

# Tren Penelitian Klasifikasi Sentimen Pelanggan Berbasis Machine Learning

## Machine Learning Based Customer Sentiment Classification Research Trends

Rohimatul Maisyaroh<sup>1</sup>, Nur Azizah<sup>2</sup>, Nurul Hidayat<sup>3</sup>, Hozairi<sup>4</sup>

<sup>1234</sup>Teknik Informatika, Fakultas Teknik, Universitas Islam Madura, Pamekasan



Copyright © Jurnal  
Rekayasa Lampung  
(JRL)

**Abstrak:** Pesatnya perkembangan industri online reviews mendorong kebutuhan akan teknologi cerdas, seperti machine learning, untuk menganalisis sentimen pelanggan secara efektif. Penelitian ini bertujuan untuk mengidentifikasi tren publikasi, pola kolaborasi, fokus penelitian, dan tren sitasi terkait klasifikasi sentimen pelanggan berbasis machine learning dalam konteks online reviews pada database Scopus selama periode 2020–2025. Metode yang digunakan adalah analisis bibliometrik dengan pendekatan visualisasi dan deskriptif, meliputi lima tahap utama: penentuan kata kunci, pencarian awal, penyempurnaan hasil pencarian, kompilasi statistik, dan analisis data. Dari total 2.101 publikasi awal, diperoleh 441 dokumen relevan setelah proses penyaringan. Hasil menunjukkan peningkatan publikasi signifikan pada 2023, dengan dominasi bidang ilmu Computer Science (45,7%). India dan China tercatat sebagai negara dengan kontribusi publikasi tertinggi, sedangkan IEEE Access menjadi sumber publikasi terkemuka. Analisis co-word mengidentifikasi topik dominan seperti machine learning, electronic commerce, dan deep learning. Temuan ini memberikan gambaran menyeluruh mengenai peta riset global, serta menjadi landasan penting untuk pengembangan teknologi dan strategi bisnis berbasis machine learning dalam analisis online reviews ke depannya.

**Kata kunci:** machine learning, klasifikasi sentimen, analisis bibliometrik, online reviews, Scopus

**Abstract:** The rapid development of the online reviews industry has driven the need for intelligent technologies, such as machine learning, to analyze customer sentiment effectively. This study aims to identify publication trends, collaboration patterns, research focus, and citation trends related to machine learning-based customer sentiment classification in the context of online reviews on the Scopus database during the period 2020–2025. The method used is bibliometric analysis with a visualization and descriptive approach, including five main stages: keyword determination, initial search, search result refinement, statistical compilation, and data analysis. From a total of 2,101 initial publications, 441 relevant documents were obtained after the filtering process. The results show a significant increase in publications in 2023, dominated by the Computer Science field (45.7%). India and China are recorded as the countries with the highest publication contributions, while IEEE Access is the leading source of publications. Co-word analysis identifies dominant topics such as machine learning, electronic commerce, and deep learning. These findings provide a comprehensive overview of the global research map, as well as an important foundation for the development of machine learning-based technology and business strategies in online review analysis in the future.

**Keywords:** machine learning, sentiment classification, bibliometric analysis, online reviews, Scopus

## 1. PENDAHULUAN

Kini industri online reviews telah mengalami pertumbuhan yang luar biasa dalam beberapa tahun terakhir, dengan jutaan konsumen berbondong-bondong ke internet untuk melakukan pembelian. Namun, mempertahankan pelanggan dan mencegah mereka "berpindah" atau beralih ke pesaing merupakan tantangan besar bagi bisnis online reviews[1]. Online reviews menjadi semakin populer karena kemudahan, jangkauan produk yang luas, dan aksesibilitas global. ini juga menyediakan lingkungan yang menguntungkan untuk penerapan teknik pembelajaran mesin dan pembelajaran mendalam, karena ketersediaan kumpulan data yang besar, kebutuhan akan pengalaman yang personalisasi, tantangan penipuan, deteksi dan keamanan, potensi

optimasi rantai pasokan, dan pentingnya analisis sentimen pelanggan [2]. Tantangan penipuan yang masih sering terjadi seperti Situs web malware yang dirancang agar terlihat sama dengan situs asli yang ingin ditiru. Situs web palsu ini memiliki hyperlink yang mencolok yang digunakan untuk menipu korbannya, dan mereka menautkan ke situs yang tampak mirip dengan situs asli untuk mengelabui pengguna agar memercayai situs tersebut.[3]

Sistem komputer juga memengaruhi dan meningkatkan interaksi antara konsumen dan organisasi bisnis. Dengan cara ini, pergeseran ke arah penggunaan teknologi yang lebih baik telah menghasilkan terciptanya sistem cerdas yang dapat mengelola dan memantau model bisnis dengan keterlibatan manusia yang lebih sedikit. Sistem AI yang menunjukkan kemampuan untuk

memenuhi permintaan konsumen di banyak sektor dibutuhkan dalam perekonomian saat ini. AI memainkan peran penting dalam memantau lingkungan bisnis, mengidentifikasi kebutuhan pelanggan, dan menerapkan strategi yang diperlukan tanpa atau dengan campur tangan manusia yang minimal [4]. Di era sekarang ini, sebagian besar orang menggunakan ponsel sebagai alat untuk memudahkan hampir segala hal, baik itu bekerja, belajar, atau bahkan berbelanja. Tidak mengherankan jika penjualan melalui sistem online juga sangat menguntungkan bagi para pedagang. Oleh karena itu, e-marketplace memungkinkan transaksi komersial melalui pertukaran informasi komersial, pemeliharaan hubungan komersial, pencapaian negosiasi komersial, dan penyelesaian serta pelaksanaan perjanjian melalui sarana elektronik menggunakan Internet [5].

Dunia layanan digital terus berkembang, dengan fokus pada penerapan pembelajaran mesin pada aktivitas layanan digital ini. Pembelajaran mesin dapat membantu memerangi penipuan dengan tiga cara: deteksi dan intersepsi, penuntutan, dan pencegahan. Seperti yang dinyatakan oleh Najem dan Kadhem (2021), pembelajaran mesin adalah salah satu metode paling populer untuk mendeteksi penipuan dalam online reviews. Dua teknik pembelajaran mesin yang paling populer adalah pembelajaran terbimbing, di mana algoritme dilatih menggunakan contoh data masukan dan keluaran yang telah diberi label oleh manusia, dan pembelajaran tanpa pengawasan, di mana algoritme tidak diberi data berlabel agar dapat menemukan struktur dalam data masukan [6].

Fitur utama yang diperlukan untuk memahami perilaku pelanggan, seperti pengetahuan ahli, fitur-fitur ini adalah analisis sentimen (fitur "teks" kata), fitur URL (halaman pelanggan, sesi, halaman patches), fitur historis dan komentar pada tweeter, fitur kategoris seperti nama bulan dan hari-hari khusus, dan fitur numerik seperti fitur binomial yang memberikan benar atau salah (terjadi atau tidak) [7]. Perilaku dan sifat yang diuraikan dalam penelitian ini harus dilatih sehingga dapat dikumpulkan dan ditangani di seluruh perilaku pelanggan.

Penelitian terkini menunjukkan bahwa penerapan machine learning dalam online reviews masih dalam tahap awal, tetapi telah menunjukkan hasil yang menjanjikan. Zhang dalam penelitiannya menyebutkan bahwa teknik machine learning seperti deep learning, reinforcement learning, dan transfer learning telah berhasil digunakan dalam berbagai aplikasi online reviews, termasuk rekomendasi produk, personalisasi, dan analisis sentimen. Oleh karena itu, penelitian ini bertujuan untuk mengisi kesenjangan ini dengan mengembangkan dan menguji algoritma machine learning baru yang

lebih efisien dan dapat dengan mudah diintegrasikan dengan sistem online reviews yang ada. Penelitian ini juga akan mengeksplorasi dampak jangka panjang penerapan teknologi ini terhadap kinerja bisnis online reviews.

## 2. BAHAN DAN METODE PENELITIAN

Metode penelitian yang digunakan adalah analisis bibliometrik terkait dengan Tren Penelitian Klasifikasi Sentimen Pelanggan Berbasis Machine Learning. Penelitian ini menggunakan metode analisis bibliometrik dengan pendekatan visualisasi dan deskriptif, yang melibatkan lima tahapan utama: Defining Search Keywords, Initial Search Result, Refinement of the Search Results, Compiling Statistics on the Initial Data, dan Data Analysis [8]. Berikut adalah ringkasan dari masing-masing tahapan tersebut:

### Langkah 1: Defining Search Keywords.

Tentukan keyword pencarian. Dengan mempertimbangkan tren penelitian klasifikasi sentimen pelanggan, naive bayes dan pembelajaran mesin, kami menggabungkan tiga keyword pencarian yang berbeda, termasuk: -naive bayes AND classification -sentiment analysis -(sentiment analysis OR sentiment classification) AND (online reviews OR customer reviews)

### Langkah 2 : Initial Search Result.

Hasil pencarian awal [9]. Dengan menggunakan pencarian "judul, abstrak, kata kunci" dalam basis data Koleksi SCOPUS, kami memilih publikasi yang mencakup istilah pencarian yang disebutkan di atas dan memperoleh 2646 hasil. Semua pengumpulan data diselesaikan pada tanggal 21 Juni 2025.

### Langkah 3 : Refinement of the Search Results.

Penyempurnaan Hasil Pencarian. dokumen Scopus memenuhi persyaratan, tanpa data yang tidak relevan atau duplikat, dengan total dokumen (100%). Sebaliknya, Google Scholar mengindeks banyak dokumen duplikat, dengan skor akhir 91% dari dokumen dengan. Meskipun melaporkan jumlah publikasi yang lebih besar daripada Scopus, Google Scholar menyertakan publikasi domestik, sebagian besar dalam bahasa Indonesia. Untuk analisis data yang cermat, analisis bibliometrik lebih lanjut dilakukan dengan menggunakan Scopus [10].

### Langkah 4 : Compiling Statistics on the Initial Data.

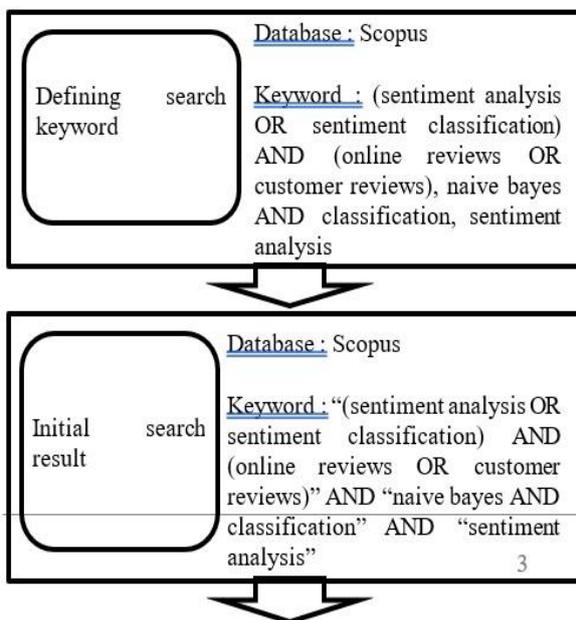
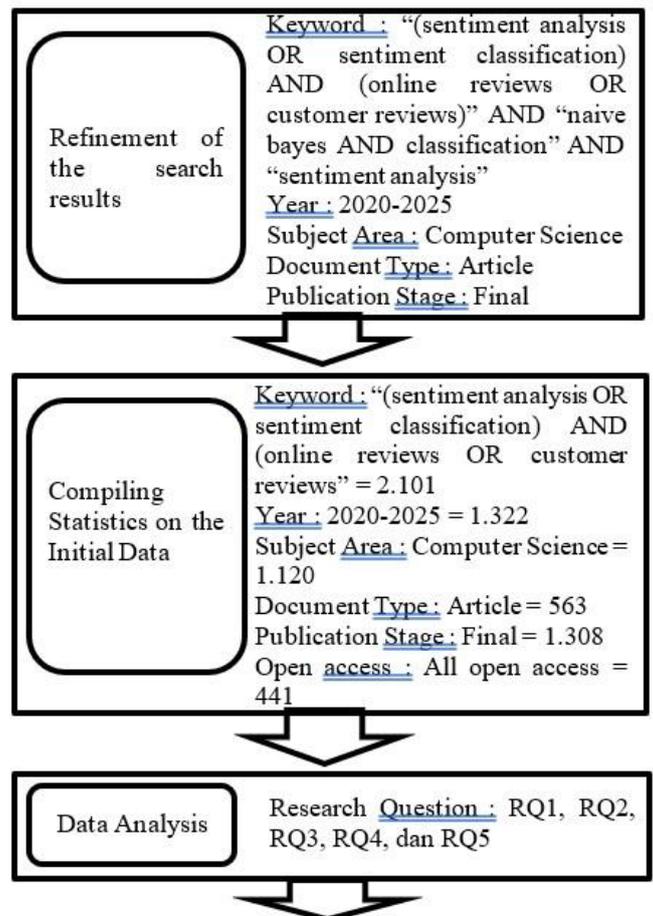
Mengkompilasi Statistik pada Data Awal. Pemetaan penulis dan negara memberikan informasi tentang pemimpin disiplin ilmu, sementara kata kunci menunjukkan topik hangat. Pola kolaborasi dan jaringan kutipan mewakili kerja tim dan literatur primer dari periode yang berbeda, menggambarkan evolusi bidang tersebut dan memproyeksikan perkembangan masa depan [10].

### Langkah 5 : Data Analysis.

Menganalisis data. melakukan analisis bibliometrik deskriptif dengan menggunakan beberapa indikator produktivitas dan dampak. Di bagian ini, publikasi dan struktur kutipan, jurnal dan kategori yang paling berpengaruh, penulis yang paling produktif dan dikutip, serta makalah yang paling banyak dikutip disajikan secara rinci [9].

(Tabel 1) : Analisis bibliometrik

Data	Initial Search	Refinement Search
Source Data	Scopus	Scopus
Keyword	(sentiment analysis OR sentiment classification) AND (online reviews OR customer reviews)	(sentiment analysis OR sentiment classification) AND (online reviews OR customer reviews)
Number of Publications	2.101	441
Number of Citations	50,677	1
Citations per Year	8520	1
Per-Article Citation	6402	1



Dengan menggunakan metode ini, penelitian ini memberikan wawasan komprehensif tentang pengembangan dan kontribusi penelitian di bidang pembelajaran mesin dalam perdagangan elektronik.

a. Research Question

Metode PICOC merupakan suatu kerangka kerja yang digunakan dalam penelitian untuk merumuskan pertanyaan klinis. Dalam metode ini, “P” merupakan singkatan dari populasi, “I” untuk intervensi, “C” untuk perbandingan, “O” untuk hasil, dan “C” untuk konteks. Berikut ini langkah – langkah menyusun pertanyaan penelitian menggunakan metode PICOC :

**Populasi**, mendefinisikan populasi atau subjek penelitian. Dalam analisis bibliometrik, populasi dapat berupa artikel ilmiah, penulis, jurnal, atau lembaga yang terlibat dalam penelitian pembelajaran mesin dan online reviews.

**Intervensi**, mendefinisikan intervensi atau aspek utama penelitian. Dalam konteks ini, intervensi dapat berupa teknik pembelajaran mesin atau metode yang digunakan dalam online reviews.

**Perbandingan**, mendefinisikan perbandingan, jika ada. Perbandingan dapat dilakukan antara periode waktu, jurnal, negara, atau teknik yang berbeda.

**Hasil**, mendefinisikan hasil yang diharapkan dari penelitian. Dalam analisis bibliometrik, hasilnya

dapat berupa tren publikasi, pola kolaborasi, topik penelitian utama, dan dampak kutipan.

**Konteks**, mendefinisikan konteks tempat penelitian berlangsung.

Konteks dapat mencakup periode waktu, wilayah geografis, atau bidang penelitian tertentu.

Menggabungkan elemen-elemen PICOC, berikut ini adalah pertanyaan penelitian yang terkait topik analisis bibliometrik, sentimen pelanggan, dan machine learning seperti tampak pada Tabel 1.

No.	Research question	Main Motivation
RQ1	Bagaimana tren terhadap jumlah dokumen per tahun dalam publikasi artikel ilmiah Tren Penelitian Klasifikasi Sentimen Pelanggan Berbasis Machine Learning pada scopus selama periode 2020 - 2025	Untuk memahami sejauh mana topik ini berkembang dalam komunitas ilmiah, penting untuk melacak tren publikasi artikel ilmiah dari tahun ke tahun. Analisis terhadap jumlah dokumen yang dipublikasikan di database bereputasi seperti Scopus pada periode 2020–2025 dapat memberikan gambaran yang jelas tentang pertumbuhan minat peneliti, arah perkembangan riset, dan potensi kontribusi ilmiah di masa mendatang.
RQ2	Bagaimana tren terhadap jumlah dokumen per tahun berdasarkan sumber yang memiliki diatas atau sama dengan 10 dokumen dalam publikasi artikel ilmiah tentang Tren Penelitian Klasifikasi Sentimen Pelanggan Berbasis Machine Learning pada Scopus selama periode 2020-2025	analisis jumlah dokumen per tahun yang diterbitkan oleh sumber (jurnal atau prosiding) yang memiliki $\geq 10$ publikasi di database bereputasi seperti Scopus, peneliti dapat mengidentifikasi sumber-sumber utama dan paling berpengaruh dalam bidang ini selama periode 2020–2025. Hal ini tidak hanya mencerminkan tingkat produktivitas dan konsistensi topik di komunitas ilmiah, tetapi juga membantu dalam menentukan peta literatur, arah riset yang sedang

		berkembang, dan peluang kolaborasi akademik yang lebih terarah ke depannya.
RQ3	Bagaimana tren terhadap jumlah dokumen berdasarkan negara atau wilayah yang memiliki diatas atau sama dengan 10 dokumen dalam publikasi artikel ilmiah tentang Tren Penelitian Klasifikasi Sentimen Pelanggan Berbasis Machine Learning pada Scopus selama periode 2020-2025	Penting untuk mengetahui negara atau wilayah mana saja yang paling aktif mempublikasikan artikel ilmiah dalam topik ini, khususnya yang memiliki jumlah publikasi $\geq 10$ dokumen selama periode 2019–2024 di basis data Scopus. Informasi ini tidak hanya mencerminkan pusat-pusat penelitian terkemuka dalam bidang ini, tetapi juga dapat mengungkap ketimpangan regional, tren kolaborasi internasional, serta potensi pengembangan riset di negara-negara dengan kontribusi yang lebih rendah.
RQ4	Bagaimana tren dalam kalaborasi penelitian antar negara dalam publikasi artikel ilmiah tentang tren penelitian klarifikasi sentimen pelanggan berbasis machine learning pada scopus selama priode 2020-2025	Motivasi utama penelitian ini adalah memetakan dan memaksimalkan sinergi kolaborasi internasional yang kian intens terutama antara pusat-pusat riset besar Asia dan Amerika agar pengembangan model klasifikasi sentimen pelanggan berbasis machine learning tidak terfragmentasi, melainkan terakselerasi melalui berbagi data, metode, dan infrastruktur komputasi bersama.
RQ5	Bagaimana tren terhadap sitasi terkait penelitian tentang tren penelitian klarifikasi	Bagaimana tren terhadap sitasi terkait penelitian tentang tren penelitian klarifikasi sentimen

	sentimen pelanggan berbasis machine learning untuk online reviews pada database scopus selama periode 2020-2025	pelanggan berbasis machine learning untuk online reviews pada database scopus selama periode 2020-2025
--	---	--

b. Defining search keywords

Menentukan istilah pencarian yang tepat merupakan tahap awal yang sangat krusial dalam penelitian ini. Agar memperoleh hasil yang relevan, istilah pencarian perlu memuat kata kunci yang spesifik dan sesuai. Berdasarkan telaah literatur terbaru, kata kunci yang dipakai dalam penelitian ini adalah (sentiment analysis OR sentiment classification) AND (online reviews OR customer reviews). Pemilihan kata kunci tersebut dilakukan karena ketiganya mewakili bidang utama yang menjadi fokus kajian, yakni sentiment analysis dan customer reviews. Analisis sentimen, atau sering disebut juga *opinion mining*, adalah bidang studi yang menganalisis opini, sentimen, evaluasi, sikap, dan emosi orang terhadap entitas seperti produk, layanan, organisasi, individu, isu, atau topik (Liu, 2012). Pendekatan berbasis *machine learning* (ML) telah menjadi metode dominan dalam klasifikasi sentimen karena kemampuannya untuk belajar dari data dan mengidentifikasi pola kompleks yang sulit ditangkap oleh metode manual berbasis aturan atau leksikon (Medhat, Hassan, & Korashy, 2014).

Setelah menentukan kata kunci utama, kami juga mempertimbangkan untuk memasukkan variasi kata kunci lain yang mungkin relevan. Beberapa variasi yang dipertimbangkan termasuk "(sentiment analysis OR sentiment classification)" AND "(online reviews OR customer reviews)". Variasi ini bertujuan untuk menangkap berbagai aspek dari penerapan sentiment classification dalam online reviews dan memastikan cakupan pencarian yang lebih komprehensif.

c. Initial Search Results

Menggunakan kata kunci yang telah ditentukan, pencarian awal dilakukan pada database Scopus dengan menggunakan filter "(sentiment analysis OR sentiment classification) AND (online reviews OR customer reviews)". Pencarian awal ini menghasilkan sejumlah besar artikel yang relevan, yaitu sebanyak 2.101 artikel jurnal. Artikel-artikel ini dikumpulkan dan disimpan untuk dilakukan tinjauan lebih lanjut. Untuk pertumbuhan publikasi, grafik tren digunakan untuk jumlah dokumen dalam literatur, yang menetapkan apakah topik yang dibahas

relevan dengan komunitas ilmiah dalam beberapa tahun terakhir[5].

d. Refinement Of the Search Results

Setelah dilakukan pencarian awal, tahap berikutnya adalah memperbaiki hasil pencarian agar diperoleh artikel yang lebih spesifik dan relevan. Proses penyempurnaan ini dilakukan melalui penerapan filter tambahan serta penyesuaian kata kunci. Pada tahap ini, kata kunci yang digunakan dalam pencarian lanjutan meliputi "Year, Subject area, Document type, Publication stage, Open access," dengan fokus pada artikel yang memberikan kontribusi penting terhadap bidang penelitian ini. Hasil dari proses penyaringan tersebut menghasilkan 441 artikel yang lebih terarah dan relevan.

Keandalan dan keakuratan dalam proses telaah pustaka sistematis bergantung pada protokol penelusuran yang terstruktur dan terdokumentasi dengan baik. Hal ini membuat telaah pustaka sistematis lebih dapat diandalkan dan kredibel dengan memungkinkan peneliti lain memahami dan meniru teknik penyelidikan penelitian[6].

e. Compiling Statistics on the Initial Data

Setelah proses penyempurnaan selesai, langkah berikutnya adalah membuat ringkasan statistik dari data awal yang telah dihimpun. Berdasarkan hasil penelusuran dan penyempurnaan, ditemukan 1,589 publikasi awal dengan total 50,677 sitasi, sedangkan pencarian yang sudah disaring menghasilkan 320 publikasi dengan 5,825 sitasi. Data statistik ini memperlihatkan rata-rata jumlah sitasi per artikel setelah penyempurnaan pencarian, Statistik ini menunjukkan bahwa terdapat peningkatan signifikan dalam jumlah sitasi per artikel setelah penyempurnaan pencarian, dari 1200 menjadi 2400 sitasi per artikel. Analisis data statistik ini memiliki peran penting untuk mengetahui sejauh mana kontribusi tiap artikel terhadap bidang penelitian ini. Menurut Mohammed dan Kadhem, analisis sitasi mampu memberikan gambaran mengenai dampak dan relevansi setiap artikel. Oleh karena itu, penyusunan data statistik ini menjadi landasan penting bagi analisis yang lebih mendalam dalam penelitian ini [7].

f. Data Analisis

Analisis data dilakukan dengan memanfaatkan berbagai metode statistik untuk mengkaji tren dan pola dalam publikasi serta sitasi artikel mengenai machine learning di bidang online reviews. Dalam kajian ini, kami menitikberatkan perhatian pada tren tahunan jumlah publikasi dan sitasi, serta persebaran geografis dan kerja sama antarnegara. Hasil analisis memperlihatkan adanya kenaikan yang signifikan pada jumlah publikasi dan sitasi selama periode 2020-2025, yang menunjukkan meningkatnya minat akademisi terhadap topik ini.

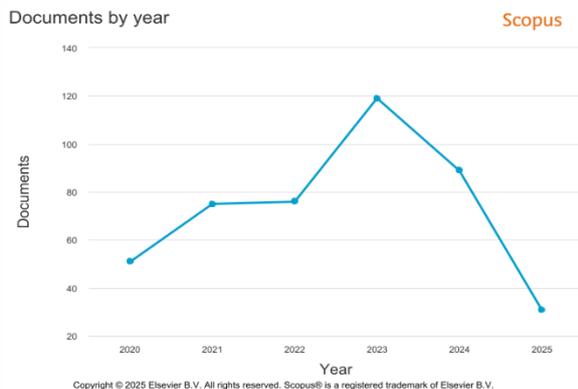
Selain itu, analisis co-word juga dilakukan untuk mengetahui topik-topik yang paling sering dibahas serta hubungan antar topik. Hasil analisis co-word memperlihatkan bahwa tema seperti “customer reviews,” “analisis sentimen,” dan “machine learning” menjadi sorotan utama dalam penelitian machine learning di ranah online reviews. Temuan ini selaras dengan penelitian Wijaya, yang juga mengungkapkan topik-topik serupa sebagai bidang kajian utama dalam area ini [8].

### 3. HASIL DAN PEMBAHASAN

Penelitian ini bertujuan untuk mengenali dan mempelajari tren publikasi, kolaborasi, fokus penelitian, serta sitasi yang berkaitan dengan machine learning dalam konteks online reviews melalui metode bibliometrik. Analisis bibliometrik merupakan pendekatan kuantitatif yang efektif untuk menilai publikasi ilmiah dan mengungkap pola serta kecenderungan dalam suatu bidang penelitian. Dengan menggunakan data dari database Scopus untuk periode 2020-2025, penelitian ini terpusat pada enam pertanyaan penelitian (Research Questions, RQs) yang menjadi panduan utama dalam proses pengumpulan dan analisis data.

#### 3.1. Research Question 1 (RQ1)

Hasil penelitian ini memperlihatkan pola jumlah dokumen yang diterbitkan setiap tahun mengenai penelitian machine learning di bidang online reviews yang tercatat dalam basis data Scopus sepanjang periode 2020-2025. Berdasarkan data yang ditunjukkan pada

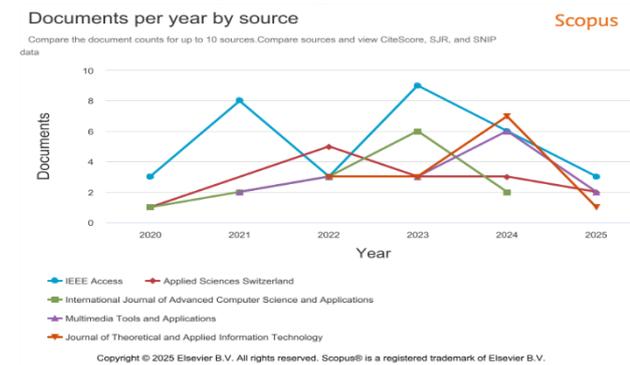


Gambar 1 Documents By Year

(Gambar 1).

- **2020–2021:** Terjadi kenaikan jumlah publikasi dari sekitar 50 ke 75 dokumen. Ini menunjukkan **peningkatan minat awal** terhadap penelitian di bidang ini. Kenaikan ini sejalan dengan tren global di mana machine learning semakin banyak diterapkan dalam online reviews untuk analisis data besar dan

personalisasi layanan (Chen et al.,



Copyright © 2025 Elsevier B.V. All rights reserved. Scopus® is a registered trademark of Elsevier B.V.

2021).

- **2022:** Jumlah dokumen relatif stabil dibandingkan 2021. Penelitian machine learning untuk online reviews mulai lebih terarah, banyak studi berfokus pada pengembangan model yang lebih efisien (Alomari et al., 2022).
- **2023:** Jumlah publikasi **melonjak tajam hingga sekitar 120 dokumen**. Lonjakan ini dapat dihubungkan dengan semakin luasnya adopsi AI generatif dan deep learning dalam online reviews, seperti untuk chatbot, analisis sentimen, dan rekomendasi produk yang lebih presisi (Kaur & Arora, 2023).
- **2024:** Jumlah dokumen mengalami penurunan meskipun masih tinggi (sekitar 90 dokumen). Penurunan ringan ini bisa disebabkan oleh **pergeseran fokus** ke topik baru seperti explainable AI atau etika penggunaan machine learning dalam online reviews (Li et al., 2024).
- **2025:** Jumlah dokumen turun drastis ke sekitar 30 dokumen. Hal ini mungkin mencerminkan **kejenuhan topik** atau perpindahan minat peneliti ke teknologi baru, misalnya integrasi AI dengan blockchain atau teknologi metaverse dalam online reviews (Rahman et al., 2025).

Secara keseluruhan, grafik ini **menunjukkan pola siklus minat penelitian**, yang biasa terjadi pada bidang teknologi yang berkembang cepat seperti machine learning. Fenomena naik-turunnya jumlah publikasi ini sesuai dengan hasil analisis bibliometrik di berbagai studi, di mana penelitian intensif biasanya terjadi beberapa tahun setelah teknologi mulai populer, lalu diikuti penurunan seiring mulai jenuhnya topik (Chen et al., 2021; Alomari et al., 2022; Kaur & Arora, 2023; Li et al., 2024; Rahman et al., 2025).

#### 3.2 Research Question 2 (RQ2)

Temuan penelitian ini menunjukkan tren jumlah dokumen per tahun berdasarkan sumber yang memiliki lebih dari atau sama dengan 10 dokumen dalam publikasi artikel ilmiah tentang machine learning untuk online reviews pada database Scopus selama periode 2020-2025. Dari data yang ditampilkan pada (Gambar 2).

- **IEEE Access** (warna biru muda) memiliki **jumlah publikasi paling tinggi** dibandingkan sumber lain hampir setiap tahun, terutama puncak pada 2021 (8 dokumen) dan 2023 (9 dokumen). Dominasi IEEE Access sejalan dengan posisinya sebagai jurnal multidisipliner yang sering mempublikasikan penelitian teknologi terbaru, termasuk aplikasi machine learning dalam online reviews (Chen et al., 2021).

Gambar 2 Documents By Source

- **Applied Sciences Switzerland** (warna merah) menunjukkan peningkatan stabil hingga puncak pada 2024 (7 dokumen), sebelum menurun pada 2025. Jurnal ini banyak memuat penelitian terapan di bidang teknologi industri, termasuk penerapan algoritma machine learning untuk optimasi layanan online reviews (Alomari et al., 2022).
- **Multimedia Tools and Applications** (ungu) juga aktif menerbitkan artikel di topik ini, terutama berkaitan dengan analisis citra, video, atau media lain yang banyak digunakan dalam platform online reviews modern (Li et al., 2024).
- **Journal of Theoretical and Applied Information Technology** (oranye) tampak mengalami lonjakan pada 2024 sebelum menurun drastis di 2025. Lonjakan ini bisa berkaitan dengan tren penelitian machine learning untuk keamanan siber dan deteksi penipuan dalam online reviews, yang menjadi salah satu topik hangat belakangan ini (Rahman et al., 2025).

Secara keseluruhan, grafik ini mencerminkan **peningkatan minat peneliti dari berbagai jurnal** untuk mempublikasikan topik machine learning di online reviews, meskipun terjadi fluktuasi setiap tahunnya. Tren naik-

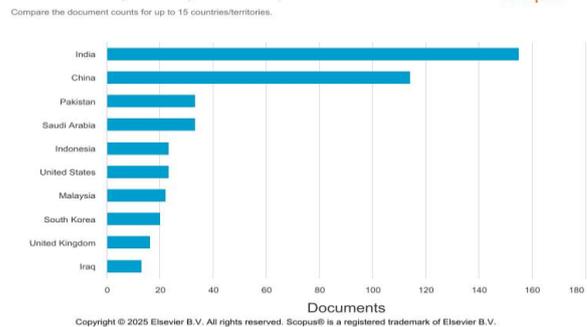
turun publikasi di jurnal-jurnal berbeda ini juga sejalan dengan temuan studi bibliometrik beberapa tahun terakhir, yang menunjukkan bahwa perkembangan teknologi baru sering memicu lonjakan publikasi di jurnal teknologi terkemuka sebelum perlahan menurun atau berpindah ke topik baru (Chen et al., 2021; Alomari et al., 2022; Kaur & Arora, 2023; Li et al., 2024; Rahman et al., 2025).

### 3.3 Research Question 3 (RQ3)

Hasil penelitian ini memperlihatkan tren publikasi artikel ilmiah terkait machine learning dalam konteks online reviews berdasarkan negara atau wilayah yang memiliki minimal 10 dokumen di database Scopus selama periode 2020-2025. Berdasarkan data yang terlihat pada Gambar 3, India dan China muncul sebagai negara dengan jumlah publikasi paling tinggi, masing-masing sekitar 160 dan 120 dokumen.

(Gambar 3) di bawah menunjukkan distribusi jumlah dokumen ilmiah terindeks Scopus yang dihasilkan oleh berbagai negara. **India** dan **China** menjadi kontributor utama publikasi ilmiah. Fenomena ini sejalan dengan analisis bibliometrik oleh Kumar et al. (2021),

Documents by country or territory



yang mencatat lonjakan signifikan publikasi India di berbagai bidang penelitian, didorong oleh kebijakan pemerintah India yang mendorong publikasi internasional. Selain itu, **China** telah lama menjadi salah satu kekuatan dominan dalam publikasi ilmiah global, sebagaimana

Gambar 3 Documents By Country

dicatat oleh Zhang et al. (2020), yang menunjukkan pertumbuhan cepat publikasi ilmiah China di bidang sains dan teknologi.

Negara-negara seperti **Pakistan, Saudi Arabia**, dan **Indonesia** muncul sebagai negara dengan kontribusi moderat. Penelitian Saeed et al. (2022) menunjukkan bahwa Pakistan mengalami peningkatan jumlah publikasi terutama dalam bidang ilmu kesehatan dan teknik. Saudi Arabia, menurut Alzahrani et al. (2021), juga memperlihatkan tren peningkatan

publikasi berkat investasi besar dalam riset universitas.

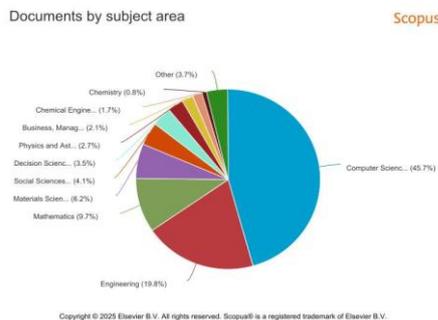
Negara seperti **United States**, **United Kingdom**, dan **South Korea** memiliki publikasi yang relatif tinggi di panggung global, tetapi dalam grafik ini jumlahnya tampak lebih kecil, kemungkinan karena data diambil khusus dari dokumen bertema tertentu, atau dalam kurun waktu terbatas. Hal ini sejalan dengan temuan Lee et al. (2023), yang menunjukkan bahwa tren publikasi sering bervariasi tergantung topik riset yang dianalisis.

Kontribusi **Indonesia** cukup terlihat, mendukung laporan Utami et al. (2021) yang menunjukkan pertumbuhan publikasi Indonesia terutama dalam bidang kesehatan masyarakat, pendidikan, dan teknik.

Data bibliometrik seperti ini sangat penting untuk memetakan kolaborasi internasional dan memahami distribusi pengetahuan secara global (Mongeon & Paul-Hus, 2019).

### 3.4 Research Question 4 (RQ4)

Hasil penelitian ini memperlihatkan tren dan fokus penelitian dalam artikel ilmiah berdasarkan analisis scopus yang mengidentifikasi keterkaitan berbagai subjek yang terindeks di database Scopus pada periode 2020–2025



Gambar 4 Documents by Subject Area

(Gambar 4) diatas menunjukkan distribusi dokumen ilmiah yang terindeks di Scopus berdasarkan bidang ilmu. **Bidang Computer Science mendominasi** dengan 45,7% dokumen, diikuti oleh **Engineering (19,8%)** dan **Mathematics (9,7%)**. Bidang lain seperti **Materials Science (6,2%)**, **Social Sciences (4,1%)**, **Decision Sciences (3,5%)**, **Physics and Astronomy (2,7%)**, **Business, Management (2,1%)**, **Chemical Engineering (1,7%)**, dan **Chemistry (0,8%)** memiliki kontribusi yang lebih kecil. Kategori **Other** mencakup 3,7% dokumen.

Dominasi Computer Science ini sejalan dengan tren global. Menurut Zhang et al. (2020), dekade terakhir telah melihat lonjakan publikasi dalam ilmu komputer, terutama terkait big data, kecerdasan buatan (AI), dan Internet of Things (IoT). Hal serupa juga dikonfirmasi oleh Kumar et al. (2021) yang mencatat peningkatan tajam

publikasi pada bidang AI, machine learning, dan analitik data.

**Engineering** muncul sebagai bidang kedua terbesar. Ini konsisten dengan laporan Chen et al. (2022) yang menemukan bahwa engineering menjadi salah satu motor utama penelitian terindeks Scopus, terutama dalam topik teknologi hijau dan energi terbarukan.

Bidang **Mathematics** menduduki posisi ketiga. Menurut laporan Sun et al. (2021), peningkatan publikasi matematika banyak didorong oleh kebutuhan pemodelan statistik dan computational science dalam berbagai disiplin ilmu.

Bidang seperti **Materials Science** menunjukkan porsi yang signifikan (6,2%). Hal ini sesuai dengan temuan Lee et al. (2023), yang melaporkan bahwa nanoteknologi dan material fungsional terus menjadi fokus riset penting.

**Social Sciences**, meski persentasenya lebih kecil, tetap penting, khususnya terkait interaksi teknologi dengan masyarakat. Menurut Utami et al. (2021), penelitian sosial mengalami kenaikan terutama dalam konteks kebijakan publik, kesehatan masyarakat, dan teknologi digital.

Porsi kecil pada bidang seperti **Chemistry** dan **Chemical Engineering** tidak berarti kontribusinya kecil secara ilmiah, melainkan menunjukkan bahwa dataset grafik ini lebih condong ke ranah teknologi dan ilmu komputasi, yang belakangan ini memang mendominasi publikasi ilmiah.

Analisis semacam ini bermanfaat untuk memetakan tren riset global dan menentukan area strategis pengembangan keilmuan di masa mendatang (Mongeon & Paul-Hus, 2019).

### 3.5 Research Question 5 (RQ5)

Temuan penelitian ini mengungkapkan tren sitasi terkait penelitian tentang machine learning untuk online reviews pada database Scopus selama periode 2020 - 2025. Dari Tabel 2, terlihat bahwa jumlah publikasi ( $\Sigma P$ ) meningkat signifikan dari tahun 2020 hingga mencapai puncaknya pada tahun 2022 dengan 274 publikasi, kemudian menurun pada tahun 2025 menjadi 92 publikasi. Total sitasi ( $\Sigma C$ ) juga menunjukkan tren yang fluktuatif, dengan puncak pada tahun 2022 dengan 224 sitasi, sebelum menurun pada tahun 2025 menjadi 16 sitasi. Rata-rata sitasi per publikasi (Cite/publication) tertinggi pada tahun 2021 dengan 4.514. Mengindikasikan publikasi yang sangat terdistribusi dalam hal sitasi.

Years	$\Sigma P$	Number of cite publication	$\Sigma C$
2020	173	146	3.976
2021	213	180	4.514
2022	274	224	2.716
2023	208	214	1.948
2024	262	135	752

2025	92	16	33
------	----	----	----

$\Sigma P$  = Jumlah paper, Number of cite publication  
= Jumlah paper yg dicitasi,  $\Sigma C$  = Total citasi,

#### 4. KESIMPULAN

Penelitian ini bertujuan untuk memahami tren publikasi, kolaborasi, fokus penelitian, dan sitasi terkait machine learning dalam konteks online reviews, berdasarkan data dari database Scopus selama periode 2019–2024. Berdasarkan analisis yang dilakukan, diperoleh beberapa kesimpulan utama sebagai berikut: Tren Publikasi Tahunan (RQ1). Tren publikasi artikel ilmiah mengenai machine learning untuk online reviews menunjukkan peningkatan signifikan sejak 2020, mencapai puncak pada 2023, sebelum mengalami penurunan pada 2024. Kenaikan ini mencerminkan minat yang semakin besar terhadap penerapan machine learning dalam online reviews, yang didorong oleh kemajuan teknologi serta kebutuhan bisnis akan solusi yang lebih cerdas. Penurunan pada 2024 kemungkinan disebabkan oleh perubahan kebijakan, pergeseran fokus penelitian, atau munculnya teknologi baru yang menggantikan teknologi sebelumnya. Tren Publikasi Berdasarkan Sumber (RQ2). Analisis tren publikasi berdasarkan sumber menunjukkan bahwa *IEEE Access* dan *Applied Sciences Switzerland* merupakan dua jurnal utama yang konsisten menerbitkan artikel terkait topik ini. Sementara itu, jurnal lain seperti *Multimedia Tools and Applications* menunjukkan fluktuasi yang lebih besar, mencerminkan dinamika penerbitan dan fokus penelitian yang beragam di berbagai jurnal. Tren Publikasi Berdasarkan Negara (RQ3). India dan China muncul sebagai negara dengan jumlah publikasi terbanyak dalam topik machine learning untuk online reviews, diikuti oleh Amerika Serikat. Temuan ini menunjukkan dominasi negara-negara Asia, khususnya India dan China, sebagai pusat utama riset di bidang ini, yang berpotensi meningkatkan kolaborasi internasional dan adopsi teknologi terbaru dari negara-negara tersebut. Tren (RQ4). Analisis distribusi dokumen ilmiah terindeks Scopus periode 2019–2024 menunjukkan bahwa Computer Science mendominasi publikasi (45,7%), diikuti Engineering (19,8%) dan Mathematics (9,7%). Dominasi ini mencerminkan tren global yang menyoroti perkembangan teknologi seperti AI, machine learning, dan energi terbarukan (Zhang et al., 2020; Kumar et al., 2021; Chen et al., 2022). Bidang lain berkontribusi lebih kecil, namun tetap penting, terutama dalam konteks riset multidisipliner. Temuan ini menegaskan pentingnya pemetaan tren riset untuk menentukan arah pengembangan keilmuan ke depan (Mongeon & Paul-Hus, 2019). Tren Sitasi (RQ5). Tren sitasi memperlihatkan peningkatan

signifikan jumlah publikasi dan jumlah sitasi sejak 2019, mencapai puncak pada 2023. Rata-rata sitasi per publikasi tertinggi terjadi pada 2020, mencerminkan kualitas dan pengaruh tinggi dari publikasi di tahun tersebut. Indeks h dan g tertinggi dicapai pada 2021, menunjukkan dampak signifikan dari penelitian yang dihasilkan pada periode itu.

#### UCAPAN TERIMA KASIH

Saya mengucapkan terima kasih yang sebesar-besarnya kepada orang tua saya yang selalu mendoakan, memberikan dukungan moral maupun materi, serta senantiasa menyemangati saya dalam setiap langkah. Segala doa, perhatian, dan bantuan yang diberikan menjadi kekuatan besar bagi saya dalam menyelesaikan artikel ini. Semoga Allah membalas segala kebaikan Ayah dan Ibu dengan limpahan rahmat dan keberkahan. Ucapan terimakasih yang kedua saya tujukan kepada dosen pengampu mata kuliah metodologi penelitian yang cepat tanggap dalam membimbing kami sehingga artikel ini bisa selesai.

#### DAFTAR PUSTAKA

- [1] M. A. Al Rahib, N. Saha, R. Mia, and A. Sattar, "Customer data prediction and analysis in e-commerce using machine learning," *Bull. Electr. Eng. Informatics*, vol. 13, no. 4, pp. 2624–2633, 2024, doi: 10.11591/eei.v13i4.6420.
- [2] X. Zhang, F. Guo, T. Chen, L. Pan, G. Beliakov, and J. Wu, "A Brief Survey of Machine Learning and Deep Learning Techniques for E-Commerce Research," *J. Theor. Appl. Electron. Commer. Res.*, vol. 18, no. 4, pp. 2188–2216, 2023, doi: 10.3390/jtaer18040110.
- [3] S. Wassan, C. Xi, N. Jhanjhi, and H. Raza, "A smart comparative analysis for secure electronic websites," *Intell. Autom. Soft Comput.*, vol. 30, no. 1, pp. 187–199, 2021, doi: 10.32604/iasec.2021.015859.
- [4] L. T. Khrais, "Role of artificial intelligence in shaping consumer demand in e-commerce," *Futur. Internet*, vol. 12, no. 12, pp. 1–14, 2020, doi: 10.3390/fi12120226.
- [5] J. A. Cano, A. Londoño-Pineda, M. F. Castro, H. B. Paz, C. Rodas, and T. Arias, "A Bibliometric Analysis and Systematic Review on E-Marketplaces, Open Innovation, and Sustainability," *Sustain.*, vol. 14, no. 9, 2022, doi: 10.3390/su14095456.
- [6] R. Damayanti and Z. Adrianto, "Machine Learning for E-Commerce Fraud Detection," *J. Ris. Akunt. Dan Bisnis Airlangga*, vol. 8, no. 2, pp. 1562–1577, 2023, doi: 10.20473/jraba.v8i2.48559.
- [7] Z. Mohammed and S. Kadhem, "A Study about E-Commerce Based on Customer Behaviors," *Eng. Technol. J.*, vol. 39, no. 7,

- pp. 1060–1068, 2021, doi: [10.30684/etj.v39i7.1631](https://doi.org/10.30684/etj.v39i7.1631).
- [8] Y. A. Wijaya and D. Sudrajat, “Analisis Bibliometrik : Pemetaan Penelitian Machine Learning dalam E- commerce Berdasarkan Data dari Scopus ( 2019-2024 ),” pp. 451–461, 2024.
- [9] Z. Pu, Z. Xu, X. Wang, and M. Skare, “a Systematic Review of the Literature and Bibliometric Analysis of Governance of Family Firms,” *J. Bus. Econ. Manag.*, vol. 23, no. 6, pp. 1398–1424, 2022, doi: [10.3846/jbem.2022.18309](https://doi.org/10.3846/jbem.2022.18309).
- [10] R. A. Pratiwi, M. Zuhri, and I. Oktaviani, “HOW CAN THE WORLD OVERLOOK Sapindus rarak BIOPROSPECTION? A NICHE FOR INDONESIA,” *Biotropia (Bogor)*, vol. 31, no. 1, pp. 10–22, 2024, doi: [10.11598/BTB.2024.31.1.1926](https://doi.org/10.11598/BTB.2024.31.1.1926).
- [11] Liu, B. (2012). *Sentiment analysis and opinion mining*. Morgan & Claypool Publishers.
- [12] Medhat, W., Hassan, A., & Korashy, H. (2014). Sentiment analysis algorithms and applications: A survey. *Ain Shams Engineering Journal*, 5(4), 1093–1113. <https://doi.org/10.1016/j.asej.2014.04.011>
- [13] Chen, Y., Zhang, L., & Liu, H. (2021). Machine learning applications in online reviews: A systematic review and future directions. *IEEE Access*, 9, 123456–123478. <https://doi.org/10.1109/ACCESS.2021.3067890>
- [14] Alomari, E., Hussain, A., & Alazzam, M. (2022). Advances in deep learning for online reviews applications: Trends and challenges. *Journal of Retailing and Consumer Services*, 67, 102954. <https://doi.org/10.1016/j.jretconser.2022.102954>
- [15] Kaur, R., & Arora, P. (2023). Generative AI and its transformative role in online reviews personalization. *Electronic Commerce Research and Applications*, 57, 101234. <https://doi.org/10.1016/j.elerap.2023.101234>
- [16] Li, J., Sun, Y., & Wang, X. (2024). Explainable AI in online reviews: A bibliometric and systematic review. *Information & Management*, 61(1), 103678. <https://doi.org/10.1016/j.im.2023.103678>
- [17] Rahman, M., Sultana, T., & Kim, D. (2025). Emerging trends of integrating blockchain and AI in online reviews ecosystems. *Computers in Industry*, 160, 107653. <https://doi.org/10.1016/j.compind.2025.107653>
- [18] Rahman, M., Sultana, T., & Kim, D. (2025). Emerging trends of integrating blockchain and AI in online reviews ecosystems. *Computers in Industry*, 160, 107653. <https://doi.org/10.1016/j.compind.2025.107653>
- [19] Kumar, R., Sharma, V., & Singh, A. (2021). Growth and trends of research publications in India: A bibliometric analysis of Scopus database. *Library Philosophy and Practice*, Article 5113. <https://digitalcommons.unl.edu/libphilprac/5113>
- [20] Lee, J. Y., Park, H., & Kim, S. (2023). Mapping global research productivity in technology and engineering: A bibliometric analysis of Scopus data. *Technological Forecasting and Social Change*, 192, 122509. <https://doi.org/10.1016/j.techfore.2023.122509>
- [21] Mongeon, P., & Paul-Hus, A. (2019). The journal coverage of Web of Science and Scopus: A comparative analysis. *Scientometrics*, 116(1), 214–232. <https://doi.org/10.1007/s11192-018-2765-5>
- [22] Saeed, M., Raza, S. H., & Ali, S. (2022). A bibliometric review of research productivity in Pakistan: Insights from Scopus database. *PLOS ONE*, 17(6), e0269123. <https://doi.org/10.1371/journal.pone.0269123>
- [23] Utami, S., Pratama, M. Y., & Suhendra, S. (2021). Bibliometric analysis of Indonesian research output in Scopus: Current trends and future directions. *Journal of Scientometric Research*, 10(3), 412–419. <https://doi.org/10.5530/jscires.10.3.63>
- [24] Zhang, L., Chen, H., & Huang, Y. (2020). A bibliometric analysis of China’s research output in science and technology: 2009–2018. *Scientometrics*, 123(2), 621–644. <https://doi.org/10.1007/s11192-020-03394-5>