, identify hidden thematic structures, and track topic dynamics over time. At the same time, significant limitations of this approach are considered, such as the assumption of topic stability in most models, poor interpretability of results, their instability and strong dependence on parameter settings. This requires a critical attitude towards the results obtained and their careful verification based on the subject knowledge of the researcher. Specific limitations analyzed include the static nature of topics in many models, weak interpretability of results, their instability, and high dependence on parameter settings. Based on the analysis, it is concluded that it is essential to balance artificial intelligence methods with traditional qualitative approaches in the humanities. The article provides concrete recommendations on the application of topic modeling in historical research to maximize the benefits while minimizing the risks.

Keywords: artificial intelligence, topic modeling, historical research, text analysis, machine learning.

1. Введение

Развитие цифровых технологий преобразовало почти все аспекты исследовательского процесса историков. В последние десятилетия наблюдается быстрый прогресс в области технологий искусственного интеллекта, открываются новые перспективы для их применения в различных сферах, включая гуманитарные науки. В исторических исследованиях особый интерес вызывает использование методов машинного обучения для анализа больших массивов текстовых данных, в частности такого метода как тематическое моделирование. Тематическое моделирование позволяет автоматически выявлять скрытые темы в больших коллекциях текстов. Эти алгоритмы машинного обучения открывают новые возможности для анализа и интерпретации исторических источников. В то же время эффективность использования автоматизированных методов анализа текстов в гуманитарной сфере имеет свои пределы. В целом ряде статей в последнее время затрагивается вопрос о сложностях и ограничениях применения тематического моделирования в исторических исследованиях [1-4].

2. Постановка задачи (Цель исследования)

Цель данной работы – проанализировать сильные и слабые стороны тематического моделирования как инструмента исторического исследования. Рассмотреть основные возможности и ограничения этого подхода, чтобы объективно оценить его потенциал и риски применения в исторической науке.

3. Метод тематического моделирования

Тематическое моделирование относится к методам машинного обучения и обработки естественного языка, позволяющим проанализировать скрытую тематическую структуру больших массивов текстовых данных без необходимости предварительной ручной категоризации. Темы в тематическом моделировании представляют собой вероятностные распределения над всем словарным запасом слов, встречающихся в анализируемых документах. На интуитивном уровне темы часто описывают как списки слов, которые часто встречаются вместе, и представляют собой набор ключевых понятий и терминов, характерных для данной предметной области.

Методы тематического моделирования изначально разрабатывались в области информационного поиска для анализа больших неструктурированных массивов текстовых данных. Однако в последнее десятилетие они стали очень популярными в

социальных и гуманитарных науках и активно применяется для решения содержательных исследовательских задач.

Наиболее распространенным алгоритмом тематического моделирования в гуманитарных науках является латентное размещение Дирихле (англ. Latent Dirichlet allocation, LDA). LDA основывается на предположении, что каждый документ в корпусе представляет собой набор слов, созданный в соответствии с темами, которые автор текста хотел затронуть. Задача тематического моделирования состоит в том, чтобы найти параметры процесса LDA, в результате которого, вероятно, был сформирован корпус. В литературе по LDA этот процесс также называют «выводом» (англ. inference), и, по сути, это задача обратного инжиниринга намерений автора при создании корпуса. Алгоритм LDA использует байесовский подход для реконструкции вероятностных распределений тем и документов, наиболее вероятно породивших исходный корпус текстов. В качестве выходных данных тематического моделирования формируется набор распределений слов по темам, описывающий вероятность принадлежности того или иного слова определенной теме, и аналогичный набор распределений тем по документам, описывающий вероятность представленности определенной темы в каждом конкретном документе корпуса.

4. Тематическое моделирование в исторических исследованиях

Первым историческим исследованием, использующим тематическое моделирование, стала статья Д. Ньюмана и Ш. Блок 2006 года «Вероятностная тематическая декомпозиция американской газеты восемнадцатого века». Авторы проанализировали темы в архиве газеты Pennsylvania Gazette [5]. В 2010-х годах тематическое моделирование получило широкое распространение среди историков как метод «дальнего чтения» и разведочного анализа больших текстовых наборов данных. Его применяли для анализа разнообразных исторических источников, включая газеты, журналы, книги, личные дневники и письма [6].

5. Преимущества тематического моделирования

Использование методов тематического моделирования обладает несколькими ключевыми преимуществами, которые объясняют растущий интерес историков к этому инструменту.

Во-первых, тематическое моделирование дает возможность анализировать гораздо большие объемы текстовых данных, чем это доступно при использовании только традиционных методов качественного анализа и «пристального чтения». Если вручную исследователь может детально проанализировать лишь небольшое количество документов, то алгоритмы тематического моделирования способны обработать тысячи текстов для выявления в них скрытых закономерностей. Это дополняет традиционный герменевтический подход новыми возможностями.

Во-вторых, поскольку тематическое моделирование использует методы неконтролируемого машинного обучения, это создает впечатление объективности таких методов, снижения влияния предвзятости и ошибок самого исследователя, которые неизбежно проявлялись бы при ручном анализе и категоризации текстов [5, 7]. Алгоритмы выявляют темы исключительно на основе статистики совместной встречаемости слов, без использования заранее заданных человеком категорий.

В-третьих, современные алгоритмы тематического моделирования, такие как Dynamic Topic Modeling дают возможность отслеживать эволюцию тем на больших временных интервалах путем включения метаданных, таких как даты создания текстов. Структурные же тематические модели позволяют количественно оценить изменение распространенности разных тем во времени или в зависимости от других переменных. Это способно выявить исторические тенденции и закономерности, сложно уловимые при традиционном качественном анализе [3].

5. Ограничения использования тематического моделирования в исторической науке

Несмотря на ряд очевидных достоинств, применение тематического моделирования в исторических исследованиях сталкивается с определенными ограничениями. Чтобы максимально эффективно использовать данный подход, необходимо быть осведомленным об этих ограничениях.

Одной из основных трудностей является предположение большинства алгоритмов тематического моделирования, в первую очередь LDA, о статичности тем во времени. Это допущение сделано ради упрощения математической модели и повышения вычислительной эффективности. Однако предположение, что темы не меняются со временем, ограничивает применимость LDA для анализа диахронических процессов.

Кроме того, исследования продемонстрировали, что темы, выделяемые LDA, часто оказываются трудно интерпретируемыми и не имеют аналогов среди категорий,

которые выделил бы в тех же текстах эксперт-человек, хорошо знакомый с материалом [8]. Это свидетельствует о том, что формальная тематическая модель может не соответствовать реальной смысловой структуре текстов.

Алгоритмы тематического моделирования также демонстрируют низкую воспроизводимость результатов [1]. Многократный запуск моделирования на одних и тех же входных данных может давать существенно различающиеся результаты. Это связано с вероятностной природой алгоритмов и чувствительностью к настройкам параметров. Такая неустойчивость ставит под вопрос надежность получаемых результатов.

Наконец, применение тематического моделирования требует от исследователя глубокого понимания моделируемого материала и контекста для правильной интерпретации результатов [1]. В отличие от человека, алгоритмы тематического моделирования неспособны понимать смысл текста и не обладают фоновыми знаниями о моделируемой предметной области. Поэтому без квалифицированной интерпретации со стороны эксперта результаты тематического моделирования могут привести к ошибочным выводам.

Важно также помнить, что тематическое моделирование является упрощенной моделью языка и не учитывает такие важные элементы, как синтаксис и контекст слов в тексте.

5. Рекомендации по применению тематического моделирования в исторических исследованиях

На основании анализа применения методов тематического моделирования в истории можно сформулировать рекомендации, которые позволят избежать типичных ошибок и максимально эффективно применять данный подход в исторические исследования:

1. Использовать тематическое моделирование как инструмент генерации гипотез, а не для получения окончательных выводов. Как показали исследования [1, 8], результаты модели всегда требуют дальнейшей интерпретации и проверки исследователем на соответствие контексту. Необходимо избегать чрезмерного доверия к статистическим моделям [7].
2. Комбинировать тематическое моделирование с традиционными методами анализа текстов для сбалансированного использования количественных и качественных подходов.

3. Учитывать временную изменчивость тем путем использования динамических или структурных тематических моделей при анализе диахронических процессов. Это позволит лучше учесть временную природу исторических текстов.
4. Тщательно интерпретировать полученные темы через призму глубоких предметных знаний о контексте изучаемой области [7-10]. Алгоритмы не способны понимать смысл текста.
5. Обязательно проводить валидацию модели [9], например, путем ручной категоризации небольших выборок для оценки интерпретируемости полученных тем.
6. Документировать всю процедуру моделирования, включая настройки, исходные данные и результаты [1], для обеспечения прозрачности и воспроизводимости.

Такой обдуманный и осторожный подход позволит максимально использовать аналитические возможности тематического моделирования в исторических исследованиях.

5. Выводы

Тематическое моделирование, с одной стороны, действительно открывает новые горизонты для анализа больших массивов текстовых данных, недоступных для обработки традиционными методами. Алгоритмы машинного обучения позволяют автоматизировать выявление тем в текстах, отслеживание динамики тем во времени. Это существенно расширяет аналитические возможности исследователей. С другой стороны, тематическое моделирование имеет ряд важных ограничений, таких как предположение о статичности тем, слабая интерпретируемость результатов, их неустойчивость и сильная зависимость от настроек параметров. Это требует критического отношения к получаемым результатам и их тщательной верификации с опорой на предметные знания исследователя. Таким образом, несмотря на большой потенциал, технологии искусственного интеллекта не могут полностью заменить традиционные методы гуманитарных наук. Их эффективное использование требует комбинирования с традиционными подходами и глубокого понимания моделируемой предметной области. Только взвешенный критический подход позволит максимально реализовать преимущества алгоритмов искусственного интеллекта и минимизировать риски ошибочных выводов при их применении в исторических исследованиях.

Список литературы

1. Oiva M. Topic Modeling Russian History / M. Oiva // The Palgrave Handbook of Digital Russia Studies. – 2021. – P. 427-442.
2. Althage M. Potenziale und Grenzen der Topic / M. Althage // Digital History: Konzepte, Methoden und Kritiken Digitaler Geschichtswissenschaft. – 2022. – Т. 6. – №. 24. – P. 255-277.
3. Villamor Martin M. The promise of machine-learning-driven text analysis techniques for historical research: topic modeling and word embedding / Martin M. Villamor, D.A. Kirsch, F. Prieto-Nañez // Management & Organizational History. – 2023. – Т. 18. – №. 1. – P. 81-96.
4. Gillings M. The interpretation of topic models for scholarly analysis: An evaluation and critique of current practice / M. Gillings, A. Hardie // Digital Scholarship in the Humanities. – 2023. – Т. 38. – №. 2. – P. 530-543.
5. Newman D.J. Probabilistic topic decomposition of an eighteenth-century American newspaper / D.J. Newman, S. Block // Journal of the American Society for Information Science and Technology. – 2006. – Т. 57. – №. 6. – P. 753-767.
6. Boyd-Graber J. Applications of Topic Models / J. Boyd-Graber, Y. Hu, D. Mimno // Foundations and Trends in Information Retrieval. – 2017. – Vol. 11. – №. 2/3. – P. 143-296.
7. Schmidt B.M. Words alone: Dismantling topic models in the humanities / B.M. Schmidt // Journal of Digital Humanities. – 2012. – Т. 2. – №. 1. – P. 49-65.
8. Mohr J. W., Bogdanov P. Introduction—Topic models: What they are and why they matter / J.W. Mohr, P. Bogdanov // Poetics. – 2013. – Т. 41. – №. 6. – P. 545-569.
9. Grimmer J. Text as data: The promise and pitfalls of automatic content analysis methods for political texts / J. Grimmer, B.M. Stewart // Political analysis. – 2013. – Т. 21. – №. 3. – P. 267-297.
10. Варламова Л.Н. Управление документацией: англо-русский аннотированный словарь стандартизированной терминологии / Л.Н. Варламова, Л.С. Баюн, К.А. Бастрикова. – Москва: Спутник+, 2017. – 398 с.