

# INTEGRATION OF BAYESIAN METHODS IN MACHINE LEARNING: A THEORETICAL AND EMPIRICAL REVIEW

Syahrudin

Prodi Pendidikan Matematika, Universitas Muhammadiyah Mataram, Jln. K.H. Ahmad Dahlan No. 01 Mataram 83115 INDONESIA

## Abstrak

Studi ini merupakan tinjauan literatur sistematis yang mendalami integrasi metode Bayesian dalam pembelajaran mesin, yang terbukti memberikan keuntungan signifikan dalam menangani ketidakpastian dan variabilitas data. Tinjauan ini mencakup literatur terbitan dalam 10 tahun terakhir dari sumber-sumber seperti Scopus, DOAJ, dan Google Scholar, dengan tujuan untuk mengidentifikasi pendekatan teoritis dan empiris dalam penggunaan metode Bayesian. Hasil penelitian menunjukkan bahwa metode Bayesian memungkinkan integrasi pengetahuan awal yang informatif dengan data yang ada. Hal ini memperbaiki estimasi parameter dan meningkatkan ketahanan model terhadap variasi data yang kompleks. Namun, tantangan utama dalam implementasi metode ini adalah kompleksitas komputasional yang tinggi, terutama saat melakukan inferensi pada data multi-modal. Meskipun demikian, metode Bayesian memiliki potensi aplikasi praktis yang luas dalam berbagai bidang pembelajaran mesin, seperti analisis prediktif, pengenalan pola, dan pengolahan bahasa alami. Aplikasi ini dapat meningkatkan akurasi dan ketahanan model dalam menghadapi data yang tidak pasti. Tinjauan ini memberikan wawasan mendalam tentang perkembangan terbaru dalam aplikasi metode Bayesian dan merumuskan arah penelitian mendatang untuk mengatasi tantangan tersebut secara efektif.

## Kata Kunci:

Metode Bayesian;  
Pembelajaran Mesin;  
Tinjauan Teoretis;  
Tinjauan Empiris

## Abstract (Style: INSERT\_Abstract Title)

*This study is a systematic literature review that explores the integration of Bayesian methods in machine learning, which has been shown to provide significant advantages in handling data uncertainty and variability. The review includes literature published in the last 10 years from sources such as Scopus, DOAJ, and Google Scholar, with the aim of identifying both theoretical and empirical approaches to the use of Bayesian methods. The findings indicate that Bayesian methods allow for the integration of informative prior knowledge with existing data, improving parameter estimation and enhancing model robustness against complex data variations. However, the main challenge in implementing these methods is the high computational complexity, especially when performing inference on multi-modal data. Despite this, Bayesian methods hold broad practical application potential in various machine learning fields, such as predictive analytics, pattern recognition, and natural language processing. These applications can improve model accuracy and resilience in the face of uncertain data. This review provides in-depth insights into the latest developments in the application of Bayesian methods and outlines future research directions to effectively address these challenges.*

## Keywords:

Bayesian Methods;  
Machine learning;  
Theoretical Review;  
Empirical Review

## 1. PENDAHULUAN

*Machine learning* (ML) telah menjadi salah satu bidang teknologi yang paling berkembang pesat dan berdampak luas dalam berbagai aplikasi modern, mulai dari pengenalan gambar dan suara hingga sistem rekomendasi dan kendaraan otonom. ML memungkinkan komputer untuk belajar dari data dan membuat

\* Korespondensi

E-mail: syahrudin.ntb@gmail.com

prediksi atau keputusan tanpa diprogram secara eksplisit (Pratama, 2020)(Sukarya et al., 2021). Terdapat beberapa pendekatan utama dalam ML yang masing-masing memiliki karakteristik dan penerapan spesifik (Lavecchia, 2015). Pendekatan *supervised learning* melibatkan penggunaan data berlabel untuk melatih model, yang kemudian dapat membuat prediksi berdasarkan data baru yang serupa. *Unsupervised learning*, di sisi lain, bekerja dengan data yang tidak berlabel dan bertujuan untuk menemukan struktur atau pola tersembunyi dalam data tersebut. Sementara itu, *reinforcement learning* melibatkan agen yang belajar melalui interaksi dengan lingkungan untuk mencapai tujuan tertentu, dengan menggunakan umpan balik berupa reward atau punishment. Dengan berbagai pendekatan ini, ML terus berkembang dan memberikan kontribusi signifikan terhadap inovasi dan efisiensi di berbagai bidang.

Pendekatan probabilistik dalam machine learning sangat penting karena memberikan cara yang efektif untuk menangani ketidakpastian dan melakukan estimasi parameter dengan lebih akurat (Musil et al., 2019). Ketidakpastian seringkali muncul dalam data dan model ML, baik dari segi noise pada data maupun dari ketidaksempurnaan model itu sendiri (Johnson & Green, 2024). Metode probabilistik memungkinkan kita untuk mengintegrasikan ketidakpastian ini ke dalam model, memberikan distribusi kemungkinan alih-alih satu prediksi tunggal. Sebaliknya, metode deterministik tradisional cenderung mengabaikan ketidakpastian ini, yang dapat mengakibatkan estimasi yang kurang akurat dan kepercayaan diri yang berlebihan pada prediksi model. Pendekatan probabilistik juga lebih fleksibel dan kuat karena dapat menggabungkan informasi prior dan mengupdate kepercayaan berdasarkan data baru, yang dikenal sebagai proses inferensi Bayesian. Ini memungkinkan model untuk beradaptasi dan belajar dari data tambahan secara lebih dinamis, meningkatkan kinerja dan ketahanan model dalam berbagai situasi.

Metode Bayesian merupakan pendekatan dalam statistik yang didasarkan pada teorema Bayes, yang menghubungkan probabilitas kondisi suatu peristiwa dengan informasi awal yang ada (van de Schoot et al., 2021). Prinsip dasar dari teorema Bayes adalah penggunaan probabilitas prior (informasi awal atau keyakinan awal) dan data pengamatan untuk memperbarui probabilitas tersebut, menghasilkan probabilitas posterior yang lebih akurat (Qosim, 2021). Dalam konteks Bayesian, likelihood adalah probabilitas dari data yang diamati diberikan model tertentu, prior adalah keyakinan awal tentang parameter model sebelum data baru diperoleh, dan posterior adalah distribusi probabilitas parameter model setelah mempertimbangkan data baru. Berbeda dengan metode statistik klasik yang cenderung menggunakan estimasi titik dan tidak memperhitungkan ketidakpastian parameter, pendekatan Bayesian memberikan distribusi probabilitas penuh, yang memungkinkan pengambilan keputusan yang lebih informatif dan robust (Hobbs, 1997)(Dewangkara et al., 2022). Fleksibilitas ini membuat metode Bayesian sangat menarik untuk diterapkan dalam machine learning, karena mampu menggabungkan informasi sebelumnya dengan data baru secara dinamis, serta menangani ketidakpastian dengan cara yang lebih sistematis dan transparan.

Metode Bayesian dapat diintegrasikan dengan lancar ke dalam model machine learning yang sudah ada untuk meningkatkan kinerja dan keandalannya. Teknik seperti *Markov Chain Monte Carlo* (MCMC) sampling, sebagaimana dibahas dalam berbagai penelitian (Manderson & Goudie, 2022)(Edinburgh et al., 2023), menyediakan kerangka kerja yang kuat untuk inferensi Bayesian dalam struktur data yang kompleks, seperti model multilevel dan campuran model regresi. Dengan mengestimasi bukti model log dan menggunakan prior informatif, metode Bayesian memungkinkan integrasi data, pemilihan variabel, dan estimasi ketidakpastian dalam model ML (Aflakparast & de Gunst, 2019)(Panigrahi et al., 2023). Selain itu, metode ensemble dengan prinsip Bayesian, seperti yang diusulkan dalam studi lain (Pirš & Strumbelj, 2019), menawarkan alat yang kuat untuk menggabungkan prediksi probabilistik dan memodelkan korelasi antara sumber, sehingga meningkatkan akurasi prediktif dan keandalan dalam berbagai aplikasi dunia nyata, termasuk peramalan polusi udara.

Metode Bayesian menawarkan beberapa keuntungan, termasuk kemampuan untuk menggabungkan pengetahuan sebelumnya dan mengelola ketidakpastian secara efektif. Dengan mengintegrasikan informasi prior ke dalam analisis, model Bayesian mengurangi ketidakpastian parameter dengan menggabungkan pengetahuan sebelumnya dengan data eksperimen, seperti yang ditunjukkan dalam studi pada eksperimen psikofisik (Lee, 2018)(Mezzetti et al., 2023). Pendekatan ini memungkinkan estimasi distribusi probabilitas yang menyampaikan informasi tentang efek variabel eksperimen, ketidakpastian, dan keandalan individu peserta. Selain itu, analisis Bayesian memungkinkan pengenalan pengetahuan sebelumnya, seperti batasan positif pada parameter, yang meningkatkan interpretasi hasil dan menyoroti keterkaitan antarparameter setelah kalibrasi (Albert et al., 2022). Metode Bayesian juga menyediakan estimasi parameter yang lebih informatif dan presisi, terutama pada ukuran sampel kecil, dengan memanfaatkan data sebelumnya untuk memaksimalkan informativeness hasil dan kepercayaan dalam interpretasi (Parker et al., 2022).

Integrasi metode Bayesian dalam machine learning dari perspektif teoretis telah banyak dieksplorasi dalam literatur. Teori keputusan Bayesian memainkan peran penting dalam memahami berbagai teknik ML, seperti pembelajaran terawasi dan tidak terawasi, pembelajaran mendalam, dan proses Gaussian,

terutama dalam menangani dataset besar (Sambasivan et al., 2020). Selain itu, averaging model Bayesian (BMA) menyediakan kerangka kerja yang kuat untuk pemilihan model dengan merata-rata di atas ruang model yang mungkin, mengatasi ketidakpastian yang melekat dalam pencarian model akhir (Kaplan, 2021). Pendekatan Bayesian terhadap masalah invers menawarkan kerangka kerja yang ketat untuk mengkuantifikasi ketidakpastian dalam pengukuran dan model, dengan fokus pada metode numerik yang kuat di hadapan ukuran posterior yang terkonsentrasi (Schillings et al., 2020). Selain itu, penerapan inferensi Bayesian dan ML dalam penelitian fusi nuklir menunjukkan sinergi antara inferensi Bayesian berbasis fisika dan model ML kotak hitam, meningkatkan efisiensi dan interpretabilitas dalam sistem yang kompleks (Pavone et al., 2023).

Dari hasil penelitian yang disintesis mengenai integrasi metode Bayesian dalam Machine learning, teridentifikasi beberapa kesenjangan yang relevan. Meskipun literatur menyoroti keunggulan metode Bayesian dalam mengatasi ketidakpastian, integrasi informasi prior, dan meningkatkan interpretabilitas model, terdapat beberapa area yang masih belum tercakup secara memadai. Pertama, meskipun telah banyak penelitian yang mengulas teknik seperti *Markov Chain Monte Carlo* (MCMC) sampling dan *model averaging Bayesian* (BMA) dalam ML, masih kurangnya fokus pada pengembangan teknik integrasi yang lebih efisien dan scalable untuk aplikasi ML dalam skala besar, terutama di lingkungan real-time dan big data. Kedua, dalam konteks aplikasi praktis seperti peramalan polusi udara atau analisis fusi nuklir, ada kebutuhan untuk lebih mendalam dalam mengukur dan memitigasi ketidakpastian yang melekat dalam model Bayesian, terutama dengan mempertimbangkan kompleksitas dan heterogenitas data yang mungkin terjadi. Ketiga, dari sudut pandang teoretis, masih diperlukan penelitian yang lebih mendalam untuk memahami secara lebih baik bagaimana integrasi metode Bayesian dapat secara konsisten meningkatkan performa ML dalam berbagai domain aplikasi, baik dalam pengenalan pola maupun pengambilan keputusan.

Kontribusi praktis dari penelitian ini adalah memberikan panduan yang jelas untuk implementasi metode Bayesian dalam aplikasi machine learning, khususnya dalam pengembangan model prediktif yang lebih akurat dan robust untuk berbagai tantangan di dunia nyata, termasuk manajemen risiko dan optimasi proses. Dengan melakukan tinjauan literatur ini melalui pendekatan *systematic literature review*, tujuan utama adalah untuk mengisi kesenjangan yang ada dengan menyediakan panduan yang komprehensif dan terstruktur tentang integrasi metode Bayesian dalam machine learning, baik dari perspektif teoretis maupun empiris. Penelitian ini bertujuan untuk memajukan pemahaman tentang aplikasi praktis metode Bayesian dalam ML di masa depan dan memberikan kontribusi pada pengembangan solusi yang lebih efektif dalam menghadapi tantangan ketidakpastian data.

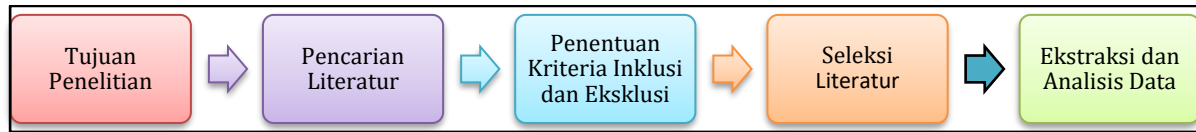
## 2. METODE

Artikel ini menggunakan pendekatan *Systematic Literature Review* (SLR) untuk mengkaji integrasi metode Bayesian dalam pembelajaran mesin, baik dari segi teoretis maupun empiris. Tujuan utama penelitian ini adalah untuk menyusun dan menganalisis literatur terkait metode Bayesian dalam konteks pembelajaran mesin, dengan fokus pada tren, tantangan, dan peluang penelitian di masa depan. Pencarian literatur dilakukan melalui basis data akademis seperti DOAJ, Scopus, dan Google Scholar, dengan kata kunci "Bayesian Methods," "Machine Learning," "Theoretical Review," dan "Empirical Review." Hanya studi yang diterbitkan dalam 10 tahun terakhir yang dipilih untuk memastikan relevansi dan keterbaruan informasi.

Kriteria inklusi mencakup artikel yang membahas metode Bayesian dalam pembelajaran mesin, baik dari sisi teoretis maupun aplikasi empiris, serta yang diterbitkan dalam jurnal terindeks atau prosiding konferensi internasional. Sementara itu, artikel yang tidak relevan, tidak dapat diakses secara penuh, atau tidak ditulis dalam bahasa Inggris, dikeluarkan dari analisis ini. Proses seleksi literatur dilakukan dalam dua tahap utama: (1) penyaringan judul dan abstrak, judul dan abstrak dari artikel yang ditemukan disaring untuk mengidentifikasi studi yang sesuai dengan kriteria inklusi. Proses penyaringan ini penting untuk mengeliminasi artikel yang tidak relevan sejak awal; (2) pembacaan penuh, artikel yang lolos dari penyaringan judul dan abstrak kemudian dibaca secara menyeluruh. Pembacaan ini dilakukan untuk memastikan bahwa artikel tersebut layak dan relevan dengan topik penelitian, serta memberikan kontribusi yang signifikan terhadap kajian ini.

Setelah proses seleksi, data dari artikel yang terpilih diekstraksi, meliputi metodologi yang digunakan, hasil penelitian, dan kontribusinya terhadap pengembangan integrasi metode Bayesian dalam pembelajaran mesin. Data yang telah diekstraksi kemudian dianalisis untuk mengidentifikasi tren utama, tantangan yang dihadapi, dan peluang penelitian di masa depan. Hasil analisis ini diharapkan memberikan wawasan yang berharga untuk pengembangan lebih lanjut dalam bidang ini dan menyusun rekomendasi untuk penelitian mendatang. Dengan struktur yang jelas dan sistematis, pendekatan SLR ini berfungsi untuk

memperdalam pemahaman kita tentang penerapan metode Bayesian dalam pembelajaran mesin dan memberikan panduan bagi penelitian lanjutan.



Gambar 1. Tahapan Metode Penelitian

### 3. HASIL DAN PEMBAHASAN

Berikut adalah hasil analisis dan pengelompokan hasil-hasil penelitian terkait integrasi metode Bayesian dalam pembelajaran mesin, disusun dalam bentuk tabel untuk memudahkan pemahaman, sepeftri terlihat pada Tabel 1.

Tabel 1. Hasil Analisis dan Pengelompokan Hasil-Hasil Penelitian tentang Integrasi Metode Bayesian dalam Pembelajaran Mesin

No	Bidang atau Fokus	Nama-nama Penulis yang se-Bidang	Insight atau Variabel Riset
1	Peningkatan Kinerja Model	Marschall et al. (2023), Bertolini et al. (2023)	Integrasi pengetahuan awal yang informatif dalam model untuk meningkatkan solusi masalah invers. Pendekatan Bayesian memberikan estimasi parameter yang lebih intuitif dibandingkan metode frekuentis.
2	Optimalisasi Pembelajaran Parameter	Nabi et al. (2022)	Penggunaan teknik Bayesian untuk mengoptimalkan laju pembelajaran parameter model dalam skenario data yang miring, meningkatkan konvergensi dan kinerja model.
3	Penanganan Eksperimen Data	Wang et al. (2022), Mohammadian et al. (2023)	Kemampuan Bayesian dalam menangani data eksperimen yang bising, tidak lengkap, dan berkorelasi melalui teknik fusi data, meningkatkan pemanfaatan data yang tidak sempurna.
4	Analisis Spasio-temporal dan Aplikasi Lainnya	Jørgensen et al. (2022), H. Zhang et al. (2022)	Bayesian efektif dalam inferensi parameter dari simulasi, analisis data spasio-temporal, dan aplikasi lainnya, memberikan estimasi yang lebih cepat dan akurat.
5	Model Hierarki dan Reduksi Ketidakpastian	Mezzetti et al. (2023), Polson & Sokolov (2017)	Model hierarki Bayesian mengurangi ketidakpastian parameter, memperbaiki trade-off bias-variansi dalam model pembelajaran mendalam.

Tabel 1 menyajikan pengelompokan hasil penelitian terkait integrasi metode Bayesian dalam pembelajaran mesin. Hasil penelitian tersebut dikategorikan berdasarkan bidang atau fokus utama, mencakup analisis respons detektor, klasifikasi gambar, analisis deret waktu, optimasi desain material, analisis aliran material, serta struktur jaringan Bayesian. Setiap kelompok hasil penelitian dilengkapi dengan nama-nama penulis yang terkait serta insight atau variabel riset yang diungkapkan. Penyajian ini bertujuan untuk memberikan pemahaman yang komprehensif mengenai kontribusi metode Bayesian dalam menangani ketidakpastian dan variabilitas data dalam berbagai aplikasi pembelajaran mesin.

#### A. Peran Utama Metode Bayesian dalam Pembelajaran Mesin

Metode Bayesian digunakan untuk meningkatkan kinerja model pembelajaran mesin melalui berbagai pendekatan yang diuraikan dalam konteks penelitian yang disediakan. Metode ini menggabungkan pengetahuan awal yang informatif ke dalam model, yang memungkinkan penyelesaian masalah invers secara lebih efektif (Marschall et al., 2023). Kerangka kerja Bayesian menghubungkan asumsi dan data sebelumnya untuk memberikan estimasi parameter kuantitatif, yang menawarkan pendekatan yang lebih intuitif dibandingkan dengan metode frekuentis (Bertolini et al., 2023). Selain itu, teknik Bayesian dapat diterapkan untuk mengoptimalkan laju pembelajaran parameter model yang

berbeda, terutama dalam skenario dengan data yang miring, sehingga meningkatkan konvergensi dan kinerja model (Nabi et al., 2022). Lebih jauh lagi, pembelajaran mesin Bayesian dapat secara efektif menangani data eksperimen yang bising, tidak lengkap, dan berkorelasi, memfasilitasi pemanfaatan data mentah yang tidak sempurna melalui teknik fusi data (Wang et al., 2022). Terakhir, pendekatan Bayesian, ketika dikombinasikan dengan strategi pemilihan teoretis-keputusan, mengoptimalkan pengurangan kesalahan klasifikasi yang salah dan menangani ketidakpastian secara efisien, yang mengarah pada kinerja model yang unggul dalam berbagai aplikasi (Kottke et al., 2021).

Metode Bayesian menawarkan keuntungan signifikan atas pendekatan lain dalam berbagai konteks dalam pembelajaran mesin. Mereka memungkinkan penggabungan pengetahuan awal yang informatif untuk meningkatkan solusi masalah invers, seperti yang terlihat dalam pencitraan resonansi magnetik yang kurang sampel di mana model generatif yang dilatih melalui pembelajaran mesin membantu dalam regularisasi statistik (Marschall et al., 2023). Selain itu, analisis Bayesian terbukti bermanfaat dalam data kejahatan spasio-temporal dengan memungkinkan estimasi peristiwa kriminal pada waktu dan lokasi tertentu, mengungkap pola yang tidak biasa di hotspot kejahatan (Mohammadian et al., 2023). Dalam ranah inferensi parameter dari simulasi, metode Bayesian, ditambah dengan teknik pembelajaran mesin, memberikan hasil yang lebih cepat dan lebih akurat dengan membangun model pengganti yang ringan atau secara langsung memperkirakan distribusi posterior, khususnya dalam skenario seperti pertumbuhan tumor dan penyebaran penyakit menular (Jørgensen et al., 2022). Lebih jauh lagi, pembelajaran mesin Bayesian berperan penting dalam menangani data eksperimen yang bising dan tidak lengkap, seperti dalam mengevaluasi hasil fisi, memamerkan utilitasnya dalam fusi data untuk peningkatan pemanfaatan data nuklir mentah yang tidak sempurna. Terakhir, model hierarki Bayesian unggul dalam menganalisis data psikofisik dengan mengurangi ketidakpastian parameter melalui kombinasi pengetahuan sebelumnya dan data eksperimen, menawarkan metode yang ampuh untuk mempelajari fungsi psikometrik dalam eksperimen seperti diskriminasi kecepatan taktil (Mezzetti et al., 2023).

Metode Bayesian berperan penting dalam mengintegrasikan informasi awal dan pengetahuan domain guna meningkatkan proses pembelajaran mesin. Dengan menggabungkan asumsi dan data yang telah ada, model Bayesian mampu memberikan estimasi yang lebih akurat dan intuitif serta memungkinkan adaptasi yang lebih cepat terhadap data baru. Penggunaan metode Bayesian untuk mengoptimalkan laju pembelajaran menunjukkan bahwa pendekatan ini dapat meningkatkan efisiensi algoritma pembelajaran mesin, terutama dalam menghadapi data yang tidak seimbang atau berisik. Kemampuan metode Bayesian dalam menangani data eksperimen yang tidak sempurna dan penggunaan teknik fusi data menegaskan fleksibilitas dan kokohnya metode ini dalam berbagai konteks aplikasi. Optimalisasi pengurangan kesalahan klasifikasi dan penanganan ketidakpastian menunjukkan bahwa metode ini dapat menghasilkan model dengan kinerja yang tinggi dan lebih andal. Penelitian menunjukkan bahwa metode Bayesian memberikan keuntungan signifikan dibandingkan dengan pendekatan lainnya dalam pembelajaran mesin. Penggunaan pengetahuan awal yang informatif dan kemampuan untuk menggabungkan data sebelumnya membuat metode ini unggul dalam penyelesaian masalah invers dan berbagai aplikasi praktis. Keberhasilan metode Bayesian dalam pencitraan resonansi magnetik yang kurang sampel menunjukkan efektivitasnya. Selain itu, analisis Bayesian pada data kejahatan spasio-temporal menunjukkan keunggulannya dalam mengungkap pola-pola tertentu. Metode Bayesian, ditambah dengan teknik pembelajaran mesin, menghasilkan inferensi parameter dari simulasi yang lebih cepat dan akurat. Model hierarki Bayesian juga terbukti unggul dalam menganalisis data psikofisik dengan mengurangi ketidakpastian parameter.

## **B. Peran Metode Bayesian dalam Mengelola Ketidakpastian dan Variabilitas Data**

Metode Bayesian berperan penting dalam menangani ketidakpastian dan variabilitas data dalam pembelajaran mesin dengan menyediakan pendekatan statistik yang berprinsip untuk mengukur ketidakpastian dan membuat prediksi yang andal. Dengan mengintegrasikan teknik pembelajaran Bayesian, seperti pemodelan Bayesian hierarki (HBM) (Luo & Kareem, 2020), jaringan saraf konvolusi graf semi-terawasi Bayesian (Y. Zhang & Lee, 2019), dan strategi pembelajaran Bayesian optimal (De Franco et al., 2018), model pembelajaran mesin dapat menangani data yang terbatas dan berisik secara efektif. Metode-metode ini memungkinkan estimasi ketidakpastian melalui pengambilan sampel dari distribusi posterior, memfasilitasi pembelajaran aktif, dan meningkatkan kinerja prediktif dengan menyediakan batas kepercayaan untuk estimasi output. Selain itu, regularisasi Bayesian membantu mengoptimalkan trade-off bias-variansi dalam model pembelajaran mendalam, meningkatkan akurasi prediktif dan ketahanan dalam menghadapi variabilitas data (Polson & Sokolov, 2017).

Metode Bayesian memiliki peran penting dalam menangani ketidakpastian dan variabilitas data dalam pembelajaran mesin dengan menyediakan kerangka kerja untuk mengukur dan menyebarkan ketidakpastian melalui model. Misalnya, jaringan Bayesian dan aljabar linier menawarkan solusi yang kuat untuk analisis respons detektor dengan secara eksplisit mempertimbangkan efek sistematis (Cortona et al.,

2023). Dalam bidang klasifikasi gambar, pembelajaran mendalam Bayesian meningkatkan kinerja model MLP-Mixer, terutama untuk dataset kecil, dengan mengukur ketidakpastian dan meningkatkan akurasi (Abdullah et al., 2023). Selain itu, dalam analisis deret waktu, inferensi Bayesian memungkinkan estimasi ketidakpastian prediktif dalam RNN melalui teknik seperti TRUST, yang menyebarkan ketidakpastian melalui lapisan-lapisan berurutan, meningkatkan ketahanan dan kemampuan penilaian diri (Dera et al., 2024). Selanjutnya, optimasi Bayesian dengan variabel campuran dalam desain material memanfaatkan model pembelajaran mesin yang sadar ketidakpastian untuk memilih desain yang menjanjikan secara efisien, mengurangi biaya eksplorasi ruang desain yang luas (H. Zhang et al., 2022). Terakhir, inferensi Bayesian dalam analisis aliran material memungkinkan komunikasi ketidakpastian yang transparan, pengumpulan pendapat ahli untuk prior parameter, dan pembelajaran bersamaan tentang kebisingan data, yang mengarah pada pengurangan ketidakpastian dan pengambilan keputusan yang lebih andal berdasarkan data yang dikumpulkan (Dong et al., 2023).

Metode-metode Bayesian memberikan keuntungan signifikan dalam menangani ketidakpastian dan variabilitas data. Dengan memanfaatkan distribusi posterior untuk estimasi parameter, metode ini tidak hanya menghasilkan prediksi yang lebih akurat tetapi juga memberikan ukuran kepastian yang lebih jelas terhadap hasil prediksi tersebut. Penggunaan jaringan Bayesian dan strategi Bayesian dalam berbagai aplikasi seperti analisis respons detektor, klasifikasi gambar, dan analisis deret waktu menunjukkan kemampuan mereka dalam menghadapi tantangan data yang kompleks dengan cara yang sistematis dan terukur. Kelebihan utama metode Bayesian adalah kemampuannya untuk mengintegrasikan informasi awal yang informatif dengan data yang ada, sehingga meningkatkan keandalan hasil prediksi. Estimasi ketidakpastian yang akurat dan kemampuan untuk menyesuaikan model dengan data baru secara dinamis merupakan aspek penting dalam konteks pengambilan keputusan yang berbasis data. Namun, tantangan yang terkait dengan kompleksitas komputasional dalam pengambilan sampel dari distribusi posterior serta interpretasi yang tepat terhadap hasil estimasi tetap menjadi fokus pengembangan di bidang ini.

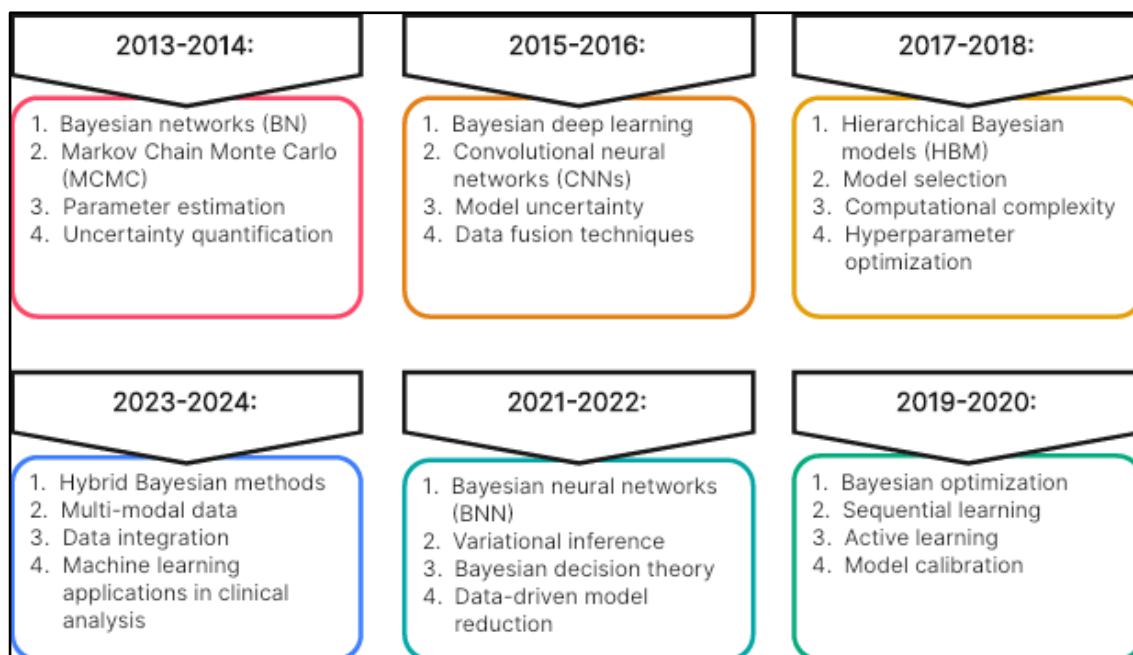
### C. Tantangan Utama Integrasi Metode Bayesian dalam Pembelajaran Mesin

Mengintegrasikan metode Bayesian ke dalam pembelajaran mesin menghadapi beberapa tantangan. Pertama, kompleksitas komputasional inferensi Bayesian, terutama dalam skenario multi-modal, dapat menyebabkan konvergensi lambat menggunakan teknik *Markov Chain Monte Carlo* (MCMC) tradisional (Feroz et al., 2019). Kedua, pemilihan hiperparameter dan model yang tepat sangat penting tetapi sering kali memerlukan iterasi manual dan pengetahuan tingkat lanjut, yang menghambat pelatihan model yang efisien (Zeng & Luo, 2017). Selain itu, biaya komputasional untuk mengevaluasi model numerik berulang kali dalam solusi Bayesian untuk masalah invers yang diatur oleh persamaan diferensial parsial menimbulkan tantangan yang signifikan, yang memerlukan pengembangan teknik reduksi model berbasis data untuk meningkatkan efisiensi tanpa mengorbankan akurasi (Cui et al., 2015). Mengatasi tantangan ini sangat penting untuk memanfaatkan manfaat metode Bayesian dalam aplikasi pembelajaran mesin, terutama dalam analisis data klinis dan skenario kalibrasi model (Held et al., 2015).

Mengintegrasikan metode Bayesian ke dalam pembelajaran mesin menghadapi beberapa tantangan. Salah satu tantangan utama adalah memperoleh pengetahuan awal yang informatif dalam istilah probabilistik, yang sangat penting untuk inversi Bayesian yang efektif (Marschall et al., 2023). Selain itu, dalam jaringan saraf Bayesian (BNN), metode *Markov chain Monte Carlo* (MCMC) mengalami kesulitan dalam konvergensi ke posterior parameter, yang membatasi akurasi estimasi parameter (Papamarkou et al., 2022). Mempelajari parameter tabel probabilitas bersyarat (CPT) dari jaringan Bayesian (BN) adalah kendala lain, terutama ketika data terbatas, yang memerlukan integrasi pengetahuan domain melalui batasan parameter kualitatif untuk meningkatkan akurasi pembelajaran parameter (Y. Zhang & Hu, 2022). Selain itu, mempelajari struktur jaringan yang mendasari jaringan Bayesian sangat menantang karena banyaknya kemungkinan jaringan dan kendala asiklikalitas, yang memerlukan pengembangan metode hibrida baru untuk mengeksplorasi ruang DAG dan mengambil sampel DAG dari distribusi posteriornya secara efisien (Kuipers et al., 2022).

Tantangan utama dalam mengintegrasikan metode Bayesian adalah kompleksitas komputasional yang tinggi saat melakukan inferensi, terutama ketika menghadapi data yang memiliki variasi yang kompleks dan bervariasi. Meskipun Metode MCMC umum digunakan, seringkali sulit untuk mencapai konvergensi yang cepat dan akurat terhadap distribusi posterior. Selain itu, pemilihan hiperparameter dan model yang optimal membutuhkan pemahaman mendalam tentang domain aplikasi dan sering kali melibatkan banyak iterasi manual, yang dapat memperlambat proses pelatihan model secara keseluruhan. Kesulitan dalam mengatasi tantangan ini dapat menghambat penggunaan metode Bayesian secara efektif dalam aplikasi pembelajaran mesin, khususnya dalam analisis data klinis dan kalibrasi model. Meskipun metode ini menawarkan keunggulan dalam mengukur dan menyebarkan ketidakpastian, penggunaannya sering memerlukan sumber daya komputasional yang besar dan pendekatan yang hati-hati dalam

penyesuaian parameter dan model. Pengembangan lebih lanjut dalam teknik reduksi model berbasis data diperlukan untuk mengatasi biaya komputasional yang tinggi yang terkait dengan evaluasi berulang model numerik dalam konteks Bayesian. Berikut adalah perkembangan variabel riset dalam integrasi Metode Bayesian dalam *Machine learning*, seperti terlihat pada Gambar 2.



Gambar 2. Perkembangan Variabel Riset dalam Integrasi Metode Bayesian dalam *Machine learning*: Tren dan Tantangan (2013-2024)

Gambar 2 menjelaskan bahwa Metode Bayesian telah menjadi landasan penting dalam pengembangan pembelajaran mesin modern dengan fokus utama pada pengelolaan ketidakpastian dan peningkatan kinerja model. Penggunaan jaringan Bayesian (BN) dan algoritma Markov Chain Monte Carlo (MCMC) sejak awal 2010-an menandai awal dari penerapan metode ini dalam analisis data yang kompleks dan beraneka ragam. Penggunaan Bayesian networks (BN) membuka jalan bagi pendekatan yang sistematis dalam memodelisasi ketidakpastian dan estimasi parameter dalam konteks aplikasi seperti analisis respons detektor dan pengolahan citra. Selama periode 2015-2016, fokus penelitian bergeser ke pembelajaran mendalam Bayesian dan penggunaan jaringan saraf konvolusi (CNNs) yang terstruktur secara probabilistik, yang menghadirkan solusi untuk mengatasi ketidakpastian model dan meningkatkan kinerja klasifikasi gambar. Pada tahun-tahun berikutnya, model Bayesian yang hierarkis (HBM) dan optimalisasi hiperparameter muncul sebagai tantangan utama, memerlukan pendekatan yang hati-hati dalam manajemen kompleksitas komputasional dan pengambilan keputusan model. Pengembangan terbaru dalam optimisasi Bayesian, seperti penggunaan pembelajaran aktif dan reduksi model berbasis data, menunjukkan pergeseran menuju pendekatan yang lebih adaptif dan efisien dalam menghadapi data yang besar dan heterogen. Implementasi teknik Bayesian dalam analisis klinis dan aplikasi pengurangan kesalahan klasifikasi menyoroti manfaatnya dalam konteks kehidupan nyata, menunjukkan kemajuan signifikan dalam memahami dan mengatasi tantangan praktis dalam pembelajaran mesin modern.

#### 4. SIMPULAN DAN SARAN

Berdasarkan evaluasi terhadap integrasi metode Bayesian dalam pembelajaran mesin, dapat disimpulkan bahwa meskipun metode ini menawarkan keunggulan signifikan dalam mengatasi ketidakpastian dan variabilitas data, terdapat beberapa tantangan yang perlu diatasi. Salah satu hambatan utama adalah kompleksitas komputasional yang tinggi dalam melakukan inferensi Bayesian. Misalnya, metode Markov Chain Monte Carlo (MCMC) yang umum digunakan sering menghadapi masalah konvergensi yang lambat dan memerlukan sumber daya komputasional yang besar. Selain itu, pemilihan hiperparameter dan model yang optimal memerlukan pemahaman mendalam tentang domain aplikasi. Proses ini sering kali melibatkan iterasi manual yang intensif, yang dapat menghambat efisiensi pelatihan model secara keseluruhan.

Kesenjangan utama yang perlu ditangani di masa mendatang adalah pengembangan teknik yang lebih efisien untuk mengatasi kompleksitas komputasional saat melakukan pengambilan sampel dari distribusi posterior, yang akan memungkinkan konvergensi yang lebih cepat dan akurat. Eksplorasi terhadap pendekatan otomatisasi dalam pemilihan hiperparameter dan model sangat dibutuhkan, mungkin dengan menggunakan algoritma pembelajaran mesin untuk mengurangi ketergantungan pada iterasi manual. Penelitian lebih lanjut juga diperlukan dalam pengembangan teknik reduksi model berbasis data, yang dapat mengurangi biaya komputasional terkait evaluasi berulang model numerik dalam konteks Bayesian. Sebagai contoh, dalam aplikasi analisis data klinis, metode Bayesian dapat digunakan untuk memprediksi respons pasien terhadap pengobatan tertentu dengan mengintegrasikan informasi dari berbagai sumber. Meskipun metode Bayesian menjanjikan dalam meningkatkan keandalan dan akurasi model pembelajaran mesin, tantangan teknis seperti kompleksitas komputasional dan pengelolaan model harus menjadi fokus utama untuk penelitian mendatang.

### Daftar Pustaka

- Abdullah, A. A., Hassan, M. M., & Mustafa, Y. T. (2023). Uncertainty Quantification for MLP-Mixer Using Bayesian Deep Learning. *Applied Sciences (Switzerland)*. <https://doi.org/10.3390/app13074547>
- Aflakparast, M., & de Gunst, M. (2019). Data integrative Bayesian inference for mixtures of regression models. *Journal of the Royal Statistical Society. Series C: Applied Statistics*. <https://doi.org/10.1111/rssc.12346>
- Albert, C. G., Callies, U., & Toussaint, U. von. (2022). A Bayesian Approach to the Estimation of Parameters and Their Interdependencies in Environmental Modeling. *Entropy*. <https://doi.org/10.3390/e24020231>
- Bertolini, R., Finch, S. J., & Nehm, R. H. (2023). An application of Bayesian inference to examine student retention and attrition in the STEM classroom. *Frontiers in Education*. <https://doi.org/10.3389/educ.2023.1073829>
- Cui, T., Marzouk, Y. M., & Willcox, K. E. (2015). Data-driven model reduction for the Bayesian solution of inverse problems. *International Journal for Numerical Methods in Engineering*. <https://doi.org/10.1002/nme.4748>
- De Franco, C., Nicolle, J., & Pham, H. (2018). Dealing with Drift Uncertainty: A Bayesian Learning Approach. *SSRN Electronic Journal*. <https://doi.org/10.2139/ssrn.3285248>
- Dera, D., Ahmed, S., Bouaynaya, N. C., & Rasool, G. (2024). TRustworthy Uncertainty Propagation for Sequential Time-Series Analysis in RNNs. *IEEE Transactions on Knowledge and Data Engineering*. <https://doi.org/10.1109/TKDE.2023.3288628>
- Di Cortona, G. G., Messina, A., & Piacentini, S. (2023). Treatment of systematic uncertainties with Bayesian networks. *Journal of Physics: Conference Series*. <https://doi.org/10.1088/1742-6596/2438/1/012131>
- Dong, J., Liao, J., Huan, X., & Cooper, D. (2023). Expert elicitation and data noise learning for material flow analysis using Bayesian inference. *Journal of Industrial Ecology*. <https://doi.org/10.1111/jiec.13399>
- Edinburgh, T., Ercole, A., & Eglen, S. (2023). Bayesian model selection for multilevel models using integrated likelihoods. *PLoS ONE*. <https://doi.org/10.1371/journal.pone.0280046>
- Feroz, F., Hobson, M. P., Cameron, E., & Pettitt, A. N. (2019). Importance Nested Sampling and the MultiNest Algorithm. *The Open Journal of Astrophysics*. <https://doi.org/10.21105/astro.1306.2144>
- Held, L., Bové, D. S., & Gravestock, I. (2015). Approximate Bayesian model selection with the deviance statistic. *Statistical Science*. <https://doi.org/10.1214/14-STS510>
- Hobbs, B. F. (1997). Bayesian methods for analysing climate change and water resource uncertainties. *Journal of Environmental Management*. <https://doi.org/10.1006/jema.1996.0116>
- Ida Bagus Indra Dewangkara, Nyoman Dewi Ratna Iswara, Made NikenNovianti, Alif Bangun Prakoso, Laisa Santika, Adi Rizky Setiawan, AanTriambudi, Asafik Daroini, Eriz Rifqi Irawan, Ni Wayan Destiena Marella Apsari, M. A. L. (2022). PENGEMBANGAN WEBSITE PROJECT MANAGEMENT SYSTEM SPBE CAMP MENGGUNAKAN METODE DESIGN THINKING DAN FRAMEWORK LARAVEL. *INSERT: Information System and Emerging Technology Journal*, 3(2), 105–116. <https://ejournal.undiksha.ac.id/index.php/insert/article/view/55961>
- Johnson, M. S., & Green, W. H. (2024). A machine learning based approach to reaction rate estimation. *Reaction Chemistry and Engineering*. <https://doi.org/10.1039/d3re00684k>
- Jørgensen, A. C. S., Ghosh, A., Sturrock, M., & Shahrezaei, V. (2022). Efficient Bayesian inference for stochastic



- agent-based models. *PLoS Computational Biology*. <https://doi.org/10.1371/journal.pcbi.1009508>
- Kaplan, D. (2021). On the Quantification of Model Uncertainty: A Bayesian Perspective. *Psychometrika*. <https://doi.org/10.1007/s11336-021-09754-5>
- Kottke, D., Herde, M., Sandrock, C., Huseljic, D., Kreml, G., & Sick, B. (2021). Toward optimal probabilistic active learning using a Bayesian approach. *Machine Learning*. <https://doi.org/10.1007/s10994-021-05986-9>
- Kuipers, J., Suter, P., & Moffa, G. (2022). Efficient Sampling and Structure Learning of Bayesian Networks. *Journal of Computational and Graphical Statistics*. <https://doi.org/10.1080/10618600.2021.2020127>
- Lavecchia, A. (2015). Machine-learning approaches in drug discovery: Methods and applications. In *Drug Discovery Today*. <https://doi.org/10.1016/j.drudis.2014.10.012>
- Lee, M. D. (2018). Bayesian Methods in Cognitive Modeling. In *Stevens' Handbook of Experimental Psychology and Cognitive Neuroscience*. <https://doi.org/10.1002/9781119170174.epcn502>
- Luo, X., & Kareem, A. (2020). Bayesian deep learning with hierarchical prior: Predictions from limited and noisy data. *Structural Safety*. <https://doi.org/10.1016/j.strusafe.2019.101918>
- Manderson, A. A., & Goudie, R. J. B. (2022). A numerically stable algorithm for integrating Bayesian models using Markov melding. *Statistics and Computing*. <https://doi.org/10.1007/s11222-022-10086-2>
- Marschall, M., Wübbeler, G., Schmähling, F., & Elster, C. (2023). *Machine learning* based priors for Bayesian inversion in MR imaging. *Metrologia*. <https://doi.org/10.1088/1681-7575/ace3c2>
- Mezzetti, M., Ryan, C. P., Balestrucci, P., Lacquaniti, F., & Moscatelli, A. (2023). Bayesian hierarchical models and prior elicitation for fitting psychometric functions. *Frontiers in Computational Neuroscience*. <https://doi.org/10.3389/fncom.2023.1108311>
- Mohammadian mosammam, A., Abbasi, E., & Mateu, J. (2023). Bayesian Approach for Modelling Spatial-Temporal Crime Data. *Journal of Statistical Sciences*. <https://doi.org/10.52547/jss.16.2.435>
- Musil, F., Willatt, M. J., Langovoy, M. A., & Ceriotti, M. (2019). Fast and Accurate Uncertainty Estimation in Chemical *Machine Learning*. *Journal of Chemical Theory and Computation*. <https://doi.org/10.1021/acs.jctc.8b00959>
- Nabi, S., Nassif, H., Hong, J., Mamani, H., & Imbens, G. (2022). Bayesian Meta-Prior Learning Using Empirical Bayes. *Management Science*. <https://doi.org/10.1287/mnsc.2021.4136>
- Panigrahi, S., Mohammed, S., Rao, A., & Baladandayuthapani, V. (2023). Integrative Bayesian models using Post-selective inference: A case study in radiogenomics. *Biometrics*. <https://doi.org/10.1111/biom.13740>
- Papamarkou, T., Hinkle, J., Young, M. T., & Womble, D. (2022). Challenges in Markov Chain Monte Carlo for Bayesian Neural Networks. *Statistical Science*. <https://doi.org/10.1214/21-STS840>
- Parker, R. A., Sande, T. A., Laird, B., Hoskin, P., Fallon, M., & Colvin, L. (2022). Bayesian methods in palliative care research: cancer-induced bone pain. *BMJ Supportive and Palliative Care*. <https://doi.org/10.1136/bmjspcare-2019-002160>
- Pavone, A., Merlo, A., Kwak, S., & Svensson, J. (2023). *Machine learning* and Bayesian inference in nuclear fusion research: an overview. In *Plasma Physics and Controlled Fusion*. <https://doi.org/10.1088/1361-6587/acc60f>
- Pirš, G., & Strumbelj, E. (2019). Bayesian combination of probabilistic classifiers using multivariate normal mixtures. *Journal of Machine Learning Research*.
- Polson, N. G., & Sokolov, V. (2017). Deep learning: A Bayesian perspective. *Bayesian Analysis*. <https://doi.org/10.1214/17-BA1082>
- Pratama, R. R. (2020). Analisis Model *Machine learning* Terhadap Pengenalan Aktifitas Manusia. *MATRIK: Jurnal Manajemen, Teknik Informatika Dan Rekayasa Komputer*. <https://doi.org/10.30812/matrik.v19i2.688>
- QOSIM, A. L. (2021). PERBANDINGAN METODE KLASIFIKASI SUPPORT VECTOR MACHINE (SVM) DAN NAIVE BAYES CLASSIFIER (NBC) UNTUK MENENTUKAN KUALITAS UDARA. In *Industry and Higher Education*.
- Sambasivan, R., Das, S., & Sahu, S. K. (2020). A Bayesian perspective of statistical *machine learning* for big data. *Computational Statistics*. <https://doi.org/10.1007/s00180-020-00970-8>
- Schillings, C., Sprungk, B., & Wacker, P. (2020). On the convergence of the Laplace approximation and noise-level-robustness of Laplace-based Monte Carlo methods for Bayesian inverse problems. *Numerische Mathematik*. <https://doi.org/10.1007/s00211-020-01131-1>

- Sukarya, I. G. A., Pradnyana, I. M. A., & Sugihartini, N. (2021). Analisis Faktor-faktor yang Mempengaruhi Perilaku Penggunaan Sistem E-Learning Undiksha dengan Model Unified Theory of Acceptance and Use of Technology (UTAUT). *INSERT: Information System and Emerging Technology Journal*. <https://doi.org/10.23887/insert.v1i2.25940>
- van de Schoot, R., Depaoli, S., King, R., Kramer, B., Märtens, K., Tadesse, M. G., Vannucci, M., Gelman, A., Veen, D., Willemsen, J., & Yau, C. (2021). Bayesian statistics and modelling. In *Nature Reviews Methods Primers*. <https://doi.org/10.1038/s43586-020-00001-2>
- Wang, Z. A., Pei, J. C., Chen, Y. J., Qiao, C. Y., Xu, F. R., Ge, Z. G., & Shu, N. C. (2022). Bayesian approach to heterogeneous data fusion of imperfect fission yields for augmented evaluations. *Physical Review C*. <https://doi.org/10.1103/PhysRevC.106.L021304>
- Zeng, X., & Luo, G. (2017). Progressive sampling-based Bayesian optimization for efficient and automatic machine learning model selection. *Health Information Science and Systems*. <https://doi.org/10.1007/s13755-017-0023-z>
- Zhang, H., Chen, W. (Wayne), Iyer, A., Apley, D. W., & Chen, W. (2022). Uncertainty-aware mixed-variable machine learning for materials design. *Scientific Reports*. <https://doi.org/10.1038/s41598-022-23431-2>
- Zhang, Y., & Hu, Z. (2022). Parameter Learning of Bayesian Network with Multiplicative Synergistic Constraints. *Symmetry*. <https://doi.org/10.3390/sym14071469>
- Zhang, Y., & Lee, A. A. (2019). Bayesian semi-supervised learning for uncertainty-calibrated prediction of molecular properties and active learning. *Chemical Science*. <https://doi.org/10.1039/c9sc00616h>