


Research Article

Tren Penelitian Machine Learning dalam Kajian Bahasa Arab: Analisis Bibliometrik

Roobiatul Adawiyah¹, Luthfatul Qibthiyah²

1. Universitas Al Amien prenduan, Indonesia
E-mail: robiatuladw2005@gmail.com 
2. Universitas Al Amien prenduan, Indonesia
E-mail: lutfatulqibthiyah@gmail.com



Copyright © 2026 by Authors, Published by AL-AFKAR: Journal For Islamic Studies. This is an open access article under the CC BY License (<https://creativecommons.org/licenses/by/4.0>).

Received : April 12, 2026
Accepted : June 13, 2026

Revised : May 14, 2026
Available online : July 11, 2026

How to Cite: Roobiatul Adawiyah and Luthfatul Qibthiyah. (2026) "Research Trends in Machine Learning for Arabic Language Studies: A Bibliometric Analysis", *al-Afkar, Journal For Islamic Studies*, 9(3), pp. 373-381. doi: 10.31943/afkarjournal.v9i3.3720.

Research Trends in Machine Learning for Arabic Language Studies: A Bibliometric Analysis

Abstract. This study aims to map the development of publications on the application of *Machine Learning* in Arabic language studies through bibliometric analysis. Research data was obtained from the Lens.org database and analyzed using VOSviewer software to identify publication trends, keyword networks, and research opportunities. The results show that the number of publications has increased significantly since 2015 and peaked in the 2020–2025 period. Keyword co-occurrence analysis shows that *Machine Learning* is a central theme closely related to deep learning, Arabic Natural Language Processing, classification, and Arabic Sentiment Analysis. Overlay visualizations show a shift in research focus from basic topics, such as translation, tagging, and text mining, to more sophisticated

themes, such as Arabic hate speech, clustering, and ensemble learning. Meanwhile, density analysis indicates that Arabic Sign Language, FastText, Ensemble Learning, and Arabic Hate Speech still have a low publication density, thus potentially becoming future research directions. The findings of this study provide a comprehensive overview of the trends and developments in *Machine Learning* research in Arabic language studies and can serve as a reference for researchers in determining research themes and developing artificial intelligence-based innovations in Arabic language learning and processing.

Keywords: Machine Learning, Bibliometrics, Arabic, VOSviewer, Arabic Natural Language Processing

Abstrak: Penelitian ini bertujuan memetakan perkembangan publikasi mengenai penerapan *Machine Learning* dalam kajian Bahasa Arab melalui analisis bibliometrik. Data penelitian diperoleh dari basis data Lens.org dan dianalisis menggunakan perangkat lunak VOSviewer untuk mengidentifikasi tren publikasi, jaringan kata kunci, serta peluang penelitian. Hasil penelitian menunjukkan bahwa jumlah publikasi mengalami peningkatan signifikan sejak tahun 2015 dan mencapai puncaknya pada periode 2020–2025. Analisis ko-okurensi kata kunci memperlihatkan bahwa *Machine Learning* menjadi tema sentral yang berkaitan erat dengan deep learning, Arabic Natural Language Processing, classification, dan Arabic Sentiment Analysis. Visualisasi overlay menunjukkan pergeseran fokus penelitian dari topik dasar, seperti translation, tagging, dan text mining, menuju tema yang lebih mutakhir, seperti Arabic hate speech, clustering, dan ensemble learning. Sementara itu, analisis density mengindikasikan bahwa topik Arabic Sign Language, FastText, Ensemble Learning, dan Arabic Hate Speech masih memiliki tingkat kepadatan publikasi yang rendah sehingga berpotensi menjadi arah penelitian selanjutnya. Temuan penelitian ini memberikan gambaran komprehensif mengenai tren dan perkembangan riset *Machine Learning* dalam kajian Bahasa Arab serta dapat menjadi referensi bagi peneliti dalam menentukan tema penelitian dan mengembangkan inovasi berbasis kecerdasan buatan pada pembelajaran maupun pengolahan Bahasa Arab.

Kata Kunci: Machine Learning, Bibliometrik, Bahasa Arab, VOSviewer, Arabic Natural Language Processing

PENDAHULUAN

Beberapa dekade terakhir ini teknologi mengalami peningkatan yang sangat pesat. Sehingga terciptalah *artificial intelligence* (AI) yang dimana istilah ini pertama kali ditemukan pada tahun 1956 oleh perintis AI seperti John Mc Carthy, Marvin Minsky, Nathaniel Rochester dan Claude Shannon pada konferensi Darmouth dan diresmikan dibidang akademis dan penelitian ilmiah.¹ AI diciptakan dengan tujuan untuk mengalahkan kecerdasan manusia dibidang kognitif di masa mendatang. Namun perkembangan AI ini sempat mengalami keterbatasan sehingga tidak berjalan secara efektif.² Kemudian pada tahun 1980-1990 AI Kembali hidup dengan adanya Teknik pembelajaran mesin (*Machine Learning*) yang menjadi salah satu cabang AI yang lebih efisien.

¹ Sudirwo S.E M.M, Abdul Hadi S.E M.Si, Loso judijanto Ssi M.M M.Stats, Nuraini Purwandari S.T MMSI, "ARTIFICIAL INTELLEGENCE teori, konsep dan implementasi dalam berbagai bidang" (Jambi, 2025).

² Tia Siti Nurazizah, "ARTIFICIAL INTELLIGENCE DAN MACHINE LEARNING DALAM KEHIDUPAN MANUSIA" (t.t.).

Machine Learning (ML) sendiri merupakan sistem yang dilakukan oleh komputer untuk mempelajari pola dari data atau pengalaman tanpa diprogram dan diarahkan langsung oleh manusia.³ Biasanya ML ini meniru Langkah-langkah yang dilakukan manusia dengan cara memprediksi. Keterkaitan antara Artificial Intelligence dan *Machine Learning* adalah dimana Artificial Intelligence (AI) merupakan tools besar yang berisi *Machine Learning (ML)* didalamnya. *Machine Learning* dapat digunakan dalam berbagai aspek termasuk Pendidikan dan linguistik. Dalam pengolahan bahasa *Machine Learning (ML)* memiliki cabang yang dapat membantu untuk memahami, menafsirkan dan juga merespon bahasa manusia secara alami baik secara lisan maupun tulisan yang disebut Natural Language Processing (NLP).⁴

Pada pembelajaran Bahasa Arab Natural Language Processing sering digunakan untuk mendeteksi gramatikal, struktur kalimat, mengidentifikasi kategori morfologis kata (misalnya *fi'il*, *isim*, dan huruf) dan menerjemahkan.⁵ Ini sangat membantu dalam proses mempelajari Bahasa Arab. Sebagai salah satu bahasa yang sulit, bahasa arab ini memiliki beberapa tantangan dalam mempelajarinya seperti keterbatasan sarana dan prasarana dan juga guru bahasa arab yang kompeten. Dengan adanya kemajuan teknologi mampu membuka ruang inovasi yang menghadirkan keragaman metode pembelajaran dengan menggunakan perangkat lunak yang disediakan oleh *Machine Learning*. Beberapa tools yang disediakan sangat memudahkan karena sumber yang lebih luas, bersifat fleksibel, mudah diakses.⁶

Penggunaan *Machine Learning* pada linguistik khususnya Bahasa Arab berkembang secara signifikan sehingga banyak peneliti yang tertarik untuk melakukan penelitian dan publikasi terkait *Machine Learning* dan Bahasa Arab dalam beberapa tahun terakhir. berbagai studi di bidang ini menyoroti fokus dan tema penelitian yang bervariasi. Mulai dari penelitian tentang penerjemah otomatis, gramatikal, dan morfologis pada Bahasa Arab. keragaman tersebut menunjukkan bahwa *Machine Learning* pada Bahasa Arab memiliki eksistensi yang cukup tinggi dibidang penelitian dan banyak dimanfaatkan dalam pengembangan studi Bahasa Arab. Namun pertumbuhan publikasi yang pesat dan beragamnya fokus penelitian menimbulkan tantangan tersendiri bagi peneliti untuk menelusuri arah perkembangan apabila dilakukan secara manual. Oleh karena itu dibutuhkan metode yang dapat memetakan dan memberikan gambaran secara menyeluruh mengenai perkembangan tren penelitian, fokus kajian yang dominan, dan peluang penelitian di masa depan dapat diketahui secara lebih sistematis.

Tujuan Penelitian ini adalah untuk menganalisis perkembangan publikasi *Machine Learning* dalam studi Bahasa Arab, mengidentifikasi kata kunci yang

³ Agung Wijoyo dkk., "Pembelajaran Machine Learning," vol.3, no. 2 (2024).

⁴ Aish Albladi dkk., "Hate Speech Detection Using Large Language Models: A Comprehensive Review," *IEEE Access*, vol.13 (2025): 20871-20892.

⁵ Lamiyatul Bayan, "Arsitektur Digital Bahasa Arab: Transformasi Komunikasi Multimodal dalam Lanskap Intelektual Artifisial," *Journal of Arabic Studies*, vol.1, no. 1 (2026).

⁶ Muhammad Azhar dkk., "Revolusi Pembelajaran Bahasa Arab Di Era Digital: Analisis Peran Artificial Intelligence Dalam Pengembangan Kompetensi Berbahasa," *AL-MA'LUMAT: JURNAL ILMU-ILMU KEISLAMAN*, vol.3, no. 2 (1 Agustus 2025): 58-81.

dominan, tema-tema penelitian yang berkembang, memetakan kontribusi penulis, institusi dan negara yang berperan pada pengembangan penelitian *Machine Learning* pada kajian Bahasa Arab.

METODE PENELITIAN

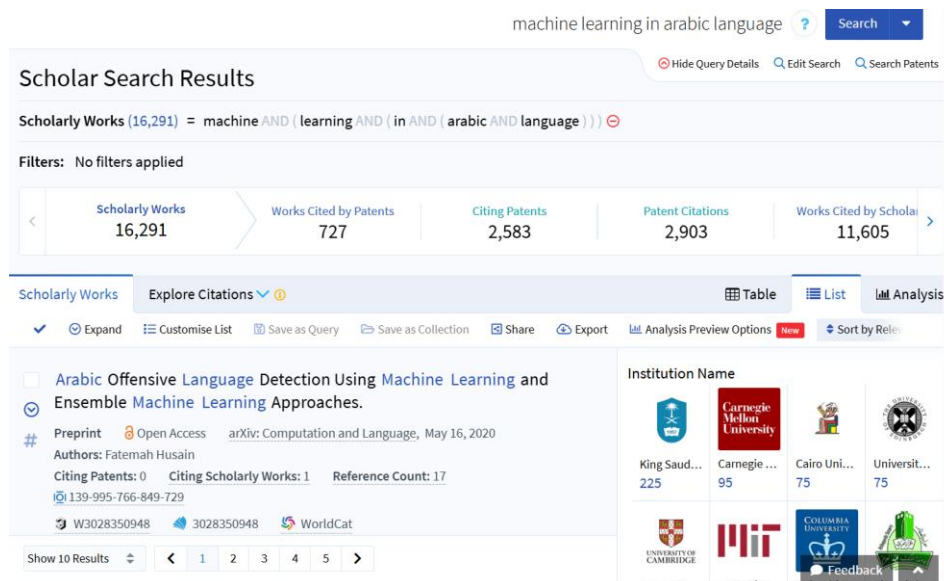
Penelitian ini menggunakan bibliometrik yaitu metode yang digunakan untuk menganalisis sejumlah data ilmiah publikasi dan sitasi.⁷ Menurut (Diodato,1999) Bibliometrik berasal dari kata *biblio* atau *bibliography* dan *metrics*. Biblio atau bibliografi berarti buku dan metrics berarti sesuatu yang berkaitan dengan mengukur. Jadi bibliometrika (bibliometrics) berarti mengukur atau menganalisis buku atau literatur dengan menggunakan pendekatan statistika.⁸ Penelitian ini dilakukan dengan cara mengumpulkan database dari Lens.org dengan kata kunci "*Machine Learning* in arabic language, kemudian data tersebut dimasukkan kedalam aplikasi yang digunakan untuk melakukan analisis bibliometrik yaitu VosViewer. Dari database yang dikumpulkan akan menghasilkan kata kunci yang paling dominan, peneliti yang sering berkontribusi pada publikasi *Machine Learning* pada bahasa arab dan tahun yang menunjukkan eksistensi penelitian mengenai *Machine Learning* pada bahasa arab. Itu ditandai dengan warna merah yang paling mencolok dan warna lain untuk yang tidak dominan. Untuk hitungan tahun dapat dilihat dari semakin terang dan besarnya warna kuning pada latar belakang biru tua. Peneliti menjadikan Lens.org sebagai database karena platform tersebut sudah menyediakan grafik-grafik frekuensi publikasi yang memudahkan peneliti.

HASIL DAN PEMBAHASAN

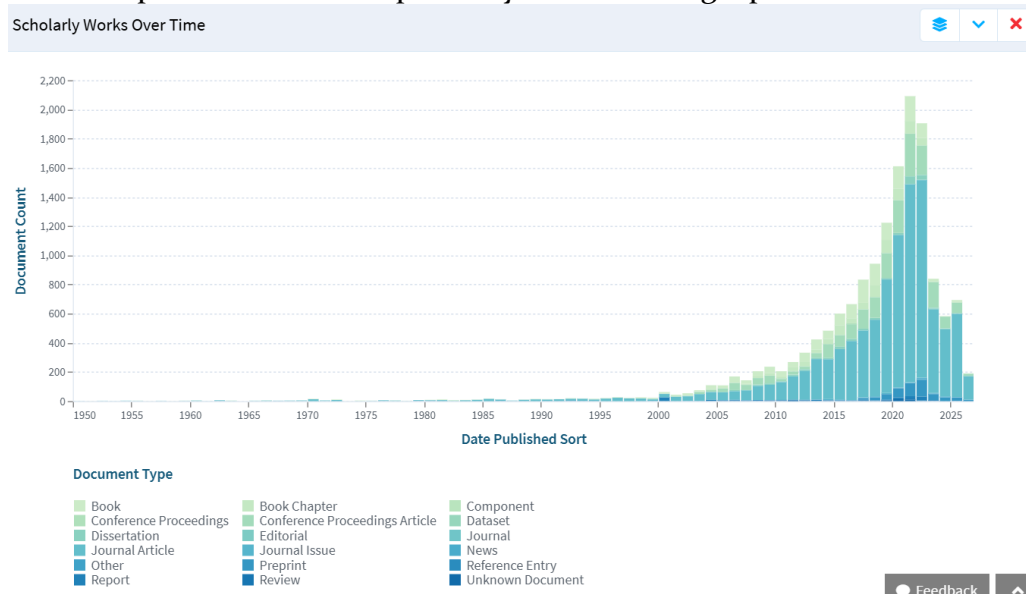
Untuk menemukan hasil penelitian, peneliti memulai dengan menumpulkan database pada platform Lens.org kemudian memilih tipe data "scholarly works" dan memasukkan kata kunci yaitu "*Machine Learning* In Arabic Language". Pada gambar 1 menunjukkan hasil yang keluar terdapat 18.291 (delapan belas ribu dua ratus Sembilan puluh satu) jumlah total artikel ilmiah mengenai *Machine Learning*, 727 (tujuh ratus dua puluh tujuh) jumlah karya ilmiah yang pernah dikutip atau digunakan sebagai referensi didalam dokumen resmi, 2.583 (dua ribu lima ratus delapan puluh tiga) jumlah total dokumen yang melakukan kutipan terhadap 727 karya ilmiah diatas, 2.903 (dua ribu Sembilan ratus tiga) jumlah total aktivitas kutipan yang dilakukan oleh dokumen-dokumen paten. 11.605 (sebelas ribu enam ratus lima) jumlah karya ilmiah dari total pencarian peneliti yang dikutip Kembali oleh peneliti lain dalam artikel ilmiah mereka.

⁷ Pisuko Herawati - dkk., "ANALISIS BIBLIOMETRIK: PERKEMBANGAN PENELITIAN DAN PUBLIKASI MENGENAI KOORDINASI PROGRAM MENGGUNAKAN VOSVIEWER," *Jurnal Pustaka Budaya*, vol.9, no. 1 (4 Januari 2022): 1-8.

⁸ Sitti Husaebah Pattah, "PEMANFAATAN KAJIAN BIBLIOMETRIKA SEBAGAI METODE EVALUASI DAN KAJIAN DALAM ILMU PERPUSTAKAAN DAN INFORMASI," vol.1, no. 1 (2013).

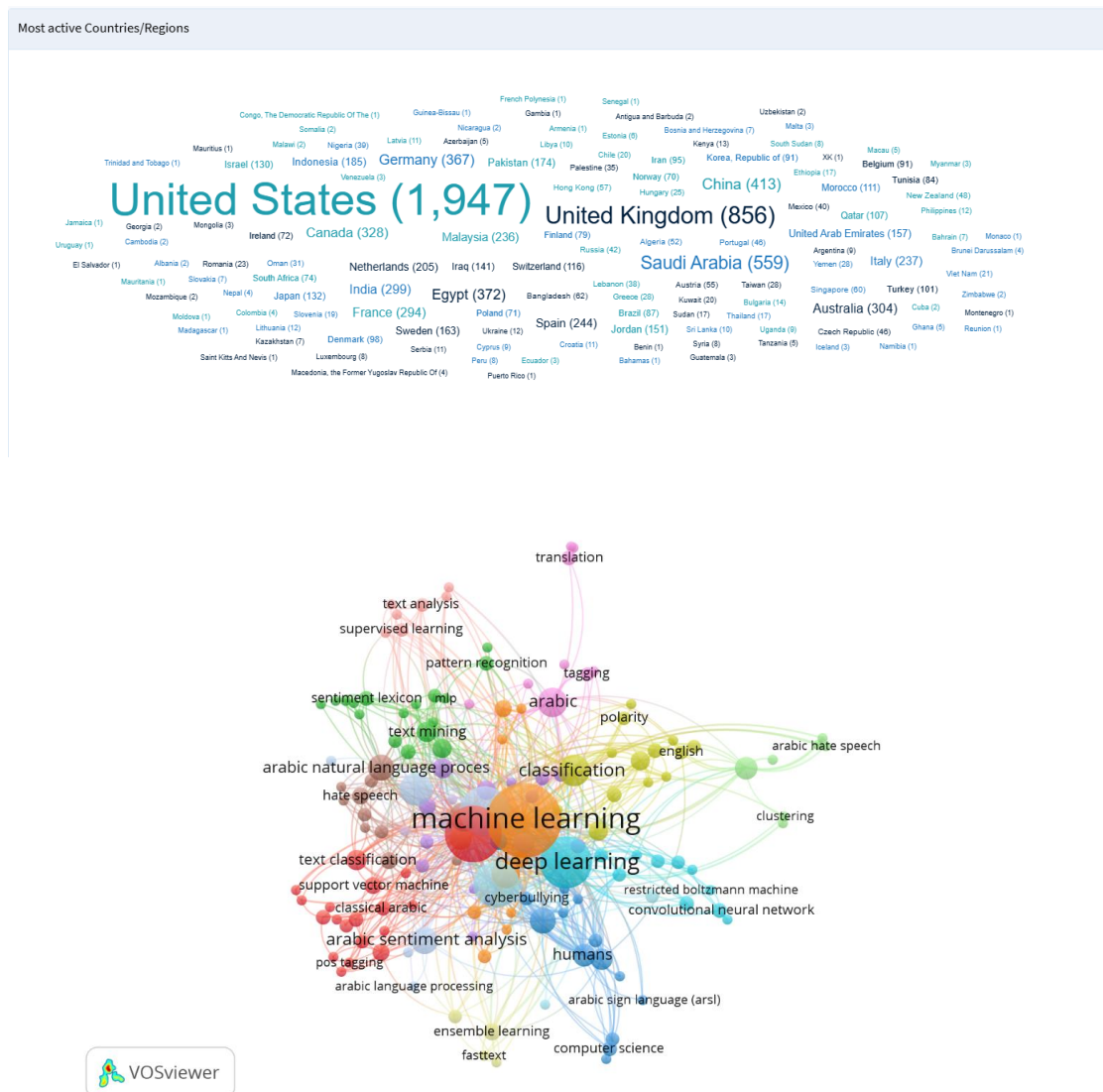


Publikasi mengenai *Machine Learning* dalam studi bahasa arab pada tahun 1950-2000 masih sangat sedikit dan belum mengalami perkembangan, kemudian pada tahun 2000-2015 sudah mulai ada peningkatan publikasi secara bertahap, pada tahun 2015-2025 peningkatan publikasi meningkat cukup pesat terutama pada tahun 2020 dan sedikit mengalami penurunan lagi pada tahun 2025, penurunan ini bisa saja terjadi karena publikasi belum sepenuhnya tercatat lengkap didalam database.



Negara-negara paling aktif (TOP 5) yang berkontribusi dalam publikasi karya ilmiah pada kata kunci *Machine Learning* in arabic language adalah:

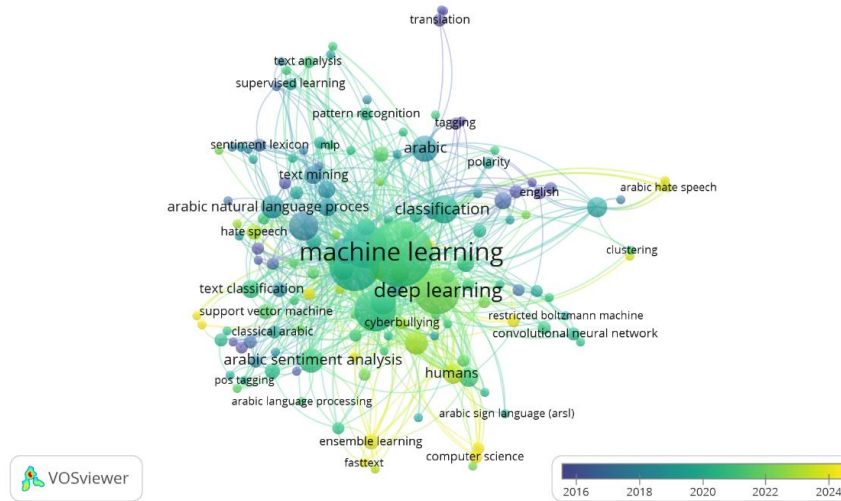
1. United States dengan jumlah dokumen 1.947 (seribu sembilan ratus empat puluh tujuh).
2. United Kingdom dengan jumlah dokumen 856 (delapan ratus lima puluh enam).
3. Saudi Arabia dengan jumlah dokumen 559 (lima ratus lima puluh sembilan).
4. China dengan jumlah dokumen 413 (empat ratus tiga belas).
5. Egypt dengan jumlah dokumen (tiga ratus tujuh puluh dua).



Berdasarkan hasil visualisasi jaringan menggunakan VOSviewer, terlihat bahwa kata kunci *Machine Learning* memiliki ukuran node terbesar dan berada di pusat jaringan. Hal ini menunjukkan bahwa *Machine Learning* merupakan topik utama yang paling banyak dibahas dalam penelitian terkait Bahasa Arab.

Kata kunci *Machine Learning* memiliki hubungan yang kuat dengan beberapa kata kunci lain, seperti *deep learning*, *classification*, *Arabic*, *Arabic Natural Language Processing*, dan *Arabic Sentiment Analysis*. Keterhubungan ini menunjukkan bahwa penerapan *Machine Learning* dalam kajian Bahasa Arab banyak digunakan untuk klasifikasi teks, analisis sentimen, dan pemrosesan bahasa alami.

Selain itu, visualisasi menunjukkan adanya beberapa kelompok penelitian yang saling berkaitan. Kelompok pertama berfokus pada pemrosesan bahasa alami Bahasa Arab (*Arabic Natural Language Processing*), kelompok kedua berfokus pada analisis sentimen Bahasa Arab, sedangkan kelompok lainnya berkaitan dengan penggunaan metode *deep learning* dalam pengolahan data bahasa Arab. Temuan ini menunjukkan bahwa penelitian *Machine Learning* dalam kajian Bahasa Arab berkembang pada berbagai bidang dan aplikasi.

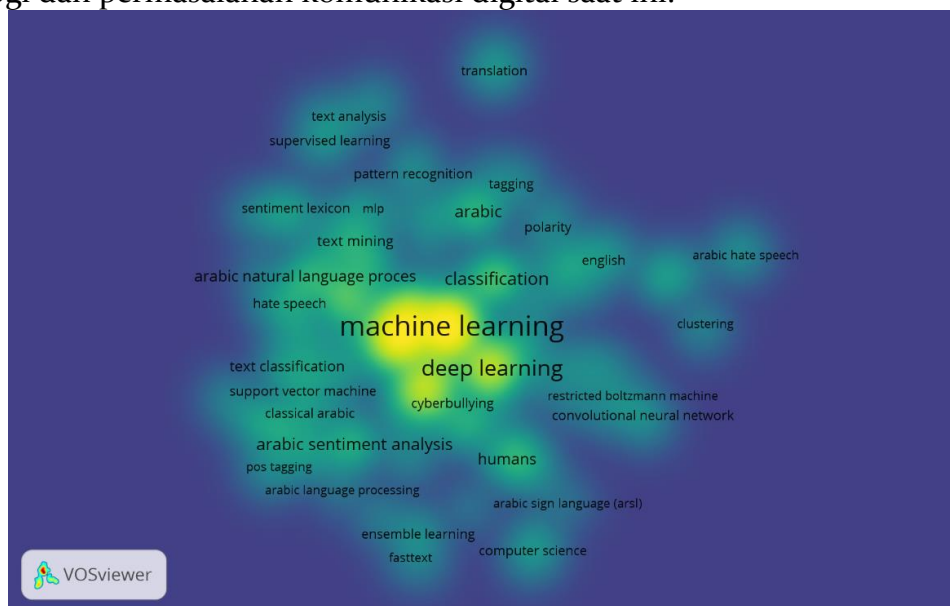


Visualisasi overlay digunakan untuk melihat perkembangan topik penelitian dari waktu ke waktu. Warna biru menunjukkan topik yang lebih lama diteliti, sedangkan warna kuning menunjukkan topik yang lebih baru.

Berdasarkan hasil visualisasi, topik seperti translation, tagging, dan text mining termasuk dalam kelompok penelitian yang lebih awal berkembang. Topik-topik tersebut menjadi dasar dalam pengembangan teknologi pengolahan Bahasa Arab.

Sementara itu, topik seperti deep learning, classification, dan Arabic Sentiment Analysis mulai berkembang pada periode berikutnya dan menjadi fokus penelitian yang cukup dominan.

Pada tahun-tahun terbaru, muncul topik seperti Arabic hate speech, clustering, dan ensemble learning. Kemunculan topik tersebut menunjukkan bahwa penelitian *Machine Learning* pada Bahasa Arab terus berkembang mengikuti kebutuhan teknologi dan permasalahan komunikasi digital saat ini.



Visualisasi densitas digunakan untuk melihat topik yang paling sering muncul dalam publikasi. Warna kuning menunjukkan tingkat kemunculan yang tinggi,

sedangkan warna hijau dan biru menunjukkan tingkat kemunculan yang lebih rendah.

Hasil visualisasi menunjukkan bahwa kata kunci *Machine Learning* dan *deep learning* memiliki tingkat kepadatan tertinggi. Hal ini menunjukkan bahwa kedua topik tersebut menjadi fokus utama dalam penelitian *Machine Learning* pada kajian Bahasa Arab.

Selain itu, kata kunci *classification*, *Arabic Natural Language Processing*, dan *Arabic Sentiment Analysis* juga memiliki tingkat kepadatan yang cukup tinggi. Temuan ini menunjukkan bahwa penelitian banyak diarahkan pada klasifikasi teks, analisis sentimen, dan pengolahan bahasa alami.

Di sisi lain, kata kunci seperti *Arabic Sign Language*, *FastText*, *Ensemble Learning*, dan *Arabic Hate Speech* memiliki tingkat kepadatan yang lebih rendah. Kondisi ini menunjukkan bahwa topik-topik tersebut masih memiliki peluang untuk dikembangkan pada penelitian selanjutnya.

KESIMPULAN

Penelitian ini berhasil memetakan perkembangan publikasi mengenai *Machine Learning* dalam kajian Bahasa Arab melalui pendekatan bibliometrik menggunakan database *Lens.org* dan perangkat lunak *VOSviewer*. Hasil penelitian menunjukkan bahwa jumlah publikasi mengalami peningkatan yang signifikan sejak tahun 2015 hingga mencapai puncaknya pada periode 2020–2025. Temuan ini mengindikasikan bahwa *Machine Learning* telah menjadi salah satu bidang yang berkembang pesat dalam studi Bahasa Arab seiring dengan kemajuan teknologi kecerdasan buatan.

Analisis jaringan kata kunci menunjukkan bahwa *Machine Learning* menjadi tema sentral yang memiliki keterkaitan kuat dengan *deep learning*, *Arabic Natural Language Processing*, *classification*, dan *Arabic Sentiment Analysis*. Selain itu, visualisasi *overlay* memperlihatkan adanya pergeseran fokus penelitian dari topik-topik dasar, seperti *translation*, *tagging*, dan *text mining*, menuju topik yang lebih mutakhir, seperti *Arabic hate speech*, *clustering*, dan *ensemble learning*. Sementara itu, analisis *density* menunjukkan bahwa topik-topik seperti *Arabic Sign Language*, *FastText*, *Ensemble Learning*, dan *Arabic Hate Speech* masih memiliki tingkat kepadatan publikasi yang relatif rendah sehingga berpotensi menjadi arah penelitian pada masa mendatang.

Secara praktis, hasil penelitian ini memberikan gambaran yang komprehensif mengenai tren, fokus, serta perkembangan riset *Machine Learning* dalam kajian Bahasa Arab. Informasi tersebut dapat menjadi referensi bagi peneliti dalam menentukan tema penelitian yang masih terbuka, memperluas kolaborasi ilmiah, serta mendukung pengembangan inovasi teknologi berbasis kecerdasan buatan untuk pembelajaran dan pengolahan Bahasa Arab. Penelitian selanjutnya disarankan menggunakan lebih dari satu basis data, seperti *Scopus*, *Web of Science*, atau *Dimensions*, sehingga mampu menghasilkan pemetaan bibliometrik yang lebih luas, komprehensif, dan representatif.

DAFTAR PUSTAKA

- Agung Wijoyo dkk., “Pembelajaran Machine Learning,” vol.3, no. 2 (2024).
- Aish Albladi dkk., “Hate Speech Detection Using Large Language Models: A Comprehensive Review,” *IEEE Access*, vol.13 (2025): 20871–20892.
- Lamiyatul Bayan, “Arsitektur Digital Bahasa Arab: Transformasi Komunikasi Multimodal dalam Lanskap Intelegensia Artifisial,” *Journal of Arabic Studies*, vol.1, no. 1 (2026).
- Muhammad Azhar dkk., “Revolusi Pembelajaran Bahasa Arab Di Era Digital: Analisis Peran Artificial Intelligence Dalam Pengembangan Kompetensi Berbahasa,” *AL-MA’LUMAT : JURNAL ILMU-ILMU KEISLAMAN*, vol.3, no. 2 (1 Agustus 2025): 58–81.
- Muhammad Rizky. (2025). Problems in the Educational Background of Arabic Language Students at the University of Jambi. *Al-Kalimat: Journal of Linguistic and Arabic Teaching*, 1(2), 102–107. <https://doi.org/10.61166/alkalimat.vii2.21>
- Pisuko Herawati - dkk., “ANALISIS BIBLIOMETRIK: PERKEMBANGAN PENELITIAN DAN PUBLIKASI MENGENAI KOORDINASI PROGRAM MENGGUNAKAN VOSVIEWER,” *Jurnal Pustaka Budaya*, vol.9, no. 1 (4 Januari 2022): 1–8.
- Sitti Husaebah Pattah, “PEMANFAATAN KAJIAN BIBLIOMETRIKA SEBAGAI METODE EVALUASI DAN KAJIAN DALAM ILMU PERPUSTAKAAN DAN INFORMASI,” vol.1, no. 1 (2013).
- Sudirwo S.E M.M, Abdul Hadi S.E M.Si, Loso judijanto Ssi M.M M.Stats, Nuraini Purwandari S.T MMSI, “ARTIFICIAL INTELLIGENCE teori, konsep dan implementasi dalam berbagai bidang” (Jambi, 2025).
- Tia Siti Nurazizah, “ARTIFICIAL INTELLIGENCE DAN MACHINE LEARNING DALAM KEHIDUPAN MANUSIA” (t.t.).