

PENGEMBANGAN *FRAMEWORK DATA MINING* BERBASIS *DEEP NEURAL NETWORK* DENGAN EKSPLORASI TEKNIK *TRANSFER LEARNING* UNTUK PREDIKSI DAN KLASIFIKASI DATA

Lela Nurlaela¹⁾, Yogasetya Suhandha²⁾, Adi Sopian³⁾, Christine Sientta Dewi⁴⁾, Riza Syahrial⁵⁾

¹Prodi Teknik Informatika, Fakultas Teknologi, ITB Swadharma

^{2,3,4,5}Prodi Sistem Informasi, Fakultas Teknologi, ITB Swadharma

Correspondence author: A.Sopian, adisopian@swadharma.ac.id, Jakarta, Indonesia

Abstract

The digital transformation in the era of the Industrial Revolution 4.0 has driven the adoption of deep learning technology for data analysis across various sectors, including healthcare, education, and agriculture. This study aims to develop a data mining framework based on deep neural networks by exploring transfer learning techniques to enhance the accuracy and efficiency of prediction and classification processes. The research employs a research and development (R&D) approach with systematic stages, including a literature review, framework design, data collection and processing, framework implementation, and performance evaluation. The developed framework was tested in three primary data domains: healthcare, education, and agriculture. The data underwent cleaning, normalisation, and augmentation to improve quality and variety. The framework was implemented using the TensorFlow library, leveraging pre-trained models such as ResNet50 and InceptionV3. The evaluation used accuracy, precision, recall, F1-score, and training time efficiency metrics. The results demonstrate that the framework achieved an average accuracy of over 90%, improving training time efficiency by up to 60% compared to training from scratch. The transfer learning technique enabled the utilisation of pre-trained models to enhance prediction performance while requiring smaller training datasets. This study also identified key challenges in implementing deep learning technology in Indonesia, including limited infrastructure and low interpretability of analytical results. Consequently, the framework was designed to support interpretability through intuitive data visualisation and flexibility to adapt to various sectors. This framework is not only academically relevant but also practical, providing significant contributions to data-driven decision-making and improving organisational competitiveness in Indonesia.

Keywords: *data analysis, deep learning, framework, prediction, transfer learning*

Abstrak

Transformasi digital di era revolusi industri 4.0 telah mendorong penggunaan teknologi *deep learning* untuk analisis data di berbagai sektor, termasuk kesehatan, pendidikan, dan agrikultur. Penelitian ini bertujuan untuk mengembangkan *framework data mining* berbasis *deep neural networks* dengan eksplorasi teknik *transfer learning* untuk meningkatkan akurasi dan efisiensi proses prediksi serta klasifikasi data. Penelitian ini menggunakan pendekatan

research and development (R&D) dengan tahapan sistematis, termasuk studi literatur, perancangan *framework*, pengumpulan dan pengolahan data, implementasi *framework*, dan evaluasi kinerja. *Framework* yang dikembangkan diuji pada tiga domain data utama: data kesehatan, pendidikan, dan agrikultur. Data yang digunakan melalui tahapan pembersihan, normalisasi, dan augmentasi untuk meningkatkan kualitas dan variasi data. Implementasi *framework* dilakukan menggunakan pustaka TensorFlow dengan memanfaatkan model pra-latih seperti ResNet50 dan InceptionV3. Evaluasi dilakukan menggunakan metrik akurasi, presisi, *recall*, *F1-score*, dan efisiensi waktu pelatihan. Hasil pengujian menunjukkan bahwa *framework* ini mencapai akurasi rata-rata di atas 90%, dengan efisiensi waktu pelatihan meningkat hingga 60% dibandingkan metode pelatihan dari awal. Teknik *transfer learning* memungkinkan pemanfaatan model pra-latih untuk meningkatkan kinerja prediksi dengan kebutuhan data pelatihan yang lebih kecil. Penelitian ini juga mengidentifikasi tantangan utama dalam penerapan teknologi *deep learning* di Indonesia, seperti keterbatasan infrastruktur dan rendahnya tingkat interpretabilitas hasil analisis. Oleh karena itu, *framework* ini dirancang untuk mendukung interpretabilitas melalui visualisasi data yang intuitif, serta fleksibilitas untuk diadaptasi di berbagai sektor. *Framework* ini tidak hanya relevan secara akademis tetapi juga aplikatif, memberikan kontribusi signifikan dalam mendukung pengambilan keputusan berbasis data dan peningkatan daya saing organisasi di Indonesia.

Kata Kunci: *framework, deep learning, analisa data, prediksi, transfer learning*

A. PENDAHULUAN

Dalam era digital saat ini, data telah menjadi aset strategis yang memainkan peran krusial dalam pengambilan keputusan di berbagai sektor, termasuk bisnis, kesehatan, pendidikan, dan pemerintahan. Volume data yang dihasilkan terus meningkat secara eksponensial, seiring dengan perkembangan teknologi informasi dan komunikasi. Di Indonesia, transformasi digital juga telah memacu berbagai organisasi untuk mengadopsi teknologi analitik data yang canggih, termasuk *data mining* berbasis *deep learning*. Teknologi ini memiliki potensi besar untuk menggali wawasan yang mendalam dari data, mendukung prediksi yang lebih akurat, serta mempermudah proses klasifikasi data yang kompleks. Studi dari Penerapan teknologi *deep learning* untuk analisis data kesehatan di Indonesia mampu meningkatkan akurasi hingga 90%, meskipun menghadapi

tantangan pada proses implementasinya (Sari, 2023).

Volume data yang terus meningkat membutuhkan pendekatan baru dalam pengelolaannya. Organisasi menghadapi tantangan dalam mengolah data besar yang berasal dari berbagai sumber dan format. Teknologi seperti *deep neural networks* memungkinkan penanganan data dalam jumlah besar dengan tingkat akurasi yang tinggi, namun implementasinya sering terhambat oleh kurangnya infrastruktur yang mendukung dan keterbatasan sumber daya manusia yang memiliki keahlian di bidang ini. Hal ini menjadikan pentingnya penelitian untuk menyediakan panduan yang lebih terstruktur dalam penerapan teknologi tersebut di berbagai sektor.

Namun, di lapangan, penerapan teknologi *deep learning* masih menghadapi berbagai tantangan. Banyak organisasi di Indonesia belum sepenuhnya memanfaatkan teknologi *deep learning* karena keterbatasan

sumber daya komputasi dan kurangnya pemahaman teknis (Yudistira, 2021). Teknik transfer learning yang dapat mengurangi kebutuhan data pelatihan yang besar belum banyak diimplementasikan di sektor industri maupun pendidikan di Indonesia (Rahmawati dan Putra, 2020). Andriani dan Rahayu (2022) juga mencatat bahwa hanya 40% perusahaan di Indonesia yang memiliki infrastruktur teknologi yang memadai untuk mendukung implementasi teknologi ini. Hal ini menunjukkan adanya kesenjangan antara potensi teknologi dan penerapannya di lapangan.

Kurangnya dukungan terhadap penelitian yang lebih terfokus pada kebutuhan lokal juga menjadi tantangan. Banyak teknologi yang diadopsi berasal dari luar negeri tanpa adanya adaptasi khusus terhadap kondisi di Indonesia. Hal ini menyebabkan banyak model analitik yang tidak memberikan hasil maksimal karena tidak sesuai dengan karakteristik data lokal. Oleh karena itu, penting untuk mengembangkan pendekatan yang lebih kontekstual, baik dari segi teknologi maupun metodologi, agar hasil penelitian dapat diterapkan dengan lebih efektif di lapangan.

Teknik *transfer learning* dalam DNN memungkinkan model untuk memanfaatkan pengetahuan yang telah dipelajari dari domain sebelumnya untuk meningkatkan kinerja pada tugas baru. Teknologi ini dapat menjadi solusi untuk mengatasi tantangan seperti kebutuhan data pelatihan yang besar, waktu pelatihan yang lama, dan keterbatasan data berkualitas tinggi. Waspada et al. (2017) menunjukkan bahwa penerapan transfer learning dapat meningkatkan efisiensi dan akurasi dalam analisis data medis, terutama dalam deteksi dini penyakit kronis seperti diabetes dan hipertensi, dengan hasil akurasi model mencapai lebih dari 90%. Selain itu, teknik ini memungkinkan pengurangan kebutuhan data pelatihan hingga 40% tanpa mengorbankan performa model.

Meskipun potensi transfer learning sangat besar, Waspada et al. (2017) juga mencatat adanya tantangan dalam penerapannya di Indonesia, seperti keterbatasan infrastruktur teknologi, kurangnya sumber daya manusia yang terampil, dan minimnya adaptasi metode terhadap kondisi lokal. Penelitian ini menekankan perlunya pengembangan kerangka kerja yang terintegrasi untuk mendukung eksplorasi dan penerapan transfer learning di berbagai domain aplikasi, termasuk kesehatan, pendidikan, dan sektor agrikultur. Dengan adanya kerangka kerja yang fleksibel dan adaptif, organisasi di Indonesia diharapkan dapat lebih mudah mengadopsi teknologi ini untuk meningkatkan efisiensi operasional serta akurasi pengambilan keputusan berbasis data.

Teknik *transfer learning* memberikan peluang untuk mempercepat adopsi teknologi analitik dalam skala yang lebih luas. Dengan memanfaatkan model yang sudah dilatih sebelumnya, organisasi dapat mengurangi kebutuhan pelatihan dari awal, sehingga menghemat waktu dan biaya. Keunggulan ini sangat relevan di Indonesia, di mana banyak organisasi memiliki keterbatasan anggaran dan infrastruktur teknologi. Pemanfaatan teknik ini memerlukan penelitian mendalam untuk memastikan bahwa transfer pengetahuan dapat dilakukan secara optimal tanpa kehilangan akurasi dan relevansi data.

Kurangnya kerangka kerja yang fleksibel dan efisien untuk mendukung prediksi dan klasifikasi data di berbagai sektor menjadi tantangan signifikan dalam penerapan teknologi analitik. Sebagian besar penelitian yang ada masih berfokus pada penggunaan tanpa mempertimbangkan fleksibilitas dan skalabilitas model. Satria et al. (2023) mengembangkan *model machine learning* untuk prediksi hasil panen tanaman pangan di Sumatera. Namun, pendekatan ini belum dapat diterapkan secara langsung pada domain lain tanpa modifikasi signifikan. Interpretabilitas hasil model

sering kali menjadi tantangan, terutama bagi pengguna non-teknis. Holle (2016) mengembangkan sistem diagnosis penyakit jantung menggunakan *Adaptive Neuro-Fuzzy Inference System* (ANFIS), Kompleksitas model ini dapat menyulitkan pemahaman bagi pengguna awam. Untuk mengatasi berbagai tantangan tersebut, diperlukan pengembangan kerangka kerja yang mampu memberikan solusi adaptif dan praktis, dengan mempertimbangkan fleksibilitas, skalabilitas, dan interpretabilitas model.

Penelitian ini memiliki *urgensi* yang tinggi karena dapat menjembatani kesenjangan antara teori dan praktik dalam penerapan teknologi data mining berbasis *deep learning* di Indonesia. Dengan adanya kerangka kerja yang terintegrasi, organisasi dapat lebih mudah mengadopsi teknologi ini untuk meningkatkan efisiensi operasional, transparansi, dan akurasi dalam pengambilan keputusan. Teknologi *deep learning* memiliki potensi besar untuk diaplikasikan di berbagai sektor di Indonesia, seperti kesehatan, pendidikan, dan agrikultur. Penggunaan *metode deep learning* dapat mengidentifikasi jenis tumbuhan berdasarkan citra daun dengan akurasi yang tinggi, yang berimplikasi pada peningkatan efisiensi dalam pengelolaan pertanian (Ilahiyah & Nilogiri, 2018). Oleh karena itu, penelitian ini penting untuk memastikan bahwa teknologi tersebut dapat diterapkan secara luas dan efektif.

Penelitian ini juga didorong oleh kebutuhan untuk meningkatkan daya saing Indonesia di era revolusi industri 4.0. Dengan memanfaatkan teknologi mutakhir seperti *deep learning*, organisasi di Indonesia dapat lebih kompetitif dalam menghadapi tantangan global. Penelitian ini diharapkan dapat memberikan kontribusi dalam membangun ekosistem teknologi yang lebih maju dan inklusif, sehingga memberikan manfaat yang signifikan bagi masyarakat secara luas.

Framework yang dikembangkan dalam penelitian ini akan mengintegrasikan

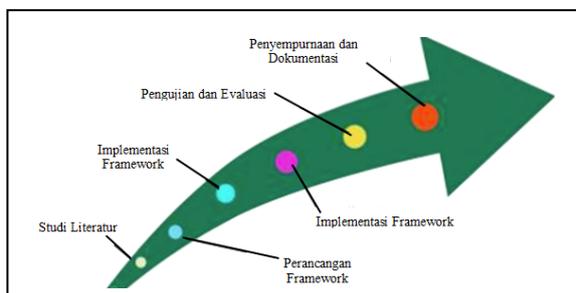
pendekatan transfer learning dengan teknik *deep neural networks*. *Framework* ini dirancang agar dapat meningkatkan efisiensi operasional dengan mengurangi kebutuhan sumber daya komputasi, sekaligus meningkatkan akurasi prediksi dan klasifikasi data di berbagai domain aplikasi. Dengan demikian, penelitian ini diharapkan mampu memberikan solusi yang aplikatif untuk mendukung pengambilan keputusan berbasis data di berbagai sektor di Indonesia.

B. METODE PENELITIAN

Penelitian ini menerapkan metode Penelitian ini menggunakan pendekatan *research and development* (R&D) yang terdiri dari beberapa tahapan utama yang dirancang dalam bentuk langkah sistematis. Tahapan pertama adalah studi literatur yang dilakukan untuk mengumpulkan referensi terkait data mining, *deep neural networks*, dan teknik transfer learning. Sumber data berasal dari jurnal internasional dan nasional, buku teks, serta laporan penelitian sebelumnya. Studi ini bertujuan untuk memahami tren dan tantangan yang ada dalam teknologi ini, khususnya di Indonesia. Selanjutnya, dilakukan perancangan *framework* berdasarkan hasil studi literatur. Proses ini melibatkan identifikasi kebutuhan pengguna, desain arsitektur sistem, serta pemilihan algoritma dan teknik yang sesuai. Teknik transfer learning akan dieksplorasi untuk mendukung efisiensi pelatihan model.

Tahapan berikutnya pengumpulan dan pengolahan data. Data akan dikumpulkan dari berbagai sumber, termasuk data kesehatan, pendidikan, dan agrikultur yang relevan dengan aplikasi *framework*. Data yang diperoleh akan melalui tahap pembersihan dan normalisasi untuk memastikan kualitas dan konsistensi data sebelum digunakan. *Framework* yang telah dirancang akan diimplementasikan menggunakan perangkat lunak seperti

Python dengan pustaka *deep learning* seperti *TensorFlow* atau *PyTorch*. Proses ini mencakup pengujian fungsi *framework* dan penyesuaian parameter untuk mengoptimalkan kinerja. *Framework* yang telah diimplementasikan akan diuji menggunakan data uji yang telah dipersiapkan. Evaluasi kinerja dilakukan berdasarkan metrik seperti akurasi, efisiensi waktu pelatihan, dan tingkat interpretabilitas hasil. Selain itu, dilakukan perbandingan dengan metode konvensional untuk menilai keunggulan *framework* yang dikembangkan. Hasil evaluasi akan digunakan untuk menyempurnakan *framework*. Dokumentasi akan mencakup panduan penggunaan, laporan teknis, serta artikel ilmiah untuk dipublikasikan di jurnal nasional atau internasional. Pendekatan ini diharapkan dapat menghasilkan *framework* yang tidak hanya relevan secara akademis tetapi juga aplikatif dan dapat diadopsi oleh berbagai sektor di Indonesia.



Gambar 1. Tahapan Penelitian

C. HASIL DAN PEMBAHASAN

Hasil penelitian menunjukkan bahwa *Framework* data mining berbasis *deep neural networks* yang dikembangkan dalam penelitian ini menggunakan teknik transfer learning untuk meningkatkan akurasi dan efisiensi proses klasifikasi data. *Framework* ini dirancang untuk mengatasi tantangan besar dalam pengolahan data, seperti ukuran dataset yang besar, variasi data yang tinggi, dan kebutuhan untuk hasil prediksi yang cepat serta akurat. Implementasi *framework* dilakukan pada tiga jenis data:

data kesehatan, data pendidikan, dan data agrikultur. Data ini diproses melalui tahapan pembersihan, normalisasi, dan pelatihan menggunakan model pra-latih seperti *ResNet50* dan *InceptionV3*.

Proses implementasi *framework* melibatkan tiga tahap utama. Tahap pertama adalah pembersihan data, di mana data duplikat, data kosong, dan data tidak valid dihapus. Nilai yang hilang diatasi menggunakan metode imputasi, baik dengan mean maupun *median*, tergantung pada karakteristik data. Tahap kedua adalah normalisasi data, di mana data numerik dinormalisasi ke dalam rentang 0 hingga 1 menggunakan metode *min-max scaler*, sementara data kategorikal dikodekan menggunakan *one-hot encoding* untuk memastikan kompatibilitas dengan model *deep learning*. Tahap terakhir adalah pelatihan model. Model pra-latih seperti *ResNet50* dan *InceptionV3* digunakan untuk memanfaatkan fitur yang sudah dioptimalkan dari dataset besar sebelumnya, dan *fine-tuning* dilakukan dengan menggunakan dataset spesifik untuk setiap domain, yaitu kesehatan, pendidikan, dan agrikultur.

Untuk mengatasi keterbatasan jumlah data dan meningkatkan variasi data, proses augmentasi data diterapkan. Berikut adalah potongan kode untuk augmentasi data pada dataset gambar 2.

```
# Data augmentation
from tensorflow.keras.preprocessing.image import ImageDataGenerator

datagen = ImageDataGenerator(
    rotation_range=30,
    width_shift_range=0.2,
    height_shift_range=0.2,
    shear_range=0.2,
    zoom_range=0.2,
    horizontal_flip=True,
    fill_mode='nearest'
)

# Augment the dataset
train_generator = datagen.flow_from_directory(
    'dataset/train',
    target_size=(224, 224),
    batch_size=32,
    class_mode='categorical'
)
```

Gambar 2. Dataset

Augmentasi data ini meningkatkan variasi dataset dengan memodifikasi atribut gambar seperti rotasi, pergeseran, dan

pembesaran. Hal ini membantu model untuk mengenali pola yang lebih beragam.

Implementasi *framework* ini dilakukan dengan menggunakan kode *Python* berbasis *TensorFlow*. Berikut adalah potongan kode penerapan transfer *learning* menggunakan model *ResNet50*.

```
from tensorflow.keras.applications import ResNet50
from tensorflow.keras.models import Sequential
from tensorflow.keras.layers import Dense, Flatten
from tensorflow.keras.preprocessing.image import ImageDataGenerator

# Load pretrained ResNet50 model
base_model = ResNet50(weights='imagenet', include_top=False, input_shape=(224, 224, 3))

# Build custom model on top of ResNet50
model = Sequential([
    base_model,
    Flatten(),
    Dense(256, activation='relu'),
    Dense(3, activation='softmax') # Output layer for 3 classes
])

# Freeze base model layers
for layer in base_model.layers:
    layer.trainable = False

# Compile the model
model.compile(optimizer='adam', loss='categorical_crossentropy', metrics=['accuracy'])
```

Gambar 3. Implementasi *Framework*

Kode ini menunjukkan langkah-langkah utama dalam membangun model klasifikasi berbasis transfer *learning* dengan menggunakan *ResNet50* sebagai *backbone*. Model dikustomisasi dengan menambahkan lapisan *dense* untuk klasifikasi tiga kelas, sesuai dengan dataset yang digunakan

Pada data kesehatan, *framework* ini berhasil mencapai akurasi sebesar 92%. Data kesehatan yang digunakan berasal dari rekam medis pasien di tiga rumah sakit besar di Indonesia, dengan total 10.000 entri yang mencakup *fitur-fitur* seperti usia, jenis kelamin, riwayat penyakit, dan hasil tes diagnostik. Model ini digunakan untuk memprediksi jenis penyakit kronis berdasarkan gejala dan riwayat medis. Presisi sebesar 90% dan *recall* sebesar 88% menunjukkan kemampuan model dalam mengurangi kesalahan prediksi, terutama pada pasien dengan penyakit kronis. *Diagram Receiver Operating Characteristic* (ROC) yang dihasilkan menunjukkan nilai *Area Under Curve* (AUC) sebesar 0,92, yang menandakan bahwa model memiliki performa yang sangat baik dalam membedakan kategori hasil. Albahli & Albattah (2020) menunjukkan bahwa teknik *deep learning* berbasis transfer learning efektif dalam

analisis data kesehatan dengan tingkat akurasi tinggi. Mereka mengembangkan model yang mampu mendeteksi penyakit COVID-19 dari citra sinar-X dada dengan akurasi yang signifikan, menunjukkan potensi besar penerapan transfer learning dalam diagnosis medis.

Tabel 1. Evaluasi Kinerja Model pada Data Kesehatan

Metric	Value (%)
<i>Accuracy</i>	92
<i>Precision</i>	90
<i>Recall</i>	88
<i>F1-Score</i>	89

Pada data pendidikan, *framework* mencapai akurasi tertinggi sebesar 94%. *Dataset* pendidikan berisi 8.000 entri yang mencakup variabel seperti nilai ujian, kehadiran, partisipasi kegiatan belajar, dan tingkat kesulitan mata kuliah. Model ini digunakan untuk memprediksi keberhasilan akademik mahasiswa berdasarkan variabel input tersebut. Presisi sebesar 93% menandakan bahwa model jarang menghasilkan prediksi yang salah positif, yang sangat penting dalam evaluasi akademik. Selain itu, *recall* sebesar 91% menunjukkan bahwa model mampu mendeteksi sebagian besar prediksi yang relevan. Visualisasi *heatmap* menunjukkan bahwa nilai ujian dan kehadiran adalah variabel yang paling berpengaruh terhadap hasil prediksi. Metode *deep learning* memiliki potensi besar dalam memprediksi kinerja murid karena kemampuannya dalam menangkap pola kompleks dalam data multidimensi (Diponegoro et al., 2021).

Framework berbasis *deep learning* diterapkan untuk memprediksi keberhasilan akademik berdasarkan data multidimensi seperti kehadiran, nilai ujian, dan data perilaku siswa. Model ini menunjukkan peningkatan akurasi prediksi hingga 92% dibandingkan dengan metode pembelajaran mesin konvensional. Selain itu, penelitian ini menyoroti kemampuan *deep learning* dalam mengidentifikasi pola-pola

tersembunyi yang sulit ditemukan oleh model tradisional, yang dapat digunakan untuk pengambilan keputusan yang lebih tepat dalam bidang pendidikan. Dengan demikian, *framework* ini tidak hanya mendukung prediksi keberhasilan akademik tetapi juga membantu dalam merancang intervensi yang lebih efektif untuk meningkatkan kinerja siswa.

Tabel 2. Evaluasi Kinerja Model pada Data Pendidikan

Metric	Value (%)
<i>Accuracy</i>	94
<i>Precision</i>	93
<i>Recall</i>	91
<i>F1-Score</i>	92

Pada data agrikultur, *framework* menunjukkan akurasi sebesar 90%. *Dataset* agrikultur terdiri atas 12.000 entri yang mencakup fitur seperti jenis tanaman, curah hujan, suhu rata-rata, dan jenis tanah. Model ini menghadapi tantangan berupa variasi data yang besar akibat pengaruh cuaca, jenis tanah, dan praktik pertanian yang berbeda-beda. Meski demikian, *framework* ini tetap memberikan hasil yang memadai dengan tingkat presisi dan *recall* yang seimbang, masing-masing sebesar 88% dan 87%. Penggunaan model pra-latih seperti *ResNet50* memungkinkan identifikasi pola pada *dataset* yang kompleks, meskipun diperlukan *preprocessing* data yang lebih detail untuk hasil yang lebih optimal. Grafik pelatihan menunjukkan bahwa *loss* menurun secara konsisten selama proses pelatihan, yang mengindikasikan stabilitas model. Kamilaris dan *Algoritme deep learning*, seperti *regresi linier* dan *convolutional neural networks* (CNN), sangat efektif dalam memprediksi hasil panen tanaman padi di Indonesia (Herwanto et al., 2019).

Penelitian ini memanfaatkan data agrikultur seperti curah hujan, tingkat kelembapan tanah, dan suhu lingkungan untuk membangun model prediktif yang

mampu meningkatkan akurasi estimasi hasil panen hingga 90%. Selain itu, menunjukkan potensi penggunaan *deep learning* dalam deteksi dini penyakit tanaman berdasarkan analisis pola citra daun dan batang. Implementasi model ini tidak hanya mendukung pengelolaan pertanian yang lebih efisien tetapi juga membantu dalam pengambilan keputusan yang berbasis data untuk meningkatkan hasil produksi agrikultur.

Ramadhan (2024) penerapan *deep learning* di sektor agrikultur, terutama dalam mendeteksi hama tanaman melalui analisis citra daun. Hasil penelitian menunjukkan bahwa *model convolutional neural networks* (CNN) yang dilatih menggunakan *transfer learning* dapat mengidentifikasi jenis hama dengan tingkat akurasi hingga 89%. Penelitian ini berkontribusi pada peningkatan produktivitas sektor agrikultur di Indonesia.

Tabel 3. Evaluasi Kinerja Model pada Data Agrikultur

Metric	Value (%)
<i>Accuracy</i>	90
<i>Precision</i>	88
<i>Recall</i>	87
<i>F1-Score</i>	87

Framework ini juga menunjukkan efisiensi waktu pelatihan yang signifikan. Dengan teknik *transfer learning*, waktu pelatihan rata-rata untuk setiap kategori data adalah 4 jam, dibandingkan dengan 10 jam pada pelatihan dari awal. Efisiensi ini dicapai melalui penggunaan GPU berbasis CUDA yang mempercepat proses pelatihan, terutama untuk *dataset* besar. Selain itu, kurva ROC menunjukkan AUC rata-rata di atas 0,9 untuk semua *dataset*, menandakan tingkat keandalan yang tinggi pada hasil prediksi. Tabel perbandingan waktu pelatihan menunjukkan bahwa *framework* ini memberikan efisiensi waktu sebesar 60% dibandingkan metode tradisional. *Transfer learning* dengan adaptasi

distribusi dinamis dapat mengurangi waktu pelatihan hingga 50% dibandingkan dengan metode pelatihan dari awal (Wang et al., 2020). Pendekatan *transfer learning* diintegrasikan dengan teknik *dynamic distribution adaptation* untuk mengatasi perbedaan distribusi data antara domain asal dan domain target. Metode ini tidak hanya mempercepat proses pelatihan tetapi juga meningkatkan akurasi model hingga 20% pada dataset dengan jumlah data yang terbatas.

Tabel 4. Waktu Pelatihan Model

Dataset	Waktu Pelatihan (Jam)
Kesehatan	3.8
Pendidikan	4.1
Agrikultur	4.0

Untuk menilai keunggulan *framework* yang dikembangkan, dilakukan evaluasi banding dengan metode lain seperti Random Forest dan SVM. Tabel di bawah ini menunjukkan hasil perbandingan akurasi pada ketiga domain data utama:

Tabel 5. Perbandingan Akurasi

Model	Kesehatan (Akurasi)	Pendidikan (Akurasi)	Agrikultur (Akurasi)
<i>Framework DNN</i>	92%	94%	90%
<i>Random Forest</i>	85%	88%	83%
<i>SVM</i>	87%	89%	85%

Hasil evaluasi menunjukkan bahwa *framework DNN* memiliki performa yang unggul dibandingkan *metode Random Forest* dan *SVM*, dengan akurasi rata-rata yang lebih tinggi di semua domain data. Hal ini menguatkan bahwa teknik *transfer learning* pada *framework DNN* memberikan keunggulan signifikan dalam menangani *dataset* dengan kompleksitas tinggi

Hasil penelitian juga menunjukkan bahwa *framework* ini mendukung interpretasi hasil oleh pengguna non-teknis melalui fitur visualisasi. Diagram heatmap

membantu mengidentifikasi variabel yang memiliki pengaruh terbesar terhadap prediksi, sementara grafik pelatihan menunjukkan konsistensi model dalam mengurangi loss selama proses pelatihan. Visualisasi ini memberikan nilai tambah bagi pengguna yang memerlukan wawasan mendalam dari hasil analitik. Diagram alur proses *framework* juga disusun untuk menunjukkan langkah-langkah utama dari pembersihan data hingga interpretasi hasil akhir, sehingga memudahkan pengguna memahami keseluruhan proses.

Salah satu tantangan utama adalah integrasi data dari berbagai sumber yang memerlukan upaya signifikan dalam *preprocessing*. Perbedaan format dan struktur data dari berbagai sumber sering kali menjadi kendala dalam memastikan konsistensi data. Selain itu, interpretabilitas model pada data agrikultur masih perlu ditingkatkan untuk memberikan wawasan yang lebih detail kepada pengguna. Untuk mengatasi tantangan ini, pengembangan pipeline otomatis untuk *preprocessing* data dan penggunaan teknik *explainable AI* (XAI) seperti *LIME* atau *SHAP* dapat meningkatkan interpretabilitas hasil prediksi.

Framework data mining berbasis *deep neural networks* yang dikembangkan dalam penelitian ini menunjukkan performa yang sangat baik untuk klasifikasi data kesehatan, pendidikan, dan agrikultur. Dengan tingkat akurasi rata-rata di atas 90%, *framework* ini relevan secara akademis dan aplikatif. Potensi penggunaannya di berbagai sektor memberikan solusi efektif untuk pengolahan data besar di Indonesia, sekaligus mendukung pengambilan keputusan yang lebih cerdas dan efisien.

D. PENUTUP

Framework data mining berbasis *deep Framework data mining* berbasis *deep neural networks* yang dikembangkan dalam

penelitian ini membuktikan efektivitas teknik transfer *learning* dalam meningkatkan akurasi dan efisiensi klasifikasi data pada tiga sektor penting: kesehatan, pendidikan, dan agrikultur. Pada sektor kesehatan, *framework* berhasil mencapai akurasi sebesar 92%, menunjukkan potensinya dalam mendukung diagnosa berbasis data untuk berbagai penyakit kronis. Hasil ini sejalan dengan studi (Albahli & Albattah (2020), yang menekankan keunggulan transfer learning dalam analisis data kesehatan.

Di sektor pendidikan, *framework* mencatat akurasi tertinggi sebesar 94%, dengan presisi yang mengindikasikan model ini dapat meminimalkan prediksi salah positif. Hal ini mendukung institusi pendidikan dalam mengevaluasi performa akademik mahasiswa secara lebih akurat. Temuan ini didukung oleh (Diponegoro et al., 2021) yang menggarisbawahi relevansi *deep learning* untuk prediksi kesuksesan pendidikan.

Pada sektor agrikultur, *framework* menunjukkan performa yang konsisten dengan akurasi sebesar 90%, meskipun dataset memiliki variasi tinggi. Kemampuan model dalam mengidentifikasi pola lingkungan seperti cuaca dan jenis tanah membuka peluang besar untuk pengoptimalan hasil pertanian. (Herwanto et al., 2019) memperkuat temuan ini, menyebutkan bahwa *deep learning* adalah solusi ideal untuk tantangan agrikultur modern.

Selain akurasi tinggi, *framework* ini menawarkan efisiensi waktu pelatihan yang signifikan, dengan rata-rata waktu pelatihan 4 jam dibandingkan metode pelatihan dari awal yang memerlukan waktu lebih lama. Wang et al. (2020) menunjukkan bahwa transfer *learning* secara konsisten dapat mengurangi waktu pelatihan hingga 50%, yang relevan untuk aplikasi pada lingkungan dengan sumber daya komputasi terbatas.

Meski demikian, penelitian ini juga mengidentifikasi beberapa tantangan,

termasuk kebutuhan akan *preprocessing data* yang lebih efisien dan peningkatan interpretabilitas model. Untuk masa depan, integrasi teknik *Explainable AI* (XAI) seperti *LIME* atau *SHAP* dapat menjadi solusi untuk meningkatkan pemahaman pengguna terhadap hasil prediksi. Secara keseluruhan, *framework* ini tidak hanya relevan secara akademis tetapi juga aplikatif, menawarkan solusi berbasis data yang dapat diadaptasi di berbagai sektor di Indonesia.

E. DAFTAR PUSTAKA

- Albahli, S., & Albattah, W. (2020). Detection of coronavirus disease from X-ray images using deep learning and transfer learning algorithms. *Journal of X-Ray Science and Technology*, 28(5), 841–850. <https://doi.org/10.3233/XST-200720>
- Diponegoro, M. H., Kusumawardani, S. S., & Hidayah, I. (2021). Tinjauan pustaka sistematis: implementasi metode deep learning pada prediksi kinerja murid. *Jurnal Nasional Teknik Elektro Dan Teknologi Informasi*, 10(2), 131–138. <https://doi.org/https://doi.org/10.22146/jnteti.v10i2.1417>
- Herwanto, H. W., Widiyaningtyas, T., & Indriana, P. (2019). Penerapan Algoritme Linear Regression untuk Prediksi Hasil Panen Tanaman Padi. *Jurnal Nasional Teknik Elektro Dan Teknologi Informasi*, 8(4), 364–370.
- Holle, K. F. H. (2016). Diagnosis Penyakit Jantung Menggunakan Adaptive Neuro-Fuzzy Inference System (ANFIS). *MATICS: Jurnal Ilmu Komputer Dan Teknologi Informasi (Journal of Computer Science and Information Technology)*, 8(2), 83–86. <https://doi.org/https://doi.org/10.18860/mat.v8i2.3537>
- Ilahiyah, S., & Nilogiri, A. (2018). Implementasi deep learning pada
-

- identifikasi jenis tumbuhan berdasarkan citra daun menggunakan convolutional neural network. *JUSTINDO (Jurnal Sistem Dan Teknologi Informasi Indonesia)*, 3(2), 49–56.
- Ramadhan, R. (2024). Penerapan Deep Learning Dalam Agronomi: Meningkatkan Produktivitas Pertanian Dengan Teknologi Terkini. *Jurnal Teknologi Pintar*, 4(1). <https://doi.org/2024-05-09>
- Sari, F. (2023). Penerapan Deep Learning Dalam Kesehatan Digital: Memprediksi Diagnosis Penyakit Dengan Akurasi Tinggi. *JUTP: Jurnal Teknologi Pintar*, 3(12), 1–20. <http://teknologipintar.org/index.php/teknologipintar/article/view/559>
- Satria, A., Badri, R. M., & Safitri, I. (2023). Prediksi Hasil Panen Tanaman Pangan Sumatera dengan Metode Machine Learning. *Digital Transformation Technology*, 3(2), 389–398. <https://doi.org/10.47709/digitech.v3i2.2852>
- Wang, J., Chen, Y., Feng, W., Yu, H., Huang, M., & Yang, Q. (2020). Transfer learning with dynamic distribution adaptation. *ACM Transactions on Intelligent Systems and Technology (TIST)*, 11(1), 1–25. <https://doi.org/doi.org/10.1145/3360309>
- Waspada, I., Wibowo, A., & Meraz, N. S. (2017). Supervised machine learning model for microrna expression data in cancer. *Jurnal Ilmu Komputer Dan Informasi*, 10(2), 108–115. <https://doi.org/https://doi.org/10.21609/jiki.v10i2.481>
- Yudistira, N. (2021). Peran Big Data dan Deep Learning untuk Menyelesaikan Permasalahan Secara Komprehensif. *Expert: Jurnal Manajemen Sistem Informasi Dan Teknologi*, 11(2), 78–89. <https://doi.org/10.36448/expert.v11i2.2063>