

OpenITI Tabanlı RAG Sistemi ile Doğal Dil İşleme

Yazar

Öğr. Gör. Dr. Muhammet TAY



© Copyright 2025

Bu kitabın, basım, yayın ve satış hakları Akademisyen Kitabevi A.Ş.'ne aittir. Anılan kuruluşun izni alınmadan kitabın tümü ya da bölümleri mekanik, elektronik, fotokopi, manyetik kağıt ve/veya başka yöntemlerle çoğaltılamaz, basılamaz, dağıtılamaz. Tablo, şekil ve grafikler izin alınmadan, ticari amaçlı kullanılamaz. Bu kitap T.C. Kültür Bakanlığı bandrolü ile satılmaktadır.

ISBN 978-625-375-855-4	Yayın Koordinatörü Yasin DİLMEN
Kitap Adı OpenITI Tabanlı RAG Sistemi ile Doğal Dil İşleme	Sayfa ve Kapak Tasarımı Akademisyen Dizgi Ünitesi
Editör Arafat ŞENTÜRK ORCID iD: 0000-0002-9005-3565	Yayıncı Sertifika No 47518
Yazar Muhammet TAY ORCID iD: 0000-0001-5569-7886	Baskı ve Cilt Vadi Matbaacılık
	Bisac Code TEC000000
	DOI 10.37609/akya.4033

Kütüphane Kimlik Kartı
Tay, Muhammet.

OpenITI Tabanlı RAG Sistemi ile Doğal Dil İşleme / Muhammet Tay; ed. Arafat Şentürk.
Ankara : Akademisyen Yayınevi Kitabevi, 2025.
109 s. : şekil. ; 135x210 mm.
Kaynakça var.
ISBN 9786253758554

GENEL DAĞITIM
Akademisyen Kitabevi A.Ş.

Halk Sokak 5 / A
Yenişehir / Ankara
Tel: 0312 431 16 33
siparis@akademisyen.com

www.akademisyen.com

ÖNSÖZ

Dijital teknolojiler, özellikle son yıllarda yapay zekâ ve Doğal Dil İşleme alanında yaşanan hızlı gelişmelerle birlikte, beşeri bilimler ve dini ilimler alanlarında da yeni imkânlar ve tartışma alanları ortaya çıkarmıştır. Büyük Dil Modelleri, metinleri anlama ve üretme konusundaki dikkat çekici başarılarına rağmen, bilgiye erişim, kaynak gösterme ve doğrulanabilirlik gibi konularda önemli metodolojik ve etik soruları da beraberinde getirmektedir. Bu kitap, söz konusu sorulara teknik olduğu kadar akademik bir perspektiften yaklaşma çabasının bir ürünüdür.

Bu çalışmanın temel amacı, OpenITI tabanlı Retrieval-Augmented Generation (RAG) mimarisinin Doğal Dil İşleme bağlamında nasıl çalıştığını ortaya koymak ve bu yaklaşımın özellikle İslami metinler üzerinde sorumlu, denetlenebilir ve akademik açıdan tutarlı biçimde nasıl kullanılabileceğini göstermektir. Kitapta yapay zekâ, bağımsız bir bilgi otoritesi olarak değil; belirli metinlere dayalı olarak çalışan, araştırmacıyı destekleyen bir araç olarak konumlandırılmıştır. Bu tercih, çalışmanın metodolojik omurgasını oluşturmaktadır.

Kitap boyunca Doğal Dil İşlemenin temel kavramlarından başlayarak, embedding tabanlı anlamsal arama, vektör veritabanları ve RAG mimarisinin teknik bileşenleri adım adım ele alınmıştır. Bu teknik altyapı, yalnızca teorik düzeyde bırakılmamış; gerçek bir sistem tasarımı ve uygulama örneği üzerinden somutlaştırılmıştır. Özellikle uygulama bölümünde, Arapça ve Türkçe İslami metinlerle çalışırken karşılaşılan dilsel ve kavramsal zorluklar, teknik çözümlerle birlikte ele alınmıştır.

Bu çalışma, klasik anahtar kelime tabanlı arama yaklaşımlarının ötesine geçerek, bağlam temelli ve anlamsal ilişkileri merkeze

alan bir bilgi erişim anlayışını savunmaktadır. Ancak bu yaklaşım, mutlak doğrular üretme iddiası taşımamaktadır. Aksine, üretilen çıktılar, araştırmacının eleştirel değerlendirmesine açık olacak şekilde tasarlanmıştır. Nihai yorum ve ilmî sorumluluk her zaman insana aittir; yapay zekâ ise bu süreci kolaylaştıran bir yardımcıdır.

Kitabın hedef kitlesi; bilgisayar mühendisliği, veri bilimi ve Doğal Dil İşleme alanlarında çalışan araştırmacıların yanı sıra, dijital araçları akademik çalışmalarında kullanmak isteyen ilahiyat ve İslami ilimler alanındaki araştırmacılarıdır. Teknik anlatım ile akademik yaklaşım arasında dengeli bir dil tercih edilmiş, mümkün olduğunca disiplinlerarası bir köprü kurulması amaçlanmıştır.

Son olarak bu kitabın, yapay zekâ teknolojilerinin beşeri bilimler ve dini ilimler alanında nasıl daha bilinçli, eleştirel ve sorumlu biçimde kullanılabileceğine dair yeni tartışmalara ve çalışmalara zemin hazırlaması temenni edilmektedir. Burada sunulan yaklaşım, tamamlanmış bir sonuçtan ziyade, geliştirilmeye açık bir başlangıç olarak görülmelidir.

Bu kitabın hazırlanma sürecinde, metinlerin dil ve imla açısından daha akıcı ve tutarlı hâle getirilmesi amacıyla yapay zekâ tabanlı bir dil modeli olan ChatGPT'den yararlanılmıştır. Kullanılan araç yalnızca dilsel düzenleme ve ifade iyileştirme amacıyla kullanılmış olup, kitabın bilimsel içeriği, yöntemleri ve sonuçları tamamen yazara aittir.

Muhammet TAY

2025 / DÜZCE

İÇİNDEKİLER

BÖLÜM 1

GİRİŞ VE KAVRAMSAL ÇERÇEVE	1
1.1. Doğal Dil İşlemenin Ortaya Çıkışı ve Yapay Zekâ ile İlişkisi	1
1.2. Kural Tabanlı Sistemler Dönemi.....	4
1.3. İstatistiksel Yaklaşımların Yükselişi	7
1.4. Derin Öğrenme ve Vektör Temsiline Geçiş	10
1.5. Transformer Mimarisi ve Büyük Dil Modelleri.....	13
1.6. Dijital Beşeri Bilimler ve İslami Metinler	16
1.7. Problemin Tanımı.....	19
1.8. Kitabın Amacı, Kapsamı ve Katkısı.....	22

BÖLÜM 2

DOĞAL DİL İŞLEMENİN TEMELLERİ	27
2.1. Dil Kavramı ve Doğal Dilin Özellikleri.....	27
2.2. Dilin Yapısal Bileşenleri (Fonoloji, Morfoloji, Sözdizim, Anlambilim, Pragmatik)	30
2.3. Metin Temsili ve Tokenization Yaklaşımları	33
2.4. Temel NLP Görevleri (POS, NER, Metin Sınıflandırma).....	36
2.5. İleri Düzey NLP Görevleri (Soru-Cevap, Özetleme, Diyalog Sistemleri)	39
2.6. NLP Sistemlerinde Değerlendirme ve Performans Ölçütleri.....	42
2.7. İslami Metinler Bağlamında NLP'nin Sınırları ve İmkânları	46

BÖLÜM 3

RAG MİMARİSİ VE SİSTEM TASARIMI	49
3.1. RAG (Retrieval-Augmented Generation) Yaklaşımına Giriş	49
3.2. RAG Mimarilerinin Temel Bileşenleri	51
3.3. Vektör Veritabanları ve Anlamsal Arama.....	53
3.4. Prompt Tasarımı ve Bağlam Yönetimi.....	55

3.5. OpenITI Tabanlı RAG Sisteminin Genel Mimarisi	57
3.6. Sistem Tasarımında Akademik ve Etik İlkeler.....	59

BÖLÜM 4

UYGULAMA VE ANALİZ..... 63

4.1. Metin Ön İşleme (Text Preprocessing) Aşaması.....	63
4.1.1. Metin Temizleme (Text Cleaning).....	64
4.1.2. Arapça Metinler için Normalizasyon.....	65
4.1.3. Tokenization ve Kelime Düzeyi İşleme.....	66
4.1.4. Ön İşleme Aşamasının RAG Sistemine Etkisi	67
4.2. Embedding Oluşturma ve Vektörleştirme Süreci.....	67
4.2.1. Embedding Kavramı ve Anlamsal Temsil	68
4.2.2. Kullanılan Embedding Modelinin Seçimi	68
4.2.3. Embedding Üretim Süreci	69
4.2.4. Embedding Boyutu ve Performans İlişkisi.....	70
4.2.5. Embedding’lerin RAG Sistemindeki Rolü.....	70
4.3. Qdrant ile Vektör Veritabanı ve Anlamsal Arama	71
4.3.1. Embedding’den Anlamsal Aramaya Geçiş	71
4.3.2. Qdrant’ın Tercih Edilme Nedenleri	72
4.3.3. Koleksiyon Yapısının Tasarımı	72
4.3.4. Embedding’lerin Veritabanına Eklenmesi (Upsert).....	73
4.3.5. Anlamsal Arama (Semantic Retrieval)	74
4.3.6. Embedding Temelli Aramanın Akademik Önemi.....	74
4.4. RAG Retrieve + Generate Akışı (retrieve_and_generate)	75
4.4.1. Sorgu Embedding’i ve Retrieval Aşaması.....	76
4.4.2. Bağlamın Oluşturulması ve Yapılandırılması.....	76
4.4.3. Prompt Tasarımı ve Modelin Sınırlandırılması	77
4.4.4. Üretken Dil Modeli ile Cevap Üretimi	78
4.4.5. retrieve_and_generate Fonksiyonunun Birleştirilmiş Hâli....	79
4.4.6. Akışın Akademik ve Etik Değerlendirmesi.....	80
4.5. Arapça NLP’ye Özgü Teknik Çözümler	80
4.6. Türkçe ve İslami Terminolojiye Yönelik İşleme Yaklaşımları	82

4.7. Uygulama Senaryosunun Tanımı ve Amaçları.....	83
4.8. Kullanılan Veri Kaynağı ve Metin Külliyatı.....	84
4.9. Sistem Akışı ve İşlem Adımları	85
4.10. Kullanıcı Arayüzü ve Etkileşim Yapısı	87
4.11. Örnek Sorgular ve Sistem Çıktıları.....	88
4.12. Değerlendirme ve Gözlemler.....	90

BÖLÜM 5

TARTIŞMA, SINIRLILIKLAR VE GELECEK

ÇALIŞMALAR.....	93
5.1. Bulguların Tartışılması.....	93
5.2. Sistem Sınırlılıkları	94
5.3. Metodolojik ve Epistemolojik Değerlendirme.....	95
5.4. Gelecek Çalışmalar İçin Öneriler.....	96

SONUÇ	97
--------------------	-----------

KAYNAKLAR	99
------------------------	-----------

KAYNAKLAR

- ALPAC. (1966). Languages and machines: Computers in translation. National Academy of Sciences.
- Alrabiah, M., Al-Salman, A. S., Atwell, E., & Alsaif, A. (2014). The design and construction of the 50-million-word KSUCCA. *International Journal of Computational Linguistics*, 5(2), 1–14.
- Bender, E. M., & Friedman, B. (2018). Data statements for natural language processing. *Transactions of the Association for Computational Linguistics*, 6, 587–604.
- Bender, E. M., & Koller, A. (2020). Climbing towards NLU: On meaning, form, and understanding in the age of data. *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, 5185–5198. <https://doi.org/10.18653/v1/2020.acl-main.463>
- Bender, E. M., Gebru, T., McMillan-Major, A., & Shmitchell, S. (2021). On the dangers of stochastic parrots: Can language models be too big? *Proceedings of the ACM Conference on Fairness, Accountability, and Transparency*, 610–623. <https://doi.org/10.1145/3442188.3445922>
- Bengio, Y., Ducharme, R., Vincent, P., & Janvin, C. (2003). A neural probabilistic language model. *Journal of Machine Learning Research*, 3, 1137–1155.
- Bengio, Y., Simard, P., & Frasconi, P. (1994). Learning long-term dependencies with gradient descent is difficult. *IEEE Transactions on Neural Networks*, 5(2), 157–166.
- Berry, D. M. (2012). *Understanding digital humanities*. Palgrave Macmillan.
- Blei, D. M., & Lafferty, J. D. (2009). Topic models. In *Text mining: Classification, clustering, and applications* (pp. 71–94).
- Brown, T. B., Mann, B., Ryder, N., Subbiah, M., Kaplan, J., Dhariwal, P., et al. (2020). Language models are few-shot learners. *Advances in Neural Information Processing Systems*, 33, 1877–1901.
- Chomsky, N. (1957). Syntactic structures. Mouton.
- Cho, K., van Merriënboer, B., Gulcehre, C., Bahdanau, D., Bougares, F., Schwenk, H., & Bengio, Y. (2014). Learning phrase representations using RNN encoder–decoder for statistical machine translation. *Proceedings of EMNLP*, 1724–1734.
- Deerwester, S., Dumais, S. T., Furnas, G. W., Landauer, T. K., & Harshman, R. (1990). Indexing by latent semantic analysis. *Journal of the American Society for Information Science*, 41(6), 391–407.
- Devlin, J., Chang, M.-W., Lee, K., & Toutanova, K. (2019). BERT: Pre-training of deep bidirectional transformers for language understanding. *Proceedings of NAACL-HLT*, 4171–4186.

- Floridi, L., Cowsls, J., Beltrametti, M., et al. (2018). AI4People—An ethical framework for a good AI society. *Minds and Machines*, 28(4), 689–707. <https://doi.org/10.1007/s11023-018-9482-5>
- Francis, W. N., & Kučera, H. (1979). *Brown corpus manual*. Brown University.
- Goldberg, Y. (2017). *Neural network methods for natural language processing*. Morgan & Claypool.
- Habash, N. (2010). *Introduction to Arabic natural language processing*. Morgan & Claypool.
- Harris, Z. S. (1954). Distributional structure. *Word*, 10(2–3), 146–162.
- Hochreiter, S., & Schmidhuber, J. (1997). Long short-term memory. *Neural Computation*, 9(8), 1735–1780.
- Hutchins, J. (2004). The Georgetown–IBM experiment demonstrated in January 1954. *Machine Translation*, 19(2), 91–114.
- Jockers, M. L. (2013). *Macroanalysis: Digital methods and literary history*. University of Illinois Press.
- Johnson, J., Douze, M., & Jégou, H. (2019). Billion-scale similarity search with GPUs. *IEEE Transactions on Big Data*, 7(3), 535–547. <https://doi.org/10.1109/TBDATA.2019.2921572>
- Jurafsky, D., & Martin, J. H. (2023). *Speech and language processing* (3rd ed. draft). Prentice Hall.
- Larkey, L. S., Ballesteros, L., & Connell, M. E. (2007). Light stemming for Arabic information retrieval. In *Arabic Computational Morphology* (pp. 221–243). Springer.
- 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. *Advances in Neural Information Processing Systems*, 33, 9459–9474.
- Liu, P., Yuan, W., Fu, J., Jiang, Z., Hayashi, H., & Neubig, G. (2023). Pre-train, prompt, and predict: A systematic survey of prompting methods in NLP. *ACM Computing Surveys*, 55(9), 1–35. <https://doi.org/10.1145/3560815>
- Manning, C. D., & Schütze, H. (1999). *Foundations of statistical natural language processing*. MIT Press.
- Manning, C. D., Raghavan, P., & Schütze, H. (2008). *Introduction to information retrieval*. Cambridge University Press.
- McGillivray, B., Hengchen, S., et al. (2019). A computational approach to lexical semantic change. *Computational Linguistics*, 45(3), 427–472.
- Mikolov, T., Chen, K., Corrado, G., & Dean, J. (2013). Efficient estimation of word representations in vector space. *arXiv preprint arXiv:1301.3781*.
- Mitchell, M., Wu, S., Zaldívar, A., et al. (2019). Model cards for model reporting. *Proceedings of FAT*, 220–229.
- Moretti, F. (2013). *Distant reading*. Verso.
- Nadeau, D., & Sekine, S. (2007). A survey of named entity recognition and classification. *Linguisticae Investigationes*, 30(1), 3–26.
- Nenkova, A., & McKeown, K. (2012). A survey of text summarization techniques. In *Mining text data* (pp. 43–76).

- Powers, D. M. W. (2011). Evaluation: From precision, recall and F-measure to ROC. *Journal of Machine Learning Technologies*, 2(1), 37–63.
- Rabiner, L. R. (1989). A tutorial on hidden Markov models. *Proceedings of the IEEE*, 77(2), 257–286.
- Rajpurkar, P., Zhang, J., Lopyrev, K., & Liang, P. (2016). SQuAD. *Proceedings of EMNLP*, 2383–2392.
- Romanov, M. (2017). Calculating similarity and relative dating of Arabic texts. *Digital Scholarship in the Humanities*, 32(1), 125–144.
- Romanov, M. (2019). Computational reading of Arabic biographical collections. *Digital Scholarship in the Humanities*, 34(2), 248–268.
- Ruder, S. (2017). An overview of multi-task learning in deep neural networks. *arXiv preprint arXiv:1706.05098*.
- Russell, S., & Norvig, P. (2021). *Artificial intelligence: A modern approach* (4th ed.). Pearson.
- Sebastiani, F. (2002). Machine learning in automated text categorization. *ACM Computing Surveys*, 34(1), 1–47.
- Sennrich, R., Haddow, B., & Birch, A. (2016). Neural machine translation of rare words with subword units. *Proceedings of ACL*, 1715–1725.
- See, A., Liu, P. J., & Manning, C. D. (2017). Get to the point. *Proceedings of ACL*, 1073–1083.
- Sokolova, M., & Lapalme, G. (2009). A systematic analysis of performance measures. *Information Processing & Management*, 45(4), 427–437.
- Tjong Kim Sang, E. F., & De Meulder, F. (2003). CoNLL-2003 shared task. *Proceedings of CoNLL*, 142–147.
- Turing, A. M. (1950). Computing machinery and intelligence. *Mind*, 59(236), 433–460.
- Vaswani, A., Shazeer, N., Parmar, N., et al. (2017). Attention is all you need. *Advances in Neural Information Processing Systems*, 30, 5998–6008.
- Wei, J., Wang, X., Schuurmans, D., et al. (2022). Chain-of-thought prompting. *NeurIPS*, 35, 24824–24837.
- Weizenbaum, J. (1966). ELIZA. *Communications of the ACM*, 9(1), 36–45.
- Winograd, T., & Flores, F. (1986). *Understanding computers and cognition*. Ablex.
- Young, S., Gašić, M., Thomson, B., & Williams, J. D. (2013). POMDP-based spoken dialog systems. *Proceedings of the IEEE*, 101(5), 1160–1179.