

Рекомендована література

1. Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, L., & Polosukhin, I. (2017). *Attention is all you need*. arXiv. <https://doi.org/10.48550/arXiv.1706.03762> (arXiv)
2. Dai, Z., Yang, Z., Yang, Y., Carbonell, J., Le, Q. V., & Salakhutdinov, R. (2019). *Transformer-XL: Attentive language models beyond a fixed-length context*. arXiv. <https://doi.org/10.48550/arXiv.1901.02860> (arXiv)
3. Lewis, P., Perez, E., Piktus, A., Petroni, F., Karpukhin, V., Goyal, N., Küttler, H., Lewis, M., Yih, W.-T., Rocktäschel, T., Riedel, S., & Kiela, D. (2020). *Retrieval-augmented generation for knowledge-intensive NLP tasks*. arXiv. <https://doi.org/10.48550/arXiv.2005.11401> (arXiv)
4. Ho, J., Jain, A., & Abbeel, P. (2020). *Denoising diffusion probabilistic models*. arXiv. <https://doi.org/10.48550/arXiv.2006.11239> (arXiv)
5. Baevski, A., Zhou, H., Mohamed, A., & Auli, M. (2020). *wav2vec 2.0: A framework for self-supervised learning of speech representations*. arXiv. <https://doi.org/10.48550/arXiv.2006.11477> (arXiv)
6. Li, Z., Kovachki, N., Azizzadenesheli, K., Liu, B., Bhattacharya, K., Stuart, A., & Anandkumar, A. (2020). *Fourier neural operator for parametric partial differential equations*. arXiv. <https://doi.org/10.48550/arXiv.2010.08895> (arXiv)
7. Dosovitskiy, A., Beyer, L., Kolesnikov, A., Weissenborn, D., Zhai, X., Unterthiner, T., Dehghani, M., Minderer, M., Heigold, G., Gelly, S., Uszkoreit, J., & Houlsby, N. (2020). *An image is worth 16×16 words: Transformers for image recognition at scale*. arXiv. <https://doi.org/10.48550/arXiv.2010.11929> (arXiv)
8. Jumper, J., Evans, R., Pritzel, A., Green, T., Figurnov, M., Ronneberger, O., Tunyasuvunakool, K., Bates, R., Žídek, A., Potapenko, A., Bridgland, A., Meyer, C., Kohl, S. A. A., Ballard, A. J., Cowie, A., Romera-Paredes, B., Nikolov, S., Jain, R., Adler, J., ... Hassabis, D. (2021). Highly accurate protein structure prediction with AlphaFold. *Nature*, 596(7873), 583–589. <https://doi.org/10.1038/s41586-021-03819-2>
9. Ouyang, L., Wu, J., Jiang, X., Almeida, D., Wainwright, C. L., Mishkin, P., Zhang, C., Agarwal, S., Slama, K., Ray, A., Schulman, J., Hilton, J., Kelton, F., Miller, L., Simens, M., Askell, A., Welinder, P., Christiano, P., Leike, J., & Lowe, R. (2022). *Training language models to follow instructions with human feedback*. arXiv. <https://doi.org/10.48550/arXiv.2203.02155> (arXiv)
10. Liang, P., Bommasani, R., Lee, T., Tsipras, D., Soylu, D., Yasunaga, M., Zhang, Y., Narayanan, D., Wu, Y., Kumar, A., Newman, B., Yuan, B., Yan, B., Zhang, C., Cosgrove, C., Manning, C. D., Ré, C., Acosta-Navas, D., Hudson, D. A., ... Koreeda, Y. (2022). *Holistic evaluation of language models*. arXiv. <https://doi.org/10.48550/arXiv.2211.09110> (arXiv)

11. Yao, S., Zhao, J., Yu, D., Du, N., Shafran, I., Narasimhan, K., & Cao, Y. (2022). *ReAct: Synergizing reasoning and acting in language models*. arXiv. <https://doi.org/10.48550/arXiv.2210.03629> (arXiv)
12. Rombach, R., Blattmann, A., Lorenz, D., Esser, P., & Ommer, B. (2021). *High-resolution image synthesis with latent diffusion models*. arXiv. <https://doi.org/10.48550/arXiv.2112.10752> (arXiv)
13. Radford, A., Kim, J. W., Xu, T., Brockman, G., McLeavey, C., & Sutskever, I. (2022). *Robust speech recognition via large-scale weak supervision*. arXiv. <https://doi.org/10.48550/arXiv.2212.04356> (arXiv)
14. Zhang, L., Rao, A., & Agrawala, M. (2023). *Adding conditional control to text-to-image diffusion models*. arXiv. <https://doi.org/10.48550/arXiv.2302.05543> (arXiv)
15. Park, J. S., O'Brien, J. C., Cai, C. J., Ringel Morris, M., Liang, P., & Bernstein, M. S. (2023). *Generative agents: Interactive simulacra of human behavior*. arXiv. <https://doi.org/10.48550/arXiv.2304.03442> (arXiv)