УДК 004.85

ИСПОЛЬЗОВАНИЕ СЕМАНТИЧЕСКОГО ПОИСКА ДЛЯ ВЫБОРА И РАНЖИРОВАНИЯ НАУЧНЫХ ГЕОЛОГИЧЕСКИХ ПУБЛИКАЦИЙ

М. И. Патук[1] [0000-0003-3036-2275], В. В. Наумова[2] [0000-0002-3001-1638]

ФГБУН Государственный геологический музей им. В.И. Вернадского РАН, Москва;

[1]patuk@mail.ru, [2]naumova_new@mail.ru

Аннотация

Агрегирование научной информации играет важную роль для комплексного анализа геологических объектов. В настоящей работе мы рассматриваем потенциал и возможности семантического поиска для выбора тематически близких геологических публикаций. Проанализированы различные языковые модели в контексте нахождения сходства и различия между текстами при описании месторождений полезных ископаемых. Показано значительное улучшение результатов поиска после дополнительной тренировки языковых моделей. Представлены два веб-сервиса, основанных на методе расчета семантической близости текстов с количественной оценкой меры близости.

Ключевые слова: искусственный интеллект, машинное обучение, обработка естественного языка, семантический поиск, геология.

ВВЕДЕНИЕ

Научные геологические публикации состоят, как правило, из трех типов информации: текста, числовых данных и пространственных данных. Основой для анализа служат, в основном, числовые данные. Реже анализу подвергаются пространственные данные, например, карты, схемы, разрезы, спутниковые изображения и др. Сам текст публикаций содержит анализ, выполненный автором. Но геологические объекты, которые описаны в публикациях, являются комплексными образованиями. И каждая отдельная статья отражает лишь некоторые свойства описываемого объекта.

© М. И. Патук, В. В. Наумова, 2024.
Данная статья распространяется на условиях международной лицензии Creative Commons License Attribution 4.0 International (CC BY 4.0).

Таким образом, необходим числовой инструмент для выбора близких по изучаемому объекту публикаций. Такой инструмент уже существует – это семантический поиск, при котором способ и технология поиска информации основаны на использовании контекстного (смыслового) значения запрашиваемых фраз, вместо словарных значений отдельных слов или выражений при поисковом запросе [1]. Этот подход активно развивается в последнее десятилетие в связи с бурным развитием методов обработки естественного языка (NLP) [2]. Одна из задач обработки естественного языка – семантическая схожесть текстов (Semantic Text Similarity, STS) [3]. Это направление достаточно хорошо освещено в литературе. Основной импульс развитию семантического поиска дал поиск информации в интернете, а именно, такие поисковые системы, как Google и Yandex.

Ключевой вехой на этом пути стал выпуск большой языковой модели BERT [4]. BERT – это языковая модель, основанная на архитектуре трансформер [5] и предназначенная для предобучения языковых представлений с целью их последующего применения для решения широкого спектра задач обработки естественного языка [6]. В настоящее время только на портале HuggingFace насчитывается 15 различных BERT моделей [7], а с учетом клонов количество BERT-подобных моделей насчитывает несколько сотен.

Анализ литературных источников

До наступления эры больших языковых моделей семантическая близость текстов определялась, в основном, на основе словарей и онтологий, формальных описаний терминов предметной области и отношений между ними. Хороший обзор этих методов приведен в работе [8].

Следующим этапом в семантическом поиске можно считать применение таких методов, как TF-IDF [9], BM25 [10] и Word2Vec [11]. От лексического подхода предыдущих методов данные методы отличает использование векторных представлений текстов. Хотя векторное представление слов в этих методах является статичным и не зависит от контекста, они до сих пор активно применяются для поиска информации [12]. Отсутствие зависимости от контекста

в методе Word2Vec может компенсироваться построением расширенного запроса с применением синонимов [13]. Мейнстримом в данной области является использование больших языковых моделей типа BERT, дополненных различными методиками выбора и обработки текста [14, 15], хотя некоторые авторы считают, что поиск с помощью больших языковых моделей следует дополнить использованием ключевых слов [16].

Методика работы

Анализ литературы по семантическому поиску информации с использованием больших языковых моделей типа BERT позволяет принять следующую схему работы:

1. На портале HuggingFace выбираем подходящую языковую модель.
2. Создаем обучающий набор данных (датасет).
3. Дообучаем (fine-tuning) модель на созданном наборе данных.
4. Производим оценку полученного результата.
5. Применяем выбранную языковую модель для поиска информации.

Метод поиска, близкий к описанному выше, используется в электронной библиотеке Elibrary [17]. Насколько можно судить, он основан на языковой модели sci-rus-tiny [18], для обучения которой были использованы данные этой библиотеки [19]. Такой поиск присутствует на основной странице в виде пункта «Нейропоиск» и на странице конкретной публикации в виде пункта «Найти близкие по тематике публикации». Проверка работы этого поиска на запросах геологической тематики показала, что он хорошо находит основную тематику запроса, но представляет слишком широкие результаты в рамках этой тематики. Несколько обобщенные результаты поиска становятся понятны, если обратить внимание на структуру обучающего набора данных. Все естественные науки составляют в нем 11%, а геология, естественно, еще меньше.

Модельными геологическими объектами в нашей работе были выбраны месторождения твердых полезных ископаемых. Подбирались статьи, относящиеся к их описанию, с сайтов интернет-библиотек Elibrary [17] и CyberLeninka

[20], на русском и английском языках. Обучающий набор данных был создан на основе пар наименований статей и пар абстрактов с указанием степени схожести в интервале от 0 (совсем не похожи) до 5 (идентичные).

Дообучение (fine-tuning) моделей выполнялось по следующей методике [21]. Необходимость дообучения языковых моделей при их применении в конкретных предметных областях неоднократно демонстрировалось в литературе [22]. Данное обстоятельство связано с тем, что первоначальное обучение моделей происходит на больших наборах данных (от сотен тысяч до сотен миллионов записей), имеющихся в интернете. Эти данные, как правило, взяты из Википедии и социальных сетей. Как следствие, информации из требуемой нам предметной области там или очень мало, или нет вообще. Есть редкие примеры, когда модель обучается на корпусе научных статей [19]. Но и в этом случае авторы стремятся охватить наиболее широкий круг областей знаний, и в результате конкретная предметная область оказывается «сжатой» в узкую полосу представлений. Для того чтобы языковая модель наиболее подробно отражала нюансы требуемой предметной области, необходимо выполнить тонкую настройку ее многочисленных параметров. Эта задача решается путем дополнительного обучения модели на наборе соответствующих примеров.

В процессе обучения языковых моделей части слов, отдельные слова, предложения или целые фрагменты обучающего текста преобразуются в числовые многомерные векторы, так называемые эмбединги (embedding) или векторные представления. При таком обучении векторное представление конкретного слова оказывается зависимым от стоящих рядом с ним слов, т. е. от контекста. К полученным векторным представлениям можно применять методы математической обработки, для измерения меры их близости. Анализ литературы показал, что в большинстве случаев для определения степени близости векторов используется косинусное сходство [23] (угол между векторами, значение которого находится в интервале [0, 1]): 1 означает, что вектора совпадают (максимальное сходство), 0 – вектора перпендикулярны (максимальное различие).

Выбор языковых моделей выполнялся по следующим критериям: поддерживаемые языки – русский и английский, тренировка на задаче – семантическая схожесть текстов (Sentence Similarity). Размер модели не более ~ 2 Гб. Модели большего размера проблематично тренировать на персональном компьютере или облачных платформах типа Yandex Cloud.

В процессе производства языковых моделей создаются, как правило, 3 версии модели: маленькая (small), основная (base) и большая (large). Но иногда, для специальных целей, ограничиваются только одной версией. Мы выбрали 4 модели разного размера: 2 маленьких, 1 основную (среднюю) и 1 большую. Первая – sci-rus-tiny – маленькая модель, обученная на текстах научных статей с сайта Elibrary.ru [18]. Вторая – rubert-tiny2 – маленькая модель, обученная на общеупотребительной лексике [24]. Третья – multilingual-e5-base – модель среднего размера, мультиязычная – 94 языка, обученная на данных из Википедии и новостных сайтов [25]. Четвертая – E5-large-en-ru – большая языковая модель [26]. Это очищенная модель – multilingual-e5-large [27] – в которой оставлены только русский и английский языки. За счет этого удалось значительно уменьшить размер модели и сократить количество параметров.

Для оценки качества моделей в литературе используются бенчмарки (benchmark) [28]. Поскольку на интересующей нас задаче – семантического сходства текстов описаний месторождений твердых полезных ископаемых – таких бенчмарков нет, то нами был создан собственный бенчмарк. Были выбраны фрагменты текстов описания золоторудных месторождений на русском и английском языках. Фрагменты описаний железорудных месторождений и фрагменты описаний месторождений меди. При использовании этого бенчмарка мы ожидали, что метрика сходства между текстами описания месторождений золота (не зависимо от языка фрагмента) будет большой – в интервале 0.5–1.0, а метрика сходства между текстами описания месторождений золота и месторождений железа и меди гораздо меньше – в интервале 0.0–0.5.

Результат проверки показал, что для исходных выбранных моделей наши ожидания не оправдались. Для удобства визуализации мы объединили результаты в 4 пары – золото–золото, золото–gold, золото–железо, золото–медь

– с усреднением результатов по каждой паре. Все 4 пары консолидированных примеров дали практически одинаковые результаты на нашем бенчмарке – в районе 0.7–0.9. Такой результат еще раз подтверждает тезис о необходимости дообучения моделей на примерах из исследуемой предметной области (в нашем случае – это Геология – Месторождения полезных ископаемых).

Далее мы выполнили дообучение выбранных моделей на нашем наборе данных, варьируя величину обучающего датасета, и повторно выполнили проверку на нашем бенчмарке. По результатам измерений выделился явный лидер – модель E5-large-en-ru. Зависимость 4 пар консолидированных примеров от величины обучающего датасета приведена на Рис. 1. Видно, что метрики разных типов месторождений значимо разделяются. Полученные результаты полностью отвечают нашим ожиданиям. Указанная модель была выбрана для дальнейшей работы.

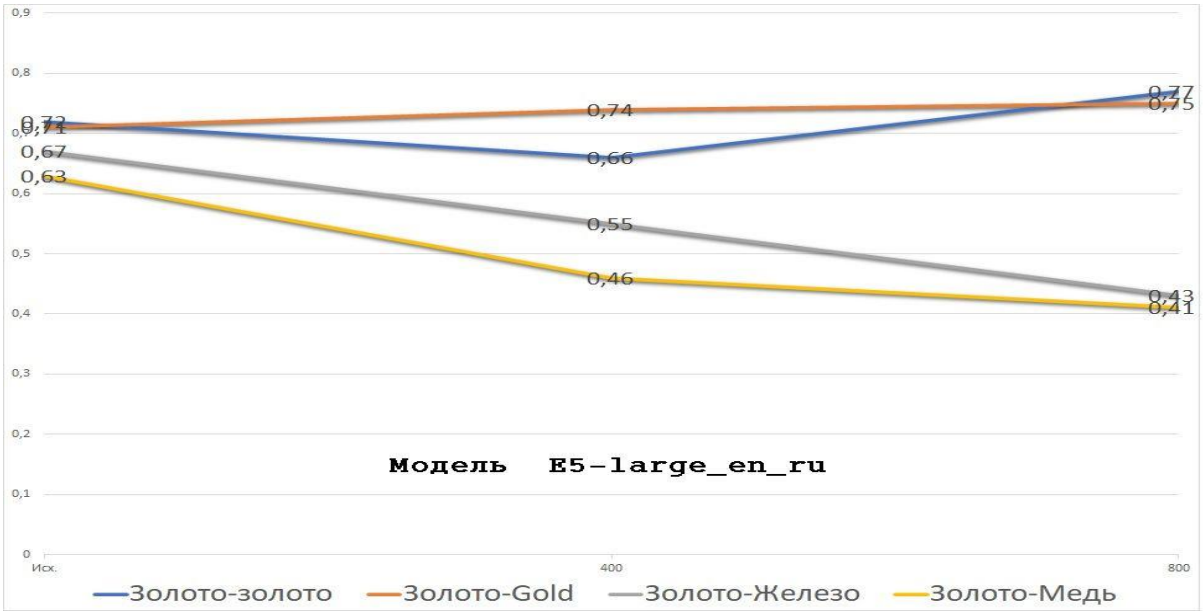

Рис. 1. Зависимость качества модели от величины обучающего датасета.

РЕЗУЛЬТАТЫ

Цель первого этапа нашей работы – выбрать семантически близкие публикации из Репозитория ГГМ РАН (https://repository.geologyscience.ru/) [29] для дальнейшего анализа. Для уменьшения временных затрат на поиск была создана база данных на PostgreSQL. В таблицу помещены наименования ста-

тей, их ссылки на архив публикаций и заранее рассчитанные векторные представления наименований статей и абстрактов. Таким образом, поиск статей по искомому запросу преобразуется в вычисление косинусной близости векторного представления запроса с векторными представлениями статей. Максимальные значения косинусной близости будут отражать искомые публикации.

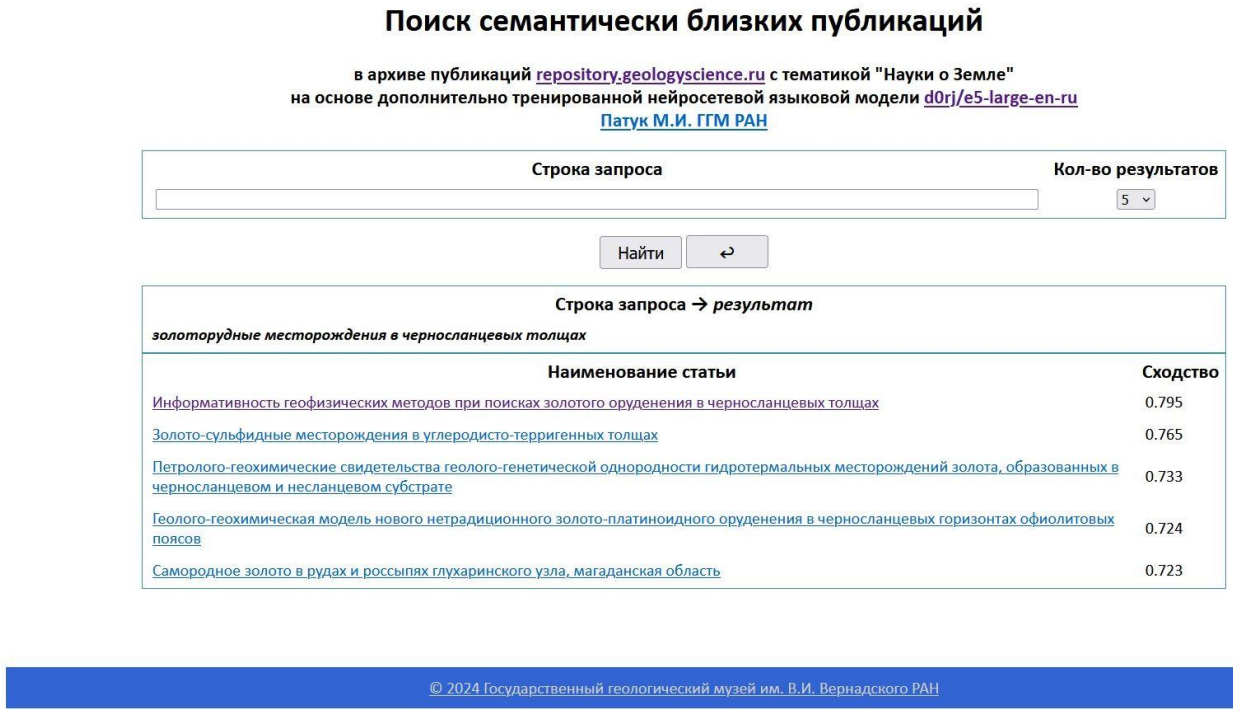

Рис. 2. Веб-сервис поиска публикаций

Описанная технология была реализована в виде веб-сервиса и доступна по адресу https://service.geologyscience.ru/ [30]. После введенного запроса и указания количества публикаций для отображения производится расчет, и результаты выводятся на экран (Рис. 2). Клик по статье в результатах открывает страницу со статьей в архиве публикаций. Чем подробнее будет составлена строка запроса, тем более точным будет результат. Объединенные текстовые данные выбранных статей будут использованы для дальнейшего анализа.

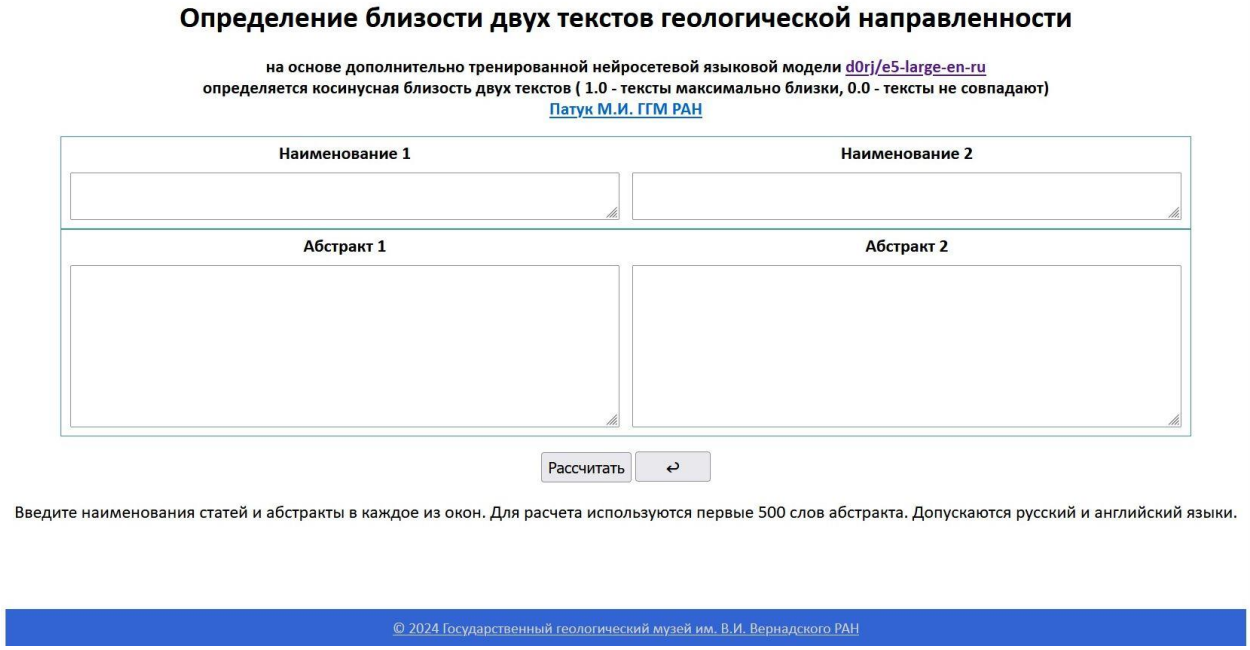

Рис. 3. Веб-сервис определения близости текстов

На основе данной модели можно решать еще одну актуальную задачу – экспресс-анализ схожести двух публикаций на основе расчета метрики сходства фрагментов их текстов, на русском или английском языках. Для этого реализован еще один веб-сервис – Определение близости двух текстов геологической направленности (Рис. 3). Реализована такая же технология работы. Выполняется расчет векторных представлений фрагментов текста и выводится результат их косинусной близости (Рис. 4).

ВЫВОДЫ

Проведенная нами работа показала, что современные методы искусственного интеллекта с успехом могут применяться для анализа текстов научных статей. Возможности семантического поиска значительно расширяют возможности поиска информации. Полученные языковые модели могут с успехом применяться в смежных областях анализа – определения схожести текстов. У приведенных методов есть свои естественные ограничения. Они хорошо работают только в той предметной области, на которой была выполнена дополнительная тренировка моделей.

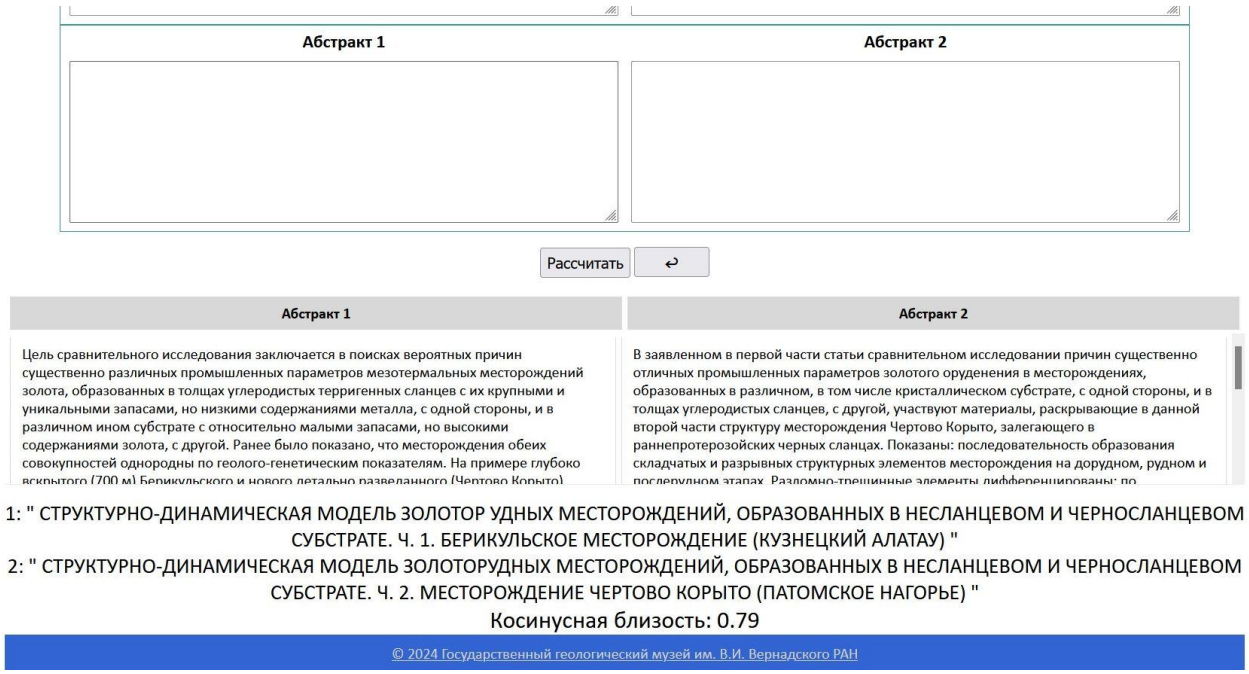

Рис. 4. Веб-сервис определения близости текстов. Результат расчета.

Для улучшения качества работы нашей языковой модели мы дополнили систему логированием запросов, чтобы выбрать направление дальнейшего совершенствования работы системы.

Работы выполняются в рамках Государственного задания ГГМ РАН по теме № 1021061009468-8-1.5.1 «Цифровая платформа интеграции и анализа геологических и музейных данных».

СПИСОК ЛИТЕРАТУРЫ

1. Семантический поиск. URL: https://ru.wikipedia.org/wiki/Семантический_поиск (дата обращения 10.09.2024)

2. Ваш путеводитель по миру NLP (обработке естественного языка), URL: https://habr.com/ru/companies/otus/articles/705482/ (дата обращения 10.09.2024)

3. Semantic similarity. URL: https://en.wikipedia.org/wiki/Semantic_similarity (дата обращения 10.09.2024)

4. *Devlin J., Chang M.W., Lee K., Toutanova K.* Bert: pre-training of deep bidirectional transformers for language understanding. URL: arxiv.org/abs/1810.04805 (2018)

5. Объясняем простым языком, что такое трансформеры. URL: https://habr.com/ru/companies/mws/articles/770202/ (дата обращения 10.09.2024)

6. BERT (языковая модель). URL: https://neerc.ifmo.ru/wiki/index.php?title= BERT_(языковая_модель) (дата обращения 10.09.2024)

7. BERT community. URL: https://huggingface.co/google-bert (дата обращения 10.09.2024)

8. *Akila D., Jayakumar C.* Semantic Similarity – A Review of Approaches and Metrics // International Journal of Applied Engineering Research. 2014. Vol. 9, No. 24. P. 27581–27600.

9. TF-IDF. URL: https://ru.wikipedia.org/wiki/TF-IDF (дата обращения 10.09.2024)

10. Okapi BM25. URL: https://ru.wikipedia.org/wiki/Okapi_BM25 (дата обращения 10.11.2024)

11. Word2vec. URL: https://ru.wikipedia.org/wiki/Word2vec (дата обращения 10.09.2024)

12. *Краснов Ф.В., Смазневич И.С., Баскакова Е.Н.* Проблема потери решений в задаче поиска схожих документов: Применение терминологии при построении векторной модели корпуса // Бизнес-информатика. 2021. Т. 15. № 2. С. 60–74. https://doi.org/10.17323/2587-814X.2021.2.60.74

13. *Атаева О.М., Серебряков В.А., Тучкова Н.П.* Модель поиска схожих документов в семантической библиотеке // Научный сервис в сети Интернет: труды XXIII Всероссийской научной конференции (20–23 сентября 2021 г., онлайн). М.: ИПМ им. М.В. Келдыша, 2021. С. 54–64. https://doi.org/10.20948/abrau-2021-24

14. *Kanerva J., Kitti H., Chang L.-H., Vahtola T., Creutz M., Ginter F.* Semantic search as extractive paraphrase span detection // Lang Resources and Evaluation. 2024. https://doi.org/10.1007/s10579-023-09715-7

15. *Denghui Yang, Dengyun Zhu, Hailong Gai, Fucheng Wan*. Semantic Similarity Calculating based on BERT // Journal of Electrical Systems. 2024. Vol. 20, No. 2. P. 73–79.

16. *Kuang M. et al.* Multi-task Learning Based Keywords Weighted Siamese Model for Semantic Retrieval // In: Kashima H., Ide T., Peng W/C. (Eds.) Advances in Knowledge Discovery and Data Mining. PAKDD 2023. Lecture Notes in Computer Science. 2023. Vol. 13937. Springer, Cham. https://doi.org/10.1007/978-3-031-33380-4_7

17. Elibrary. URL: https://www.elibrary.ru/defaultx.asp (дата обращения 10.11.2024)

18. Sci-rus-tiny. URL: https://huggingface.co/mlsa-iai-msu-lab/sci-rus-tiny (дата обращения 10.09.2024)

19. ruSciBench — бенчмарк для оценки эмбеддингов научных текстов. URL: https://habr.com/ru/articles/781032/ (дата обращения 10.09.2024).

20. Научная электронная библиотека «КиберЛенинка». URL: https://cyberleninka.ru/ (дата обращения 10.09.2024)

21. Fine-tuning BERT for Semantic Textual Similarity with Transformers in Python. URL: https://thepythoncode.com/article/finetune-bert-for-semantic-textual-similarity-in-python (дата обращения 10.09.2024)

22. *Lawley C.J.M., Raimondo S., Chen T., Brin L., Zakharov A., Kur D., Hui J., Newton G., Burgoyne S.L., Marquis G.* Geoscience language models and their intrinsic evaluation // Applied Computing and Geosciences. 2022. Vol. 14, 100084. P. 1–10. https://doi.org/10.1016/j.acags.2022.100084

23. Cosine similarity. URL: https://en.wikipedia.org/wiki/Cosine_similarity (дата обращения 10.09.2024)

24. Rubert-tiny2. URL: https://huggingface.co/cointegrated/rubert-tiny2 (дата обращения 10.09.2024)

25. Multilingual-e5-base. URL: https://huggingface.co/intfloat/multilingual-e5-base (дата обращения 10.09.2024)

26. E5-large-en-ru. URL: https://huggingface.co/d0rj/e5-large-en-ru (дата обращения 10.09.2024)

27. Multilingual-e5-large.

URL: https://huggingface.co/intfloat/multilingual-e5-large (дата обращения 10.09.2024)
28.Тест производительности.
URL: https://ru.wikipedia.org/wiki/Тест_производительности (дата обращения 10.09.2024)

29. *Патук М.И., Наумова В.В., Ерёменко В.С.* Цифровой репозиторий "geologyscience.ru": открытый доступ к научным публикациям по геологии России // Электронные библиотеки. 2020. Т. 23, № 6. С. 1324–1338. https://doi.org/10.26907/1562-5419-2020-23-6-1324-1338

30. Geologyscience.ru – Виртуальный ассистент – Сервисы с использованием ИИ – Сервисы нейросетевого анализа геологических текстов. URL: https://service.geologyscience.ru/ (дата обращения 10.09.2024)

USING SEMANTIC SEARCH TO SELECT AND RANK GEOLOGICAL PUBLICATIONS

Mikhail I. Patuk[1] [0000-0003-3036-2275], Vera V. Naumova[2][0000-0002-3001-1638]

State Geological Museum named after Vladimir Vernadsky of RAS, Moscow

[1]patuk@mail.ru; [2]Naumova_new@mail.ru

Abstract

The aggregation of scientific information is essential for a comprehensive analysis of geological objects. This paper explores the potential and possibilities of semantic search to select thematically similar publications in the geological domain. Various language models are examined in the context of identifying similarities and differences in texts describing mineral deposits. After additional training of language models, a significant improvement in search results is demonstrated. Two web services are presented, based on a method for calculating the semantic similarity between texts and providing a quantitative assessment of their similarity.

***Keywords**:* artificial intelligence, machine learning, natural language processing, semantic search, geology.

REFERENCES

1. Semantic search. URL: https://en.wikipedia.org/wiki/Semantic_search (date of access 10.09.2024)

2. Your guide to the world of NLP (Natural Language Processing). URL: https://habr.com/ru/companies/otus/articles/705482/ (date of access 10.09.2024)

3. Semantic similarity. URL: https://en.wikipedia.org/wiki/Semantic_similarity (date of access 10.09.2024)

4. *Devlin J., Chang M.W., Lee K., Toutanova K.* Bert: pre-training of deep bidirectional transformers for language understanding. arxiv.org/abs/1810.04805 (2018).

5. We explain in simple terms what transformers are. URL: https://habr.com/ru/companies/mws/articles/770202/ (date of access 10.09.2024)

6. BERT (language model). URL: https://neerc.ifmo.ru/wiki/index.php?title=BERT_(языковая_модель) (date of access 10.11.2024)

7. BERT community. URL: https://huggingface.co/google-bert (date of access 10.09.2024)

8. *AkilaD., Jayakumar C.* Semantic Similarity- A Review of Approaches and Metrics // International Journal of Applied Engineering Research. 2014. Vol. 9, No. 24. P. 27581–27600.

9. TF-IDF. URL: https://en.wikipedia.org/wiki/Tf-idf (date of access 10.09.2024)

10. Okapi BM25. URL: https://en.wikipedia.org/wiki/Okapi_BM25 (date of access 10.09.2024)

11. Word2vec. URL: https://en.wikipedia.org/wiki/Word2vec (date of access 10.09.2024).

12. *Krasnov F., Smaznevich I., Baskakova E.* The problem of loss of solutions in the task of searching similar documents: Applying terminology in the construction of a corpus vector model// Business Informatics. 2021. Vol. 15. 60–74. https://doi.org/10.17323/2587-814X.2021.2.60.74.

13. *Ataeva O.M., Serebryakov V.A., Tuchkova N.P.* Search model for similar documents in the semantic library // Scientific service in Internet: Proceedings of the XXIII All-Russian Scientific Conference (20–23 September 2021). M.: Keldysh Institute of Applied Mashematics, 2021. P. 54–64 (in Russian). https://doi.org/10.20948/abrau-2021-24

14. *Kanerva J., Kitti H., Chang L.-H., Vahtola T., Creutz M., Ginter F.* Semantic search as extractive paraphrase span detection // Lang Resources and Evaluation. 2024. https://doi.org/10.1007/s10579-023-09715-7

15. *Denghui Yang, Dengyun Zhu, Hailong Gai, Fucheng Wan.* Semantic Similarity Caculating based on BERT // Journal of Electrical Systems. 2024. Vol. 20, No. 2. P. 73–79.

16. *Kuang M. et al.* Multi-task Learning Based Keywords Weighted Siamese Model for Semantic Retrieval // In: Kashima H., Ide T., Peng W/C. (Eds.) Advances in Knowledge Discovery and Data Mining. PAKDD 2023. Lecture Notes in Computer Science. 2023. Vol. 13937. Springer, Cham. https://doi.org/10.1007/978-3-031-33380-4_7

17. Elibrary. URL: https://www.elibrary.ru/defaultx.asp (date of access 10.09.2024).

18. Sci-rus-tiny. URL: https://huggingface.co/mlsa-iai-msu-lab/sci-rus-tiny (date of access 10.09.2024).

19. ruSciBench — A benchmark for evaluating the quality of embeddings for scientific texts. URL: https://habr.com/ru/articles/781032/ (date of access 10.09.2024)

20. Scientific electronic Library «CyberLeninka». https://cyberleninka.ru/ (date of access 10.09.2024).

21. Fine-tuning BERT for Semantic Textual Similarity with Transformers in Python.

URL: https://thepythoncode.com/article/finetune-bert-for-semantic-textual-similarity-in-python (date of access 10.09.2024).

22. *Lawley C.J.M., Raimondo S., Chen T., Brin L., Zakharov A., Kur D., Hui J., Newton G., Burgoyne S.L., Marquis G.* Geoscience language models and their intrinsic evaluation // Applied Computing and Geosciences. 2022. Vol. 14. 100084. P. 1–10. https://doi.org/10.1016/j.acags.2022.100084

23. Cosine similarity. URL: https://en.wikipedia.org/wiki/Cosine_similarity (date of access 10.09.2024).

24. Rubert-tiny2. URL: https://huggingface.co/cointegrated/rubert-tiny2c (date of access 10.09.2024)

25. Multilingual-e5-base.
URL: https://huggingface.co/intfloat/multilingual-e5-base (date of access 10.09.2024).

26. E5-large-en-ru. URL: https://huggingface.co/d0rj/e5-large-en-ru (date of access 10.09.2024).

27. Multilingual-e5-large.
URL: https://huggingface.co/intfloat/multilingual-e5-large (date of access 10.09.2024).

28. Benchmark (computing).
URL: https://en.wikipedia.org/wiki/Benchmark_ (computing) (date of access 10.09.2024).

29. *Patuk M.I., Naumova V.V., Eryomenko V.S.* Digital repository "geology-science.ru": open access to scientific publications on russian geology // Russian Digital Library Journal. 2020. Vol. 23, no. 6. P. 1324–1338 (in Russian).
https://doi.org/10.26907/1562-5419-2020-23-6-1324-1338

30. Geologyscience.ru – Virtual Assistant – Services using AI – Neural network analysis services for geological texts, https://service.geologyscience.ru/ (date of access 10.09.2024).

СВЕДЕНИЯ ОБ АВТОРАХ

ПАТУК Михаил Иванович – к. г.-м. н., н. с., научный отдел Государственного геологического музея им. В.И. Вернадского РАН, Москва, Россия.

Michail I. PATUK – PhD, scientific department, Vernadsky State Geological Museum RAS, Moscow (Russia).

Email: patuk@mail.ru

ORCID: 0000-0003-3036-2275

НАУМОВА Вера Викторовна – д. г.-м. н., г. н. с., зав. Научным отделом Государственного геологического музея им. В.И. Вернадского РАН, Москва, Россия.

Vera V. NAUMOVA – Prof., head of SGM scientific department, Vernadsky State Geological Museum RAS, Moscow (Russia).

Email: naumova_new@mail.ru

ORCID: 0000-0002-3001-1638

Материал поступил в редакцию 24 сентября 2024 года