Title: Cross-Modal Text Representation for Ranking

In recent years, there has been a growing interest in developing effective methods for ranking text documents. Traditional approaches mainly focus on using textual features to represent and rank documents. However, with the increasing availability of multimodal data, there is a need to explore cross-modal text representation techniques that can leverage information from different modalities to improve ranking performance.

One approach to cross-modal text representation for ranking is the use of machine translation to pair original English queries with relevant and non-relevant passages in non-English languages [REF0]. By training models on these newly constructed triples, researchers have been able to achieve improved retrieval performance in a zero-shot setting [REF0]. This approach allows for the retrieval of content in non-English languages using English queries, which is particularly useful in scenarios where users may not be proficient in the target language.

Another strategy for cross-modal text representation involves assessing style strength and content preservation using a series of classifiers [REF1]. These classifiers are trained to evaluate the style and content of transformed text, ensuring that the transformed text remains compatible with the original text [REF1]. By incorporating these classifiers into the ranking process, researchers have been able to achieve more comprehensive evaluations of text representations.

Memes, which combine text and images, have gained significant popularity on the internet [REF2]. Detecting and analyzing memes can provide valuable insights into users' intentions and opinions. However, identifying memes from non-memes and understanding their underlying meanings pose challenges in online marketing campaigns. Researchers have been exploring techniques to automatically identify and analyze memes, contributing to the automated identification of opinions related to specific user groups [REF2].

Concreteness and abstractness are important factors in understanding the meaning of words and identifying linguistic metaphors [REF3]. Researchers have developed models that compute concreteness indexes for words in different languages, allowing for cross-lingual metaphor identification [REF3]. By training classifiers on abstract and concrete words, these models can determine the degree of concreteness of a word, facilitating metaphor identification in various languages [REF3].

In the context of document ranking, the selection and labeling of relevant words play a crucial role in improving retrieval performance [REF4]. Researchers have explored different approaches, such as selecting words based on their lemma form or using specific layers of pre-trained models like BERT [REF4]. By carefully selecting and labeling words, researchers have been able to enhance the effectiveness of document ranking systems.

While there have been advancements in cross-modal text representation for ranking, there are still areas that require further exploration. For instance, there is a need to investigate the application of similar techniques to other tasks and adapt advanced neural models to information retrieval [REF5]. Additionally, researchers are encouraged to extend their work by incorporating a wide range of other information retrieval cues [REF5].

In conclusion, cross-modal text representation techniques have shown promise in improving ranking performance by leveraging information from different modalities. By incorporating machine translation, classifiers, and other innovative approaches, researchers have been able to enhance the effectiveness of document retrieval systems. However, there is still room for further exploration and improvement in this field.

References given to GPT:

[REF0] - paperID: d1ccffb8eb1b7a99cd586723074b82fa5399bdd2 Title: Transfer Learning Approaches for Building Cross-Language Dense Retrieval Models Chunk of text: Since our focus here is using English queries to retrieve content in non-English languages, we pair the original English queries with machine translations of relevant and non-relevant MS MARCO passages to form new triples.4 We then train 4 If we had wanted to experiment with using non-English queries to find English content, we could have instead translated only the MS MARCO queries.6 S. Nair et al. Table 1: Test collection statistics for the CLEF and HC4 newswire collections. Collection HC4 HC4 CLEF CLEF CLEF CLEF CLEF Chinese Persian French German Italian Russian Spanish #documents 646K 486K 129k 294k 157k 16k 454k #passages 3.6M 3.1M 0.7M 1.6M 0.8M 0.1M 2.7M #queries 50 50 200 200 200 62 160 ColBERT-X on these newly constructed triples in the same manner as ColBERT. Figure 1 shows these two pipelines. The key difference is that in the zero-shot setting we have a single ColBERT-X model for a given query language (in this case English) that is used for retrieval in multiple document languages.

[REF1] - paperID: 336e531a59cafbe215b950fd749bca866b89cea0 Title: SNK @ DANKMEMES: Leveraging Pretrained Embeddings for Multimodal Meme Detection (short paper) Chunk of text: For our task, we propose a fully automatic strategy based on a series of classifiers to assess style strength and content preservation. For style, we train a single classifier (main). For content, we train two classifiers that perform two ‘sanity checks’: one ensures that the two headlines (original and transformed) are still compatible (HH classifier); the other ensures that the headline is still compatible with the original article (AH classifier). See also Figure 1b. In what follows we describe these classifiers in237 (a) Overall data splits EVALUATION train & test main R+A3+A1 HH A1 + random pairs AH R+A3+A1 TASK train R+A3 test A2 (b) Training/test sets Figure 1: Data splits and their use in the different training sets more detail. When discussing baseline results, we will show how the contribution of each classifier is crucial towards a comprehensive evaluation. Main classifier The main classifier uses a pretrained BERT (Devlin et al., 2019) encoder with a linear classifier on top fine-tuned with a batch size of 256 and sequences truncated at 32 tokens for 6 epochs with learning rate 1e-05.

[REF2] - paperID: 336e531a59cafbe215b950fd749bca866b89cea0 Title: SNK @ DANKMEMES: Leveraging Pretrained Embeddings for Multimodal Meme Detection (short paper) Chunk of text: Copyright © 2020 for this paper by its authors. Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0). information. As the Internet and the online social interactions evolved, certain image templates emerged and gained global popularity, contributing to a de facto standardization of joint textimage usage, and thus leading to the creation of memes. Memes can be humorous, satirical, offensive, or hateful, therefore encapsulating a wide range of emotions and beliefs. Properly identifying memes from non-memes, and then analyzing them to detect the users’ intentions is becoming a stringent task in online marketing campaigns by targeting the automated identification of opinions pertaining to certain groups of users. The DANKMEMES competition from EVALITA 2020

[REF3] - paperID: 336e531a59cafbe215b950fd749bca866b89cea0 Title: SNK @ DANKMEMES: Leveraging Pretrained Embeddings for Multimodal Meme Detection (short paper) Chunk of text: The paradigm words are automatically selected from the MRC Psycholinguistic Database Machine Usable Dictionary (Coltheart, 1981), a collection of 4,295 English words rated with degrees of abstractness by human subjects in psycholinguistic experiments. Tsvetkov et al. (2013) also compute the concreteness indexes of English words by using a distributional semantic model and the MRC database. They train a logistic regression classifier on 1,225 most abstract and 1,225 most concrete words from MRC; the degree of concreteness of a word is the posterior probability produced by the classifier. The Tsvetkov et al. system for metaphor identification with concreteness indexes is based on cross-lingual model transfer, when the model is trained on English data, and then the classification features are translated into other languages by means of an electronic dictionary. 335 [yuliya.badryzlova@gmail.com](mailto:yuliya.badryzlova@gmail.com) Badryzlova (2020) explores concreteness and abstractness indexes for linguistic metaphor identification in Russian and English. The paradigm words are selected in a semi-automatic fashion: the Russian paradigm is derived from the Open Semantics of the Russian Language, the semantically annotated dataset of the KartaSlov database (Kulagin, 2019); the English paradigm is selected from the MRC database (Coltheart, 1981). The indexes of concreteness and abstractness are computed for large sets of Russian and English words (about 18,000 and 17,000 lexemes, respectively).

[REF4] - paperID: 336e531a59cafbe215b950fd749bca866b89cea0 Title: SNK @ DANKMEMES: Leveraging Pretrained Embeddings for Multimodal Meme Detection (short paper) Chunk of text: We limited the number of uses to 200 for computational efficiency reasons. Then, for each occurrence, we extracted and averaged the token vectors of (i) the last four layers of BERT, and (ii) the first and last layer. For our first submission (‘Last Four, 7’) we labeled those 7 words with ‘1’ that achieved the highest APD scores in layer combination (i). For our second submission (‘First + Last, 7’) we labeled those 7 words with ‘1’ that achieved the highest APD scores in layer combination (ii). In (i) and (ii) the same 9 words had the highest APD scores. Therefore, in our third submission (‘Average, 9’) exactly these 9 words were labeled with ‘1’. And for our last submission (Lemma, Average, 6’) we extracted only sentences in which the target words were present in their lemma form.

[REF5] - paperID: e052d22cba4eb069e8edf8ee39cbef81cc3eb84b Title: MarkedBERT: Integrating Traditional IR Cues in Pre-trained Language Models for Passage Retrieval Chunk of text: This study is encouraging future work on (a) using the same marking technique for other tasks, (b) further adapting advanced neural models to IR. We are looking forward to extending our work using a wide range of other IR cues.

Figure 1: Confusion matrix for a development set run with a macro-average F1 score of 0.95. 4.2 Classification on cross-genre data The runs submitted for the second sub-task are based on samples coming from a cross-genre, outof-domain test data set. These samples are a subset of the documents collected for the Epistolario project (Tonelli et al., 2020), an ongoing effort to create a digital archive of Alcide De Gasperi’s private and public correspondence. CLASS PRECISION RECALL F1 1901-1918 0.583 0.7 0.636 1919-1926 1.0 0.15 0.261 1927-1942 0.0 0.0 0.0 1943-1947 0.6 0.75 0.667 1948-1954 0.354 0.85 0.5 Table 5: Per-class results of the best test run for sub-task 2. As expected, despite scoring above the baseline, cross-genre results are significantly lower than those obtained in the same-genre task.

[REF7] - paperID: f6d69afebcebcbd3e511faf19375f71dd679cdcb Title: A passage-based approach to learning to rank documents Chunk of text: The ordered list of features is: ESA (p,15), SDM unigrams (d,4), SDM biterms (d,2), SW1 (d,2), Ent (d,1), SW2 (d,1), SDM bigrams (d,1), MaxPDSim (p,1), LengthRatio (p,1), SynonymsOverlap (p,1), pLocation (p,1), Entity (p,1). Thus, as was the case for the SVM-based feature weight analysis from above, ESA which is a passage feature and SDM unigrams which is a document feature are the most important. More generally, the list contains both document and passage features. We note that while the removal of each of the document features resulted in at least one case of statistically significant drop, for quite a few passage features this was not the case; i.e., there is redundancy between the passage features. We next turn to present feature analysis for the SMPD approach 13 . SMPD uses the same document features as JPDs, but different passagebased features: mainly those which quantify the rank positions of the document’s passages in the passage ranking. The results of an ablation test, as that performed above, are: max (p,5), SW2 (d,4), SDM unigrams (d,3), SDM biterms (d,2), avg (p,2), numPsg (p,2), Ent (d,1), SW1 (d,1), SDM bigrams (d,1), min (p,1), std (p,1), top50 (p,1).

[REF8] - paperID: f6d69afebcebcbd3e511faf19375f71dd679cdcb Title: A passage-based approach to learning to rank documents Chunk of text: Then, DLT R is re-ranked using the document retrieval methods from Section 3 that utilize G(DLT R ). We use MAP and p@10 to evaluate document retrieval performance. Baselines. Recall that DLT R was attained by re-ranking Dini t using an LTR approach; i.e., the set of documents in these two lists is the same. All the baselines we describe and our passage-based document retrieval methods from Section 3 are used to rank this document set. The initial language-model-based ranking ofDini t , denoted LM, is the first baseline. The second is the initial LTR-based ranking of DLT R, init-LTR.

[REF9] - paperID: df79bcba3a92605d21ef71faa0e703b7422ead22 Title: Speech Summarization using Essence Vector Modeling Chunk of text: Using speech summarization, one can efficiently digest the amount of information present in the spoken documents with minimal human interference. Our project focuses on summarizing a large amount of spoken data using natural language processing techniques and thus minimizing the time taken for manually extracting meaningful information from large speech-related data. Our project aims to create an Automatic Speech Summarization System using state of the art neural network architectures to summarize spontaneous speech into text. 2. LITERATURE SURVEY 2.1 An Information Distillation Framework for Extractive Speech Summarization In the context of text summarization, Kuan-Yu Chen and Shih-Hung Liu proposed a novel unsupervised paragraph embedding method, named the Essence Vector (EV) model , which aims at extracting the most representative information from a paragraph. The proposed method eliminates the general background information to produce a more informative lowdimensional vector representation for the input paragraph. According to , Classical paragraph embedding methods infer the vector representation by considering all of the words occurring in the paragraph, and Therefore the stop words that occur frequently in the paragraph might drift the theme of the content into producing an irrelevant summary.

........................................................................................................................................................................................................

Title: Deep Learning Models for Text Representation in Ranking

Deep learning models have gained significant attention in the field of natural language processing (NLP) due to their ability to capture complex patterns and representations in textual data. In the context of ranking, text representation plays a crucial role in determining the relevance and importance of documents. This section explores the use of deep learning models for text representation in ranking, drawing insights from various studies and approaches.

One approach that has been explored is the use of hypergraphs for text representation [REF0]. Hypergraphs offer a more flexible and expressive way to represent textual data compared to traditional graph models. They allow for the inclusion of synonyms, contextual information, and various types of relations between terms and entities. By incorporating these additional dimensions, hypergraphs can capture more nuanced semantic relationships, which can potentially improve the ranking performance.

Another study proposed a natural language interface to a graph-based bibliographic information retrieval system [REF1]. This approach utilized named entity recognition and dependency parsing to generate a graph query that could interpret natural language queries. The graph database stored bibliographic data, and the query was processed using named entity recognition to obtain nodes for the graph query. Dependency parsing was applied to extract relations between tokens, which were then adapted to the database schema. This approach demonstrates the potential of graph-based models in enhancing natural language understanding and question answering.

To address the complexity of graph traversals in entity-oriented search, the hypergraph-of-entity model was proposed [REF2]. This model reduced the number of hyperedges for the number of nodes by simplifying relations and promoting the grouping of multiple nodes. By controlling how (hyper)edges scale, the model aimed to improve performance and reduce complexity. This highlights the importance of efficient representation models in achieving effective ranking.

Evaluation methods and resources are crucial for assessing the quality of retrieval models [REF4]. Traditional information retrieval follows an empirical approach, relying on experimentation over test collections. Evaluation forums such as TREC, INEX, and CLEF provide datasets and evaluation moments for researchers to assess their models. These evaluation approaches are also applicable to entity-oriented search tasks, where relevance judgments are specific to documents or entities. The availability of test collections and evaluation forums facilitates the assessment of ranking tasks in the context of deep learning models.

In the evaluation of deep learning models for text representation in ranking, the retrieval effectiveness and efficiency are key considerations [REF6]. The random walk score is applied as a universal ranking function, and different versions of the hypergraph-of-entity representation model are tested. Various retrieval tasks, including ad hoc document retrieval leveraging entities, ad hoc entity retrieval, and entity list completion, are evaluated. The scalability of the model is also explored, aiming to index complete datasets rather than subsets. This comprehensive evaluation approach provides insights into the performance and trade-offs of different configurations.

In summary, deep learning models offer promising avenues for text representation in ranking. The use of hypergraphs, graph-based models, and efficient representation techniques can enhance the capture of semantic relationships and improve ranking performance. Evaluation methods and resources play a crucial role in assessing the effectiveness and efficiency of these models. By considering these factors, researchers can further advance the field of deep learning models for text representation in ranking.

References:

[REF0] - As we can see, the total number of hyperedges is significantly lower (almost half) than the number of nodes...

[REF1] - proposed a natural language interface to a graph-based bibliographic information retrieval system...

[REF2] - This resulted in an increased number of nodes, reaching nearly 1 million, however the number of edges was significantly reduced...

[REF4] - Traditionally, information retrieval follows an empirical approach to research, relying on experimentation over test collections...

[REF6] - In this chapter, we focus on evaluating the retrieval effectiveness of the hypergraph-of-entity...

[REF9] - [10, 11], which is used as part of term frequency normalization...

References given to GPT:

[REF0] - paperID: d121c33a5a0d8b6615d8581cfee8a941ebc7daed Title: Graph-based entity-oriented search Chunk of text: As we can see, the total number of hyperedges is significantly lower (almost half) than the number of nodes. This is the opposite behavior that we had found in the graph-of-entity, which didn’t even include synonyms or contextual information. Most of the nodes in the hypergraph are used to represent terms, closely followed by entities. Most of the hyperedges are directed, specifically used to link terms and entities. Out of the undirected hyperedges, most are used to establish context — we might consider increasing the acceptance threshold for contextually similar terms, when building the word2vec similarity network, in order to lower the number of context hyperedges. Relations of synonymy and contextual similarity were responsible for establishing new connections between documents, which in turn had the potential to improve recall over the base model. We analyzed the base model with synonyms and we found that synonyms established 6,968 new paths between documents, with 219.90 documents linked on average per synonym, with each synonym ranging between 1 and 12,839 linked documents.

���2.2 graph-based models pendency parsing, they were able to generate a graph query that was capable of correctly interpreting 39 out of 40 natural language queries of varied complexities. The approach relied on a graph database to store the bibliographic data. A natural language query was then processed using named entity recognition to obtain the nodes for the graph query to be issued over the graph database. Dependency parsing was applied to extract relations between tokens (including entities), which were then adapted to the database schema, for instance adding missing nodes (e.g., the dependency hpapers, happy universityi might be translated into h?author, paperi, linked by a :writes relation, and h?author, happy universityi, linked by a :is\_affiliated\_with relation). This abstract graph query could then be instantiated into a graph query language available for the graph database, where ?author is a node of type #author. Despite the identified domain-dependent limitations of the model, this contributed to the application of graphs as a tool for natural language understanding and question answering.

[REF2] - paperID: d121c33a5a0d8b6615d8581cfee8a941ebc7daed Title: Graph-based entity-oriented search Chunk of text: This resulted in an increased number of nodes, reaching nearly 1 million, however the number of edges was significantly reduced to nearly 10 million — 10.1 edges were created per node. Finally, in order to lower the complexity of graph traversals in the graph-of-entity, so that we could explore index extensions like synonymy or contextual similarity, we proposed the hypergraph-ofentity. This model was not only stored in memory instead of relying on a graph database, but it also significantly reduced the number of hyperedges for the number of nodes. This was possible by simplifying relations and promoting the grouping of multiple nodes. As a result, for 607,213 nodes, only 253,154 hyperedges were created, which means that less than one hyperedge was created per node — two hyperedges were created for every five nodes. Since the index size of a collection of documents is bound by the size of its vocabulary — at the very least the unique set of all words, or a selection of keywords, must be a part of the index — being able to control how (hyper)edges scale is a essencial to control performance and reduce complexity. It is a matter of how much information can we cross-reference within a given representation model, to better solve an information need, without losing the complexities provided by captured relations.

[REF3] - paperID: d121c33a5a0d8b6615d8581cfee8a941ebc7daed Title: Graph-based entity-oriented search Chunk of text: It describes it as a common reference model, able to represent controlled vocabularies, taxonomies, thesauri, faceted classification and ontologies. Yi . Compared thesaurus based information retrieval with topic maps based information retrieval, finding topic maps to outperform thesauri. (Continued on next page) 334c.3 graph-based models Table C.3. Graph-based models models for entity-oriented search.

[REF4] - paperID: d121c33a5a0d8b6615d8581cfee8a941ebc7daed Title: Graph-based entity-oriented search Chunk of text: 2.3 evaluation methods and resources Traditionally, information retrieval follows an empirical approach to research, relying on experimentation over test collections to assess the quality of retrieval models . Evaluation forums, such as TREC, INEX (INitiative for the Evaluation of XML Retrieval), or CLEF (Conference and Labs of the Evaluation Forum) also bring the community together to prepare these datasets, offering multiple evaluation moments, under the same conditions, for registered researchers. Each event usually has a list of tracks, available to participants — tracks approach specific retrieval tasks, usually providing a dataset or system for evaluation. In this section, we begin by covering contributions that illustrate archetypal evaluation approaches, and we cover some of the most relevant test collections and evaluation forums for the assessment of information retrieval tasks, in particular focusing on entity-oriented search. 2.3.1 Evaluation approaches In entity-oriented search, evaluation approaches are in line with the overall information retrieval research methodology, relying on test collections, where topics can either be used to build keyword or entity queries, and relevance judgments are specific to the tasks, where either documents or entities are graded. In this section, we illustrate two evaluation approaches for entity ranking tasks. Komninos and Arampatzis

[REF5] - paperID: d121c33a5a0d8b6615d8581cfee8a941ebc7daed Title: Graph-based entity-oriented search Chunk of text: on pp. 129, 206, 336). G. A. Miller. “WordNet: A Lexical Database for English”. In: Commun. ACM 38.11 (1995), pp. 39–41.

[REF6] - paperID: d121c33a5a0d8b6615d8581cfee8a941ebc7daed Title: Graph-based entity-oriented search Chunk of text: In this chapter, we focus on evaluating the retrieval effectiveness of the hypergraph-of-entity, applying the random walk score to multiple entity-oriented search tasks, while testing a wide range of configurations both for the representation model and the ranking function. Although our goal is to improve effectiveness, or a the very least ensure a minimum level of effectiveness, we also measure the efficiency of each run, in order to understand the overall impact and tradeoff of each configuration. We begin by measuring the performance over a single retrieval task, ad hoc document retrieval (leveraging entities), in order to focus on different versions of the hypergraph-of-entity representation model. We then expand this line of research to other retrieval tasks, also evaluating ad hoc entity retrieval, and entity list completion. We do this over a common index data structure and using the random walk score as a universal ranking function. Since the different tasks cannot be compared among each other, we also attempt to scale the model, so that it can index the complete INEX 2009, as opposed to just the subsets that we rely on for the first experiments. We do this by indexing the top keywords for each document, reducing complexity by partially lowering the number of nodes and, indirectly, the number of hyperedges linking terms to entities.

[REF7] - paperID: bbaa9599d10b5f29546a8c52eeb34b38ef4e3596 Title: Learning Better Representations for Neural Information Retrieval with Graph Information Chunk of text: GEPS (T+G): A graph-based model that utilizes graph embeddings from an external unsupervised graph model to guide the representation learning for product search. We set the dimension of the graph embedding as 128. Compared with our models, this method only uses the fixed pre-trained graph embedding which may be incompatible with a specific task. • BERT (Pretraining+T): Pretrained language model BERT that has been shown to be very effective on the document ranking task .

[REF8] - paperID: 017386502557c27d4ffd575b17ed7c2aafed2d95 Title: Item Tagging for Information Retrieval: A Tripartite Graph Neural Network based Approach Chunk of text: So it is necessary to stack at least two propagation layers. • Stacking too much (larger than 3) layers will not bring additional promotion. Compared with TagGNN-2, only TagGNN3 got a little gain (0.2%) in P@5 (Paritial Tags of KDDCup2012), verifying that two layers are enough for TagGNN. Too many layers may lead to redundancy that hurts performance. 3.6.2 Effect of Types of GNN. To verify the superiority of the propagation design of TagGNN, we replace the TagGNN with some other popular GNN models, e.g., GCN, GraphSAGE and GAT.

[REF9] - paperID: d121c33a5a0d8b6615d8581cfee8a941ebc7daed Title: Graph-based entity-oriented search Chunk of text: [10, 11], which is used as part of term frequency normalization. It mitigates the impact of document length in relevance ranking, without completely discarding information on the length of the original document. The underlying concepts of term frequency, inverse document frequency and pivoted document length normalization are transversal to most retrieval models. For 1 Semantic search as a task either refers to the semantically informed retrieval of documents, or to the retrieval of entities or relations over RDF graphs. We cover work on either approach, as both tasks are entity-oriented, using semantic search indiscriminately in both cases. 2 Karen Ida Boalth Spärck Jones was an influential information retrieval scientist, responsible for the creation of the inverse document frequency (IDF), one of the three fundamental concepts in information retrieval, the other being term frequency and document length normalization. See [https://irsg.bcs](https://irsg.bcs/).

........................................................................................................................................................................................................

Title: Text Representation Techniques for Ranking in Information Retrieval

Text representation plays a crucial role in information retrieval systems as it directly impacts the ranking of documents. In this section, we will discuss various text representation techniques that have been proposed and used for ranking in information retrieval. These techniques aim to capture the semantic meaning and relevance of textual content to improve the accuracy and effectiveness of ranking algorithms.

One popular approach for text representation is the use of convolutional neural networks (CNNs). Gamb ¨ ack and Sikdar [REF0] applied CNNs to classify hate speech, demonstrating the effectiveness of this technique in capturing important features from text. CNNs have also been used in the context of end-to-end architectures for ranking. For example, in the work by [REF1], three different architectures were proposed, differing in the point at which the extracted features are joined. These architectures leverage the power of CNNs in capturing local patterns and combining them with recurrent neural networks (RNNs) for capturing sequential dependencies.

Another approach for text representation in ranking is the use of joint embeddings of textual and image representations. Engilberge et al. [REF2] and Portaz et al. [REF2] employed RNNs and multilingual text to enhance cross-modal retrieval. This approach leverages the complementary information from both textual and visual modalities to improve the ranking accuracy. Neural image retrieval from the joint space has also been applied to fetch a group of associated images, further enhancing the retrieval performance.

In addition to deep learning-based techniques, other methods have been explored for text representation in ranking. For instance, ColBERT-PRF [REF4] introduced a unique pseudo-relevance feedback approach for dense retrieval, which differs from existing works that primarily function as rerankers. This approach utilizes neural models to augment documents before indexing, providing a novel perspective on improving ranking accuracy.

Furthermore, various weak supervision signals have been leveraged for text representation in ranking. Zheng et al. [REF7] and Dehghani et al. [REF7] proposed the use of pseudo relevance labels generated by unsupervised retrieval methods and title-document pairs as weak supervision signals. These signals approximate query-document relevance and help improve the ranking performance of neural information retrieval models.

In summary, text representation techniques for ranking in information retrieval have evolved significantly in recent years. Deep learning-based approaches, such as CNNs and joint embeddings, have shown promising results in capturing semantic meaning and improving ranking accuracy. Additionally, the utilization of weak supervision signals and pseudo-relevance feedback has further enhanced the performance of ranking algorithms. These techniques continue to be an active area of research, with ongoing efforts to develop more effective and efficient text representation methods for information retrieval.

References:

[REF0] Bjorn Gamb ¨ ack and Utpal Kumar Sikdar. 2017. Using convolutional neural networks to classify hate speech. In Proceedings of the first workshop on abusive language online, pages 85–90.

[REF1] Based on that premise, we propose three different end-to-end architectures that basically differ on the point in which the two CRNN models are joined: (i) the PreRNN one, which joins the extracted features by each model right before the recurrent block; (ii) the InterRNN one, which performs this process after the first recurrent layer; and (iii) the PostRNN one, which gathers both sources of information after the recurrent block. These proposals are graphically shown in Figure 3.

[REF2] Recent studies have followed the two-path architecture [45, 46], in which the encoder consists of a joint embedding of textual and image representations extracted from both the images and corresponding caption.

[REF4] To the best of our knowledge, this is a unique feature of ColBERT-PRF among PRF approaches. 4.5 Discussion To the best of our knowledge ColBERT-PRF is the first investigation of pseudo-relevance feedback for dense retrieval.

[REF7] Moreover, they may often be overly confident and more unstable in the learning process (Qiao et al., 2019). A promising direction to alleviate the dependence of Neu-IR models on large-scale relevance supervision is to leverage weak supervision signals that are noisy but available at mass quantity (Zheng et al., 2019b; Dehghani et al., 2017; Yu et al., 2020).

References given to GPT:

[REF0] - paperID: bd23ce64a6422c1f73acf51675e53b7a06547da3 Title: UOBIT @ TAG-it: Exploring a Multi-faceted Representation for Profiling Age, Topic and Gender in Italian Texts Chunk of text: Bjorn Gamb ¨ ack and Utpal Kumar Sikdar. 2017. Us- ¨ ing convolutional neural networks to classify hatespeech. In Proceedings of the first workshop on abusive language online, pages 85–90. Sepp Hochreiter and Jurgen Schmidhuber. 1997. ¨ Long short-term memory.

[REF1] - paperID: 79c573c54d4c1e6d2b678695c7802df7b4d380db Title: Exploiting the Two-Dimensional Nature of Agnostic Music Notation for Neural Optical Music Recognition Chunk of text: Based on that premise, we propose three different end-to-end architectures that basically differ on the point in which the two CRNN models are joined: (i) the PreRNN one, which joins the extracted features by each model right before the recurrent block; (ii) the InterRNN one, which performs this process after the first recurrent layer; and (iii) the PostRNN one, which gathers both sources of information after the recurrent block. These proposals are graphically shown in Figure 3. Convolutional block RecL1 RecL2 Convolutional block RecL1 RecL2 Concat Dense ( ) ΣS Dense ( ) ΣH Dense ( ) ΣT z Shape Height RecL1 RecL2 (a) PreRNN model: features extracted by each specialized model are joined after the convolutional block. Convolutional block RecL1 RecL2 Convolutional block RecL1 RecL2 Concat Dense ( ) ΣS Dense ( ) ΣH Dense ( ) ΣT z Shape Height RecL2 (b) InterRNN model: features extracted by each specialized model are joined after the first recurrent layer. Convolutional block RecL1 RecL2 Convolutional block RecL1 RecL2 Concat Dense ( ) ΣS Dense ( ) ΣH Dense ( ) ΣT z Shape Height (c) PostRNN model: features extracted by each specialized model are joined after the recurrent block. Figure 3.

[REF2] - paperID: 63483c9387d17e44eeb70c7321ad0dbb59b994fc Title: Universal Multimodal Representation for Language Understanding Chunk of text: Recent studies have followed the two-path architecture [45, 46], in which the encoder consists of a joint embedding of textual and image representations extracted from both the images and corresponding caption. Notably, Engilberge et al. adopts RNN to encode sentence embeddings in the same space with extracted image representations from CNN. Portaz et al. enhances cross-modal retrieval using multilingual text. Inspired by the previous success of visual-semantic embeddings, we apply neural image retrieval from the joint space to fetch a group of associated images. 3 UNIVERSAL REPRESENTATION FRAMEWORK This section overviews our universal representation framework.

[REF3] - paperID: 4deed74a3eee7e629dce2b8ef1e437ca74b2e64a Title: Efficiently Teaching an Effective Dense Retriever with Balanced Topic Aware Sampling Chunk of text: The biggest variation is on the nDCG@10 metric of TREC-DL’20, however the recall shows a lower variance than the recall on TREC-DL’19. This result gives us great confidence in the efficacy of our TAS-Balanced training. 5.2 Comparing to Baselines In this section we focus on standalone BERTDOT retrieval results from different training methods and compare our results with related work to answer: RQ3 How does our TAS-Balanced approach compare to other dense retrieval training methods? We present the dense retrieval results for models trained on the MSMARCO collection in Table 6, first the baselines and then our TAS-Balanced results using different training batch size settings. Important for the comparison of different BERTDOT training techniques is the number of Transformer encoder layers, which linearly Table 5: Random-robustness analysis of five instances of TAS-Balanced dual-supervision each using different sampling orders across clusters, queries, and passage pairs. Stat. sig. difference w/ paired t-test (p < 0.05)

[REF4] - paperID: 44772b24ae2f68b77476c814b0607370f7195ddb Title: Pseudo-Relevance Feedback for Multiple Representation Dense Retrieval Chunk of text: To the best of our knowledge, this is a unique feature of ColBERT-PRF among PRF approaches. 4.5 Discussion To the best of our knowledge ColBERT-PRF is the first investigation of pseudo-relevance feedback for dense retrieval. Existing works on neural pseudo-relevance feedback, such as Neural PRF and BERT-QE only function as rerankers. Other approaches such as DeepCT and doc2query [24, 25] use neural models to augment documents before indexing using a traditional inverted index.

[REF5] - paperID: 7715d2fc795a6406151b94924d9276939671f919 Title: TabSim: A Siamese Neural Network for Accurate Estimation of Table Similarity Chunk of text: PMC Jaccard 93.10 94.66 Cosine 95.58 95.68 Google Fusion 94.51 95.04 RF 90.53 92.03 LR 92.11 93.13 TabSim 93.76 94.57 arXiv Jaccard 40.53 41.09 Cosine 35.03 36.18 Google Fusion 29.17 32.11 RF 81.07 82.26 LR 62.25 72.48 TabSim 74.15 82.71 Wikipedia Jaccard 91.38 91.45 Cosine 91.06 91.14 Google Fusion 90.13 90.28 RF 96.46 96.50 LR 97.18 97.20 TabSim 97.28 97.32 VI. CONCLUSION We presented TabSim, a new method for assessing table similarity which uses Siamese neural networks to learn a similarity measure from a gold standard corpus of table pairs. We showed that, in comparison to five other methods of which three are also rooted in applications based on table similarity, TabSim attains considerably higher precision, recall, F1-score, and accuracy measures on three different corpora. Our results also demonstrate that, among different configurations of TabSim, the model which uses self-attention neural networks achieve the highest performance, probably because it is, different from the 2d-based CNN or the sequence-based Bi-LSTM, invariant to row or column permutations. As part of our research, we also created the first specific gold standard corpus for table similarity research, containing 1500 table pairs manually scored regarding their semantic similarity. Although the corpus was created in a way that gives methods relying on cosine similarity a competitive advantage, TabSim also leads the field on this corpus.

[REF6] - paperID: 8a6125562341d9a839006a23b48c870504810a27 Title: SDR: Efficient Neural Re-ranking using Succinct Document Representation Chunk of text: Following Nogueira and Cho (2019), 6628we use the automatic by-article annotations variant, which considers all paragraphs within the same article as relevant. The dataset consists of 30M passages, making storage requirements a more significant challenge compared to the MSMARCO task. The test query set consists of 2,254 queries with an average of 2.74 positive passages per query. We use the MAP@1K official metric. For both datasets, in addition to the quality metrics, we also measure the Compression Ratio (CR) as the amount of storage required to store the token embeddings when compared to the baseline model. E.g., CR = 10 implies storage size that is one tenth of the baseline vectors. 4.2 Baseline – BERTSPLIT Our algorithm is based on the late-interaction architecture (MacAvaney et al., 2020; Gao et al., 2020a; Nie et al., 2020; Chen et al., 2020; Cao et al., 2020).

[REF7] - paperID: 65c2d2ffe45569101860a7defc7cccbd36b3602a Title: Few-Shot Text Ranking with Meta Adapted Synthetic Weak Supervision Chunk of text: Moreover, they may often be overly confident and more unstable in the learning process (Qiao et al., 2019). A promising direction to alleviate the dependence of Neu-IR models on large-scale relevance supervision is to leverage weak supervision signals that are noisy but available at mass quantity (Zheng et al., 2019b; Dehghani et al., 2017; Yu et al., 2020). Through IR history, various weak supervision sources have been used to approximate querydocument relevance signals, e.g., pseudo relevance labels generated by unsupervised retrieval methods (Dehghani et al., 2017; Zheng et al., 2019b), and title-document pairs (MacAvaney et al., 2019). Recently, Zhang et al. (2020b) treat paired anchor texts and linked pages as weak relevance signals and propose a reinforcement-based data selection method ReInfoSelect, which learns to filter noisy anchor signals with trial-and-error policy gradients. Despite their convincing results, anchor signals are only available in web domains. Directly applying them to non-web domains may suffer from suboptimal outcomes due to domain gaps. To obtain weak supervision that adapts arbitrary domains, Ma et al.

[REF8] - paperID: bd23ce64a6422c1f73acf51675e53b7a06547da3 Title: UOBIT @ TAG-it: Exploring a Multi-faceted Representation for Profiling Age, Topic and Gender in Italian Texts Chunk of text: CHILab (Gambino and Pirrone, 2020) experimented transformer encoders in the first run and depth-wise Separable Convolution techniques in the second one. Moreover, some teams explored classical machine learning approaches such as No Place For Hate Speech (dos S. R. da Silva and T. Roman, 2020), TextWiller (Ferraccioli et al., 2020), UR NLP (Hoffmann and Kruschwitz, 2020) and Montanti (Bisconti and Montagnani, 2020). Finally, Venses (Delmonte, 2020), based on the parser for Italian ItGetaruns, applied six different rule-based classifiers. Features and Lexical Resources Various features are tested and explored by participants. Morphosyntactic features are exploited by CHILab, using Part-of-Speech tags as additional input. To adapt the POS tagging model provided by Python’s spaCy library to social media language, they added emoticons, emojis, hashtags and URLs to vocabulary.

[REF9] - paperID: 21ee4b66ce53de6b7b23c23cae0885bf5c96ad78 Title: PRADA: Practical Black-Box Adversarial Attacks against Neural Ranking Models Chunk of text: ACM Transactions on Information Systems, Vol. 41, No. 4, Article 89. Publication date: April 2023.89:16 C. Wu et al. Table 2. Comparisons between PRADA and the Baselines under the Automatic Evaluation Method MS-MARCO-Doc MS-MARCO-Pas SR PP SSdoc SSsen SR PP SSdoc SSsen First+RR 65.9 13.0 90.9 92.0 9.3 24.5 78.1 79.7 First+NR 41.9 13.0 94.3 94.9 14.8 24.5 85.9 86.3 Last+RR 10.7 13.0 91.1 91.9 20.7 24.5 78.7 81.5 Last+NR 8.2 13.0 94.7 95.1 22.7 24.5 86.2 87.1 Tf-idf+RR 48.1 13.0 90.5 90.5 8.8 24.5 80.5 80.5 Tf-idf+NR 43.8 13.0 93.1 93.0 10.1 24.5 81.5 81.1 TextRank+RR 55.4 13.0 87.4 88.8 8.7 24.5 74.2 73.7 TextRank+NR 37.5 13.0 90.8 92.7 13.9 24.5 84.2 83.6 TSrep 93.1 12.8 87.9 89.1 99.5 24.0 85.6 87.5 TSsti 70.9 12.9 91.2 91.7 59.9 24.3 86.8 87.0 PRADA 96.7∗ 4.0∗ 95.2∗ 96.2∗ 91.4 7.8∗ 93.2∗ 93.1∗ ∗Indicates

........................................................................................................................................................................................................