

السيرة الذاتية
د. علي صفايا
alisafaya@gmail.com
HyperbeeAI

غيت هب: alisafaya
تويتر: ali_safaya
سيمانتك سكولار: Ali Safaya
أوركيد: 0000-0003-0945-8897
الموقع الشخصي: asafaya.me

الاهتمامات البحثية

فهم اللغة الطبيعية، ونماذج اللغة الكبيرة، وتعزيز الذاكرة العصبية.

الخبرة المهنية

تموز 2025 –	المنصب:	عالم أبحاث في الذكاء الاصطناعي
حتى الآن	الجهة:	HyperbeeAI، كاليفورنيا
كانون الثاني 2024 –	المنصب:	عالم أبحاث في الذكاء الاصطناعي
تموز 2025	الجهة:	Codeway، إسطنبول
نيسان 2021 –	المنصب:	عالم أبحاث في معالجة اللغات الطبيعية
كانون الثاني 2024	الجهة:	HyperbeeAI، كاليفورنيا
تشرين الثاني 2017 –	المنصب:	مهندس تعلم آلي ومعالجة لغات طبيعية
أيلول 2020	الجهة:	Casemice Digital، إسطنبول

التعليم

أيار 2020 –	الدرجة:	دكتوراه في علوم الحاسوب
أيلول 2026	الجامعة:	جامعة كوتش، إسطنبول
	المعدل:	3.75 من 4.00
	المشرف:	أ. د. دنيز يوريت
	الأطروحة:	

Vector Retrieval-Augmented Language Models: What Retrieval Can and Cannot Replace

زميل أبحاث الدكتوراه في مركز KUIS للذكاء الاصطناعي.

آب 2015 –	الدرجة:	بكالوريوس العلوم في هندسة الحاسوب
حزيران 2019	الجامعة:	جامعة سكاريا، سكاريا
	المعدل:	3.89 من 4.00

الأول على الدفعة وخطيب حفل التخرج، تخرج بأعلى معدل تراكمي طوال البرنامج الدراسي.

الأبحاث

أيلول 2022 –
2026

المشروع: نماذج اللغة طويلة المدى وتعزيز الذاكرة العصبية
الجهة: مركز KUIS للذكاء الاصطناعي، جامعة كوتش، إسطنبول
المشرف: أ. د. دنيز يوريت

المساهمات: عملتُ على الفصل بين الذاكرة الدلالية والذاكرة العرضية في نماذج اللغة، إضافةً إلى تطوير معماريات ذاكرة محسّنة لنمذجة التسلسلات الطويلة، وكان ذلك محور تركيزي البحثي الأساسي في رسالة الدكتوراه.

كانون الثاني 2021 –
كانون الثاني 2024

المشروع: مستودع البيانات التركية (Turkish Data Depository)
الجهة: مركز KUIS للذكاء الاصطناعي، جامعة كوتش، إسطنبول
المشرف: أ. د. دنيز يوريت

المساهمات: قادتُ عدة فرق بحثية لتطوير مجموعات بيانات وأدوات لمعالجة اللغة التركية. شمل ذلك تسليم مدقق إملائي تركي جديد لفريق LibreOffice، وبناء موقع إلكتروني يجمع مجموعات بيانات المعالجة التركية المتاحة بصورة مفتوحة وموثقة. ومن أبرز المشاريع Mukayese: منصة شاملة لقياس أداء معالجة اللغة التركية توفر مجموعات اختبار ونماذج أساس لمعظم المهام. انظر <https://github.com/tdd-ai> و <https://huggingface.co/mukayese>.

آب 2019 –
آب 2020

المشروع: مشروع رفاهية الأسواق الناشئة الممول من ERC
الجهة: جامعة كوتش، إسطنبول
المشرف: د. علي حُرّيت أوغلو

المساهمات: عملت مهندس مشروع على بناء قاعدة بيانات لأحداث الاحتجاجات باستخدام معالجة اللغة الطبيعية وتعلّم الآلة، واكتسبت خبرة في العمل ضمن مشروع بحثي طويل الأمد والتعاون مع أكاديميين من تخصصات متعددة. كما شاركت في تنظيم ورش عمل في مؤتمرات رائدة مثل ACL و LREC.

التدريس

2020 –
2026

المنصب: مساعد تدريس
الجهة: قسم علوم الحاسوب، جامعة كوتش

مساعد تدريس لمقررات هياكل البيانات والخوارزميات (COMP-202)، والخوارزميات والتعقيد (COMP-305)، والتعلّم العميق (COMP-541)، ومعالجة اللغة الطبيعية (COMP-542).

المنشورات

- Ali Safaya and Deniz Yuret. "Neurocache: Efficient Vector Retrieval for Long-range Language Modeling". In: *Proceedings of the 2024 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies (Volume 1: Long Papers)*. Mexico City, Mexico: Association for Computational Linguistics, June 2024, pp. 870–883. URL: <https://aclanthology.org/2024.naacl-long>. 50
- Aarohi Srivastava et al. "Beyond the Imitation Game: Quantifying and extrapolating the capabilities of language models". In: *Transactions on Machine Learning Research* (2023). URL: <https://openreview.net/forum?id=uyTL5Bvosj>
- Ali Safaya and Engin Erzin. "HuBERT-TR: Reviving Turkish Automatic Speech Recognition with Self-supervised Speech Representation Learning". In: *Computing Research Repository* arXiv:2210.07323v2 (2022).

DOI: 10.48550/arXiv.2210.07323. URL: <http://arxiv.org/abs/2210.07323v2>

- Ali Safaya et al. “Mukayese: Turkish NLP Strikes Back”. In: *Findings of the Association for Computational Linguistics: ACL 2022*. Dublin, Ireland: Association for Computational Linguistics, May 2022, pp. 846–863. DOI: 10.18653/v1/2022.findings-acl.69. URL: <https://aclanthology.org/2022.findings-acl.69>
- Ali Hürriyetoğlu et al. “Event Coreference Resolution for Contentious Politics Events”. In: *Computing Research Repository* arXiv:2203.10123 (2022). DOI: 10.48550/ARXIV.2203.10123. URL: <http://arxiv.org/abs/2203.10123>
- Ali Safaya, Moutasem Abdullatif, and Deniz Yuret. “KUISAIL at SemEval-2020 Task 12: BERT-CNN for Offensive Speech Identification in Social Media”. In: *Proceedings of the Fourteenth Workshop on Semantic Evaluation*. Barcelona (online): International Committee for Computational Linguistics, 2020, pp. 2054–2059. URL: <https://aclanthology.org/2020.semeval-1.271>
- Ali Hürriyetoğlu et al. “Automated Extraction of Socio-political Events from News (AESPEN): Workshop and Shared Task Report”. In: *Proceedings of the Workshop on Automated Extraction of Socio-political Events from News 2020*. Marseille, France: European Language Resources Association (ELRA), 2020, pp. 1–6. ISBN: 979-10-95546-50-4. URL: <https://aclanthology.org/2020.aespen-1.1>
- Ali Hürriyetoğlu et al. “COVCOR20 at WNUT-2020 Task 2: An Attempt to Combine Deep Learning and Expert rules”. In: *Proceedings of the Sixth Workshop on Noisy User-generated Text (W-NUT 2020)*. Online: Association for Computational Linguistics, 2020, pp. 495–498. DOI: 10.18653/v1/2020.wnut-1.75. URL: <https://aclanthology.org/2020.wnut-1.75>
- Ali Safaya. “Event Sentence Detection Task using Attention model”. In: *Working Notes of CLEF 2019 - Conference and Labs of the Evaluation Forum, 07*. Lugano, Switzerland: CEUR Workshop Proceedings, 2019. URL: http://ceur-ws.org/Vol-2380/paper_122.pdf

العروض التقديمية

- Ali Safaya and Taner Sezer. *Turkish Data Depository Project: Towards a Unified Turkish NLP Research Platform*. Seminar presented at KUIS AI meeting, Istanbul (online). 2022
- Ali Safaya. *Reproducing Single Headed Attention - Recurrent Neural Networks*. Seminar presented at KUIS AI meeting, Istanbul (online). 2020

اللغات

- العربية — لغة أم، إتقان تام
- الإنجليزية — إتقان عملي
- التركية — لغة أم، إتقان تام

الخبرة التقنية

PyTorch, Jax, CUDA

الجوائز والتكريمات

- زمالة مركز KUIS للذكاء الاصطناعي (2020 — 2026)
- المركز الأول في هاثون معالجة اللغة الطبيعية الوطني في تركيا (2021)
- منحة Inzva للدكتوراه في هندسة الحاسوب (2020 — 2021)
- التخرج بمرتبة الشرف من برنامج البكالوريوس (الأول على دفعة 2019)
- المنحة التركية الحكومية (2015 — 2019)