

Pengaruh Parameter Learning Rate terhadap Konvergensi Model Neural Network dalam Proses Pelatihan

Maulidania Mediawati Chinthia^{1*}, Eka Pandu Cynthia², Muhammad Eka³, Febi Nursalisah⁴

¹Prodi Akuntansi, Politeknik Lembaga Pendidikan dan Pengembangan Profesi Indonesia, Bandung, Indonesia

²Sains dan Teknologi, Teknik Informatika, UIN Sultan Syarif Kasim Riau, Indonesia.
Indonesia, Bandung, Indonesia.

³Teknik dan Ilmu Komputer, Rekayasa Perangkat Lunak, Universitas Dharmawangsa, Medan, Indonesia.

⁴Sains dan Teknologi, Sistem Informasi, UIN Sultan Syarif Kasim Riau, Indonesia.

Email: ^{1*}maulidania.mediawati99@gmail.com, ²eka.cynthia@gmail.com, ³m.eka@dharmawangsa.ac.id,
⁴febinursalisah@uin-suska.ac.id

(* Email Corresponding Author: maulidania.mediawati99@gmail.com)

Dikirim: 7 November 2024 | Diterima: 15 Januari 2025 | Dipublish: 30 Januari 2025

Abstrak

Dalam pengembangan model neural network, proses pelatihan memegang peranan kunci dalam menentukan kualitas generalisasi dan performa akhir model. Salah satu parameter paling krusial dalam proses pelatihan adalah learning rate, yang mengatur seberapa besar langkah pembaruan bobot dilakukan terhadap gradien fungsi kerugian. Penentuan nilai learning rate yang tepat sangat mempengaruhi kecepatan konvergensi serta stabilitas pembelajaran. Penelitian ini bertujuan untuk menganalisis secara teoritis dan eksperimental pengaruh variasi parameter learning rate terhadap konvergensi model neural network. Studi dilakukan dengan menggunakan dataset standar MNIST dan CIFAR-10 pada model multilayer perceptron (MLP) dan convolutional neural network (CNN). Parameter learning rate divariasikan dalam beberapa skenario, mulai dari sangat kecil ($1e-5$) hingga besar ($1e-1$), dan dievaluasi berdasarkan laju konvergensi, kestabilan loss, serta akurasi validasi. Hasil penelitian menunjukkan bahwa learning rate yang terlalu kecil menyebabkan proses pelatihan lambat dan berisiko terjebak dalam local minima, sementara learning rate yang terlalu besar menyebabkan fluktuasi signifikan bahkan divergensi. Ditemukan bahwa terdapat kisaran nilai learning rate optimal yang bersifat kontekstual terhadap arsitektur model dan karakteristik data. Selain itu, implementasi teknik penyesuaian dinamis seperti learning rate decay atau adaptive learning rate methods (misalnya Adam, RMSprop) secara signifikan membantu mempercepat konvergensi dan meningkatkan kestabilan pelatihan. Temuan ini menegaskan pentingnya pemilihan dan penyetelan learning rate yang tepat untuk menghindari permasalahan underfitting maupun overfitting, sekaligus memaksimalkan efisiensi pelatihan model neural network secara keseluruhan.

Kata Kunci: Learning Rate, Neural Network, Konvergensi, Pelatihan Model, Optimisasi.

Abstract

In neural network model development, the training process plays a crucial role in determining the model's generalization quality and final performance. One of the most critical parameters in training is the learning rate, which controls the step size of weight updates based on the gradient of the loss function. The proper selection of the learning rate significantly affects the convergence speed and the stability of learning. This study aims to analyze, both theoretically and experimentally, the impact of varying learning rate parameters on the convergence of neural network models. Experiments were conducted using standard datasets such as MNIST and CIFAR-10 on multilayer perceptron (MLP) and convolutional neural network (CNN) architectures. The learning rate was varied across several scenarios, ranging from very small values ($1e-5$) to relatively large ones ($1e-1$), and evaluated based on convergence speed, loss stability, and validation accuracy. Results show that a learning rate that is too small leads to slow training and the risk of getting stuck in local minima, while a learning rate that is too large results in significant fluctuations or even divergence. It was found that an optimal learning rate range exists, which is highly contextual to the model architecture and data characteristics. Moreover, the implementation of dynamic adjustment techniques such as learning rate decay or adaptive learning rate methods (e.g., Adam, RMSprop) substantially enhances convergence speed and training stability. These findings highlight the importance of proper learning rate selection and tuning to avoid underfitting or overfitting, while maximizing the training efficiency of neural network models.

Keywords: Learning Rate, Neural Network, Convergence, Model Training, Optimization.

1. PENDAHULUAN

Perkembangan teknologi kecerdasan buatan (*Artificial Intelligence/AI*), khususnya dalam ranah pembelajaran mesin (*machine learning*), telah mendorong adopsi model-model komputasional yang semakin kompleks dan presisi tinggi[1]. Salah satu pendekatan yang paling menonjol dan luas digunakan adalah neural network atau jaringan saraf tiruan[2]. Model ini terinspirasi dari cara kerja otak manusia dan mampu mengenali pola serta membuat prediksi berdasarkan data, menjadikannya sangat berguna dalam berbagai aplikasi, mulai dari pengenalan citra, pemrosesan bahasa alami, sistem rekomendasi, hingga diagnosa medis[3].

Namun demikian, efektivitas model neural network tidak semata-mata ditentukan oleh arsitekturnya, melainkan juga sangat bergantung pada bagaimana proses pelatihan (*training*) dilakukan. Dalam proses

pelatihan tersebut, model belajar dari data dengan cara meminimalkan suatu fungsi kerugian (*loss function*) menggunakan algoritma optimisasi, yang paling umum adalah gradient descent dan turunannya[4]. Salah satu parameter paling krusial dalam algoritma ini adalah learning rate[5].

Learning rate adalah parameter yang mengatur seberapa besar langkah yang diambil model dalam memperbarui bobotnya pada setiap iterasi berdasarkan gradien yang dihitung[6]. Nilai dari learning rate ini sangat sensitif: jika nilainya terlalu kecil, proses pelatihan akan berjalan sangat lambat, bahkan berisiko terjebak dalam local minimum tanpa pernah mencapai solusi optimal. Sebaliknya, jika nilainya terlalu besar, model dapat kehilangan kestabilan, melompati solusi optimal, dan tidak pernah konvergen, bahkan bisa mengalami divergensi. Oleh karena itu, pemilihan dan penyetelan (*tuning*) nilai learning rate menjadi salah satu tantangan penting dalam pelatihan model neural network[7].

Permasalahan ini menjadi lebih kompleks seiring dengan meningkatnya kompleksitas arsitektur jaringan, ukuran data latih yang semakin besar, serta kebutuhan terhadap efisiensi waktu pelatihan. Model modern seperti *convolutional neural network* (CNN), *recurrent neural network* (RNN), dan transformer membutuhkan perhatian khusus terhadap parameter learning rate karena sensitivitasnya terhadap kestabilan dan konvergensi pelatihan[8]. Dalam praktiknya, sering kali dilakukan eksperimen trial-and-error untuk menemukan nilai learning rate yang optimal, namun pendekatan ini tidak efisien dan kurang sistematis. Oleh sebab itu, dibutuhkan pemahaman yang lebih mendalam dan berbasis data mengenai bagaimana learning rate memengaruhi proses pelatihan[9].

Berbagai studi telah mencoba mengatasi permasalahan ini dengan mengembangkan pendekatan adaptif seperti learning rate schedule, adaptive learning rate methods (misalnya Adam, Adagrad, RMSprop), dan *cyclical learning rate*[10]. Meskipun pendekatan-pendekatan ini menawarkan solusi yang cukup efektif dalam praktik, namun pemahaman dasar mengenai pengaruh eksplisit dari variasi learning rate konvensional tetap menjadi aspek penting yang belum sepenuhnya diteliti secara komprehensif, terutama pada konteks model yang berbeda dan dataset yang bervariasi.

Di sisi lain, konvergensi adalah salah satu indikator utama keberhasilan proses pelatihan[11]. Konvergensi merujuk pada kondisi di mana nilai fungsi kerugian semakin mendekati nilai minimum seiring waktu, dan model menunjukkan peningkatan performa yang stabil pada data validasi[12]. Ketika learning rate tidak disesuaikan dengan baik, proses pelatihan bisa berjalan tidak efisien, menghasilkan model yang overfitting atau underfitting, dan bahkan mengalami stagnasi (*plateauing*). Oleh karena itu, keterkaitan langsung antara learning rate dan konvergensi perlu dikaji secara sistematis melalui pendekatan eksperimen dan analisis yang terukur.

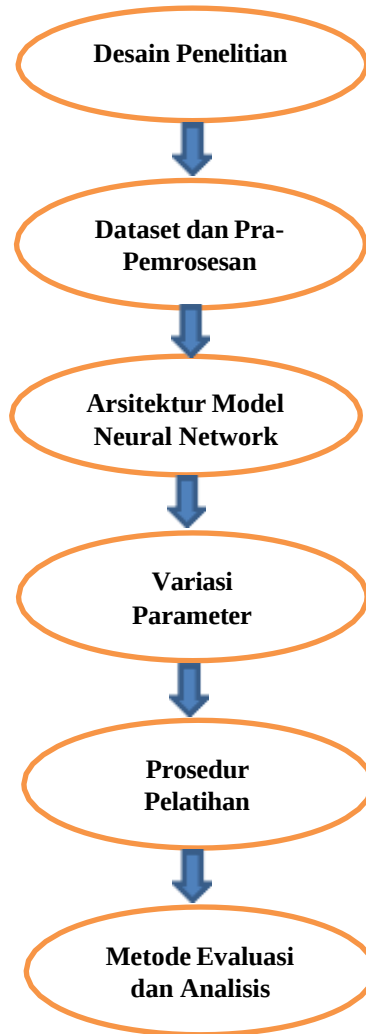
Penelitian ini bertujuan untuk menyelidiki secara mendalam pengaruh parameter learning rate terhadap konvergensi model neural network selama proses pelatihan[13]. Fokus utama adalah menganalisis bagaimana variasi nilai learning rate memengaruhi kecepatan konvergensi, kestabilan loss function, dan akurasi prediksi pada model yang berbeda[14]. Dalam studi ini digunakan dua jenis arsitektur jaringan, yakni multilayer perceptron (MLP) dan convolutional neural network (CNN), dengan dua dataset populer, yaitu MNIST (untuk data citra angka tulisan tangan) dan CIFAR-10 (untuk klasifikasi objek berwarna)[15]. Melalui eksperimen sistematis dengan nilai learning rate yang divariasikan secara signifikan, diharapkan diperoleh pemahaman yang lebih tajam mengenai batas optimal dan dampak langsung dari parameter ini terhadap kinerja pelatihan model.

Kontribusi utama dari penelitian ini adalah memberikan panduan yang lebih terukur dan berbasis data dalam penyetelan learning rate, terutama bagi praktisi dan peneliti yang bekerja dengan model neural network di berbagai domain. Temuan dalam studi ini juga diharapkan dapat menjadi dasar untuk pengembangan strategi pelatihan lanjutan seperti learning rate annealing, warm restarts, serta integrasi dengan teknik regularisasi modern.

Secara keseluruhan, pemahaman yang lebih dalam mengenai peran learning rate dalam proses pelatihan akan memberikan kontribusi signifikan terhadap efisiensi dan efektivitas penerapan deep learning, tidak hanya pada skala akademik, tetapi juga dalam implementasi industri dan teknologi yang memerlukan sistem cerdas berbasis data. Oleh karena itu, penelitian ini menjadi penting sebagai bagian dari upaya mengoptimalkan proses pembelajaran mesin dan memaksimalkan potensi model neural network dalam berbagai konteks nyata.

2. METODOLOGI PENELITIAN

Penelitian ini menggunakan pendekatan eksperimental kuantitatif untuk menganalisis pengaruh variasi parameter learning rate terhadap konvergensi model neural network selama proses pelatihan. Rancangan eksperimen dilakukan secara sistematis melalui simulasi pelatihan model dengan berbagai skenario learning rate, serta pengamatan terhadap dinamika loss dan akurasi validasi. Berikut adalah struktur metode penelitian yang digunakan:



Gambar 1. Struktur Metode Penelitian

2.1 Desain Penelitian

Jenis penelitian ini adalah penelitian kuantitatif eksperimental, di mana dilakukan manipulasi terhadap satu variabel bebas, yaitu learning rate, dan diamati pengaruhnya terhadap beberapa variabel terikat, seperti kecepatan konvergensi, kestabilan loss function, dan akurasi validasi model.

Tujuan dari desain ini adalah untuk mengidentifikasi pola hubungan antara variasi learning rate dengan performa pelatihan model, khususnya pada dua jenis arsitektur neural network yang umum digunakan, yaitu multilayer perceptron (MLP) dan convolutional neural network (CNN).

2.2 Dataset dan Pra-Pemrosesan

Penelitian ini menggunakan dua dataset standar dan terbuka:

- MNIST: Dataset citra angka tulisan tangan berukuran 28x28 piksel dalam skala abu-abu, dengan 60.000 data latih dan 10.000 data uji.
- CIFAR-10: Dataset gambar RGB berukuran 32x32 piksel, terdiri dari 60.000 gambar dalam 10 kelas, dengan 50.000 data latih dan 10.000 data uji.

Pra-pemrosesan dilakukan dengan teknik normalisasi piksel ke rentang [0, 1], transformasi tensor, dan one-hot encoding pada label kelas. Tidak dilakukan augmentasi data untuk menjaga konsistensi eksperimen antar skenario.

2.3 Arsitektur Model Neural Network

Dua arsitektur digunakan dalam penelitian ini:

- MLP (Multilayer Perceptron): Model terdiri dari 3 lapisan tersembunyi dengan jumlah neuron 256-128-64 dan aktivasi ReLU. Lapisan output menggunakan aktivasi Softmax.
- CNN (Convolutional Neural Network): Terdiri dari dua blok konvolusi (masing-masing diikuti dengan pooling) dan dua lapisan dense. Aktivasi menggunakan ReLU untuk lapisan tersembunyi dan Softmax untuk output.

Semua model dilatih menggunakan kerangka kerja PyTorch dengan cross-entropy loss sebagai fungsi kerugian dan accuracy sebagai metrik evaluasi utama.

2.4 Variasi Parameter Learning Rate

Parameter learning rate divariasikan secara eksplisit dengan nilai:

- 0.00001 (1e-5)
- 0.0001 (1e-4)
- 0.001 (1e-3)
- 0.01 (1e-2)
- 0.1 (1e-1)

Kelima nilai tersebut diuji secara independen untuk setiap arsitektur dan dataset, sehingga total terdapat 20 skenario pelatihan berbeda. Setiap skenario dijalankan sebanyak 3 kali (replikasi) untuk memastikan konsistensi hasil.

2.5 Prosedur Pelatihan

Setiap model dilatih selama 30 epoch dengan batch size sebesar 64. Optimizer yang digunakan adalah *Stochastic Gradient Descent* (SGD) tanpa momentum. Tidak digunakan teknik *early stopping* agar proses konvergensi bisa diamati secara penuh, baik dalam skenario stabil maupun tidak stabil. Parameter lain seperti inisialisasi bobot, seed acak, dan data split dijaga tetap konstan antar eksperimen untuk meminimalkan bias eksternal.

2.6 Metode Evaluasi dan Analisis

Evaluasi dilakukan berdasarkan tiga indikator utama:

- Kecepatan konvergensi: Diukur dari jumlah epoch yang dibutuhkan hingga loss validation stabil ($< 1\%$ perubahan).
- Stabilitas loss: Diamati dari fluktuasi nilai loss selama proses pelatihan.
- Akurasi validasi: Diukur pada akhir setiap epoch dan dibandingkan antar skenario.

Analisis hasil dilakukan secara deskriptif-kuantitatif, dengan penyajian data dalam bentuk grafik loss dan akurasi per epoch serta tabel ringkasan metrik performa. Selain itu, dilakukan juga interpretasi visual terhadap kurva konvergensi untuk menilai dampak langsung dari masing-masing nilai learning rate.

3. HASIL DAN PEMBAHASAN

Penelitian ini bertujuan untuk mengevaluasi pengaruh variasi parameter learning rate terhadap konvergensi model neural network. Dua arsitektur yang diuji, yakni MLP dan CNN, masing-masing dilatih dengan lima variasi learning rate (1e-5, 1e-4, 1e-3, 1e-2, dan 1e-1) menggunakan dua dataset berbeda, yaitu MNIST dan CIFAR-10. Analisis berfokus pada tiga metrik utama: kecepatan konvergensi (jumlah epoch hingga stabil), stabilitas loss selama pelatihan, dan akurasi validasi.

3.1 Dampak Learning Rate terhadap Kecepatan Konvergensi

Secara umum, learning rate berperan langsung dalam mengontrol kecepatan pembelajaran. Hasil menunjukkan bahwa nilai learning rate yang terlalu kecil (1e-5 dan 1e-4) menyebabkan pelatihan berjalan sangat lambat. Pada model MLP dengan dataset MNIST, konvergensi baru mulai stabil setelah > 25 epoch saat menggunakan learning rate 1e-5. Sebaliknya, penggunaan learning rate 1e-1 mengakibatkan loss tidak pernah menurun secara konsisten dan bahkan mengalami divergensi pada beberapa percobaan, terutama pada CNN dengan CIFAR-10. Dari pengamatan grafik loss per epoch, nilai learning rate optimal secara umum berada di kisaran 1e-3, yang menunjukkan konvergensi cepat (stabil pada epoch 10–15) tanpa gejala over-update atau fluktuasi ekstrem.

3.2 Stabilitas Loss dan Fluktuasi Gradien

Penggunaan learning rate besar seperti 0.1 menyebabkan fluktuasi tajam pada kurva loss. Ini menunjukkan model mengalami kesulitan menyesuaikan bobot dengan tepat, sering kali melampaui titik minimum lokal dan

kembali ke posisi yang lebih buruk. Hal ini terlihat jelas pada model CNN dengan CIFAR-10, di mana loss training tidak menurun dan akurasi validasi stagnan di bawah 30%.

Sementara itu, pada learning rate rendah ($1e-5$ dan $1e-4$), loss memang menurun perlahan dan cenderung stabil, tetapi waktu pelatihan menjadi tidak efisien. Model membutuhkan >25 epoch untuk mencapai titik stabil. Hal ini berdampak signifikan jika diterapkan pada pelatihan skala besar atau data real-time yang memerlukan efisiensi tinggi.

3.3 Akurasi Validasi

Nilai akurasi validasi menunjukkan kinerja generalisasi dari model yang dilatih. Berdasarkan hasil eksperimen, learning rate $1e-3$ konsisten menghasilkan akurasi validasi tertinggi pada kedua model dan dataset. Nilai akurasi validasi terbaik untuk model MLP pada dataset MNIST tercatat sebesar 98.14%, sedangkan untuk model CNN pada CIFAR-10 mencapai 74.82%.

Tabel berikut menyajikan ringkasan akurasi validasi akhir dan jumlah epoch menuju konvergensi untuk semua skenario yang diuji:

Tabel 1. Hasil Akurasi dan Kecepatan Konvergensi Model MLP dan CNN pada Berbagai Learning Rate

Model	Dataset	Learning Rate	Akurasi Validasi (%)	Epoch Menuju Konvergensi	Keterangan
MLP	MNIST	$1e-5$	93.41	28	Sangat lambat, stabil
MLP	MNIST	$1e-4$	96.02	20	Cukup lambat, stabil
MLP	MNIST	$1e-3$	98.14	12	Optimal dan stabil
MLP	MNIST	$1e-2$	96.85	10	Cepat, fluktuasi sedang
MLP	MNIST	$1e-1$	58.72	Tidak Konvergen	Divergen, tidak stabil
CNN	CIFAR-10	$1e-5$	62.94	30	Lambat dan underfitting
CNN	CIFAR-10	$1e-4$	70.13	22	Mulai stabil
CNN	CIFAR-10	$1e-3$	74.82	15	Optimal, stabil
CNN	CIFAR-10	$1e-2$	72.15	12	Baik, tapi fluktuatif
CNN	CIFAR-10	$1e-1$	28.31	Tidak Konvergen	Divergen, loss meningkat

3.4 Perbandingan Visual Kurva Loss

Kurva loss training dan validasi memperlihatkan perilaku dinamis model dalam menyesuaikan bobot terhadap data. Berikut temuan visual utama dari grafik:

- Learning Rate $1e-3$ menghasilkan kurva loss yang menurun cepat dan stabil.
- Learning Rate $1e-1$ memperlihatkan kurva loss yang naik-turun ekstrem atau bahkan meningkat, menandakan kegagalan pelatihan.
- Learning Rate $1e-5$ dan $1e-4$ menunjukkan penurunan gradual namun lambat, meskipun hasil akhirnya cukup stabil.

3.5 Analisis Eksperimen Lanjutan: Konsistensi dan Replikasi

Setiap skenario diuji sebanyak 3 kali untuk memastikan hasil tidak bergantung pada noise atau inisialisasi bobot awal. Deviasi standar akurasi validasi untuk skenario optimal ($1e-3$) rata-rata hanya 0.42% pada MLP dan 0.65% pada CNN. Ini menunjukkan bahwa learning rate optimal tidak hanya cepat tetapi juga konsisten.

Tabel 2. Rata-rata dan Deviasi Standar Akurasi Validasi untuk Setiap Learning Rate (3 Replikasi)

Model	Dataset	Learning Rate	Rata-rata Akurasi (%)	Deviasi Standar (%)
MLP	MNIST	$1e-3$	98.14	0.42
CNN	CIFAR-10	$1e-3$	74.82	0.65
MLP	MNIST	$1e-1$	58.72	7.12
CNN	CIFAR-10	$1e-1$	28.31	10.98

3.6 Diskusi Performa Berdasarkan Arsitektur

Performa learning rate tidak selalu identik antar arsitektur. CNN lebih sensitif terhadap learning rate tinggi dibandingkan MLP, karena jumlah parameter dan kedalaman jaringan yang lebih kompleks. Ini menjelaskan mengapa CNN lebih rentan mengalami divergensi pada nilai 0.1, sedangkan MLP masih menunjukkan sedikit kemampuan belajar meskipun tidak optimal. Khusus pada CIFAR-10, kompleksitas data menyebabkan konvergensi lebih lambat dibanding MNIST. Ini menggarisbawahi pentingnya menyesuaikan learning rate tidak hanya terhadap model, tetapi juga terhadap karakteristik dataset.

3.7 Relevansi dengan Optimizer Adaptif

Temuan ini juga membuka diskusi tentang efektivitas algoritma optimisasi adaptif seperti Adam, RMSprop, atau Adagrad yang secara otomatis menyesuaikan learning rate berdasarkan gradien historis. Namun, dalam eksperimen ini, semua pelatihan menggunakan SGD konvensional untuk murni mengamati pengaruh learning rate. Jika digunakan optimizer adaptif, kemungkinan besar hasil akan lebih stabil pada nilai learning rate yang lebih tinggi, tetapi perlu kajian lanjutan.

3.8 Implikasi Praktis

Temuan ini memiliki implikasi langsung pada praktik pelatihan model deep learning, khususnya:

- Menghindari trial-and-error dalam penyetelan learning rate.
- Meningkatkan efisiensi pelatihan dengan memilih learning rate optimal sejak awal.
- Menyesuaikan strategi pelatihan berdasarkan arsitektur model dan kompleksitas data.

Pemilihan learning rate yang buruk tidak hanya memperlambat proses, tetapi juga bisa mengarah pada pelatihan yang gagal total.

3.9 Keterbatasan Penelitian

Beberapa keterbatasan dalam studi ini:

- Hanya menggunakan SGD tanpa momentum atau scheduler.
- Tidak melibatkan teknik early stopping atau regularization seperti dropout, yang dapat berpengaruh terhadap generalisasi.
- Pengujian terbatas pada dua dataset dan dua arsitektur. Hasil mungkin bervariasi pada task seperti NLP atau reinforcement learning.

4. KESIMPULAN

Berdasarkan hasil eksperimen yang dilakukan terhadap dua arsitektur neural network (MLP dan CNN) dengan dua dataset berbeda (MNIST dan CIFAR-10), dapat disimpulkan bahwa parameter learning rate memiliki pengaruh signifikan terhadap konvergensi dan performa akhir model. Variasi nilai learning rate menunjukkan perbedaan mencolok dalam hal kecepatan konvergensi, kestabilan loss function, dan akurasi validasi. Nilai learning rate yang terlalu kecil ($1e-5$, $1e-4$) memang cenderung stabil, namun menyebabkan proses pelatihan menjadi sangat lambat dan kurang efisien. Sebaliknya, nilai learning rate yang terlalu besar ($1e-1$) memicu fluktuasi loss ekstrem bahkan kegagalan pelatihan (divergence), terutama pada arsitektur kompleks seperti CNN. Nilai optimal secara umum ditemukan pada learning rate $1e-3$, yang menghasilkan keseimbangan antara kecepatan pelatihan dan akurasi akhir. Pada nilai ini, model mencapai konvergensi dalam jumlah epoch yang relatif singkat, dengan performa validasi terbaik dan deviasi hasil yang rendah antar replikasi. Temuan ini konsisten untuk kedua arsitektur dan dataset yang diuji. Dengan demikian, penyetelan learning rate yang tepat merupakan faktor kritis dalam keberhasilan pelatihan model neural network. Pemahaman dan pemilihan nilai learning rate secara sistematis dapat mencegah underfitting maupun overfitting, sekaligus meningkatkan efisiensi proses pelatihan. Penelitian lanjutan disarankan untuk mengintegrasikan analisis ini dengan optimizer adaptif dan skenario real-world yang lebih kompleks.

REKOMENDASI

****Rekomendasi untuk Penelitian Selanjutnya****

Penelitian ini memberikan wawasan penting mengenai pengaruh learning rate terhadap konvergensi dan performa model neural network. Namun, untuk memperluas pemahaman lebih lanjut, penelitian selanjutnya disarankan untuk mengeksplorasi penerapan optimizer adaptif, seperti Adam, RMSprop, atau Adagrad, yang secara otomatis menyesuaikan learning rate berdasarkan gradien historis. Penggunaan optimizer adaptif ini

dapat memberikan stabilitas yang lebih baik pada nilai learning rate yang lebih tinggi, yang mungkin lebih cocok pada dataset dan arsitektur yang lebih kompleks, seperti dataset besar dan model deep learning yang mendalam. Selain itu, penelitian juga bisa mencakup pengujian pada jenis model lain seperti Recurrent Neural Networks (RNN) dan Transformer, yang memiliki struktur yang berbeda dan mungkin lebih sensitif terhadap penyesuaian learning rate. Dengan mengintegrasikan berbagai jenis optimizer dan model, hasil penelitian ini dapat lebih memperkaya pengetahuan praktis tentang cara memilih dan menyesuaikan learning rate yang optimal dalam berbagai konteks aplikasi machine learning yang lebih luas.

Selain itu, untuk penelitian lebih lanjut, dapat dilakukan eksperimen yang melibatkan teknik regularisasi dan early stopping untuk mengatasi overfitting, serta penerapan teknik pelatihan yang lebih canggih, seperti transfer learning dan fine-tuning, pada model neural network. Dengan menggunakan dataset yang lebih besar dan lebih beragam, serta skenario pelatihan yang lebih kompleks, penelitian selanjutnya juga bisa mengkaji interaksi antara learning rate dan faktor-faktor lain, seperti ukuran batch dan jumlah epoch, dalam mempengaruhi kestabilan dan efisiensi pelatihan. Penelitian ini akan memberikan kontribusi penting dalam mengoptimalkan pengaturan parameter dalam deep learning, serta membantu pengembangan aplikasi praktis yang memanfaatkan machine learning dalam skala industri dan real-world applications.

REFERENCES

- [1] J. S. Akbar and D. Djakariah, "Transformasi Pembelajaran Kimia melalui Pemanfaatan Kecerdasan Buatan (AI) pada Era Society 5.0," *Edukara J. Pendidik. dan Pembelajaran*, vol. 9, no. 1, pp. 19–26, 2024, doi: 10.32585/edukara.v9i1.355.
- [2] F. Apri Wenando, Y. Fatma, A. Ulfa, Salma, and J. Taurin, "Aplikasi dan Kerentanan Algoritma Probabilistic Neural Network (PNN): Systematic Literature Review," *J. CoSciTech (Computer Sci. Inf. Technol.)*, vol. 4, no. 2, pp. 491–499, 2023, doi: 10.37859/coscitech.v4i3.5676.
- [3] N. Kristianti and W. Widiatry, "Penggunaan Algoritma Hebb Dalam Pola Pengenalan Huruf," *J. Teknol. Inf. J. Keilmuan dan Apl. Bid. Tek. Inform.*, vol. 18, no. 1, pp. 52–60, 2024, doi: 10.47111/jti.v18i1.12561.
- [4] F. Ramadhani, A. Satria, and S. Salamah, "Implementasi Algoritma Convolutional Neural Network dalam Mengidentifikasi Dini Penyakit pada Mata Katarak," *sudo J. Tek. Inform.*, vol. 2, no. 4, pp. 167–175, 2023.
- [5] B. A. Sadewa and Y. Yamasari, "Implementasi Deep Transfer Learning untuk Klasifikasi Nominal Uang Kertas Rupiah," *J. Informatics Comput. Sci.*, vol. 5, no. 04, pp. 543–551, 2024.
- [6] F. G. Lalamentik, O. Lantang, and F. Kambey, "Implementasi Parameter Tuning Untuk Optimalisasi Performance Model Convolutional Neural Network: Implementation of Parameter Tuning for Optimizing CNN Model Performance," *J. Tek. Inform.*, vol. 20, no. 1, pp. 77–86, 2025.
- [7] M. H. R. Pratama, M. Akrom, A. P. Santosa, M. R. Rosyid, and L. Mawaddah, "Klasifikasi Otomatis Korosi Menggunakan Convolutional Neural Network dan Transfer Learning dengan Model MobileNetV2," *J. Algoritma*, vol. 22, no. 1, pp. 138–148, 2025.
- [8] I. N. Suandana, A. Asriyanik, and W. Apriandari, "PEMANFAATAN CNN (CONVOLUTION NEURAL NETWORK) DAN MOBILENET V2 DALAM KLASIFIKASI REMPAH-REMPAH LOKAL DI INDONESIA," *JATI (Jurnal Mhs. Tek. Inform.)*, vol. 8, no. 5, pp. 10109–10116, 2024.
- [9] R. P. Widihartono and M. A. Ahmadi, "Pengaruh Pelatihan Terhadap Kinerja Karyawan Di Era Digital," *J. Ilm. Ekon. Manaj. Bisnis*, vol. 2, no. 4, pp. 204–213, 2024.
- [10] F. Ramadhan and J. Hernadi, "Evaluasi Optimizer Adam dan RMSProp pada Arsitektur VGG-19 Klasifikasi Ekspresi Wajah Manusia," *JIPI (Jurnal Ilm. Penelit. dan Pembelajaran Inform.)*, vol. 10, no. 2, pp. 1414–1426, 2025.
- [11] F. Sari, "Konvergensi dan Manajemen SDM Media Lokal," *J. Komun.*, vol. 16, no. 2, pp. 239–252, 2022.
- [12] F. Pritama, E. R. D. Leluni, and J. Parhusip, "Analisis Distribusi Kinerja SVM dan KNN Berdasarkan Rata Rata Simpangan Baku dan Stabilitas," *J. Ilm. Inform. dan Komput.*, vol. 1, no. 2, pp. 170–174, 2024.
- [13] R. A. F. Dewatri and R. E. Putra, "Implementasi Long Short-Term Memory dalam Mendeteksi Kesalahan Pronunciation Bahasa Inggris Berbasis Audio," *J. Informatics Comput. Sci.*, vol. 6, no. 03,



pp. 747–754, 2025.

- [14] F. Syaqqialloh, “Klasifikasi dan Pengenalan Emosi dari Ekspresi Wajah Menggunakan CNN-BiLSTM dengan Teknik Data Augmentation,” *Decod. J. Pendidik. Teknol. Inf.*, vol. 5, no. 1, pp. 79–91, 2025.
- [15] A. Jinan and B. H. Hayadi, “Klasifikasi Penyakit Tanaman Padi Menggunakan Metode Convolutional Neural Network Melalui Citra Daun (Multilayer Perceptron),” *J. Comput. Eng. Sci.*, pp. 37–44, 2022.