BiTe-REx: An Explainable Bilingual Text Retrieval System in the Automotive Domain

Viju Sudhi
Sabine Wehnert*
viju.sudhi@st.ovgu.de
sabine.wehnert@ovgu.de
Otto von Guericke University
Magdeburg, Germany

Norbert Michael Homner
Sebastian Ernst
Mark Gonter
<firstname>.<lastname>@audi.de
Audi AG
Ingolstadt, Germany

Andreas Krug
Ernesto William De Luca*
andreas.krug@ovgu.de
ernesto.deluca@ovgu.de
Otto von Guericke University
Magdeburg, Germany

ABSTRACT

To satiate the comprehensive information need of users, retrieval systems surpassing the boundaries of language are inevitable in the present digital space in the wake of an ever-rising multilingualism. This work presents the first-of-its-kind **Bi**lingual **Te**xt **R**etrieval **Ex**planations (BiTe-REx) aimed at users performing competitor or wage analysis in the automotive domain. BiTe-REx supports users to gather a more comprehensive picture of their query by retrieving results regardless of the query language and enables them to make a more informed decision by exposing how the underlying model judges the relevance of documents. With a user study, we demonstrate statistically significant results on the understandability and helpfulness of the explanations provided by the system.

CCS CONCEPTS

• **Information systems** → **Multilingual and cross-lingual retrieval**; • **Computing methodologies** → *Natural language processing*.

KEYWORDS

bilingual text retrieval, static and contextual embeddings, explainable artificial intelligence

ACM Reference Format:
Viju Sudhi, Sabine Wehnert, Norbert Michael Homner, Sebastian Ernst, Mark Gonter, Andreas Krug, and Ernesto William De Luca. 2022. BiTe-REx: An Explainable Bilingual Text Retrieval System in the Automotive Domain. In *Proceedings of the 45th International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR '22), July 11–15, 2022, Madrid, Spain.* ACM, New York, NY, USA, 5 pages. https://doi.org/10.1145/3477495.3531665

*Also with Leibniz Institute for Educational Media | Georg Eckert Institute

SIGIR '22, July 11–15, 2022, Madrid, Spain

ACM ISBN 978-1-4503-8732-3/22/07…$15.00
https://doi.org/10.1145/3477495.3531665

1 INTRODUCTION

The information need of a user is usually expressed in their mother tongue although they can speak and search for information in different languages. When they pose queries in different languages even when the queries semantically mean the same, common search engines yield a different number and different order of relevant search results. This points towards the direction of *the digital language divide*[1] which rightly asserts that "the internet, far from our common notion of it being infinite, is just as big as the language we speak". The issue becomes more significant in a domain-specific professional search engine where the user searches for relevant information in a huge number of documents with less regard to the query language. This calls for the need of multilingual retrieval systems (inherently harder than cross-lingual systems [23]).

Nowadays, we can train complex language models (with over billions of parameters) to learn different languages from large corpora of text [28] with extended computational resources [4]. These models preform relatively better at various downstream tasks, however, at the cost of their explainability to the end user. Recently, with transparency identified as a key factor in labelling an Artificial Intelligence system trustworthy [5], more effort has been put into explaining black-box models to build users' trust [18].

Luo et al. [15] observed that there has been significant work ongoing in Natural Language Processing in the field of explaining model predictions to the user - particularly in question-answering systems, sentiment analysis and text classification tasks [1, 7, 8, 14, 16, 24, 25]. However, Verma and Ganguly [26] argued that there has been relatively less consensus in explaining an information retrieval system. It is even more intriguing to note that explaining multilingual retrieval has not been particularly explored yet.

We present BiTe-REx[2], which to the best of our knowledge is the first of its kind to explain a domain-specific bilingual retrieval task. The system aids users in the automotive domain to perform an effective competitor or wage analysis. For a given query, the users are presented relevant bilingual results from various automotive web pages. This enables them to comprehensively understand the new business ideas and strategies being adopted in the competing organizations and also benefits in computing the overall cost of a project. The explanations presented by the system further help the users to make a more informed decision on these tasks empowering them not to blindly accept the retrieved search results.

[1] http://labs.theguardian.com/digital-language-divide/
[2] available at https://github.com/vijusudhi/bite-rex

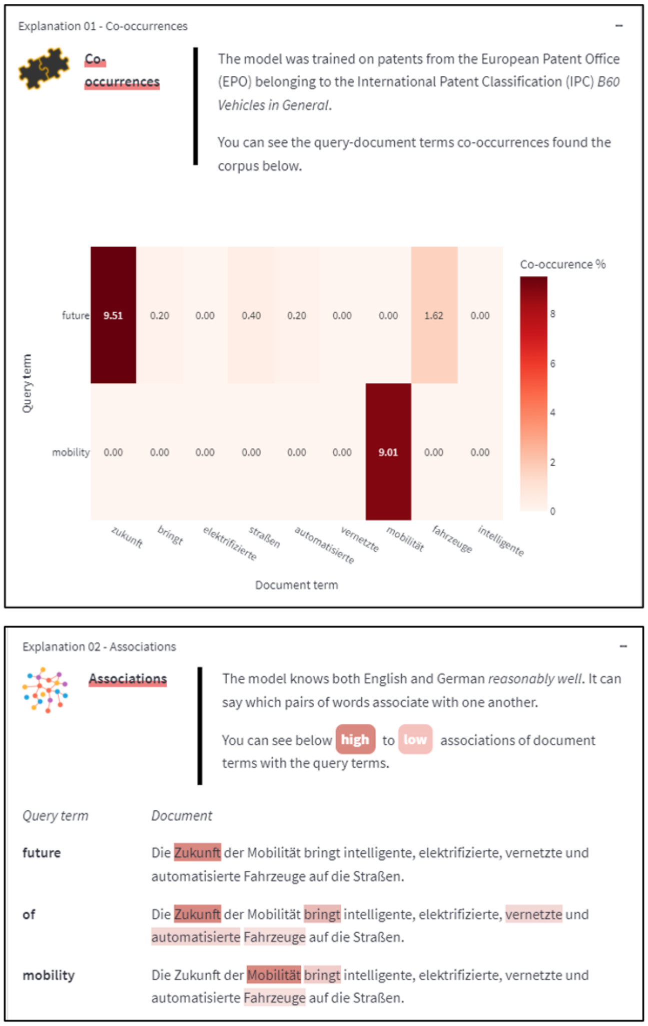


Figure 1: Exp01 and Exp02 for a German document retrieved for the English query *"future of mobility"*

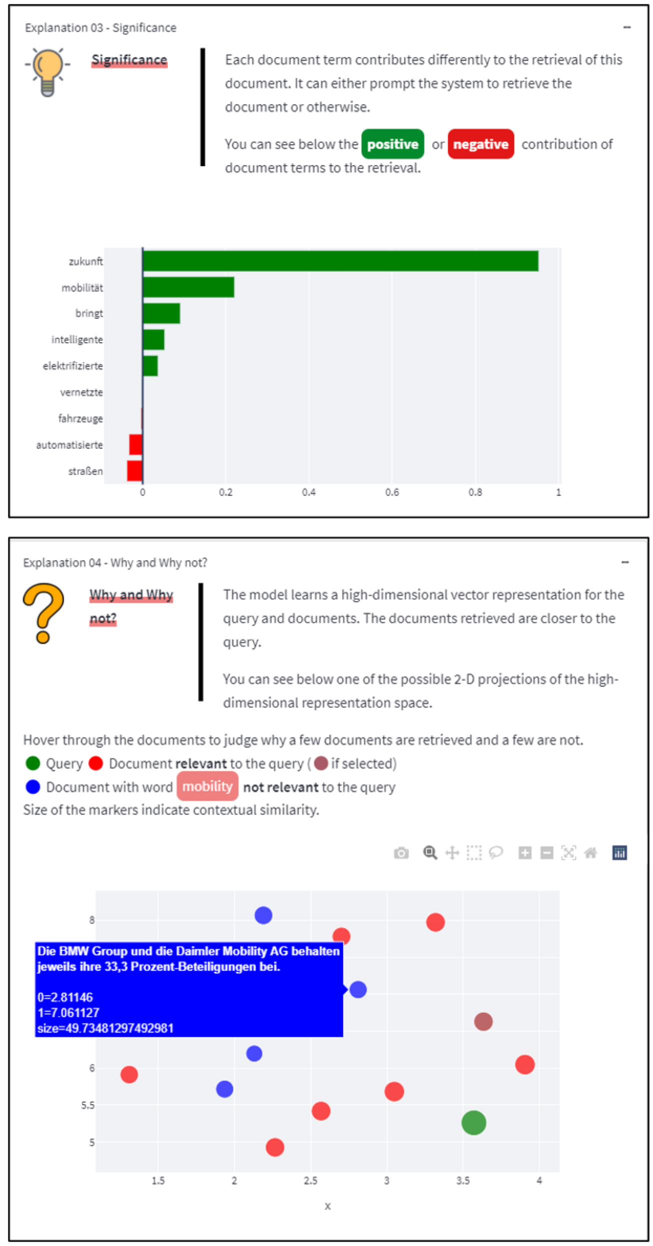


Figure 2: Exp03 and Exp04 for a German document retrieved for the English query *"future of mobility"*

2 RELATED WORK

To explain neural information retrieval (IR) models, Singh and Anand [22] extended the idea of Local Interpretable Model-Agnostic Explanations (LIME) [20], which explains black-box model predictions by approximating the complex model to a more interpretable linear one. They proposed techniques to convert the ranking scores obtained from a neural IR model to probabilities and then feed them to LIME for explanations. Verma and Ganguly [26] noted that the words sampled for producing local explanations affect the quality of the explanations and suggested sampling methods including uniform, biased and masked sampling to train the explanation model.

Sen et al. [21] approximated the frequency of a term in the document and the corpus, and the length of the document to explain why different documents are ranked differently by different ranking models. Chios and Verberne [3] exploited the gating network of the neural IR Deep Relevance Matching Model [10] to explain the importance of query words. Fernando et al. [9] tried explaining IR systems with the help of a model-specific approach DeepSHAP [13], which computes shapley values as the contribution of an input feature, and found the explanations to be considerably different from those that LIME generated.

Almost all of these works consider monolingual retrieval and to the best of our knowledge, there is none that explores the bilingual premise. We address this gap by explaining why a set of bilingual documents was retrieved regardless of the query language.

3 SYSTEM OVERVIEW

The system consists of three modules - (i) an Embedder encoding text into a dense vector representation, (ii) a Retriever taking a user query as input and returning the relevant documents and (iii) an Explainer explaining to the user why a document was retrieved.

3.1 Embedder

The Embedder transforms text from natural language to dense vector representations. Though static embeddings [2] capture word semantics by exploiting the distributional hypothesis, they fail in handling polysemous words [27] where contextual embeddings [6] could help. Hereby, we train two models (a static and a contextual model) to bring the vector representation of words in English and German into a common representation space.

Data: To build automotive domain-specific models, we used patents from the European Patent Office[3] with the International Patent Classification *B60 Vehicles in General* since 2010. Since not all of the content is sentence-aligned, we only used the 365k English-German parallel sentences from the claims of the patents.

Static Embeddings: Monolingual static embeddings in English and German were trained individually with fastText[4] [2] and then aligned in a common representation space with unsupervised learning[5] [12]. This model is further denoted as *STAT*.

Contextual Embeddings: We also further pre-trained Multilingual BERT [6] with Masked Language Modelling and Next Sentence Prediction (NSP) tasks for 2 epochs. In NSP, the model was trained to predict the label *IsNext* if it sees the first sentence as either an English or German claim and the second sentence as its parallel claim in the other language. The training data had a total of 732K instances. This model is further denoted as *CONT*.

3.2 Retriever

The Retriever takes a user query in either English or German as input and with the embeddings from the Embedder returns relevant documents from the bilingual collection of documents scraped from different automotive web pages. To judge the relevance of documents, we adopt different strategies for the different models. While using *STAT*, cosine similarity[6] is computed between the query and document vectors. The documents with higher similarity scores are then returned as relevant. When using *CONT*, the query is passed as the first sentence and documents are passed as second sentences to the model. The prediction probabilities from the NSP are then used to judge the relevance of the documents. Initially, we treat *STAT* and *CONT* scores individually to retrieve the relevant documents. However, as discussed later in Section 4.1, to improve the retrieval performance we use these scores sequentially.

3.3 Explainer

The Explainer presents to the user why a document (or, a set of documents) was retrieved by the Retriever module for a given query. The explanations provide the user with a detailed view into different retrieval aspects, from the statistics of the underlying training corpus to the final relevance judgement of the Retriever module.

Query-Document Term(s) Co-occurrences (Exp01): It is intuitive to reason about the model behaviour using the underlying training data. We therefore present training corpus co-occurrences of query and document terms to the user, often omitted by related work. The co-occurrence of a query term t_q with a document term t_d is computed according to Eqn. 1, where $|s_x|$ is the number of sentences with the term x in the set of concatenated parallel sentences in the training corpus.

$$co-occurrence(t_q, t_d) = \frac{|s_{t_q,t_d}|}{|s_{t_q}|} \cdot 100 \tag{1}$$

Query-Document Term(s) Associations (Exp02): The model can learn associations between words from the co-occurrences in the corpus which serve as a reflection of how well the model captures the multilingual semantics of the training corpus. The association of a query term t_q with a document term t_d is computed as per Eqn. 2, where $\vec{x}$ is the vector representation of the term x, $d_{\neg x}$ is the perturbed document without the term x and $sim(\vec{a},\vec{b})$ is the cosine similarity between the vectors $\vec{a}$ and $\vec{b}$.

$$association(t_q, t_d) = \frac{sim(\vec{t_q},\vec{d}) - sim(\vec{t_q},\vec{d_{\neg t_d}})}{sim(\vec{t_q},\vec{d})} \cdot 100 \tag{2}$$

Document Terms Significance (Exp03): Once the user is convinced of how the model associates query and document terms with each other, we accentuate its effect on the retrieval premise. Given a query, the model should discriminate words in the document that contribute positively and negatively to the retrieval of the document. The significance of a document term t_d with the query q is computed according to Eqn. 3, where $sim(\vec{a},\vec{b})$ is the cosine similarity just as in Eqn. 2.

$$significance(q, t_d) = \frac{sim(\vec{q},\vec{d}) - sim(\vec{q},\vec{d_{\neg t_d}})}{sim(\vec{q},\vec{d})} \cdot 100 \tag{3}$$

Visualization of Representation Space (Exp04): Finally, we present a 2D-projected common representation space (using Uniform Manifold Approximation and Projection (UMAP) [17]) of the high-dimensional query and document vectors to the users. We present the retrieved documents along with the documents which have words similar to those in the query but are not retrieved. Such common spaces expose the contextual differences, otherwise unseen, between the documents. The user is given the option to hover over the documents to draw this judgement by himself.

As illustrated in Fig. 1 and Fig. 2, Query-Document Term(s) Associations and Document Term Significance can be considered as *local* explanations while the Representation Space presents *global* explanation to the retrieval premise. The agreement between associations and significance can help the users to evaluate the underlying model both on intrinsic word representations learned by the model and the extrinsic task of retrieving bilingual documents. They help the users to draw analogies between words in a language familiar to them (presumably the query language) with the foreign words (in the retrieved documents). The users in light of the presented

[3]https://www.epo.org/searching-for-patents/data/bulk-data-sets/text-analytics
[4]https://radimrehurek.com/gensim/models/fasttext.html
[5]https://github.com/facebookresearch/MUSE
[6]using sklearn.metrics.pairwise.cosine_similarity

explanations can make a more informed decision on accepting or rejecting the retrieved results by judging the factors that contributed to the retrieval. The proposed system can also be used in conjunction with other language models since the explanations are largely *model-agnostic*.

4 EVALUATION

A two-phased evaluation is carried out to evaluate the performances of Retriever and Explainer modules and then to understand user preferences and expectations while using an explainable search system. As mentioned earlier in Section 2, we have no other explainable bilingual IR models to compare our results with.

4.1 Evaluation of the individual modules

As part of evaluating the individual modules, we initially conducted an annotation activity with over 25 users.

Relevance of results returned by the Retriever: The activity involved the users to mark a document retrieved by the Retriever as relevant to the presented query or as irrelevant otherwise. The users marked 67% of the results retrieved by *STAT* as relevant while only 60% of the results retrieved by *CONT* as relevant.

Significance of explanations returned by the Explainer: Additionally, if the user marked a document relevant in the activity, they were asked to select the words in the document that they think are the most significant to the query. The user-annotated significant words were considered as the *relevant* words and the words that the Explainer returned as the most significant were considered as the *retrieved* words. This helped us employ the standard evaluation metrics for IR [11, 19] to evaluate the significance of explanations. Table 1 shows the macro average of these computed scores. As different users annotate different words as significant, we also propose the metric *completeness* which takes into account the variances in user annotations.

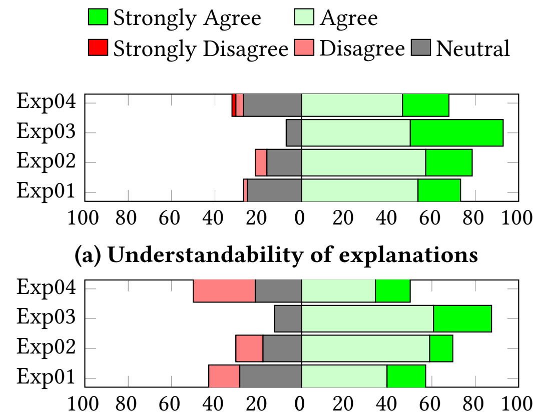
To compute this metric, we determine the set of all words annotated by the users, $W_{user,q}$ for each query q. Now, we determine the factor $factor_w$ as the relative number of users who marked the word w significant. With these values, the completeness of explanations of each query is computed according to Eqn. 4, where $comp_w = 1$, if the word w is in $W_{user,q}$ and $comp_w = 0$, otherwise.

$$completeness_q = \frac{\sum_w factor_w \cdot comp_w}{\sum_w factor_w} \tag{4}$$

Table 1: Significance of explanations

Model	Precision	Recall	F1-score	Completeness
STAT	**0.408**	**0.335**	**0.353**	**0.647**
CONT	0.332	0.263	0.272	0.574

It is evident that *STAT* outperforms *CONT* in both the relevance of retrieved documents and significance of explanations. However, to make the modules more robust in handling queries which are different from the nature of queries in the above evaluation (30 pairs of bilingual queries), we set the Retriever to initially retrieve 500 documents with *STAT* and then re-rank them using *CONT*.



(a) Understandability of explanations

(b) Helpfulness of explanations

Figure 3: User responses for usefulness and helpfulness of explanations presented by the system. Exp01: Query-Document Term(s) Co-occurrences, Exp02: Query-Document Term(s) Associations, Exp03: Document Terms Significance, Exp04: Visualizing the representation space.

4.2 Evaluation of the user interface

To evaluate user preferences while using explainable search interfaces, an extended anonymous user study was carried out. The users were not involved in the design process of the evaluation and thereby we ensured a fair and unbiased evaluation. From a set of three English and three German queries, the 15 participants (92% proficient in English and 30% proficient in German) selected one query for each language and were asked to search for the selected queries in a regular search engine and then in our system.

Comparison with regular search engines: The users marked that an English query yields 77% English search results and a German query yields 89% German search results in the first result page, while using a regular search engine. This is, however, not the case with our system which gives the same number of results in both languages regardless of the query language. A Related-Samples Wilcoxon Signed Rank Test (WSR) test showed that the users agree (p<0.001, $\alpha = 0.05$) that our system explains the retrieval premise better than a regular search engine.

Understandability and Helpfulness of explanations: The users were also asked to rate (i) how understandable are the explanations aided by the visualizations and descriptions (as shown in Fig. 1 and Fig. 2) and (ii) how helpful are the explanations to perceive the retrieval of the document. As illustrated in Fig. 3, most of the users tend to agree that the explanations provided by the system are easily comprehensible and helpful in identifying the factors affecting the retrieval. The users rated an average of 4.01 out of 5 (σ=0.76) for understandability and 3.73 out of 5 (σ=0.90) for helpfulness. With a One-Sample WSR test for understandability (η=4, p=0.496, α=0.05) and helpfulness (η=3.90, p=0.113, α=0.05), we conclude the results are statistically significant.

Across the evaluation phases, we observed that the users from the automotive domain found the proposed explanations more understandable and helpful than the other users with a background in information retrieval. It is also worth noting that most of the users preferred to see the explanation *Document Terms Significance* further in search systems, while very less of them preferred the explanation *Visualization of representation space*. To conclude the evaluation, the users unanimously responded that an explainable search system *improves* the user's trust in the underlying system.

5 CONCLUSION

In this work, we presented an explainable bilingual text retrieval system focusing on users in the automotive domain performing competitor or wage analysis. The system makes use of both static and contextual embeddings to retrieve the relevant documents given user queries. We presented that the users and our system identified similar significant words in the documents. We also showed that the system yields statistically significant ratings on both the understandability and helpfulness facets by conducting an extended user study. In the future, the system can be extended to other niche domains and to the broader spectrum of multilingual retrieval involving different languages. We also expect to investigate the benefits of the proposed explanations with extended qualitative evaluations in the future.

ACKNOWLEDGMENTS

This work is supported by Audi AG and the research project on 'Data Driven Business Models'. We also thank all the participants of our evaluation activities.

REFERENCES

[1] Dominika Basaj, Barbara Rychalska, Przemyslaw Biecek, and Anna Wróblewska. 2018. How much should you ask? On the question structure in QA systems. *CoRR* abs/1809.03734 (2018). arXiv:1809.03734 http://arxiv.org/abs/1809.03734

[2] Piotr Bojanowski, Edouard Grave, Armand Joulin, and Tomás Mikolov. 2017. Enriching Word Vectors with Subword Information. *Trans. Assoc. Comput. Linguistics* 5 (2017), 135–146. https://transacl.org/ojs/index.php/tacl/article/view/999

[3] Ioannis Chios and Suzan Verberne. 2021. Helping results assessment by adding explainable elements to the deep relevance matching model. *CoRR* abs/2106.05147 (2021). arXiv:2106.05147 https://arxiv.org/abs/2106.05147

[4] Jaeho Choi, Seunghyeok Oh, and Joongheon Kim. 2020. The Useful Quantum Computing Techniques for Artificial Intelligence Engineers. In *2020 International Conference on Information Networking, ICOIN 2020, Barcelona, Spain, January 7-10, 2020.* IEEE, 1–3. https://doi.org/10.1109/ICOIN48656.2020.9016555

[5] European Commission. 2019. *Artificial Intelligence: Commission Takes Forward its Work on Ethics Guidelines.* http://europa.eu/rapid/press-release_IP-19-1893_en.htm

[6] Jacob Devlin, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova. 2019. BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. In *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics:*, Jill Burstein, Christy Doran, and Thamar Solorio (Eds.). Association for Computational Linguistics, 4171–4186. https://doi.org/10.18653/v1/n19-1423

[7] Yanzhuo Ding, Yang Liu, Huanbo Luan, and Maosong Sun. 2017. Visualizing and Understanding Neural Machine Translation. In *Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics, ACL 2017, Vancouver, Canada, July 30 - August 4, Volume 1: Long Papers.* Association for Computational Linguistics, 1150–1159. https://doi.org/10.18653/v1/P17-1106

[8] Shi Feng, Eric Wallace, Alvin Grissom II, Mohit Iyyer, Pedro Rodriguez, and Jordan Boyd-Graber. 2018. Pathologies of neural models make interpretations difficult. *arXiv preprint arXiv:1804.07781* (2018).

[9] Zeon Trevor Fernando, Jaspreet Singh, and Avishek Anand. 2019. A study on the Interpretability of Neural Retrieval Models using DeepSHAP. In *Proceedings of the 42nd International ACM SIGIR Conference on Research and Development in Information Retrieval.* ACM, 1005–1008. https://doi.org/10.1145/3331184.3331312

[10] Jiafeng Guo, Yixing Fan, Qingyao Ai, and W. Bruce Croft. 2016. A Deep Relevance Matching Model for Ad-hoc Retrieval. In *Proceedings of the 25th ACM International Conference on Information and Knowledge Management, CIKM 2016, Indianapolis, IN, USA, October 24-28, 2016.* ACM, 55–64. https://doi.org/10.1145/2983323.2983769

[11] Raju Korra, Pothula Sujatha, Sidige Chetana, and Madarapu Naresh Kumar. 2011. Performance evaluation of multilingual information retrieval (MLIR) system over information retrieval (IR) system. In *2011 International Conference on Recent Trends in Information Technology (ICRTIT).* IEEE, 722–727.

[12] Guillaume Lample, Alexis Conneau, Ludovic Denoyer, and Marc'Aurelio Ranzato. 2018. Unsupervised Machine Translation Using Monolingual Corpora Only. In *6th International Conference on Learning Representations, ICLR 2018, Vancouver, BC, Canada, April 30 - May 3, 2018, Conference Track Proceedings.* OpenReview.net. https://openreview.net/forum?id=rkYTTf-AZ

[13] Scott M. Lundberg and Su-In Lee. 2017. A Unified Approach to Interpreting Model Predictions. In *Advances in Neural Information Processing Systems 30: Annual Conference on Neural Information Processing Systems 2017*, Isabelle Guyon, Ulrike von Luxburg, Samy Bengio, Hanna M. Wallach, Rob Fergus, S. V. N. Vishwanathan, and Roman Garnett (Eds.). 4765–4774. https://proceedings.neurips.cc/paper/2017/hash/8a20a8621978632d76c43dfd28b67767-Abstract.html

[14] Ling Luo et al. 2018. Beyond Polarity: Interpretable Financial Sentiment Analysis with Hierarchical Query-driven Attention. In *Proceedings of the Twenty-Seventh International Joint Conference on Artificial Intelligence, IJCAI 2018, July 13-19, 2018, Stockholm, Sweden*, Jérôme Lang (Ed.). ijcai.org, 4244–4250. https://doi.org/10.24963/ijcai.2018/590

[15] Siwen Luo, Hamish Ivison, Soyeon Caren Han, and Josiah Poon. 2021. Local Interpretations for Explainable Natural Language Processing: A Survey. *CoRR* abs/2103.11072 (2021). arXiv:2103.11072 https://arxiv.org/abs/2103.11072

[16] Qianren Mao, Jianxin Li, Senzhang Wang, Yuanning Zhang, Hao Peng, Min He, and Lihong Wang. 2019. Aspect-Based Sentiment Classification with Attentive Neural Turing Machines. In *Proceedings of the Twenty-Eighth International Joint Conference on Artificial Intelligence, IJCAI 2019, Macao, China, August 10-16, 2019*, Sarit Kraus (Ed.). ijcai.org, 5139–5145. https://doi.org/10.24963/ijcai.2019/714

[17] Leland McInnes and John Healy. 2018. UMAP: Uniform Manifold Approximation and Projection for Dimension Reduction. *CoRR* abs/1802.03426 (2018). arXiv:1802.03426 http://arxiv.org/abs/1802.03426

[18] Pradeepta Mishra. 2022. Explainability for NLP. In *Practical Explainable AI Using Python.* Springer, 193–227.

[19] Carol Peters, Martin Braschler, and Paul D. Clough. 2012. *Multilingual Information Retrieval - From Research To Practice.* Springer. https://doi.org/10.1007/978-3-642-23008-0

[20] Marco Túlio Ribeiro, Sameer Singh, and Carlos Guestrin. 2016. "Why Should I Trust You?": Explaining the Predictions of Any Classifier. In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, San Francisco, CA, USA, August 13-17, 2016.* ACM, 1135–1144. https://doi.org/10.1145/2939672.2939778

[21] Procheta Sen, Debasis Ganguly, Manisha Verma, and Gareth JF Jones. 2020. The Curious Case of IR Explainability: Explaining Document Scores within and across Ranking Models. In *Proceedings of the 43rd International ACM SIGIR Conference on Research and Development in Information Retrieval*, Jimmy Huang, Yi Chang, Xueqi Cheng, Jaap Kamps, Vanessa Murdock, Ji-Rong Wen, and Yiqun Liu (Eds.). ACM, 2069–2072. https://doi.org/10.1145/3397271.3401286

[22] Jaspreet Singh and Avishek Anand. 2019. EXS: Explainable Search Using Local Model Agnostic Interpretability. In *Proceedings of the Twelfth ACM International Conference on Web Search and Data Mining, WSDM 2019, Melbourne, VIC, Australia, February 11-15, 2019.* ACM, 770–773. https://doi.org/10.1145/3289600.3290620

[23] Philipp Sorg and Philipp Cimiano. 2012. Exploiting Wikipedia for cross-lingual and multilingual information retrieval. *Data Knowl. Eng.* 74 (2012), 26–45. https://doi.org/10.1016/j.datak.2012.02.003

[24] Alona Sydorova, Nina Pörner, and Benjamin Roth. 2019. Interpretable Question Answering on Knowledge Bases and Text. In *Proceedings of the 57th Conference of the Association for Computational Linguistics, ACL 2019, Florence, Italy, July 28-August 2, 2019, Volume 1: Long Papers.* Association for Computational Linguistics, 4943–4951. https://doi.org/10.18653/v1/p19-1488

[25] Ming Tu et al. 2020. Select, Answer and Explain: Interpretable Multi-Hop Reading Comprehension over Multiple Documents. In *The Thirty-Fourth AAAI Conference on Artificial Intelligence, AAAI 2020, The Thirty-Second Innovative Applications of Artificial Intelligence Conference, IAAI 2020, The Tenth AAAI Symposium on Educational Advances in Artificial Intelligence, EAAI 2020, New York, NY, USA, February 7-12, 2020.* AAAI Press, 9073–9080. https://aaai.org/ojs/index.php/AAAI/article/view/6441

[26] Manisha Verma and Debasis Ganguly. 2019. LIRME: Locally Interpretable Ranking Model Explanation. In *Proceedings of the 42nd International ACM SIGIR Conference on Research and Development in Information Retrieval, SIGIR 2019, Paris, France, July 21-25, 2019*, Benjamin Piwowarski, Max Chevalier, Éric Gaussier, Yoelle Maarek, Jian-Yun Nie, and Falk Scholer (Eds.). ACM, 1281–1284. https://doi.org/10.1145/3331184.3331377

[27] Yuxuan Wang, Yutai Hou, Wanxiang Che, and Ting Liu. 2020. From static to dynamic word representations: a survey. *Int. J. Mach. Learn. Cybern.* 11, 7 (2020), 1611–1630. https://doi.org/10.1007/s13042-020-01069-8

[28] Abdulrahman Yarali. 2022. *Big Data and Artificial Intelligence.* 299–326. https://doi.org/10.1002/9781119685265.ch17