

---

# PAPER REVIEW

Kelas 3K

---

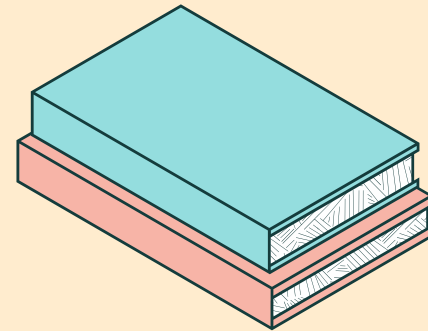
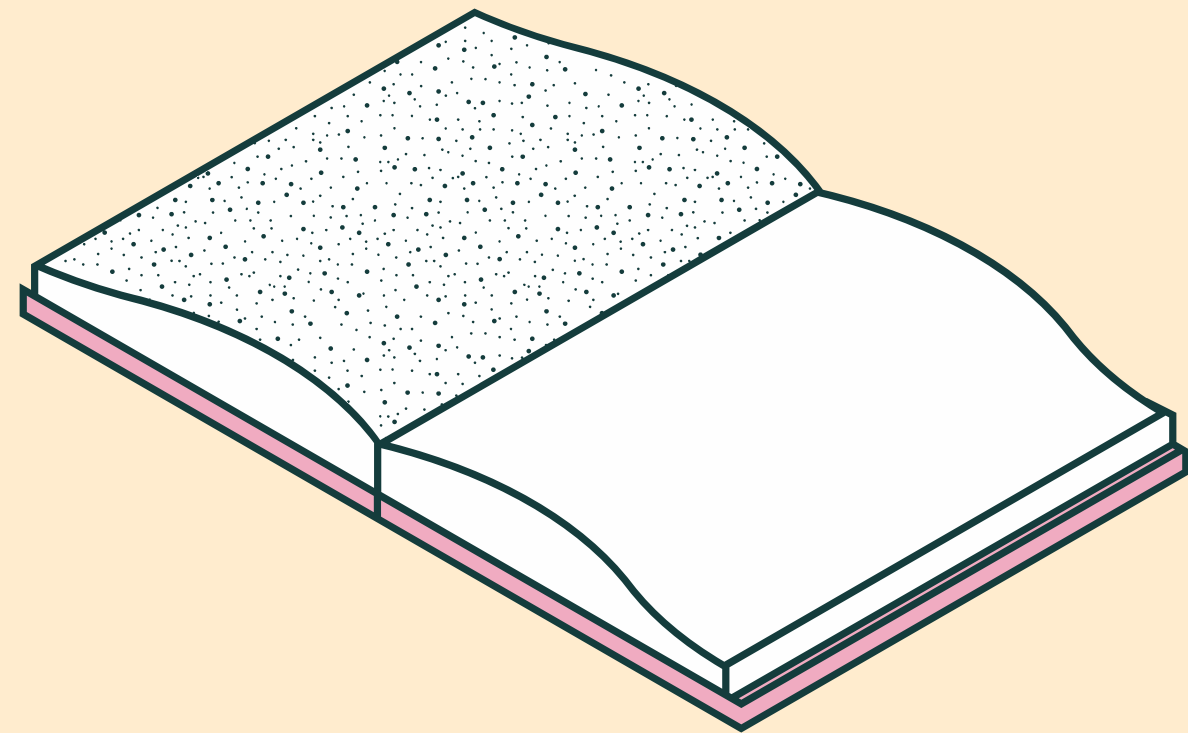
Taufik Anwar (2041720010/28)

Wahyu Rizky A. (2041720191/30)

# 1. Classification of Diabetic Retinopathy Disease

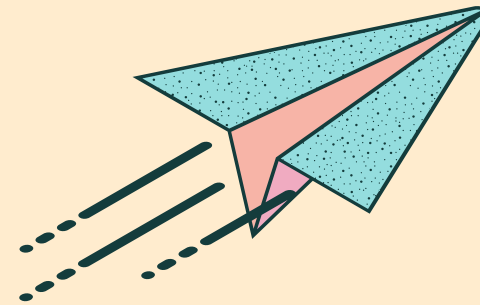
- Using Convolutional Neural Network

---



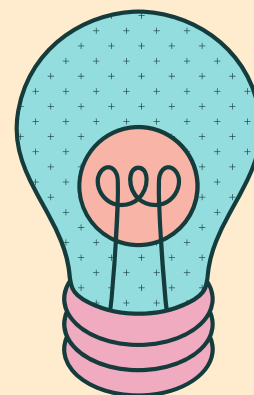
## Journal Name

International Journal on Informatics  
Visualization



## Publisher

Society of Visual Informatics, Padang



## Journal Level

Q4

# Abstract Review

---

Penelitian ini memaparkan penerapan sebuah metode deep learning menggunakan model Efficientnet-b7 untuk mengenali penyakit Retinopati Diabetik (DR) secara otomatis melalui gambar fundus retina mata. Penelitian ini menerapkan tiga teknik preprocessing dan melakukan tuning hyperparameter untuk menemukan parameter terbaik untuk model EfficientNet-B7. Teknik preprocessing Usuyama mencapai akurasi tertinggi di antara percobaan dengan tingkat akurasi 89% pada data latih dan 84% pada data uji. Penelitian ini juga menunjukkan bahwa teknik augmentasi dapat mengurangi overfitting dan meningkatkan kinerja model. Dengan demikian, model yang diusulkan dapat memberikan deteksi penyakit Retinopati Diabetik yang efektif dan akurat, sehingga mengurangi kebutuhan akan ahli oftalmologi berpengalaman.

# Abstract Framework

---

**Purpose and Scope:** Penelitian ini memaparkan penerapan sebuah metode deep learning menggunakan model Efficientnet-b7 untuk mengenali penyakit Retinopati Diabetik (DR) secara otomatis melalui gambar fundus retina mata. **Overview of Method:** Penelitian ini menerapkan tiga teknik preprocessing dan melakukan tuning hyperparameter untuk menemukan parameter terbaik untuk model EfficientNet-B7. Teknik preprocessing Usuyama mencapai akurasi tertinggi di antara percobaan dengan tingkat akurasi 89% pada data latih dan 84% pada data uji. **Research Findings:** Penelitian ini juga menunjukkan bahwa teknik augmentasi dapat mengurangi overfitting dan meningkatkan kinerja model. **Conclusion:** Dengan demikian, model yang diusulkan dapat memberikan deteksi penyakit Retinopati Diabetik yang efektif dan akurat, sehingga mengurangi kebutuhan akan ahli oftalmologi berpengalaman.

# Introduction Review

---

## A. Research Object

Obyek pada penelitian ini adalah ***Diabetic Retinopathy***

Diabetic Retinopathy is a disease of the eye's retina arising from the effects of diabetes. Diabetic Retinopathy is characterized by one of the conditions in the damage to the blood vessels at the back of the retina [1]. Patients with

## B. Existing Method

Penelitian sebelumnya dengan topik yang sama menggunakan metode CNN dengan model Inceptionv3, Deep Learning dengan EfficentNet model, dan RNN

Sichuan Province. In this study, researchers trained a modified model of the pre-trained inceptionV3 model with the help of the data processing and data augmentation stage. The proposed result achieved an accuracy of 88.72% [8].

Further research related to Diabetic Retinopathy disease was conducted by Kwasigroch et al. [9] that propose the Deep Convolutional Neural Network (CNN) method for detecting and classifying Diabetic Retinopathy. The datasets used in

for detecting Diabetic Retinopathy disease. In a subsequent study conducted by Eman Abdel Maksoud et al.in 2020, this study proposed the deep learning (DL) method with the EfficientNet architecture model. The dataset used in this study

Research related to Diabetic Retinopathy disease was conducted by Harikrishnan et al. [11] that propose the Recurrent Neural Network (RNN) method for classifying diabetic retinopathy disease stage. The dataset used in this

# Introduction Review

---

## C. The Advantages and Disadvantages of Existing Methods

Tidak disebutkan secara spesifik akan tetapi terdapat perbandingan akurasi pada sub bab *Compare and Analysis Model*.

This study proposed several scenario models with different preprocessing methods and the effect of implementing the augmentation process run on the EfficientNet-B7 model. The 3rd test scenario was the best compared to other scenarios in terms of accuracy of 84%, loss of 0.43, the precision of 83%, recall of 84%, and F1-score with a value of 83%. This scenario proves that the proposed model's use of Usuyama [19] preprocessing techniques and augmentation implementations can improve performance.

## D. Problem with Selected Method

Permasalahan yang ditemukan adalah menentukan kombinasi parameter yang tepat untuk model ***EfficientNet-B7***

In this study, the EfficientNet-B7 architecture model was implemented, then combined several preprocessing techniques and the hyperparameter Tuning method to obtain the results of the right combination of parameters used in the CNN method with EfficientNet-B7 architecture model.

retinopathy datasets. In this study, selecting the right optimizer could make the EfficientNet-B7 model optimal in conducting data training. Then, the second method could be focused on finding the best parameters for dense value and dropout on the fully connected layer with the best optimizer that has been obtained from the first method trial. The

# Introduction Review

---

## E. Solution Repair Method

Solusi yang ditawarkan adalah dengan menerapkan model ***EfficientNet-B7*** yang memiliki kemampuan terbaik diantara model pada *Existing Method*

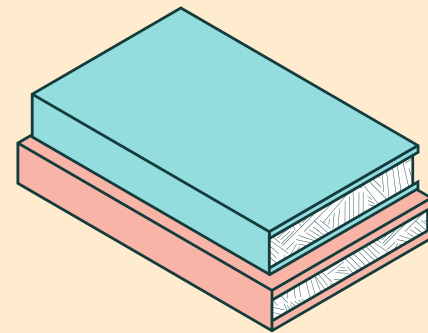
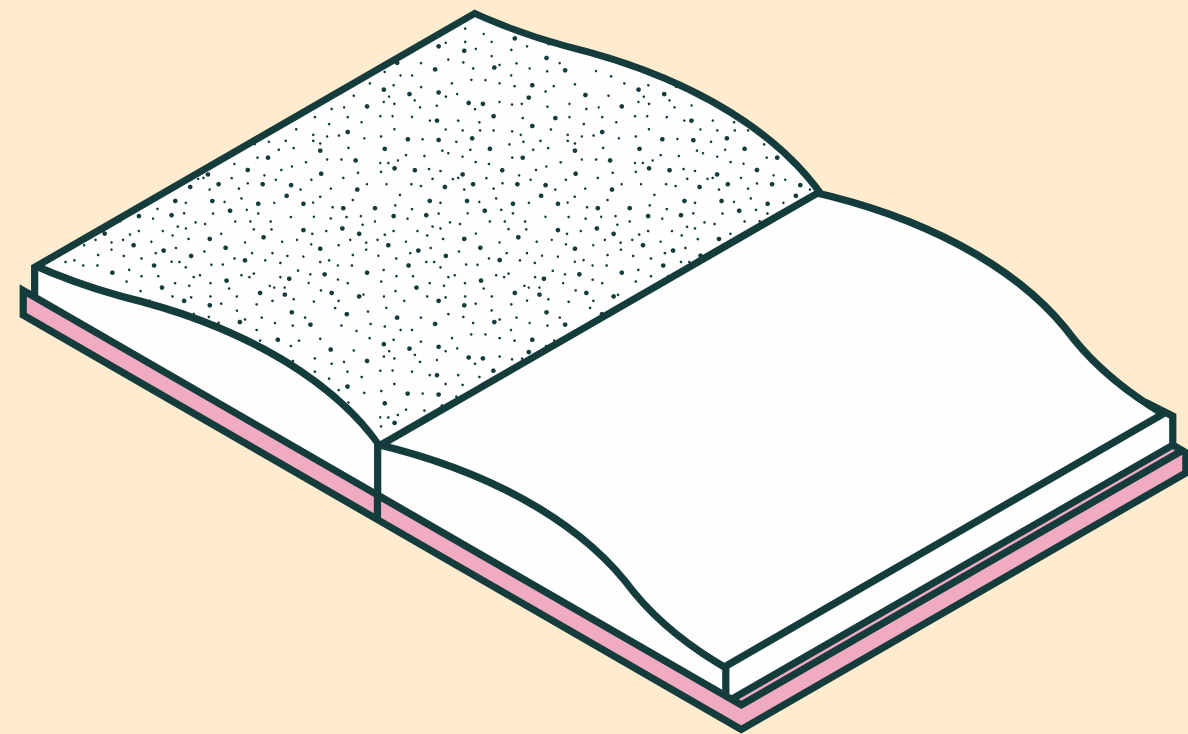
Efficientnet-B7. The EfficientNet-B7 Architecture Model is proposed because it has optimal performance compared to other architectures. In the preprocessing stage, the data is done augmentation process on the image using augmentation technique from the hard library.

## F. Summary of research objectives

Pada penelitian ini, model arsitektur ***EfficientNet-B7*** diimplementasikan, kemudian mengkombinasikan beberapa teknik preprocessing dan metode hyperparameter tuning untuk mendapatkan hasil kombinasi parameter yang tepat

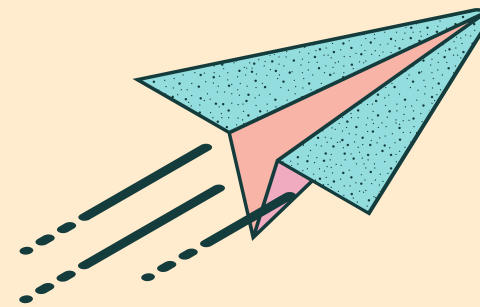
# 2. Twitter-based classification for integrated source data of weather observations

---



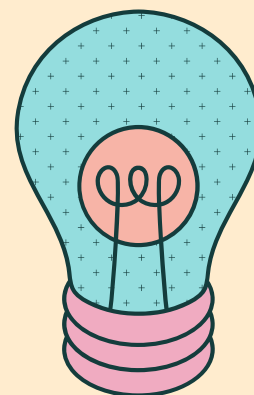
## Journal Name

IAES International Journal of Artificial Intelligence



## Publisher

IAES



## Journal Level

Q3



# Abstract Review

---

Penelitian ini bertujuan untuk benchmarking dan meningkatkan akurasi ramalan cuaca dengan menganalisis tweet terkait cuaca di Indonesia menggunakan teknik machine learning dan deep learning. Peneliti melakukan analisis terhadap tiga algoritma machine learning dan model pretrained BERT yang telah dioptimalkan melalui beberapa lapisan, dan BERT menunjukkan akurasi yang lebih tinggi (F1-score sebesar 99%) daripada metode lainnya. Hasilnya dimasukkan ke dalam sistem informasi cuaca berbasis web dan dipetakan menggunakan aplikasi Esri Maps berdasarkan lokasi datanya. Dapat disimpulkan bahwa menggunakan BERT dapat membantu meningkatkan akurasi klasifikasi tweet terkait cuaca, yang dapat bermanfaat untuk memprediksi kondisi iklim di masa depan dan membuat keputusan yang lebih baik.

# Abstract Framework

---

**Purpose and Scope:** Penelitian ini bertujuan untuk benchmarking dan meningkatkan akurasi ramalan cuaca dengan menganalisis tweet terkait cuaca di Indonesia menggunakan teknik machine learning dan deep learning. **Overview of Method:** Peneliti melakukan analisis terhadap tiga algoritma machine learning dan model pretrained BERT yang telah dioptimalkan melalui beberapa lapisan, dan BERT menunjukkan akurasi yang lebih tinggi (F1-score sebesar 99%) daripada metode lainnya. **Research Findings:** Hasilnya dimasukkan ke dalam sistem informasi cuaca berbasis web dan dipetakan menggunakan aplikasi Esri Maps berdasarkan lokasi datanya. **Conclusion:** Dapat disimpulkan bahwa menggunakan BERT dapat membantu meningkatkan akurasi klasifikasi tweet terkait cuaca, yang dapat bermanfaat untuk memprediksi kondisi iklim di masa depan dan membuat keputusan yang lebih baik.

# Introduction Review

---

## A. Research Object

Obyek pada penelitian ini adalah prediksi cuaca berdasarkan lokasi dari akun tiwitter

stay informed about what happens around them, especially in relation to relevant events [8]. Twitter is used by people worldwide to access different types of information, including all kinds of information on Twitter. Monitoring topics and events is made easier with a structured combination of search parameters on a Twitter channel. We implemented the geolocation by using available application programming interface (APIs) and web services. Using existing APIs, location-specific terms were detected in a tweet. Social media platforms are continually generating and delivering information in real-time from various sources to users. Topics,

## B. Existing Method

Penelitian sebelumnya dengan topik yang relevan menggunakan metode KNN, MNB, dan MLR

Using the K-nearest neighbor (KNN) algorithm, a potential company's employees can be identified by their personalities. KNN identified the Myers-Briggs type indicator (MBTI) categories based on character classifications for potential employees from tweets [13].

A machine learning classification method was developed by the author using the support vector machine (SVM) technique before the BERT method, which provided 93% accuracy. As a result, other machine learning algorithms, such as multinomial Naive Bayes (MNB) and multinomial logistic regression (MLR), did not achieve highly accurate predictions when applied to Twitter data about 1weather conditions [22].

# Introduction Review

## C. The Advantages and Disadvantages of Existing Methods

Tidak disebutkan secara spesifik akan tetapi terdapat perbandingan akurasi pada bab *RESULT AND DISCUSSION*.

Table 3. Classifier evaluation using machine learning approaches (%)

Model	Precision	Recall	F1-score	Accuracy	Error rate
SVM	90.6	87.3	88.1	87.3	12.7
MLR	90.3	83.3	85.6	83.3	16.7
MNB	86.3	73.6	77.5	73.5	26.5

Table 4. Classifier evaluation using BERT method (%)

Model	Precision	Recall	F1-score	Accuracy	Error rate
BERT	99.1	99	99	99	1

## D. Problem with Selected Method

Permasalahan yang muncul saat menggunakan algoritma **BERT** adalah ketika diterapkan pada dataset berukuran besar

Due to its time requirements and costs, bidirectional encoder representations from transformers (BERT) presents a challenge when used to classify large datasets, but it is generally still used because it is relatively inexpensive to train.

# Introduction Review

---

## E. Solution Repair Method

Solusi yang ditawarkan adalah dengan tetap menerapkan algoritma **BERT** dengan mempertimbangkan dataset dan keunggulannya dibandingkan model *machine learning* pada *Existing Method*

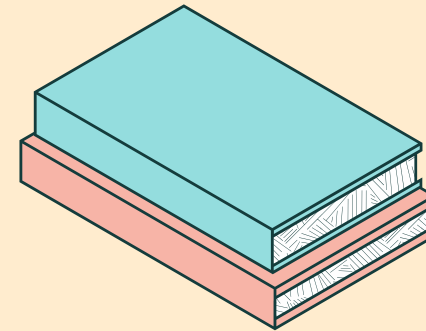
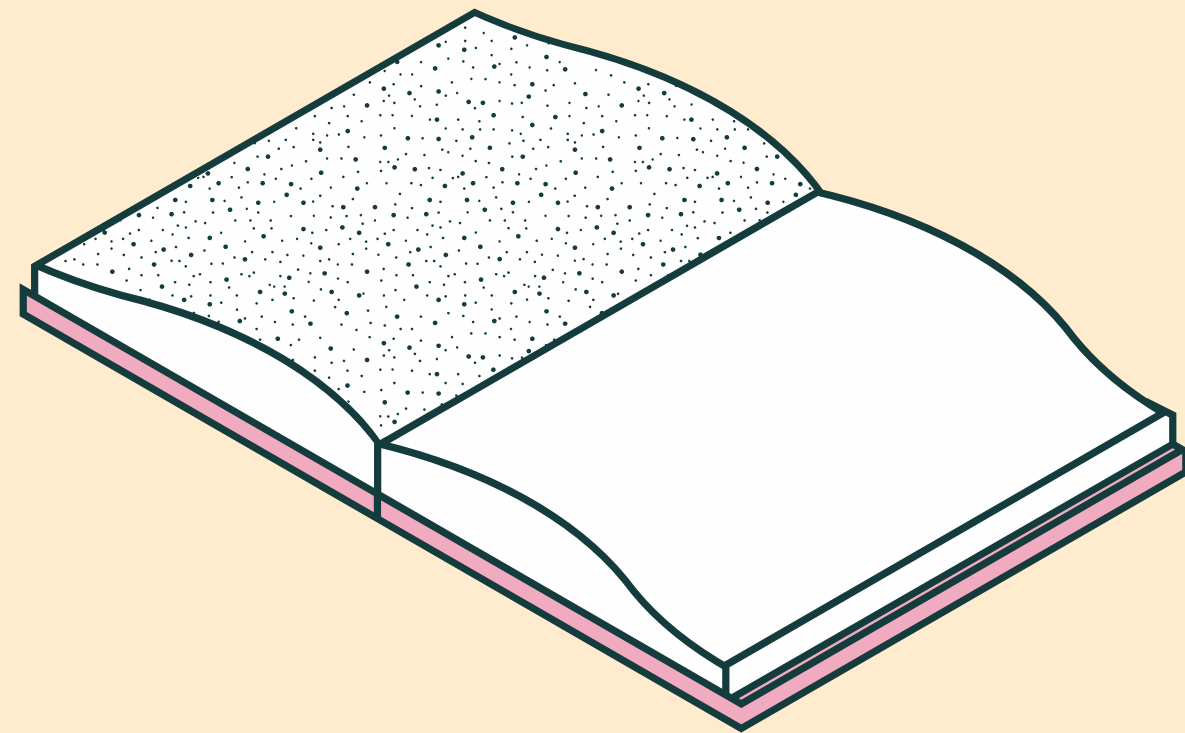
In this study, we investigated whether sentiment analysis in texts can be classified using BERT-base. Using the Pontiki dataset, known as the laptop dataset [21], BERT, developed by Alexander Rietzler and fine-tuned with several layers, has been successful in detecting sentiment. As a result, other machine learning algorithms, such as multinomial Naive Bayes (MNB) and multinomial logistic regression (MLR), did not achieve highly accurate predictions when applied to Twitter data about weather conditions [22].

## F. Summary of research objectives

Pada penelitian ini penulis mengusulkan sebuah metode *deep learning* dengan model **BERT** untuk mengintegrasikan data cuaca di Indonesia secara real-time. Data dari Twitter digunakan sebagai dasar untuk proses deep learning ini.

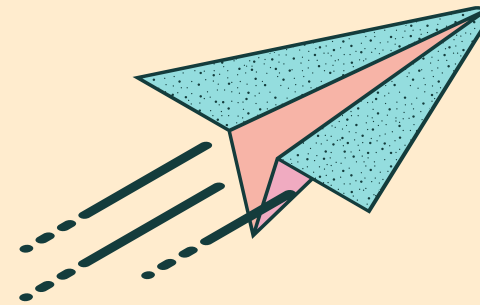
# 3. Predicting Breast Cancer from Risk Factors Using SVM and Extra-Trees-Based Feature Selection Method

---



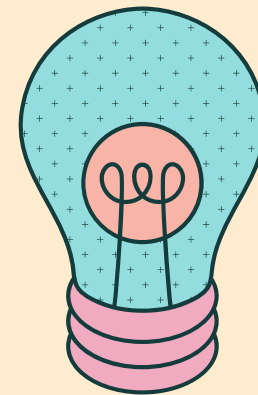
**Journal Name**

Computers



**Publisher**

MDPI AG



**Journal Level**

Q2

# Abstract Review

---

Mengembangkan model prediksi dari faktor risiko dapat memberikan metode yang efisien untuk mengenali kanker payudara. Penggunaan support vector machine (SVM) yang dikombinasikan dengan pengklasifikasi pohon yang sangat acak (extra-trees) untuk mendiagnosis kanker payudara pada tahap awal berdasarkan faktor risiko. Pengklasifikasi pohon ekstra digunakan untuk menghilangkan fitur yang tidak relevan, sementara SVM digunakan untuk mendiagnosis status kanker payudara. Model gabungan SVM dan extra-trees yang diusulkan mencapai akurasi tertinggi sebesar 80,23%, yang secara signifikan lebih baik daripada model pembelajaran mesin lainnya. Dengan menerapkan seleksi fitur berbasis extra-trees, akurasi prediksi rata-rata meningkat hingga 7,29%. Model yang diusulkan diharapkan dapat meningkatkan efisiensi diagnosis kanker payudara berdasarkan faktor risiko dan meningkatkan sistem pendukung keputusan diagnostik.



# Abstract Framework

---

**Purpose and Scope** : Mengembangkan model prediksi dari faktor risiko dapat memberikan metode yang efisien untuk mengenali kanker payudara. **Overview of Method** : Penggunaan SVM yang dikombinasikan dengan pengklasifikasi pohon (extra-trees) untuk mendiagnosis kanker payudara pada tahap awal berdasarkan faktor risiko. Pengklasifikasi pohon ekstra digunakan untuk menghilangkan fitur yang tidak relevan, sementara SVM digunakan untuk mendiagnosis status kanker payudara. **Research Findings** : Model gabungan SVM dan extra-trees yang diusulkan mencapai akurasi tertinggi sebesar 80,23%, secara signifikan lebih baik daripada model lainnya. **Conclusion** : Dengan menerapkan seleksi fitur berbasis extra-trees, akurasi prediksi rata-rata meningkat hingga 7,29%. Model yang diusulkan diharapkan dapat meningkatkan efisiensi diagnosis kanker payudara berdasarkan faktor risiko dan meningkatkan sistem pendukung keputusan diagnostik.



# Introduction Review

---

## A. Research Object

Obyek pada penelitian ini adalah breast cancer disease

help analyze the medical data so that potential health issues can be identified [1–4]. One of the major global health problems and a major cause of mortality in breast cancer. The most prevalent cancer in women and one of the leading causes of mortality among them is breast cancer [5]. The World Health Organization (WHO) reported that three out of every ten

## B. Existing Method

Penelitian sebelumnya dengan topik yang sama menggunakan metode logistic regression (LR), multi-layer perceptron (MLP), decision tree (DT), K-nearest neighbor (KNN), random forest (RF), naïve Bayes (NB), eXtreme Gradient Boosting (XGBoost), adaptive boosting (AdaBoost) dan SVM

We assessed how well the machine learning model performed and how feature selection affected the model's accuracy. The proposed SVM with extra-trees was compared with other data-driven models to predict breast cancer using known risk factors. The ML algorithms, namely logistic regression (LR), multi-layer perceptron (MLP), decision tree (DT), K-nearest neighbor (KNN), random forest (RF), naïve Bayes (NB), eXtreme Gradient Boosting (XGBoost), and adaptive boosting (AdaBoost) were employed as breast cancer prediction models. Averaging over 10 iterations of stratified 10-fold CV, the metrics for

Support vector machine (SVM) is an ML model that divides instances of each class from the others by locating the linear optimum hyperplane after nonlinearly mapping the original data into a high-dimensional feature space. SVMs have demonstrated superior performance for breast cancer detection as compared to conventional models [15–18]. Additionally,

# Introduction Review

## C. The Advantages and Disadvantages of Existing Methods

Tidak disebutkan secara spesifik akan tetapi terdapat perbandingan akurasi dari metode yang diusulkan dengan metode lainnya.pada bab hasil dan pembahasan

Table 3. Performance metrics for breast cancer prediction.

Method	Accuracy (%)	Precision (%)	Specificity (%)	Sensitivity/Recall (%)	AUC
MLP	66.82	66.48	66.60	66.60	0.67
LR	57.58	59.90	56.14	56.14	0.56
KNN	65.91	65.73	65.52	65.52	0.66
DT	67.96	70.04	67.83	67.83	0.68
NB	57.73	62.11	59.79	59.79	0.60
RF	67.80	70.14	67.02	67.02	0.67
AdaBoost	74.09	76.51	73.43	73.43	0.73
XGBoost	75.08	75.78	74.31	82.62	0.74
Proposed model	80.23	82.71	78.57	78.57	0.78

## D. Problem with Selected Method

Pada bab introduction tidak tertulis masalah pada metode yang di usulkan tetapi masalah yang ingin diselesaikan dengan jurnal ini adalah mencari metode yang efisien untuk mengenali kanker payudara.

**Abstract:** Developing a prediction model from risk factors can provide an efficient method to recognize breast cancer.

# Introduction Review

---

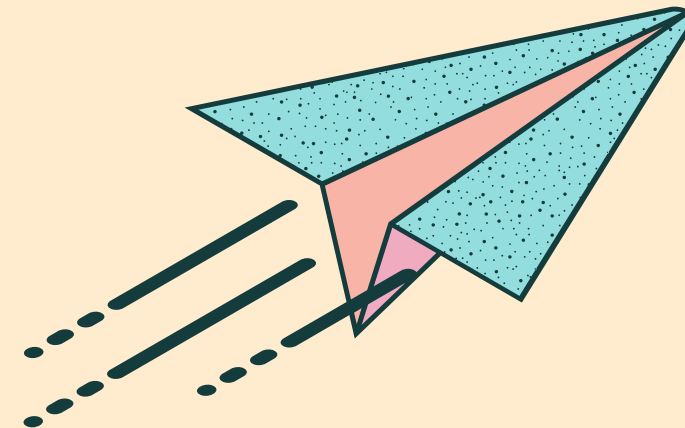
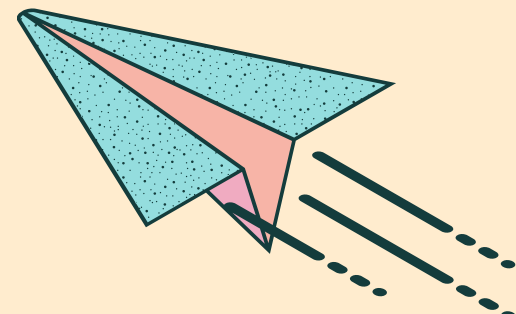
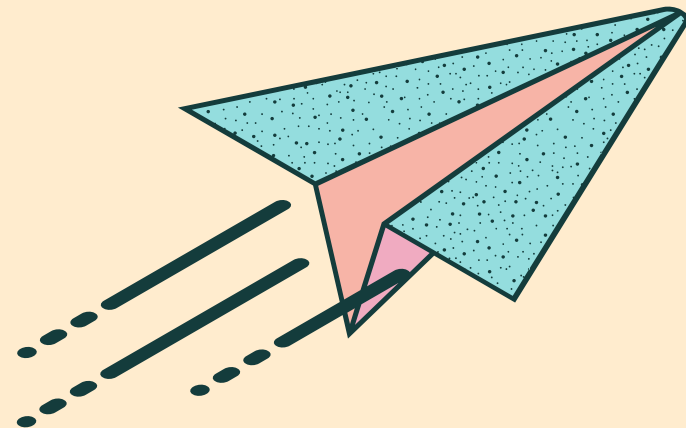
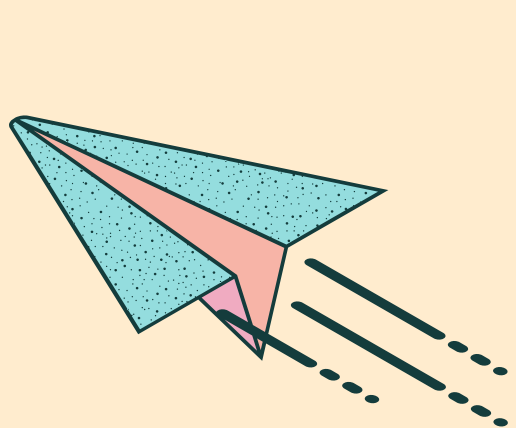
## E. Solution Repair Method

Setelah membandingkan beberapa metode, penulis mengusulkan menggunakan metode SVM dengan extra-trees untuk mengatasi masalah.

The final features were extracted based on extra-trees, and SVM was used as a classifier. The result showed that the proposed model delivers 90.76% prediction accuracy. By integrating extra-trees as feature selection and an ensemble of extreme learning machines (ELM).

## F. Summary of research objectives

Tujuan penelitian ini adalah untuk mengembangkan model prediksi kanker payudara berdasarkan SVM dan extra-trees menggunakan dataset yang menggabungkan faktor risiko kanker payudara. Penelitian ini juga bertujuan untuk membandingkan performa model yang diusulkan dengan beberapa model prediksi machine learning lainnya dan menunjukkan bahwa model yang diusulkan memiliki akurasi yang lebih baik daripada model lainnya. Peneliti juga membangun sebuah aplikasi berbasis web yang mengimplementasikan metode yang diusulkan.



**TERIMAKASIH**

