[11] M. Grbovic, N. Djuric, V. Radosavljevic, F. Silvestri, and N. Bhamidipati. Context-and content-aware embeddings for query rewriting in sponsored search. In *Proc. SIGIR*, pages 383–392. ACM, 2015.

[12] J. Guo, Y. Fan, Q. Ai, and W. B. Croft. A deep relevance matching model for ad-hoc retrieval. In *Proc. CIKM*, 2016.

[13] G. E. Hinton, J. L. McClelland, and D. E. Rumelhart. Parallel distributed processing: Explorations in the microstructure of cognition, vol. 1. chapter Distributed Representations, pages 77–109. MIT Press, Cambridge, MA, USA, 1986.

[14] T. Hofmann. Probabilistic latent semantic indexing. In *Proc. SIGIR*, pages 50–57. ACM, 1999.

[15] B. Hu, Z. Lu, H. Li, and Q. Chen. Convolutional neural network architectures for matching natural language sentences. In *Proc. NIPS*, pages 2042–2050, 2014.

[16] P.-S. Huang, X. He, J. Gao, L. Deng, A. Acero, and L. Heck. Learning deep structured semantic models for web search using clickthrough data. In *Proc. CIKM*, pages 2333–2338. ACM, 2013.

[17] K. Järvelin and J. Kekäläinen. Cumulated gain-based evaluation of IR techniques. *ACM TOIS*, 20(4):422–446, 2002.

[18] K. Järvelin and J. Kekäläinen. Cumulated gain-based evaluation of ir techniques. *TOIS*, 20(4):422–446, 2002.

[19] R. Jozefowicz, O. Vinyals, M. Schuster, N. Shazeer, and Y. Wu. Exploring the limits of language modeling. *arXiv preprint arXiv:1602.02410*, 2016.

[20] Y. LeCun, Y. Bengio, and G. Hinton. Deep learning. *Nature*, 521(7553):436–444, 2015.

[21] X. Liu, J. Gao, X. He, L. Deng, K. Duh, and Y.-Y. Wang. Representation learning using multi-task deep neural networks for semantic classification and information retrieval. *Proc. NAACL, May 2015*.

[22] Z. Lu and H. Li. A deep architecture for matching short texts. In *Advances in Neural Information Processing Systems*, pages 1367–1375, 2013.

[23] D. Metzler and W. B. Croft. A markov random field model for term dependencies. In *Proc. SIGIR*, pages 472–479. ACM, 2005.

[24] T. Mikolov, I. Sutskever, K. Chen, G. S. Corrado, and J. Dean. Distributed representations of words and phrases and their compositionality. In *Proc. NIPS*, pages 3111–3119, 2013.

[25] B. Mitra. Exploring session context using distributed representations of queries and reformulations. In *Proc. SIGIR*, pages 3–12. ACM, 2015.

[26] B. Mitra and N. Craswell. Query auto-completion for rare prefixes. In *Proc. CIKM*. ACM, 2015.

[27] B. Mitra, E. Nalisnick, N. Craswell, and R. Caruana. A dual embedding space model for document ranking. *arXiv preprint arXiv:1602.01137*, 2016.

[28] E. Nalisnick, B. Mitra, N. Craswell, and R. Caruana. Improving document ranking with dual word embeddings. In *Proc. WWW*, 2016.

[29] H. Palangi, L. Deng, Y. Shen, J. Gao, X. He, J. Chen, X. Song, and R. Ward. Deep sentence embedding using long short-term memory networks: Analysis and application to information retrieval. *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, 24(4):694–707, 2016.

[30] L. Pang, Y. Lan, J. Guo, J. Xu, and X. Cheng. A study of matchpyramid models on ad-hoc retrieval. In *Neu-IR '16 SIGIR Workshop on Neural Information Retrieval*, 2016.

[31] L. Pang, Y. Lan, J. Guo, J. Xu, S. Wan, and X. Cheng. Text matching as image recognition. In *Proc. AAAI*, 2016.

[32] J. M. Ponte and W. B. Croft. A language modeling approach to information retrieval. In *Proc. SIGIR*, pages 275–281. ACM, 1998.

[33] S. Robertson and H. Zaragoza. *The probabilistic relevance framework: BM25 and beyond*. Now Publishers Inc, 2009.

[34] D. Roy, D. Paul, M. Mitra, and U. Garain. Using word embeddings for automatic query expansion. In *Neu-IR '16 SIGIR Workshop on Neural Information Retrieva*, 2016.

[35] A. Severyn and A. Moschitti. Learning to rank short text pairs with convolutional deep neural networks. In *Proc. SIGIR*, pages 373–382. ACM, 2015.

[36] Y. Shen, X. He, J. Gao, L. Deng, and G. Mesnil. A latent semantic model with convolutional-pooling structure for information retrieval. In *Proc. CIKM*, pages 101–110. ACM, 2014.

[37] Y. Shen, X. He, J. Gao, L. Deng, and G. Mesnil. Learning semantic representations using convolutional neural networks for web search. In *Proc. WWW*, pages 373–374, 2014.

[38] A. Sordoni, Y. Bengio, H. Vahabi, C. Lioma, J. Grue Simonsen, and J.-Y. Nie. A hierarchical recurrent encoder-decoder for generative context-aware query suggestion. In *Proc. CIKM*, pages 553–562. ACM, 2015.

[39] X. Wei and W. B. Croft. Lda-based document models for ad-hoc retrieval. In *Proc. SIGIR*, pages 178–185. ACM, 2006.

[40] D. Yu, A. Eversole, M. Seltzer, K. Yao, Z. Huang, B. Guenter, O. Kuchaiev, Y. Zhang, F. Seide, H. Wang, et al. An introduction to computational networks and the computational network toolkit. Technical report, Tech. Rep. MSR, Microsoft Research, 2014, http://codebox/cntk, 2014.

[41] G. Zheng and J. Callan. Learning to reweight terms with distributed representations. In *Proc. SIGIR*, pages 575–584. ACM, 2015.

[42] Y. Zhou and W. B. Croft. Document quality models for web ad hoc retrieval. In *Proc. CIKM*, pages 331–332. ACM Press, 2005.