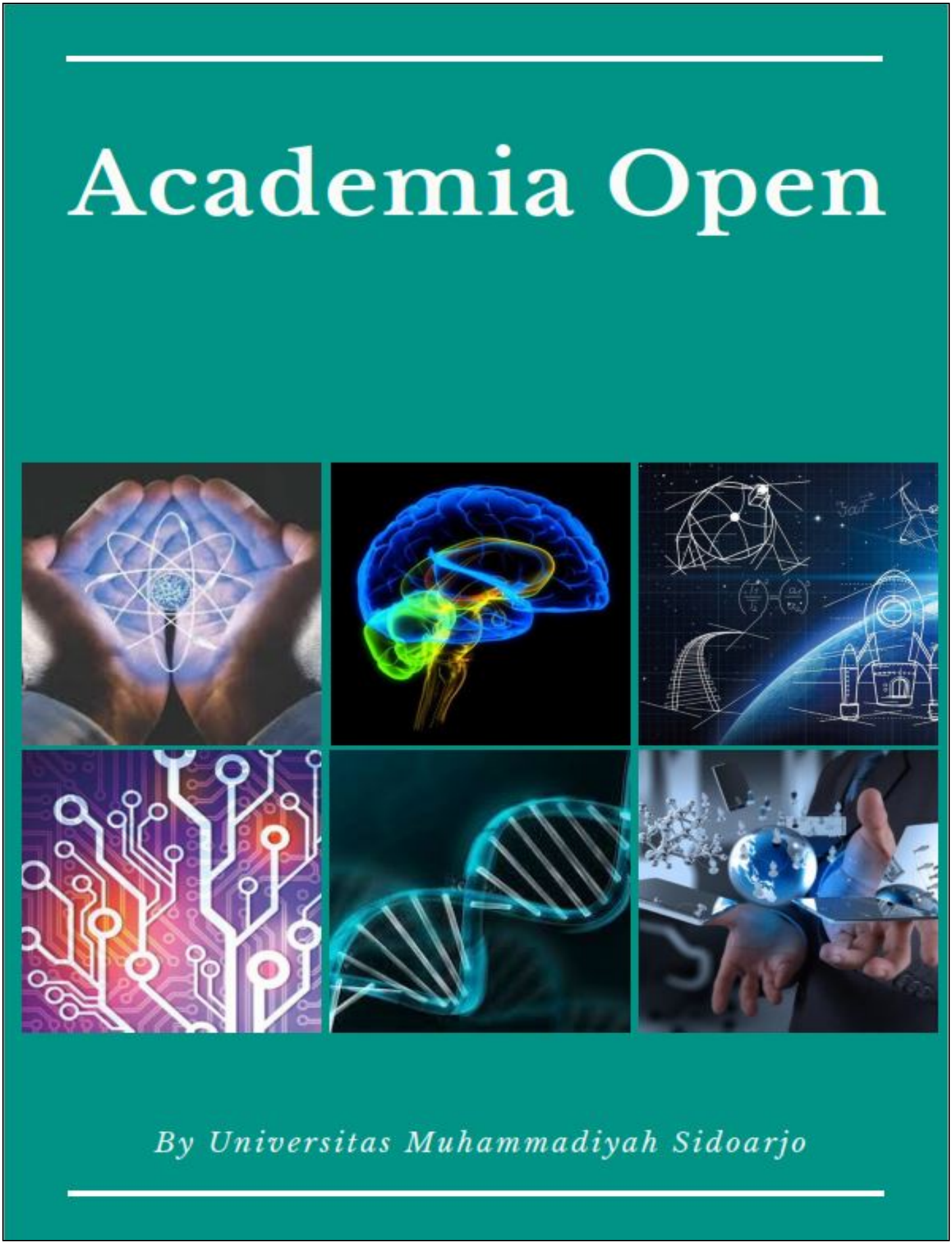
Academia Open
Vol 10 No 1 (2025): June (In Progress)
DOI: 10.21070/acopen.9.2024.10451 . Article type: (Computer Science)

Table Of Content

Journal Cover 2
Author[s] Statement 3
Editorial Team 4
Article information 5
Check this article update (crossmark) 5
Check this article impact 5
Cite this article 5
Title page 6
Article Title 6
Author information 6
Abstract 6
Article content 7

ISSN 2714-7444 (online), https://acopen.umsida.ac.id, published by Universitas Muhammadiyah Sidoarjo
Copyright © Author(s). This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY).

Academia Open
Vol 10 No 1 (2025): June (In Progress)
DOI: 10.21070/acopen.9.2024.10451 . Article type: (Computer Science)



ISSN 2714-7444 (online), https://acopen.umsida.ac.id, published by Universitas Muhammadiyah Sidoarjo
Copyright © Author(s). This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY).

Academia Open
Vol 10 No 1 (2025): June (In Progress)
DOI: 10.21070/acopen.9.2024.10451 . Article type: (Computer Science)

Originality Statement

The author[s] declare that this article is their own work and to the best of their knowledge it contains no materials previously published or written by another person, or substantial proportions of material which have been accepted for the published of any other published materials, except where due acknowledgement is made in the article. Any contribution made to the research by others, with whom author[s] have work, is explicitly acknowledged in the article.

Conflict of Interest Statement

The author[s] declare that this article was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright Statement

Copyright © Author(s). This article is published under the Creative Commons Attribution (CC BY 4.0) licence. Anyone may reproduce, distribute, translate and create derivative works of this article (for both commercial and non-commercial purposes), subject to full attribution to the original publication and authors. The full terms of this licence may be seen at http://creativecommons.org/licences/by/4.0/legalcode

Copyright © Author(s). This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY).

Academia Open
Vol 10 No 1 (2025): June (In Progress)
DOI: 10.21070/acopen.9.2024.10451 . Article type: (Computer Science)

EDITORIAL TEAM

Editor in Chief

Mochammad Tanzil Multazam, Universitas Muhammadiyah Sidoarjo, Indonesia

Managing Editor

Bobur Sobirov, Samarkand Institute of Economics and Service, Uzbekistan

Editors

Fika Megawati, Universitas Muhammadiyah Sidoarjo, Indonesia

Mahardika Darmawan Kusuma Wardana, Universitas Muhammadiyah Sidoarjo, Indonesia

Wiwit Wahyu Wijayanti, Universitas Muhammadiyah Sidoarjo, Indonesia

Farkhod Abdurakhmonov, Silk Road International Tourism University, Uzbekistan

Dr. Hindarto, Universitas Muhammadiyah Sidoarjo, Indonesia

Evi Rinata, Universitas Muhammadiyah Sidoarjo, Indonesia

M Faisal Amir, Universitas Muhammadiyah Sidoarjo, Indonesia

Dr. Hana Catur Wahyuni, Universitas Muhammadiyah Sidoarjo, Indonesia

Complete list of editorial team (link)

Complete list of indexing services for this journal (link)

How to submit to this journal (link)

ISSN 2714-7444 (online), https://acopen.umsida.ac.id, published by Universitas Muhammadiyah Sidoarjo
Copyright © Author(s). This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY).

Academia Open
Vol 10 No 1 (2025): June (In Progress)
DOI: 10.21070/acopen.9.2024.10451 . Article type: (Computer Science)

Article information

Check this article update (crossmark)



Check this article impact (*)















Save this article to Mendeley



(*) Time for indexing process is various, depends on indexing database platform

ISSN 2714-7444 (online), https://acopen.umsida.ac.id, published by Universitas Muhammadiyah Sidoarjo
Copyright © Author(s). This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY).

Academia Open
Vol 10 No 1 (2025): June (In Progress)
DOI: 10.21070/acopen.9.2024.10451 . Article type: (Computer Science)

Machine Learning: Algorithms, Real World Applications and Research Directions

Pembelajaran Mesin: Algoritme, Aplikasi Dunia Nyata, dan Arah Penelitian

Muthna Nehad Al-Tameemi, Muthna@gmail.com, (1)
, Iraq

(1) Corresponding author

Abstract

In the era of the Fourth Industrial Revolution (4IR), vast amounts of data are generated from sources such as IoT, cybersecurity, mobile, business, social media, and healthcare. Artificial intelligence (AI), particularly machine learning (ML), is essential for analyzing this data and developing smart, automated applications. Various machine learning algorithms, including supervised, unsupervised, semi-supervised, and reinforcement learning, are used to enhance application intelligence. Additionally, deep learning, a subset of ML, can process large-scale data effectively. This paper provides an overview of these algorithms and their applications across fields like cybersecurity, smart cities, healthcare, e-commerce, and agriculture. It also discusses the challenges and research opportunities in the field, aiming to be a valuable resource for both academics and industry professionals, as well as decision-makers in various real-world scenarios.

Highlights:

1. Post-modern ceramics reflect features and connotations of body language.
2. Research explores theoretical framework, variables, and applications in ceramic sciences.
3. Results reveal key indicators, conclusions, and supporting references.

Keywords: Machine learning, Deep learning, Artificial intelligence, Data science, Data-driven decision-making, Predictive analytics, Intelligent applications

Published date: 2025-01-02 00:00:00

ISSN 2714-7444 (online), https://acopen.umsida.ac.id, published by Universitas Muhammadiyah Sidoarjo
Copyright © Author(s). This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY).

Academia Open
Vol 10 No 1 (2025): June (In Progress)
DOI: 10.21070/acopen.9.2024.10451 . Article type: (Computer Science)

Introduction

We live in the age of data, where everything around us is connected to a data source, and everything in our lives is digitally recorded [21, 103]. For instance, the current elec- tronic world has a wealth of various kinds of data, such as the Internet of Things (IoT) data, cybersecurity data, smart city data, business data, smartphone data, social media data, health data, COVID-19 data, and many more. The data can be structured, semi-structured, or unstructured, discussed briefly in Sect. “Types of Real-World Data and Machine Learning Techniques”, which is increasing day-by-day. Extracting insights from these data can be used to build various intelligent applications in the relevant domains. For instance, to build a data-driven automated and intelligent cybersecurity system, the relevant cybersecurity data can be used [105]; to build personalized context-aware smart mobile applications, the relevant mobile data can be used [103], and so on. Thus, the data management tools and tech- niques having the capability of extracting insights or useful knowledge from the data in a timely and intelligent way is urgently needed, on which the real-world applications are based.

Artificial intelligence (AI), particularly, machine learning (ML) have grown rapidly in recent years in the context of data analysis and computing that typically allows the appli- cations to function in an intelligent manner [95]. ML usually provides systems with the ability to learn and enhance from experience automatically without being specifically pro- grammed and is generally referred to as the most popular latest technologies in the fourth industrial revolution (4IR or Industry 4.0) [103, 105]. “Industry 4.0” [114] is typically the ongoing automation of conventional manufacturing and industrial practices, including exploratory data processing, using new smart technologies such as machine learning automation. Thus, to intelligently analyze these data and to develop the corresponding real-world applications, machine learning algorithms is the key. The learning algorithms can be categorized into four major types, such as supervised, unsupervised, semi-supervised, and reinforcement learn- ing in the area [75], discussed briefly in Sect. “Types of Real-World Data and MachineLearning Techniques”. The popularity of these approaches to learning is increasing day-by-day, which is shown in Fig. 1, based on data col- lected from Google Trends [4] over the last five years. The x-axis of the figure indicates the specific dates and the corresponding popularity score within the range of 0 (minimum) to 100 (maximum) has been shown in y-axis. According to Fig. , the popularity indication values for these learning types are low in 2015 and are increasing day by day. These statistics motivate us to study on *machine learning* in this paper, which can play an important role in the real-world through Industry 4.0 automation.

In general, the effectiveness and the efficiency of a machine learning solution depend on the nature and char- acteristics of data and the performance of the learning algorithms. In the area of machine learning algorithms, classification analysis, regression, data clustering, fea- ture engineering and dimensionality reduction, associa- tion rule learning, or reinforcement learning techniques exist to effectively build data-driven systems [41, 125]. Besides, deep learning originated from the artificial neural network that can be used to intelligently analyze data, which is known as part of a wider family of machine learning approaches [96]. Thus, selecting a proper learn- ing algorithm that is suitable for the target application in a particular domain is challenging. The reason is that the purpose of different learning algorithms is different, even the outcome of different learning algorithms in a similar category may vary depending on the data characteristics [106]. Thus, it is important to understand the principles of various machine learning algorithms and their appli- cability to apply in various real-world application areas, such as IoT systems, cybersecurity services, business and recommendation systems, smart cities, healthcare and COVID-19, context-aware systems, sustainable agricul- ture, and many more that are explained briefly in Sect. “Applications of Machine Learning”.

Based on the importance and potentiality of “Machine Learning” to analyze the data mentioned above, in this paper, we provide a comprehensive view on various types of machine learning algorithms that can be applied to enhance the intelligence and the capabilities of an appli- cation. Thus, the key contribution of this study is explain- ing the principles and potentiality of different machine learning techniques, and their applicability in various real- world application areas mentioned earlier. The purpose of this paper is, therefore, to provide a basic guide for those academia and industry people who want to study, research, and develop data-driven automated and intelli- gent systems in the relevant areas based on machine learn- ing techniques.

The key contributions of this paper are listed as follows:

1. To define the scope of our study by taking into account the nature and characteristics of various types of real- world data and the capabilities of various learning tech- niques.

2. To provide a comprehensive view on machine learning algorithms that can be applied to enhance the intelligence and capabilities of a data-driven application.

3. To discuss the applicability of machine learning-based solutions in various real-world application domains.

4. To highlight and summarize the potential research direc- tions within the scope of our study for intelligent data analysis and services.

ISSN 2714-7444 (online), https://acopen.umsida.ac.id, published by Universitas Muhammadiyah Sidoarjo
Copyright © Author(s). This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY).

Academia Open
Vol 10 No 1 (2025): June (In Progress)
DOI: 10.21070/acopen.9.2024.10451 . Article type: (Computer Science)

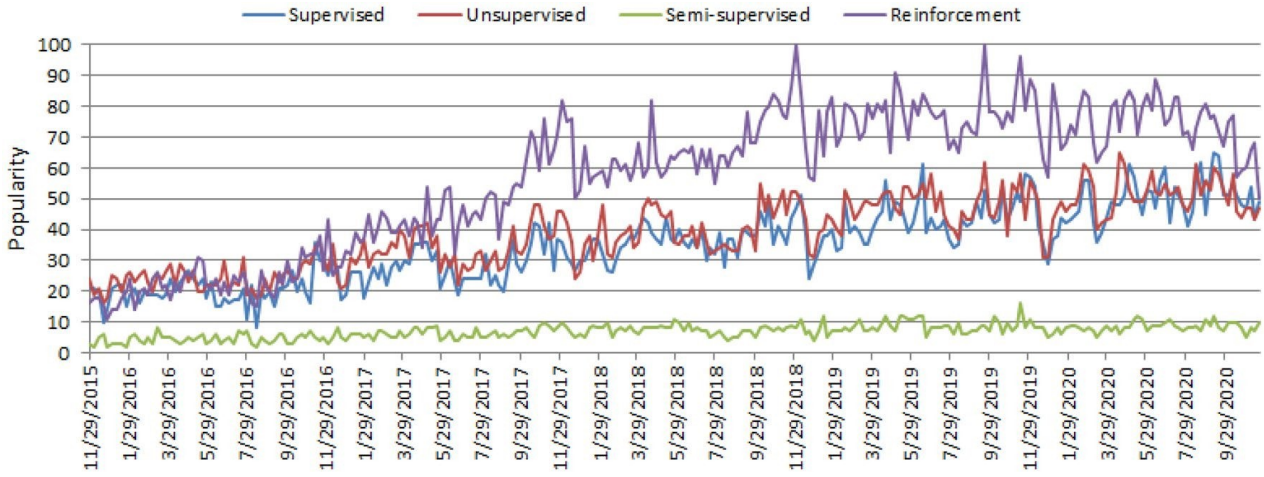


Figure 1. *The worldwide popularity score of various types of ML algorithms (supervised, unsupervised, semi-supervised, and reinforcement) in a range of 0 (min) to 100 (max) over time where x-axis represents the timestamp information and y-axis represents the corresponding score*

The rest of the paper is organized as follows. The next sec- tion presents the types of data and machine learning algo- rithms in a broader sense and defines the scope of our study. We briefly discuss and explain different machine learning algorithms in the subsequent section followed by which various real-world application areas based on machine learning algorithms are discussed and summarized. In the penultimate section, we highlight several research issues and potential future directions, and the final section concludes this paper.

Result and Discussion

Types of Real-World Data and Machine Learning Techniques

Machine learning algorithms typically consume and process data to learn the related patterns about individuals, business processes, transactions, events, and so on. In the following, we discuss various types of real-world data as well as cat- egories of machine learning algorithms.

Types of Real-World Data

Usually, the availability of data is considered as the key to construct a machine learning model or data-driven real-world systems [103, 105]. Data can be of various forms, such as structured, semi-structured, or unstructured [41, 72]. Besides, the “metadata” is another type that typically repre- sents data about the data. In the following, we briefly discuss these types of data.

1. Structured: It has a well-defined structure, conforms to a data model following a standard order, which is highly organized and easily accessed, and used by an entity or a computer program. In well-defined schemes, such as rela-tional databases, structured data are typically stored, i.e., in a tabular format. For instance, names, dates, addresses, credit card numbers, stock information, geolocation, etc. are examples of structured data.

2. Unstructured: On the other hand, there is no pre-defined format or organization for unstructured data, making it much more difficult to capture, process, and analyze, mostly containing text and multimedia material. For example, sensor data, emails, blog entries, wikis, and word processing documents, PDF files, audio files, videos, images, presentations, web pages, and many other types of business documents can be considered as unstructured data.

3. Semi-structured: Semi-structured data are not stored in a relational database like the structured data mentioned above, but it does have certain organizational proper- ties that make it easier to analyze. HTML, XML, JSON documents, NoSQL databases, etc., are some examples of semi-structured data.

4. Metadata: It is not the normal form of data, but “data about data”. The primary difference between “data” and “metadata” is that data are simply the material that can classify, measure, or even document something relative to an organization’s data properties. On the other hand, metadata describes the relevant data information, giving it more significance for data users. A basic example of a document’s metadata might be the author, file size, date generated by the document, keywords to define the docu- ment, etc.

In the area of machine learning and data science, research- ers use various widely used datasets for different purposes. These are, for example, cybersecurity datasets such as NSL-KDD [119], UNSW-NB15 [76], ISCX’12 [1], CICDDoS2019 [2], Bot-IoT [59], etc., smartphone datasets such as phone call logs [84, 101], SMS Log [29], mobile

ISSN 2714-7444 (online), https://acopen.umsida.ac.id, published by Universitas Muhammadiyah Sidoarjo
Copyright © Author(s). This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY).

Academia Open
Vol 10 No 1 (2025): June (In Progress)
DOI: 10.21070/acopen.9.2024.10451 . Article type: (Computer Science)

applica- tion usages logs [137] [117], mobile phone notification logs [73] etc., IoT data [16, 57, 62], agriculture and e-commerce data [120, 138], health data such as heart disease [92], dia- betes mellitus [83, 134], COVID-19 [43, 74], etc., and many more in various application domains. The data can be in dif- ferent types discussed above, which may vary from applica- tion to application in the real world. To analyze such data in a particular problem domain, and to extract the insights or useful knowledge from the data for building the real-world intelligent applications, different types of machine learning techniques can be used according to their learning capabili- ties, which is discussed in the following.

Types of Machine Learning Techniques

Machine Learning algorithms are mainly divided into four categories: Supervised learning, Unsupervised learning, Semi-supervised learning, and Reinforcement learning [75], as shown in Fig. 2. In the following, we briefly discuss each type of learning technique with the scope of their applicabil- ity to solve real-world problems. Fig. 2 Various types of machine learning techniques

1. Supervised: Supervised learning is typically the task of machine learning to learn a function that maps an input to an output based on sample input-output pairs [41]. It uses labeled training data and a collection of training examples to infer a function. Supervised learning is car- ried out when certain goals are identified to be accomplished from a certain set of inputs [105], i.e., a task- driven approach. The most common supervised tasks are “classification” that separates the data, and “regression” that fits the data. For instance, predicting the class label or sentiment of a piece of text, like a tweet or a product review, i.e., text classification, is an example of super- vised learning.

2. Unsupervised: Unsupervised learning analyzes unla- beled datasets without the need for human interference, i.e., a data-driven process [41]. This is widely used for extracting generative features, identifying meaningful trends and structures, groupings in results, and explora- tory purposes. The most common unsupervised learning tasks are clustering, density estimation, feature learn- ing, dimensionality reduction, finding association rules, anomaly detection, etc.

3. Semi-supervised: Semi-supervised learning can be defined as a hybridization of the above-mentioned supervised and unsupervised methods, as it operates on both labeled and unlabeled data [41, 105]. Thus, it falls between learning “without supervision” and learn- ing “with supervision”. In the real world, labeled data could be rare in several contexts, and unlabeled data are numerous, where semi-supervised learning is useful [75]. The ultimate goal of a semi-supervised learning model is to provide a better outcome for prediction than that produced using the labeled data alone from the model. Some application areas where semi-supervised learning is used include machine translation, fraud detection, labe- ling data and text classification.

4. Reinforcement: Reinforcement learning is a type of machine learning algorithm that enables software agents and machines to automatically evaluate the opti- mal behavior in a particular context or environment to improve its efficiency [52], i.e., an environment-driven approach. This type of learning is based on reward or penalty, and its ultimate goal is to use insights obtained from environmental activists to take action to increase the reward or minimize the risk [75]. It is a powerful tool for training AI models that can help increase automation or optimize the operational efficiency of sophisticated systems such as robotics, autonomous driving tasks, manufacturing and supply chain logistics, however, not preferable to use it for solving the basic or straightfor- ward problems.

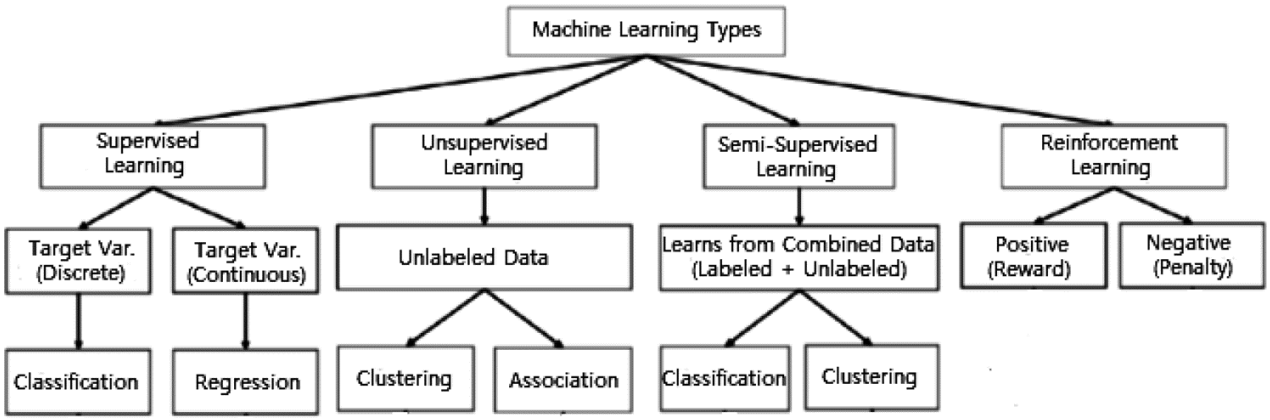


Figure 2. *Various types of machine learning techniques*

Thus, to build effective models in various application areas different types of machine learning techniques can play a significant role according to their learning capabilities, depending on the nature of the data discussed earlier, and the target outcome. In Table 1, we summarize various types of machine learning techniques with examples. In the follow- ing, we provide a comprehensive view of machine learning algorithms that can be applied to enhance the intelligence and capabilities of a data-driven application.

ISSN 2714-7444 (online), https://acopen.umsida.ac.id, published by Universitas Muhammadiyah Sidoarjo
Copyright © Author(s). This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY).

Academia Open
Vol 10 No 1 (2025): June (In Progress)
DOI: 10.21070/acopen.9.2024.10451 . Article type: (Computer Science)

Learning type	Model building	Examples
Supervised	Algorithms or models learn from labeled data (task-driven approach)	Classification, regression
Unsupervised	Algorithms or models learn from unlabeled data (Data-Driven Approach)	Clustering, associa- tions, dimensionalityreduction
Semi-supervised	Models are built using combined data (labeled + unlabeled)	Classification, clustering
Reinforcement	Models are based on reward or penalty (environment-driven approach)	Classification, control

Table 1. *Various types of machine learning techniques with examples*

Machine Learning Tasks and Algorithms

In this section, we discuss various machine learning algo- rithms that include classification analysis, regression analysis, data clustering, association rule learning, feature engineering for dimensionality reduction, as well as deep learning methods. A general structure of a machine learning- based predictive model has been shown in Fig. 3, where the model is trained from historical data in phase 1 and the outcome is generated in phase 2 for the new test data.

Classification Analysis

Classification is regarded as a supervised learning method in machine learning, referring to a problem of predictive modeling as well, where a class label is predicted for a given example [41]. Mathematically, it maps a function (f) from input variables (X) to output variables (Y) as target, label or categories. To predict the class of given data points, it can be carried out on structured or unstructured data. For example, spam detection such as “spam” and “not spam” in email service providers can be a classification problem. In the fol- lowing, we summarize the common classification problems.

1. Binary classification: It refers to the classification tasks having two class labels such as “true and false” or “yes and no” []. In such binary classification tasks, one class could be the normal state, while the abnormal state could be another class. For instance, “cancer not detected” is the normal state of a task that involves a medical test, and “cancer detected” could be considered as the abnormal state. Similarly, “spam” and “not spam” in the above example of email service providers are con- sidered as binary classification.

2. Multiclass classification: Traditionally, this refers to those classification tasks having more than two class labels []. The multiclass classification does not have the principle of normal and abnormal outcomes, unlike binary classification tasks. Instead, within a range of specified classes, examples are classified as belonging to one. For example, it can be a multiclass classification task to classify various types of network attacks in the NSL-KDD [] dataset, where the attack categories are classified into four class labels, such as DoS (Denial of Service Attack), U2R (User to Root Attack), R2L (Root to Local Attack), and Probing Attack.

3. Multi-label classification: In machine learning, multi- label classification is an important consideration where an example is associated with several classes or labels. Thus, it is a generalization of multiclass classification, where the classes involved in the problem are hierar- chically structured, and each example may simultane- ously belong to more than one class in each hierarchical level, e.g., multi-level text classification. For instance, Google news can be presented under the categories of a “city name”, “technology”, or “latest news”, etc. Multi- label classification includes advanced machine learn- ing algorithms that support predicting various mutually non-exclusive classes or labels, unlike traditional clas- sification tasks where class labels are mutually exclu- sive [].

ISSN 2714-7444 (online), https://acopen.umsida.ac.id, published by Universitas Muhammadiyah Sidoarjo
Copyright © Author(s). This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY).

Academia Open
Vol 10 No 1 (2025): June (In Progress)
DOI: 10.21070/acopen.9.2024.10451 . Article type: (Computer Science)

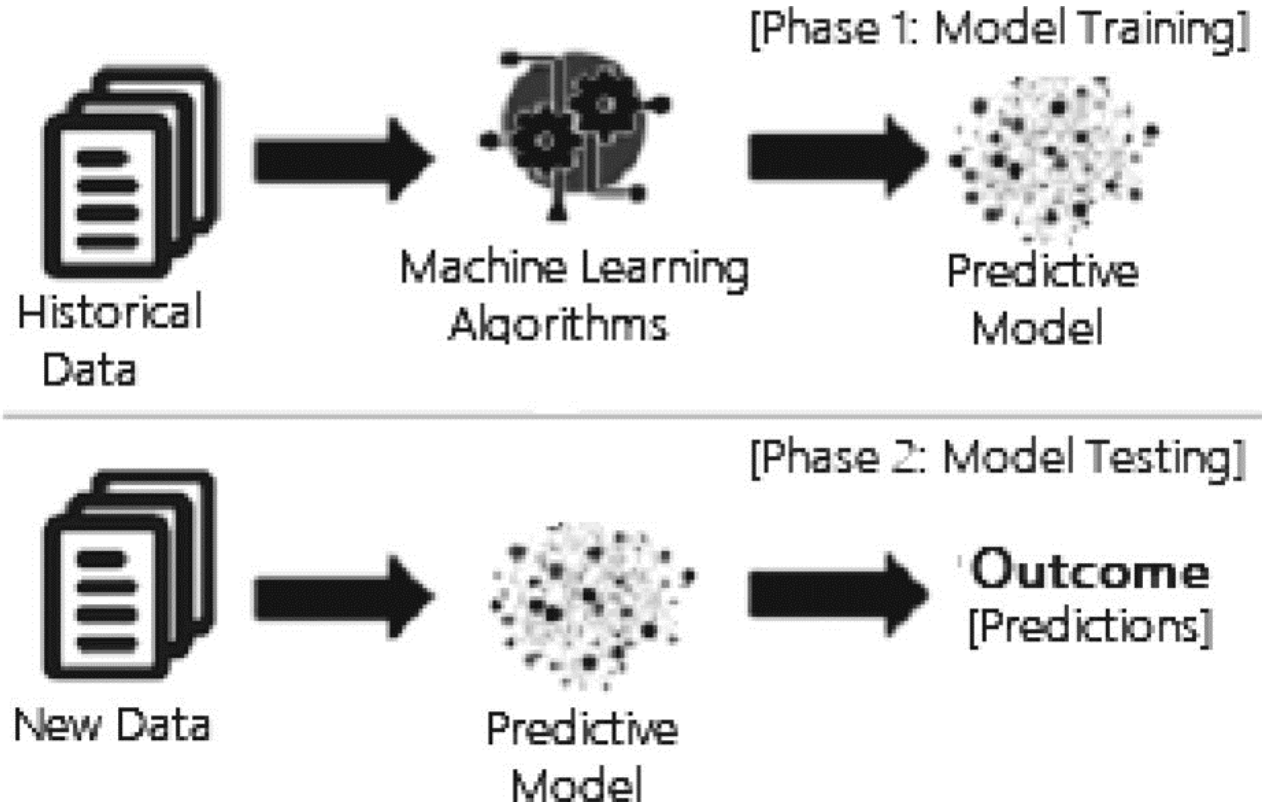


Figure 3. *A general structure of a machine learning based predictive model considering both the training and testing phase*

Many classification algorithms have been proposed in the machine learning and data science literature [41, 125]. In the following, we summarize the most common and popular methods that are used widely in various application areas.

– *Naive Bayes (NB):* The naive Bayes algorithm is based on the Bayes' theorem with the assumption of independence between each pair of features []. It works well and can be used for both binary and multi-class categories in many real-world situations, such as document or text classification, spam filtering, etc. To effectively classify the noisy instances in the data and to construct a robust prediction model, the NB classifier can be used []. The key benefit is that, compared to more sophisticated approaches, it needs a small amount of training data to estimate the necessary parameters and quickly []. However, its performance may affect due to its strong assumptions on features independence. Gaussian, Multi- nomial, Complement, Bernoulli, and Categorical are the common variants of NB classifier [].

– *Linear Discriminant Analysis (LDA):* Linear Discrimi- nant Analysis (LDA) is a linear decision boundary clas- sifier created by fitting class conditional densities to data and applying Bayes' rule [,]. This method is also known as a generalization of Fisher's linear discriminant, which projects a given dataset into a lower-dimensional space, i.e., a reduction of dimensionality that minimizes the complexity of the model or reduces the resulting model's computational costs. The standard LDA model usually suits each class with a Gaussian density, assum- ing that all classes share the same covariance matrix []. LDA is closely related to ANOVA (analysis of variance) and regression analysis, which seek to express one dependent variable as a linear combination of other features or measurements.
– *Logistic regression (LR):* Another common probabilistic based statistical model used to solve classification issues in machine learning is Logistic Regression (LR) []. Logistic regression typically uses a logistic function to estimate the probabilities, which is also referred to as the mathematically defined sigmoid function in Eq. . It can overfit high-dimensional datasets and works well when the dataset can be separated linearly. The regulari- zation (L1 and L2) techniques [] can be used to avoid over-fitting in such scenarios. The assumption of linear- ity between the dependent and independent variables is considered as a major drawback of Logistic Regression. It can be used for both classification and regression prob- lems, but it is more commonly used for classification

$$g(z) = \frac{1}{1 + \exp(-z)}. \quad (1)$$

Figure 4.

ISSN 2714-7444 (online), https://acopen.umsida.ac.id, published by Universitas Muhammadiyah Sidoarjo
Copyright © Author(s). This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY).

- *K-nearest neighbors (KNN):* K-Nearest Neighbors (KNN) [] is an “instance-based learning” or non-gen- eralizing learning, also known as a “lazy learning” algo- rithm. It does not focus on constructing a general internal model; instead, it stores all instances corresponding to training data in *n*-dimensional space. KNN uses data and classifies new data points based on similarity measures (e.g., Euclidean distance function) []. Classification is computed from a simple majority vote of the *k* nearest neighbors of each point. It is quite robust to noisy train- ing data, and accuracy depends on the data quality. The biggest issue with KNN is to choose the optimal number of neighbors to be considered. KNN can be used both for classification as well as regression.

- *Support vector machine (SVM):* In machine learning, another common technique that can be used for classi- fication, regression, or other tasks is a support vector machine (SVM) []. In high- or infinite-dimensional space, a support vector machine constructs a hyper-plane or set of hyper-planes. Intuitively, the hyper-plane, which has the greatest distance from the nearest training data points in any class, achieves a strong separation since, in general, the greater the margin, the lower the classifier’s generalization error. It is effective in high-dimensional spaces and can behave differently based on different mathematical functions known as the kernel. Linear, polynomial, radial basis function (RBF), sigmoid, etc. are the popular kernel functions used in SVM classifier []. However, when the data set contains more noise, such as overlapping target classes, SVM does not per- form well.

– *Decision tree (DT):* Decision tree (DT) [] is a well- known non-parametric supervised learning method. DT learning methods are used for both the classifica- tion and regression tasks []. ID3 [], C4.5 [], and CART [] are well known for DT algorithms. Moreover, recently proposed BehavDT [], and IntrudTree [] by Sarker et al. are effective in the relevant application domains, such as user behavior analytics and cyberse- curity analytics, respectively. By sorting down the tree from the root to some leaf nodes, as shown in Fig. , DT classifies the instances. Instances are classified by check- ing the attribute defined by that node, starting at the root node of the tree, and then moving down the tree branch corresponding to the attribute value. For splitting, the most popular criteria are “gini” for the Gini impurity and “entropy” for the information gain that can be expressed mathematically as [].

$$\text{Entropy}: H(x) = -\sum_{i=1} p(x_i)\log_2 p(x_i) \quad (2)$$

$$\text{Gini}(E) = 1 - \sum_{i=1} p_i^2. \quad (3)$$

Figure 5.

– *Random forest (RF):* A random forest classifier [] is well known as an ensemble classification technique that is used in the field of machine learning and data science in various application areas. This method uses “parallel ensembling” which fits several decision tree classifiers in parallel, as shown in Fig. , on different data set sub- samples and uses majority voting or aver- ages for the outcome or final result. It thus minimizes the over-fitting problem and increases the prediction accuracy and control []. Therefore, the RF learning model with multiple decision trees is typically more accurate than a single decision tree based model []. To build a series of decision trees with controlled vari- ation, it combines bootstrap aggregation (bagging) [] and random feature selection []. It is adaptable to both classification and regression problems and fits well for both categorical and continuous values.

ISSN 2714-7444 (online), https://acopen.umsida.ac.id, published by Universitas Muhammadiyah Sidoarjo
Copyright © Author(s). This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY).

Academia Open
Vol 10 No 1 (2025): June (In Progress)
DOI: 10.21070/acopen.9.2024.10451 . Article type: (Computer Science)

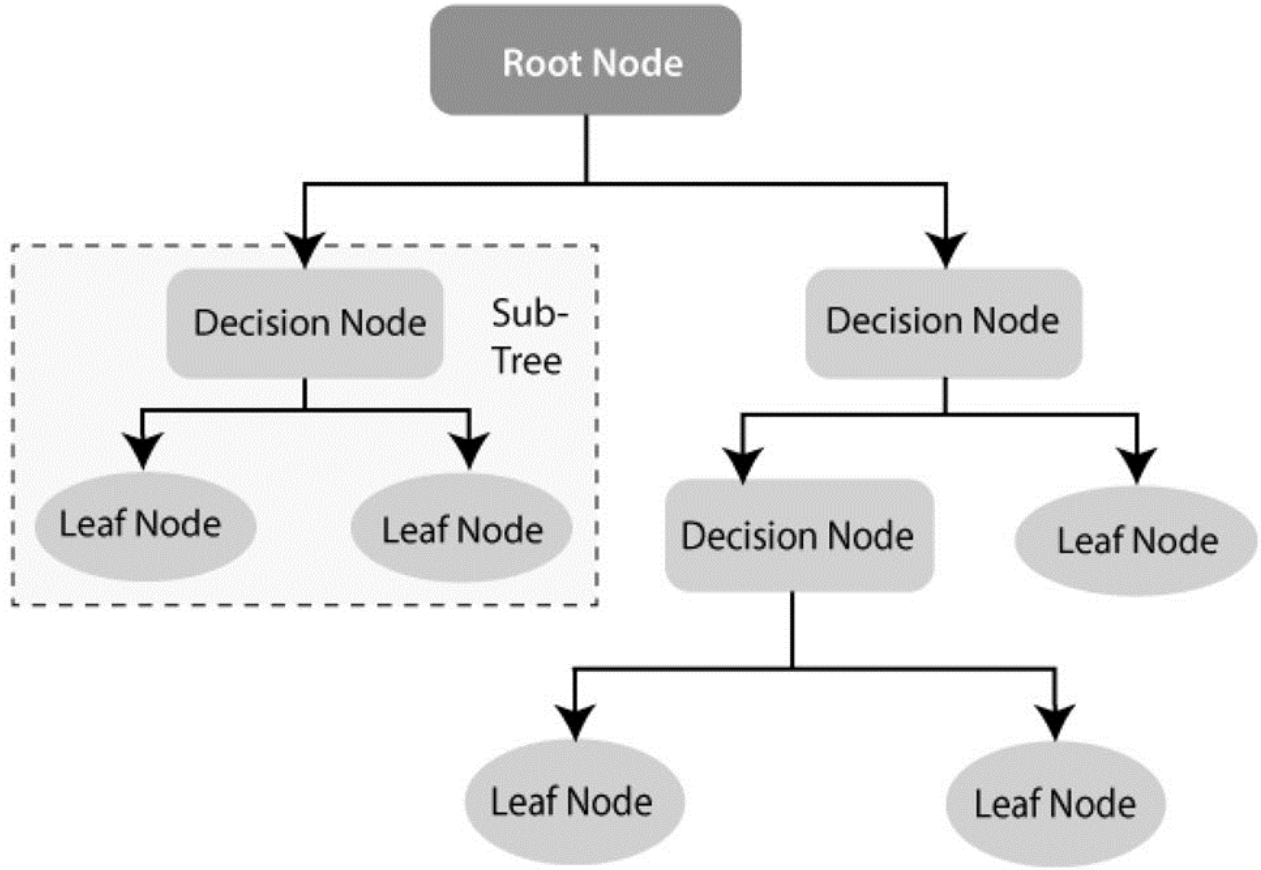
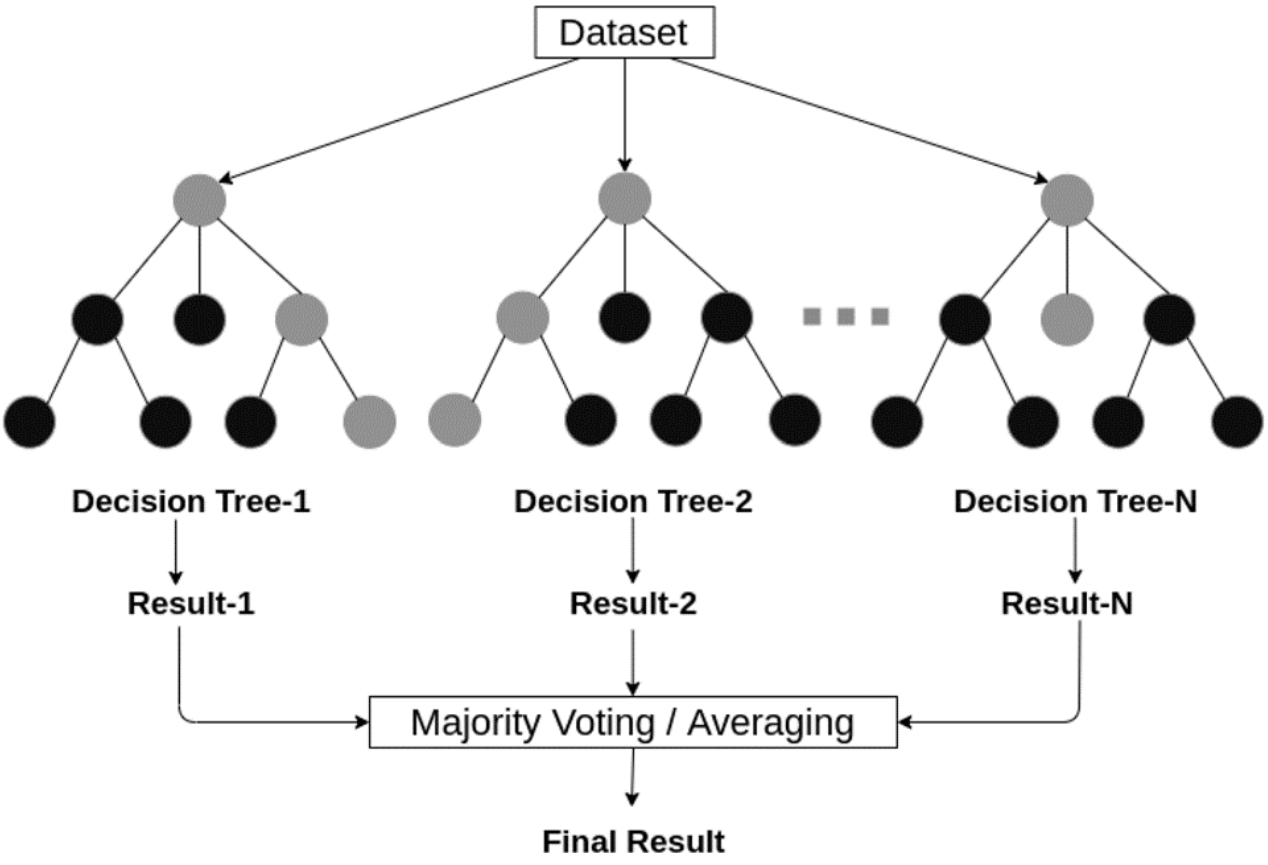


Figure 6. *An example of a decision tree structure*



ISSN 2714-7444 (online), https://acopen.umsida.ac.id, published by Universitas Muhammadiyah Sidoarjo
Copyright © Author(s). This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY).

Academia Open
Vol 10 No 1 (2025): June (In Progress)
DOI: 10.21070/acopen.9.2024.10451 . Article type: (Computer Science)

Figure 7. *An example of a random forest structure considering multiple decision trees*

- *Adaptive Boosting (AdaBoost):* Adaptive Boosting (AdaBoost) is an ensemble learning process that employs an iterative approach to improve poor classi- fiers by learning from their errors. This is developed by Yoav Freund et al. [] and also known as “meta- learning”. Unlike the random forest that uses parallel ensembling, Adaboost uses “sequential ensembling”. It creates a powerful classifier by combining many poorly performing classifiers to obtain a good classi- fier of high accuracy. In that sense, AdaBoost is called an adaptive classifier by significantly improving the efficiency of the classifier, but in some instances, it can trigger overfits. AdaBoost is best used to boost the performance of decision trees, base estimator [], on binary classification problems, however, is sensitive to noisy data and outliers.

– *Extreme gradient boosting (XGBoost):* Gradient Boost- ing, like Random Forests [] above, is an ensemble learning algorithm that generates a final model based on a series of individual models, typically decision trees. The gradient is used to minimize the loss func- tion, similar to how neural networks [] use gradient descent to optimize weights. Extreme Gradient Boosting XGBoost) is a form of gradient boosting that takes more detailed approximations into account when determining the best model []. It computes second-order gradients of the loss function to minimize loss and advanced regu- larization (L1 and L2) [], which reduces over-fitting, and improves model generalization and performance. XGBoost is fast to interpret and can handle large-sized datasets well.

- *Stochastic gradient descent (SGD):* Stochastic gradient descent (SGD) [] is an iterative method for optimiz- ing an objective function with appropriate smoothness properties, where the word ‘stochastic’ refers to random probability. This reduces the computational burden, par- ticularly in high-dimensional optimization problems, allowing for faster iterations in exchange for a lower convergence rate. A gradient is the slope of a function that calculates a variable’s degree of change in response to another variable’s changes. Mathematically, the Gradi- ent Descent is a convex function whose output is a partial derivative of a set of its input parameters. Let, is the learning rate, and *Ji* is the training example cost of *i*th, then Eq. () represents the stochastic gradient descent weight update method at the j^{th} iteration. In large-scale and sparse machine learning, SGD has been successfully applied to problems often encountered in text classifi- cation and natural language processing []. However, SGD is sensitive to feature scaling and needs a range of hyperparameters, such as the regularization parameter and the number of iterations

$$w_{i} := w_{j} - \alpha \frac{\partial}{\partial w_{ij}} \qquad (4)$$

Figure 8.

- *Rule-based classification*: The term rule-based classifi- cation can be used to refer to any classification scheme that makes use of IF-THEN rules for class prediction. Several classification algorithms such as Zero-R [], One-R [], decision trees [,], DTNB [], Ripple Down Rule learner (RIDOR) [], Repeated Incremen- tal Pruning to Produce Error Reduction (RIPPER) [] exist with the ability of rule generation. The decision tree is one of the most common rule-based classifica- tion algorithms among these techniques because it has several advantages, such as being easier to interpret; the ability to handle high-dimensional data; simplicity and speed; good accuracy; and the capability to produce rules for human clear and understandable classification [] []. The decision tree-based rules also provide signifi- cant accuracy in a prediction model for unseen test cases []. Since the rules are easily interpretable, these rule- based classifiers are often used to produce descriptive models that can describe a system including the entities and their relationships

Regression Analysis

Regression analysis includes several methods of machine learning that allow to predict a continuous (y) result variable based on the value of one or more (x) predictor variables [41]. The most significant distinction between classification and regression is that classification predicts distinct class labels, while regression facilitates the prediction of a con- tinuous quantity. Figure 6 shows an example of how classifi- cation is different with regression models. Some overlaps are often found between the two types of machine learning algo- rithms. Regression models are now widely used in a variety of fields, including financial forecasting or prediction, cost estimation, trend analysis, marketing, time series estimation, drug response modeling, and many more. Some of the familiar types of regression algorithms are linear, polynomial, lasso and ridge regression, etc, which are explained briefly in the following.

• *Simple and multiple linear regression:* This is one of the most popular ML modeling techniques as well as a well-known regression technique. In this technique, the dependent variable is continuous, the independent variable(s) can be continuous or discrete, and the form of the regression line is linear. Linear regression creates a relationship between the dependent variable (*Y*) and one or more independent variables (*X*) (also known as regression line) using the best fit straight line []. It is defined by the following equations:

Copyright © Author(s). This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY).

Academia Open
Vol 10 No 1 (2025): June (In Progress)
DOI: 10.21070/acopen.9.2024.10451 . Article type: (Computer Science)

$$y = a + bx + e \quad (5)$$

$$y = a + b_1x_1 + b_2x_2 + \ldots + b_nx_n + \epsilon. \quad (6)$$

Figure 9.

where a is the intercept, b is the slope of the line, and e is the error term. This equation can be used to predict the value of the target variable based on the given predictor variable(s). Multiple linear regression is an extension of simple linear regression that allows two or more predic- tor variables to model a response variable, y, as a linear function [] defined in Eq. , whereas simple linear regression has only 1 independent variable, defined in Eq. .

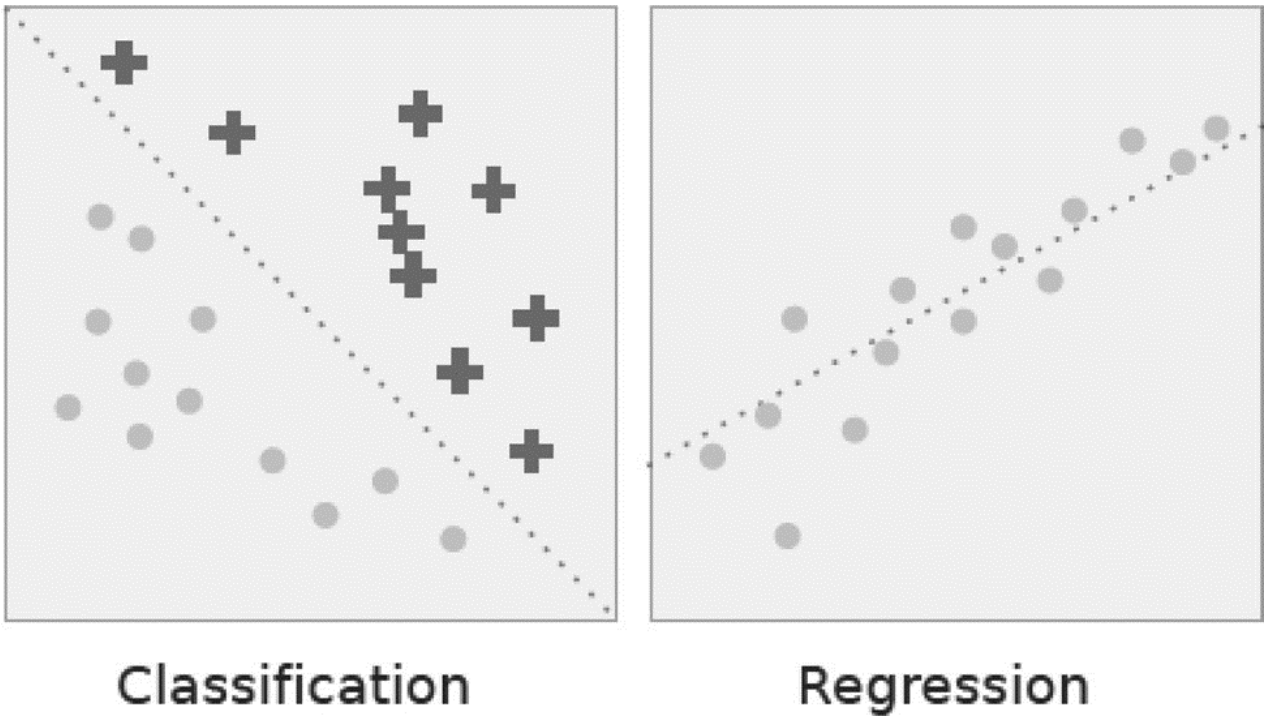


Figure 10. *Classification vs. regression. In classification the dotted line represents a linear boundary that separates the two classes; in regres- sion, the dotted line models the linear relationship between the two variables*

• *Polynomial regression:* Polynomial regression is a form of regression analysis in which the relationship between the independent variable x and the dependent variable y is not linear, but is the polynomial degree of n^{th} in x []. The equation for polynomial regression is also derived from linear regression (polynomial regression of degree 1) equation, which is defined as below:

$$y = b_0 + b_1x + b_2x^2 + b_3x^3 + \cdots + b_nx^n + \epsilon. \quad (7)$$

Figure 11.

Here, y is the predicted/target output, $b0$, $b1$, ...bn are the regression coefficients, x is an independent/ input vari- able. In simple words, we can say that if data are not dis- tributed linearly, instead it is n^{th} degree of polynomial then we use polynomial regression to get desired output.

• *LASSO and ridge regression:* LASSO and Ridge regres- sion are well known as powerful techniques which are typically used for building learning models in presence of a large number of features, due to their capability to preventing over-fitting and reducing the complexity of the model. The LASSO (least absolute shrinkage and selection operator) regression model uses $L1$ regulariza- tion technique [] that uses shrinkage, which penalizes "absolute value of magnitude of coefficients" ($L1$ pen- alty). As a result, LASSO appears to render coefficients to absolute zero. Thus, LASSO regression aims to find the subset of predictors that minimizes the prediction error for a quantitative response variable. On the other hand, ridge regression uses $L2$ regularization [], which is the "squared magnitude of coefficients" ($L2$ penalty). Thus, ridge regression forces the weights to be small but never sets the coefficient value to zero, and does a non- sparse solution. Overall, LASSO regression is useful to obtain a subset of predictors by eliminating less impor- tant features, and ridge regression is useful when a data set has "multicollinearity" which refers to the predictors that are correlated with other predictors

ISSN 2714-7444 (online), https://acopen.umsida.ac.id, published by Universitas Muhammadiyah Sidoarjo
Copyright © Author(s). This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY).

Academia Open
Vol 10 No 1 (2025): June (In Progress)
DOI: 10.21070/acopen.9.2024.10451 . Article type: (Computer Science)

Cluster Analysis

Cluster analysis, also known as clustering, is an unsuper- vised machine learning technique for identifying and group- ing related data points in large datasets without concern for the specific outcome. It does grouping a collection of objects in such a way that objects in the same category, called a cluster, are in some sense more similar to each other than objects in other groups [41]. It is often used as a data analy- sis technique to discover interesting trends or patterns in data, e.g., groups of consumers based on their behavior. In a broad range of application areas, such as cybersecurity, e-commerce, mobile data processing, health analytics, user modeling and behavioral analytics, clustering can be used. In the following, we briefly discuss and summarize various types of clustering methods.

1. Partitioning methods: Based on the features and simi- larities in the data, this clustering approach categorizes the data into multiple groups or clusters. The data sci- entists or analysts typically determine the number of clusters either dynamically or statically depending on the nature of the target applications, to produce for the methods of clustering. The most common clustering algorithms based on partitioning methods are K-means [69], K-Mediods [80], CLARA [55] etc.

2. Density-based methods: To identify distinct groups or clusters, it uses the concept that a cluster in the data space is a contiguous region of high point density iso- lated from other such clusters by contiguous regions of low point density. Points that are not part of a cluster are considered as noise. The typical clustering algorithms based on density are DBSCAN [32], OPTICS [12] etc. The density-based methods typically struggle with clus- ters of similar density and high dimensionality data.

3. Hierarchical-based methods: Hierarchical clustering typically seeks to construct a hierarchy of clusters, i.e., the tree structure. Strategies for hierarchical cluster- ing generally fall into two types: (i) Agglomerative—a “bottom-up” approach in which each observation begins in its cluster and pairs of clusters are combined as one, moves up the hierarchy, and (ii) Divisive—a “top-down” approach in which all observations begin in one cluster and splits are performed recursively, moves down the hierarchy, as shown in Fig 7. Our earlier proposed BOTS technique, Sarker et al. [102] is an example of a hierar- chical, particularly, bottom-up clustering algorithm.

4. Grid-based methods: To deal with massive datasets, grid- based clustering is especially suitable. To obtain clusters, the principle is first to summarize the dataset with a grid representation and then to combine grid cells. STING [122], CLIQUE [6], etc. are the standard algorithms of grid-based clustering.

5. Model-based methods: There are mainly two types of model-based clustering algorithms: one that uses statistical learning, and the other based on a method of neural network learning [130]. For instance, GMM [89] is an example of a statistical learning method, and SOM [22] [96] is an example of a neural network learning method.

6. Constraint-based methods: Constrained-based clustering is a semi-supervised approach to data clustering that uses

ISSN 2714-7444 (online), https://acopen.umsida.ac.id, published by Universitas Muhammadiyah Sidoarjo
Copyright © Author(s). This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY).

Academia Open
Vol 10 No 1 (2025): June (In Progress)
DOI: 10.21070/acopen.9.2024.10451 . Article type: (Computer Science)

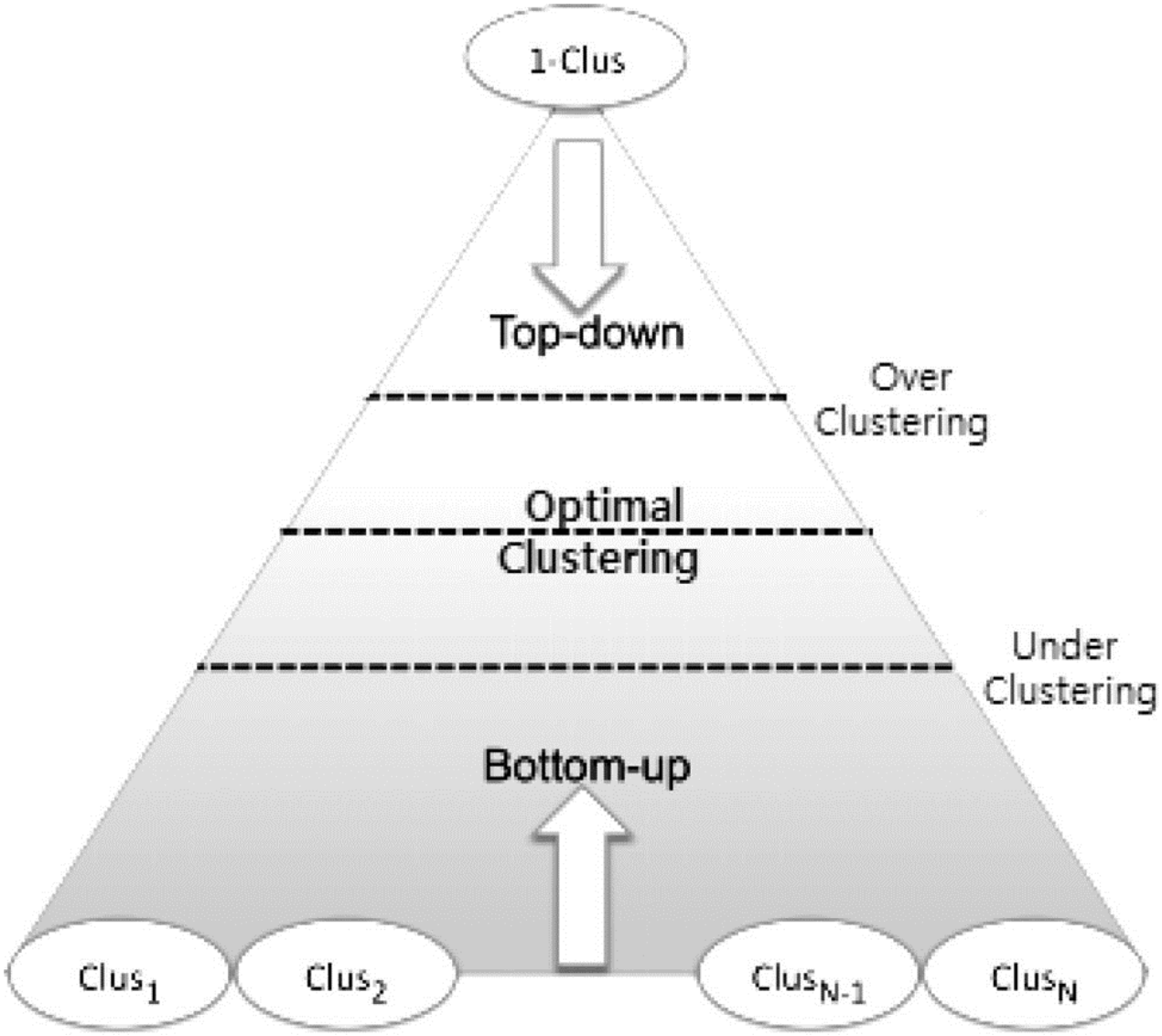


Figure 12. *A graphical interpretation of the widely-used hierarchical clus- tering (Bottom-up and top-down) technique*

constraints to incorporate domain knowledge. Applica- tion or user-oriented constraints are incorporated to perform the clustering. The typical algorithms of this kind of clustering are COP K-means [121], CMWK-Means [27], etc.

Many clustering algorithms have been proposed with the ability to grouping data in machine learning and data science literature [41, 125]. In the following, we summarize the popular methods that are used widely in various application areas.

1. K-means clustering: K-means clustering [69] is a fast, robust, and simple algorithm that provides reliable results when data sets are well-separated from each other. The data points are allocated to a cluster in this algorithm in such a way that the amount of the squared distance between the data points and the centroid is as small as possible. In other words, the K-means algorithm identi- fies the k number of centroids and then assigns each data point to the nearest cluster while keeping the centroids as small as possible. Since it begins with a random selection of cluster centers, the results can be inconsistent. Since extreme values can easily affect a mean, the K-means clustering algorithm is sensitive to outliers. K-medoids clustering [91] is a variant of K-means that is more robust to noises and outliers.

2. Mean-shift clustering: Mean-shift clustering [37] is a nonparametric clustering technique that does not require prior knowledge of the number of clusters or constraints on cluster shape. Mean-shift clustering aims to discover “blobs” in a smooth distribution or density of samples [82]. It is a centroid-based algo- rithm that works by updating centroid candidates to be the mean of the points in a given region. To form the final set of centroids, these candidates are filtered in a post-processing stage to remove near-duplicates. Clus- ter analysis in computer vision

ISSN 2714-7444 (online), https://acopen.umsida.ac.id, published by Universitas Muhammadiyah Sidoarjo
Copyright © Author(s). This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY).

and image processing are examples of application domains. Mean Shift has the disadvantage of being computationally expensive. Moreover, in cases of high dimension, where the num- ber of clusters shifts abruptly, the mean-shift algorithm does not work well.

3. DBSCAN: Density-based spatial clustering of applica- tions with noise (DBSCAN) [32] is a base algorithm for density-based clustering which is widely used in data mining and machine learning. This is known as a non-parametric density-based clustering technique for sepa- rating high-density clusters from low-density clusters that are used in model building. DBSCAN's main idea is that a point belongs to a cluster if it is close to many points from that cluster. It can find clusters of various shapes and sizes in a vast volume of data that is noisy and contains outliers. DBSCAN, unlike k-means, does not require a priori specification of the number of clus- ters in the data and can find arbitrarily shaped clusters. Although k-means is much faster than DBSCAN, it is efficient at finding high-density regions and outliers, i.e., is robust to outliers.

4. GMM clustering: Gaussian mixture models (GMMs) are often used for data clustering, which is a distribution-based clustering algorithm. A Gaussian mixture model is a probabilistic model in which all the data points are produced by a mixture of a finite number of Gaussian distributions with unknown parameters [82]. To find the Gaussian parameters for each cluster, an optimization algorithm called expectation-maximization (EM) [82] can be used. EM is an iterative method that uses a sta- tistical model to estimate the parameters. In contrast to k-means, Gaussian mixture models account for uncer- tainty and return the likelihood that a data point belongs to one of the k clusters. GMM clustering is more robust than k-means and works well even with non-linear data distributions.

5. Agglomerative hierarchical clustering: The most com- mon method of hierarchical clustering used to group objects in clusters based on their similarity is agglom- erative clustering. This technique uses a bottom-up approach, where each object is first treated as a singleton cluster by the algorithm. Following that, pairs of clus- ters are merged one by one until all clusters have been merged into a single large cluster containing all objects. The result is a dendrogram, which is a tree-based repre- sentation of the elements. Single linkage [115], Complete linkage [116], BOTS [102] etc. are some examples of

such techniques. The main advantage of agglomerative hierarchical clustering over k-means is that the tree-structure hierarchy generated by agglomerative clustering is more informative than the unstructured collection of flat clusters returned by k-means, which can help to make better decisions in the relevant application areas.

Dimensionality Reduction and Feature Learning

In machine learning and data science, high-dimensional data processing is a challenging task for both researchers and application developers. Thus, dimensionality reduc- tion which is an unsupervised learning technique, is important because it leads to better human interpretations, lower computational costs, and avoids overfitting and redundancy by simplifying models. Both the process of feature selec- tion and feature extraction can be used for dimensionality reduction. The primary distinction between the selection and extraction of features is that the "feature selection" keeps a subset of the original features [97], while "feature extrac- tion" creates brand new ones [98]. In the following, we briefly discuss these techniques.

1. Feature selection: The selection of features, also known as the selection of variables or attributes in the data, is the process of choosing a subset of unique features (vari- ables, predictors) to use in building machine learning and data science model. It decreases a model's complexity by eliminating the irrelevant or less important features and allows for faster training of machine learning algorithms. A right and optimal subset of the selected features in a problem domain is capable to minimize the overfitting problem through simplifying and generalizing the model as well as increases the model's accuracy [97]. Thus, "feature selection" [66, 99] is considered as one of the primary concepts in machine learning that greatly affects the effectiveness and efficiency of the target machine learning model. Chi-squared test, Analysis of variance (ANOVA) test, Pearson's correlation coefficient, recur- sive feature elimination, are some popular techniques that can be used for feature selection.

2. Feature extraction: In a machine learning-based model or system, feature extraction techniques usually provide a better understanding of the data, a way to improve pre- diction accuracy, and to reduce computational cost or training time. The aim of "feature extraction" [66, 99] is to reduce the number of features in a dataset by generating new ones from the existing ones and then discarding the original features. The majority of the information found in the original set of features can then be summa- rized using this new reduced set of features. For instance, principal components analysis (PCA) is often used as a dimensionality-reduction technique to extract a lower dimensional space creating new brand components from the existing features in a dataset [98].

Many algorithms have been proposed to reduce data dimen- sions in the machine learning and data science literature [41, 125]. In the following, we summarize the popular methods that are used widely in various application areas.

1. Variance threshold: A simple basic approach to feature selection is the variance threshold [82]. This excludes all features of low variance, i.e., all features whose variance does not exceed the threshold. It eliminates all zero-variance characteristics by default, i.e., characteristics that have the same value in all samples. This feature selection algorithm looks only at the (X) features, not the (y) outputs needed, and can, therefore, be used for

ISSN 2714-7444 (online), https://acopen.umsida.ac.id, published by Universitas Muhammadiyah Sidoarjo
Copyright © Author(s). This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY).

Academia Open
Vol 10 No 1 (2025): June (In Progress)
DOI: 10.21070/acopen.9.2024.10451 . Article type: (Computer Science)

unsupervised learning

2. Pearson correlation: Pearson's correlation is another method to understand a feature's relation to the response variable and can be used for feature selection [99]. This method is also used for finding the association between the features in a dataset. The resulting value is [−1, 1], where −1 means perfect negative correlation, +1 means perfect positive correlation, and 0 means that the two variables do not have a linear correlation. If two random variables represent X and Y, then the correlation coef- ficient between X and Y is defined as [41]

$$r(X,Y) = \frac{\sum_{i=1}^{n}(X-\bar{X})(Y-\bar{Y})}{\sqrt{\sum_{i=1}^{n}(X_i-\bar{X})^2}\sqrt{\sum_{i=1}^{n}(Y_i-\bar{Y})^2}} \quad (8)$$

Figure 13.

3. ANOVA: Analysis of variance (ANOVA) is a statisti- cal tool used to verify the mean values of two or more groups that differ significantly from each other. ANOVA assumes a linear relationship between the variables and the target and the variables' normal distribution. To sta- tistically test the equality of means, the ANOVA method utilizes *F* tests. For feature selection, the results 'ANOVA *F* value' [] of this test can be used where certain fea- tures independent of the goal variable can be omitted

4. Chi square: The chi-square 32 [82] statistic is an estimate of the difference between the effects of a series of events or variables observed and expected frequencies. The magnitude of the difference between the real and observed values, the degrees of freedom, and the sample size depends on 32. The chi-square 32 is commonly used for testing relationships between categorical variables. If Oi represents observed value and Ei represents expected value, then

$$\chi^2 = \sum_{i=1}^{n}\frac{(O-E)^2}{E_i} \quad (9)$$

Figure 14.

5. Recursive feature elimination (RFE): Recursive Feature Elimination (RFE) is a brute force approach to feature selection. RFE [82] fits the model and removes the weakest feature before it meets the specified number of features. Features are ranked by the coefficients or feature significance of the model. RFE aims to remove dependencies and collinearity in the model by recursively removing a small number of features per iteration.

6. Model-based selection: To reduce the dimensionality of the data, linear models penalized with the L1 regularization can be used. Least absolute shrinkage and selection operator (Lasso) regression is a type of linear regression that has the property of shrinking some of the coefficients to zero [82]. Therefore, that feature can be removed from the model. Thus, the penalized lasso regression method, often used in machine learning to select the subset of variables. Extra Trees Classifier [82] is an example of a tree-based estimator that can be used to compute impu- rity-based function importance, which can then be used to discard irrelevant features.

7. Principal component analysis (PCA): Principal compo- nent analysis (PCA) is a well-known unsupervised learn ing approach in the field of machine learning and data science. PCA is a mathematical technique that transforms a set of correlated variables into a set of uncorrelated variables known as principal components [48, 81]. Fig- ure 8 shows an example of the effect of PCA on various dimensions space, where Fig. 8a shows the original fea- tures in 3D space, and Fig. 8b shows the created prin- cipal components PC1 and PC2 onto a 2D plane, and 1D line with the principal component PC1 respectively. Thus, PCA can be used as a feature extraction technique that reduces the dimensionality of the datasets, and to build an effective machine learning model [98]. Techni- cally, PCA identifies the completely transformed with the highest eigenvalues of a covariance matrix and then uses those to project the data into a new subspace of equal or fewer dimensions [82]

ISSN 2714-7444 (online), https://acopen.umsida.ac.id, published by Universitas Muhammadiyah Sidoarjo
Copyright © Author(s). This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY).

Academia Open
Vol 10 No 1 (2025): June (In Progress)
DOI: 10.21070/acopen.9.2024.10451 . Article type: (Computer Science)



(a) An example of the original features in a 3D space. (b) Principal components in 2D and 1D space.

Figure 15. *An example of a principal component analysis (PCA) and created principal components PC1 and PC2 in different dimension space*

Association Rule Learning

Association rule learning is a rule-based machine learn- ing approach to discover interesting relationships, "IF-THEN" statements, in large datasets between variables [7]. One example is that "if a customer buys a computer or laptop (an item), s/he is likely to also buy anti-virus soft- ware (another item) at the same time". Association rules are employed today in many application areas, including IoT services, medical diagnosis, usage behavior analytics, web usage mining, smartphone applications, cybersecurity applications, and bioinformatics. In comparison to sequence mining, association rule learning does not usually take into account the order of things within or across transactions. A common way of measuring the usefulness of association rules is to use its parameter, the 'support' and 'confidence', which is introduced in [7].

In the data mining literature, many association rule learn- ing methods have been proposed, such as logic dependent [34], frequent pattern based [8, 49, 68], and tree-based [42]. The most popular association rule learning algorithms are summarized below.

1. AIS and SETM: AIS is the first algorithm proposed by Agrawal et al. [7] for association rule mining. The AIS algorithm's main downside is that too many candidate itemsets are generated, requiring more space and wasting a lot of effort. This algorithm calls for too many passes over the entire dataset to produce the rules. Another approach SETM [49] exhibits good performance and stable behavior with execution time; however, it suffers from the same flaw as the AIS algorithm.

2. Apriori: For generating association rules for a given data- set, Agrawal et al. [8] proposed the Apriori, Apriori-TID, and Apriori-Hybrid algorithms. These later algorithms outperform the AIS and SETM mentioned above due to the Apriori property of frequent itemset [8]. The term 'Apriori' usually refers to having prior knowledge of frequent itemset properties. Apriori uses a "bottom-up" approach, where it generates the candidate itemsets. To reduce the search space, Apriori uses the property "all subsets of a frequent itemset must be frequent; and if an itemset is infrequent, then all its supersets must also be infrequent". Another approach predictive Apriori [108] can also generate rules; however, it receives unexpected results as it combines both the support and confidence. The Apriori [8] is the widely applicable techniques in mining association rules.

3. ECLAT: This technique was proposed by Zaki et al. [131] and stands for Equivalence Class Clustering and bottom- up Lattice Traversal. ECLAT uses a depth-first search to find frequent itemsets. In contrast to the Apriori [8] algo- rithm, which represents data in a horizontal pattern, it represents data vertically. Hence, the ECLAT algorithm is more efficient and scalable in the area of association rule learning. This algorithm is better suited for small and medium datasets whereas the Apriori algorithm is used for large datasets.

4. FP-Growth: Another common association rule learning technique based on the frequent-pattern tree (FP-tree) proposed by Han et al. [42] is Frequent Pattern Growth, known as FP-Growth. The key difference with Apriori is that while generating rules, the Apriori algorithm [8] generates frequent candidate itemsets; on the other hand, the FP-growth algorithm [42] prevents candidate genera- tion and thus produces a tree by the successful strategy of 'divide and conquer' approach. Due to its sophistication, however, FP-Tree is challenging to use in an interactive mining environment [133]. Thus, the FP-Tree would not fit into memory for massive data sets, making it challenging to process big data as well. Another solution is RARM (Rapid Association Rule Mining) proposed by Das et al. [26] but faces a related FP-tree issue [133].

5. ABC-RuleMiner: A rule-based machine learning method, recently proposed in our earlier paper, by Sarker et al. [104], to discover the interesting non-redundant rules to provide real-world intelligent services. This algorithm effectively identifies the redundancy in associations by taking into account the impact or precedence of the related contextual features and discovers a set of non- redundant association rules. This algorithm first con- structs an association generation tree (AGT), a top-down approach, and then extracts the association rules through traversing

ISSN 2714-7444 (online), https://acopen.umsida.ac.id, published by Universitas Muhammadiyah Sidoarjo
Copyright © Author(s). This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY).

the tree. Thus, ABC-RuleMiner is more potent than traditional rule-based methods in terms of both non-redundant rule generation and intelligent decision- making, particularly in a context-aware smart comput- ing environment, where human or user preferences are involved.

Among the association rule learning techniques discussed above, Apriori [8] is the most widely used algorithm for discovering association rules from a given dataset [133]. The main strength of the association learning technique is its comprehensiveness, as it generates all associations that satisfy the user-specified constraints, such as minimum sup- port and confidence value. The ABC-RuleMiner approach [104] discussed earlier could give significant results in terms of non-redundant rule generation and intelligent decision- making for the relevant application areas in the real world.

Reinforcement Learning

Reinforcement learning (RL) is a machine learning tech- nique that allows an agent to learn by trial and error in an interactive environment using input from its actions and experiences. Unlike supervised learning, which is based on given sample data or examples, the RL method is based on interacting with the environment. The problem to be solved in reinforcement learning (RL) is defined as a Markov Deci- sion Process (MDP) [86], i.e., all about sequentially making decisions. An RL problem typically includes four elements such as Agent, Environment, Rewards, and Policy.

RL can be split roughly into Model-based and Model- free techniques. Model-based RL is the process of infer- ring optimal behavior from a model of the environment by performing actions and observing the results, which include the next state and the immediate reward [85]. Alp- haZero, AlphaGo [113] are examples of the model-based approaches. On the other hand, a model-free approach does not use the distribution of the transition probability and the reward function associated with MDP. Q-learn- ing, Deep Q Network, Monte Carlo Control, SARSA (State–Action–Reward–State–Action), etc. are some exam- ples of model-free algorithms [52]. The policy network, which is required for model-based RL but not for model- free, is the key difference between model-free and model-based learning. In the following, we discuss the popular RL algorithms.

1. Monte Carlo methods: Monte Carlo techniques, or Monte Carlo experiments, are a wide category of computational algorithms that rely on repeated random sampling to obtain numerical results [52]. The underlying concept is to use randomness to solve problems that are determinis- tic in principle. Optimization, numerical integration, and making drawings from the probability distribution are the three problem classes where Monte Carlo techniques are most commonly used.

2. Q-learning: Q-learning is a model-free reinforcement learning algorithm for learning the quality of behaviors that tell an agent what action to take under what condi- tions [52]. It does not need a model of the environment (hence the term “model-free”), and it can deal with sto- chastic transitions and rewards without the need for adap- tations. The ‘Q’ in Q-learning usually stands for qual- ity, as the algorithm calculates the maximum expected rewards for a given behavior in a given state.

3. Deep Q-learning: The basic working step in Deep Q-Learning [52] is that the initial state is fed into the neural network, which returns the Q-value of all pos- sible actions as an output. Still, when we have a reason- ably simple setting to overcome, Q-learning works well. However, when the number of states and actions becomes more complicated, deep learning can be used as a func- tion approximator.

Reinforcement learning, along with supervised and unsu- pervised learning, is one of the basic machine learning para- digms. RL can be used to solve numerous real-world prob- lems in various fields, such as game theory, control theory, operations analysis, information theory, simulation-based optimization, manufacturing, supply chain logistics, multi- agent systems, swarm intelligence, aircraft control, robot motion control, and many more.

Artificial Neural Network and Deep Learning

Deep learning is part of a wider family of artificial neural networks (ANN)-based machine learning approaches with representation learning. Deep learning provides a computa- tional architecture by combining several processing layers, such as input, hidden, and output layers, to learn from data [41]. The main advantage of deep learning over traditional machine learning methods is its better performance in sev- eral cases, particularly learning from large datasets [105, 129]. Figure 9 shows a general performance of deep learning over machine learning considering the increasing amount of data. However, it may vary depending on the data character- istics and experimental set up.

The most common deep learning algorithms are: Multi- layer Perceptron (MLP), Convolutional Neural Network (CNN, or ConvNet), Long Short-Term Memory Recurrent Neural Network (LSTM-RNN) [96]. In the following, we discuss various types of deep learning methods that can be used to build effective data-driven models for various purposes.

ISSN 2714-7444 (online), https://acopen.umsida.ac.id, published by Universitas Muhammadiyah Sidoarjo
Copyright © Author(s). This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY).

Academia Open
Vol 10 No 1 (2025): June (In Progress)
DOI: 10.21070/acopen.9.2024.10451 . Article type: (Computer Science)



Figure 16. *Machine learning and deep learning performance in general with the amount of data*

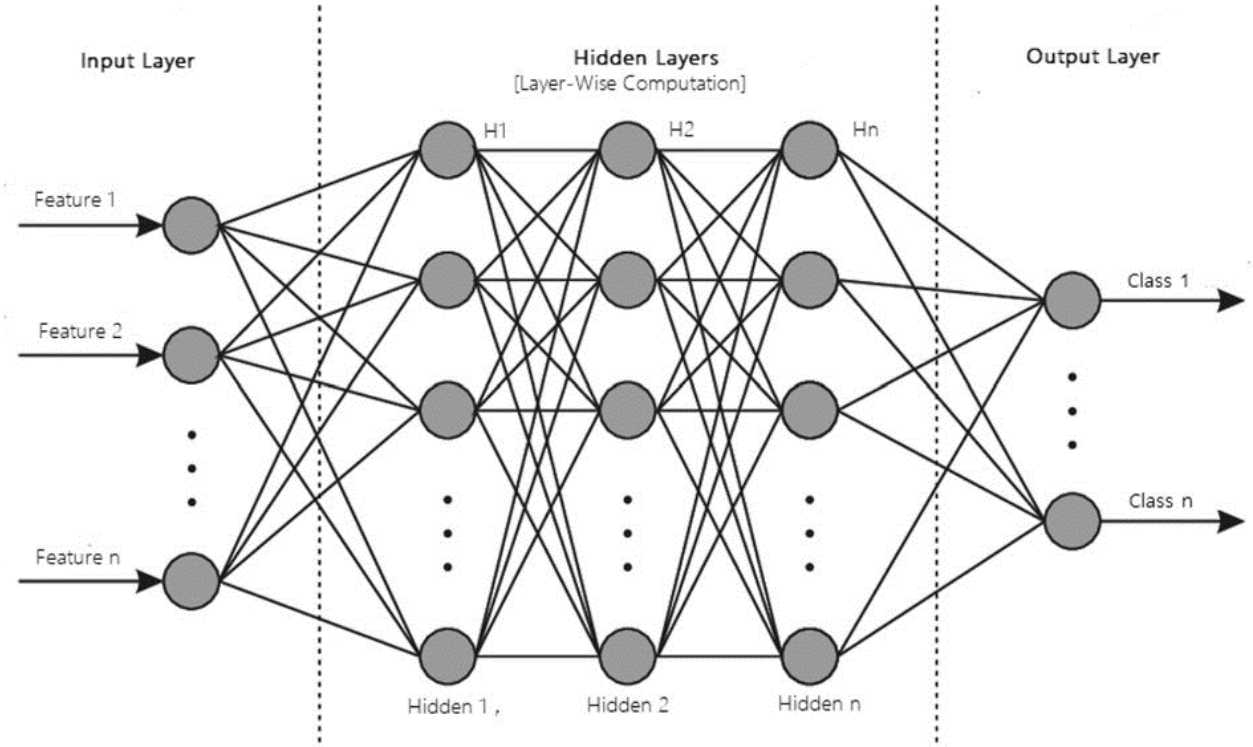


Figure 17. *A structure of an artificial neural network modeling with multiple processing layers*

ISSN 2714-7444 (online), https://acopen.umsida.ac.id, published by Universitas Muhammadiyah Sidoarjo
Copyright © Author(s). This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY).

Academia Open
Vol 10 No 1 (2025): June (In Progress)
DOI: 10.21070/acopen.9.2024.10451 . Article type: (Computer Science)

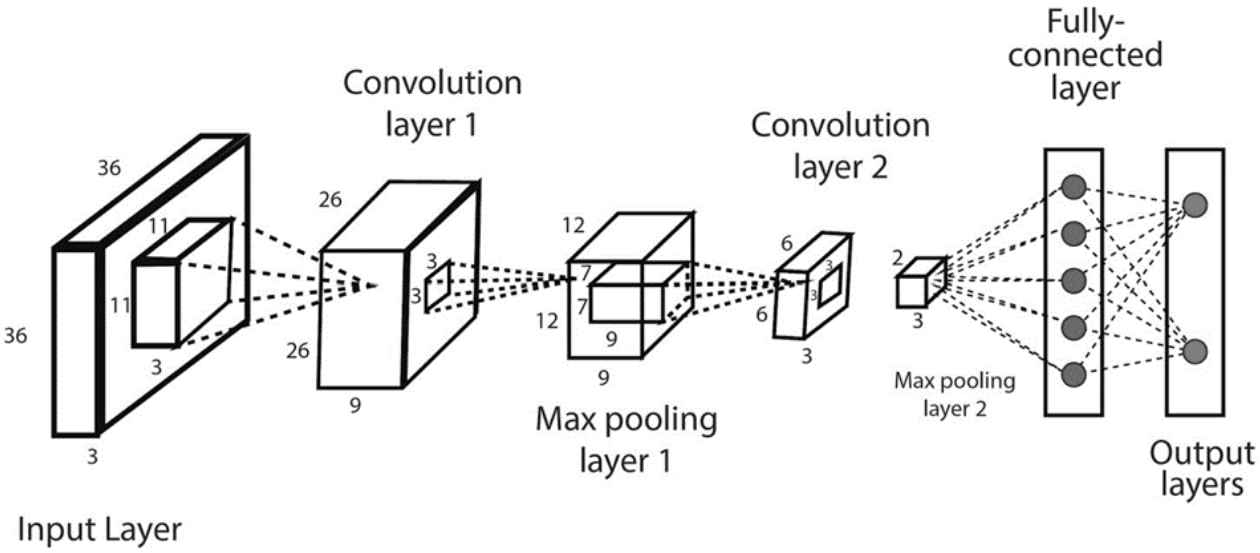


Figure 18. *An example of a convolutional neural network (CNN or Con- vNet) including multiple convolution and pooling layers*

- MLP: The base architecture of deep learning, which is also known as the feed-forward artificial neural network, is called a multilayer perceptron (MLP) [82]. A typical MLP is a fully connected network consisting of an input layer, one or more hidden layers, and an output layer, as shown in Fig. 10. Each node in one layer connects to each node in the following layer at a certain weight. MLP utilizes the "Backpropagation" technique [41], the most "fundamental building block" in a neural network, to adjust the weight values internally while building the model. MLP is sensitive to scaling features and allows a variety of hyperparameters to be tuned, such as the number of hidden layers, neurons, and iterations, which can result in a computationally costly model.

- CNN or ConvNet: The convolution neural network (CNN) [65] enhances the design of the standard ANN, consisting of convolutional layers, pooling layers, as well as fully connected layers, as shown in Fig. 11. As it takes the advantage of the two-dimensional (2D) structure of the input data, it is typically broadly used in several areas such as image and video recognition, image processing and classification, medical image analysis, natural language processing, etc. While CNN has a greater computational burden, without any manual intervention, it has the advantage of automati- cally detecting the important features, and hence CNN is considered to be more powerful than conventional ANN. A number of advanced deep learning models based on CNN can be used in the field, such as AlexNet [60], Xception [24], Inception [118], Visual Geometry Group (VGG) [44], ResNet [45], etc.

- LSTM-RNN: Long short-term memory (LSTM) is an artificial recurrent neural network (RNN) architec- ture used in the area of deep learning [38]. LSTM has feedback links, unlike normal feed-forward neural networks. LSTM networks are well-suited for analyz- ing and learning sequential data, such as classifying, processing, and predicting data based on time series data, which differentiates it from other conventional networks. Thus, LSTM can be used when the data are in a sequential format, such as time, sentence, etc., and commonly applied in the area of time-series analysis, natural language processing, speech recognition, etc.

In addition to these most common deep learning methods discussed above, several other deep learning approaches [96] exist in the area for various purposes. For instance, the self-organizing map (SOM) [58] uses unsu- pervised learning to represent the high-dimensional data by a 2D grid map, thus achieving dimensionality reduc- tion. The autoencoder (AE) [15] is another learning tech- nique that is widely used for dimensionality reduction as well and feature extraction in unsupervised learning tasks. Restricted Boltzmann machines (RBM) [46] can be used for dimensionality reduction, classification, regression, collaborative filtering, feature learning, and topic mod- eling. A deep belief network (DBN) is typically composed of simple, unsupervised networks such as restricted Boltz- mann machines (RBMs) or autoencoders, and a backprop- agation neural network (BPNN) [123]. A generative adver- sarial network (GAN) [39] is a form of the network for deep learning that can generate data with characteristics close to the actual data input. Transfer learning is currently very common because it can train deep neural networks with comparatively low data, which is typically the re-use of a new problem with a pre-trained model [124]. A brief discussion of these artificial neural networks (ANN) and deep learning (DL) models are summarized in our earlier paper Sarker et al. [96].

Overall, based on the learning techniques discussed above, we can conclude that various types of machine learning techniques, such as classification analysis, regres- sion, data clustering, feature selection and extraction, and dimensionality reduction, association rule learning, rein- forcement learning, or deep learning techniques, can play a significant role for various purposes according to their capabilities. In the following section, we discuss several application areas based on machine learning algorithms.

ISSN 2714-7444 (online), https://acopen.umsida.ac.id, published by Universitas Muhammadiyah Sidoarjo
Copyright © Author(s). This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY).

Academia Open
Vol 10 No 1 (2025): June (In Progress)
DOI: 10.21070/acopen.9.2024.10451 . Article type: (Computer Science)

Applications of Machine Learning

In the current age of the Fourth Industrial Revolution (4IR), machine learning becomes popular in various application areas, because of its learning capabilities from the past and making intelligent decisions. In the following, we summarize and discuss ten popular application areas of machine learning technology.

To detect various types of cyber-attacks or intrusions machine learning classification models by taking into account the impact of security features are useful [97]. Various deep learning-based security models can also be used on the large scale of security datasets [96, 129]. Moreover, security policy rules generated by associa- tion rule learning techniques can play a significant role to build a rule-based security system [105]. Thus, we can say that various learning techniques discussed in Sect. Machine Learning Tasks and Algorithms, can enable cybersecurity professionals to be more proac- tive inefficiently preventing threats and cyber-attacks.

- Predictive analytics and intelligent decision-making: A major application field of machine learning is intelli- gent decision-making by data-driven predictive analyt- ics [21, 70]. The basis of predictive analytics is captur- ing and exploiting relationships between explanatory variables and predicted variables from previous events to predict the unknown outcome [41]. For instance, identifying suspects or criminals after a crime has been committed, or detecting credit card fraud as it happens. Another application, where machine learn- ing algorithms can assist retailers in better understand- ing consumer preferences and behavior, better manage inventory, avoiding out-of-stock situations, and opti- mizing logistics and warehousing in e-commerce. Vari- ous machine learning algorithms such as decision trees, support vector machines, artificial neural networks, etc. [106, 125] are commonly used in the area. Since accu- rate predictions provide insight into the unknown, they can improve the decisions of industries, businesses, and almost any organization, including government agen- cies, e-commerce, telecommunications, banking and financial services, healthcare, sales and marketing, transportation, social networking, and many others.

- Cybersecurity and threat intelligence: Cybersecurity is one of the most essential areas of Industry 4.0. [114], which is typically the practice of protecting networks, systems, hardware, and data from digital attacks [114]. Machine learning has become a crucial cybersecurity technology that constantly learns by analyzing data to identify patterns, better detect malware in encrypted traffic, find insider threats, predict where bad neighborhoods are online, keep people safe while browsing, or secure data in the cloud by uncovering suspicious activity. For instance, clustering techniques can be used to identify cyber-anomalies, policy violations, etc. To detect various types of cyber-attacks or intrusions machine learning classification models by taking into account the impact of security features are useful []. Various deep learning-based security models can also be used on the large scale of security datasets [,]. Moreover, security policy rules generated by associa- tion rule learning techniques can play a significant role to build a rule-based security system []. Thus, we can say that various learning techniques discussed in Sect. , can enable cybersecurity professionals to be more proac- tive inefficiently preventing threats and cyber-attacks

- Internet of things (IoT) and smart cities: Internet of Things (IoT) is another essential area of Industry 4.0. [114], which turns everyday objects into smart objects by allowing them to transmit data and automate tasks without the need for human interaction. IoT is, therefore, considered to be the big frontier that can enhance almost all activities in our lives, such as smart governance, smart home, education, communication, transportation, retail, agriculture, health care, business, and many more [70]. Smart city is one of IoT's core fields of application, using technologies to enhance city services and residents' liv- ing experiences [132, 135]. As machine learning utilizes experience to recognize trends and create models that help predict future behavior and events, it has become a crucial technology for IoT applications [103]. For exam- ple, to predict traffic in smart cities, parking availabil- ity prediction, estimate the total usage of energy of the citizens for a particular period, make context-aware and timely decisions for the people, etc. are some tasks that can be solved using machine learning techniques accord- ing to the current needs of the people.

- Traffic prediction and transportation: Transportation systems have become a crucial component of every country's economic development. Nonetheless, several cities around the world are experiencing an excessive rise in traffic volume, resulting in serious issues such as delays, traffic congestion, higher fuel prices, increased CO2 pollution, accidents, emergencies, and a decline in modern society's quality of life [40]. Thus, an intelligent transportation system through predicting future traffic is important, which is an indispensable part of a smart city. Accurate traffic prediction based on machine and deep learning modeling can help to minimize the issues [17, 30, 31]. For example, based on the travel history and trend of traveling through various routes, machine learning can assist transportation companies in predict- ing possible issues that may occur on specific routes and recommending their customers to take a different path. Ultimately, these learning-based data-driven models help improve traffic flow, increase the usage and efficiency of sustainable modes of transportation, and limit real-world disruption by modeling and visualizing future changes.

- Healthcare and COVID-19 pandemic: Machine learning can help to solve diagnostic and prognostic problems in a variety of medical domains, such as disease prediction, medical knowledge extraction, detecting regularities in data, patient management, etc. [33, 77, 112]. Coronavirus disease (COVID-19) is an infectious disease caused by a newly discovered coronavirus, according to the World Health Organization (WHO) [3]. Recently, the learning techniques have become popular in the battle against COVID-19 [61, 63]. For the COVID-19 pandemic, the learning

ISSN 2714-7444 (online), https://acopen.umsida.ac.id, published by Universitas Muhammadiyah Sidoarjo
Copyright © Author(s). This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY).

techniques are used to classify patients at high risk, their mortality rate, and other anomalies [61]. It can also be used to better understand the virus's ori- gin, COVID-19 outbreak prediction, as well as for dis- ease diagnosis and treatment [14, 50]. With the help of machine learning, researchers can forecast where and when, the COVID-19 is likely to spread, and notify those regions to match the required arrangements. Deep learn- ing also provides exciting solutions to the problems of medical image processing and is seen as a crucial tech- nique for potential applications, particularly for COVID- 19 pandemic [10, 78, 111]. Overall, machine and deep learning techniques can help to fight the COVID-19 virus and the pandemic as well as intelligent clinical decisions making in the domain of healthcare.

- E-commerce and product recommendations: Product rec- ommendation is one of the most well known and widely used applications of machine learning, and it is one of the most prominent features of almost any e-commerce website today. Machine learning technology can assist businesses in analyzing their consumers' purchasing his- tories and making customized product suggestions for their next purchase based on their behavior and prefer- ences. E-commerce companies, for example, can easily position product suggestions and offers by analyzing browsing trends and click-through rates of specific items. Using predictive modeling based on machine learning techniques, many online retailers, such as Amazon [71], can better manage inventory, prevent out-of-stock situa- tions, and optimize logistics and warehousing. The future of sales and marketing is the ability to capture, evaluate, and use consumer data to provide a customized shopping experience. Furthermore, machine learning techniques enable companies to create packages and content that are tailored to the needs of their customers, allowing them to maintain existing customers while attracting new ones.

- NLP and sentiment analysis: Natural language pro- cessing (NLP) involves the reading and understanding of spoken or written language through the medium of a computer [79, 103]. Thus, NLP helps computers, for instance, to read a text, hear speech, interpret it, analyze sentiment, and decide which aspects are significant, where machine learning techniques can be used. Virtual personal assistant, chatbot, speech recognition, document description, language or machine translation, etc. are some examples of NLP-related tasks. Sentiment Analy- sis [90] (also referred to as opinion mining or emotion AI) is an NLP sub-field that seeks to identify and extract public mood and views within a given text through blogs, reviews, social media, forums, news, etc. For instance, businesses and brands use sentiment analysis to under- stand the social sentiment of their brand, product, or service through social media platforms or the web as a whole. Overall, sentiment analysis is considered as a machine learning task that analyzes texts for polarity, such as "positive", "negative", or "neutral" along with more intense emotions like very happy, happy, sad, very sad, angry, have interest, or not interested etc.

- Image, speech and pattern recognition: Image recogni- tion [36] is a well-known and widespread example of machine learning in the real world, which can identify an object as a digital image. For instance, to label an x-ray as cancerous or not, character recognition, or face detec- tion in an image, tagging suggestions on social media, e.g., Facebook, are common examples of image recog- nition. Speech recognition [23] is also very popular that typically uses sound and linguistic models, e.g., Google Assistant, Cortana, Siri, Alexa, etc. [67], where machine learning methods are used. Pattern recognition [13] is defined as the automated recognition of patterns and regularities in data, e.g., image analysis. Several machine learning techniques such as classification, feature selec- tion, clustering, or sequence labeling methods are used in the area.

- Sustainable agriculture: Agriculture is essential to the survival of all human activities [109]. Sustainable agri- culture practices help to improve agricultural productiv- ity while also reducing negative impacts on the environ- ment [5, 25, 109]. The sustainable agriculture supply chains are knowledge-intensive and based on informa- tion, skills, technologies, etc., where knowledge transfer encourages farmers to enhance their decisions to adopt sustainable agriculture practices utilizing the increas- ing amount of data captured by emerging technologies, e.g., the Internet of Things (IoT), mobile technologies and devices, etc. [5, 53, 54]. Machine learning can be applied in various phases of sustainable agriculture, such as in the pre-production phase - for the prediction of crop yield, soil properties, irrigation requirements, etc.; in the production phase—for weather prediction, disease detec- tion, weed detection, soil nutrient management, livestock management, etc.; in processing phase—for demand esti- mation, production planning, etc. and in the distribution phase - the inventory management, consumer analysis, etc.

- User behavior analytics and context-aware smartphone applications: Context-awareness is a system's abil- ity to capture knowledge about its surroundings at any moment and modify behaviors accordingly [28, 93]. Context-aware computing uses software and hardware to automatically collect and interpret data for direct responses. The mobile app development environment has been changed greatly with the power of AI, particu- larly, machine learning techniques through their learning capabilities from contextual data [103, 136]. Thus, the developers of mobile apps can rely on machine learning to create smart apps that can understand human behavior, support, and entertain users [107, 137, 140]. To build various personalized data-driven context-aware systems, such as smart interruption management, smart mobile recommendation, context-aware smart searching, deci- sion-making that intelligently assist end mobile phone users in a pervasive computing environment, machine learning techniques are applicable. For example, context- aware association rules can be used to build an intelligent phone call application [104]. Clustering approaches are useful in capturing users' diverse behavioral activities by taking into account data in time series [102]. To predict the future events in various contexts, the classification methods can be used [106, 139]. Thus, various learning techniques discussed in Sect. "Machine Learning Tasks and Algorithms" can help to build context-aware adap- tive and smart applications according to the preferences of the mobile phone users.

ISSN 2714-7444 (online), https://acopen.umsida.ac.id, published by Universitas Muhammadiyah Sidoarjo
Copyright © Author(s). This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY).

In addition to these application areas, machine learning- based models can also apply to several other domains such as bioinformatics, cheminformatics, computer networks, DNA sequence classification, economics and banking, robot- ics, advanced engineering, and many more.

Challenges and Research Directions

Our study on machine learning algorithms for intelligent data analysis and applications opens several research issues in the area. Thus, in this section, we summarize and discuss the challenges faced and the potential research opportunities and future directions.

In general, the effectiveness and the efficiency of a machine learning-based solution depend on the nature and characteristics of the data, and the performance of the learn- ing algorithms. To collect the data in the relevant domain, such as cybersecurity, IoT, healthcare and agriculture dis- cussed in Sect. “Applications of Machine Learning” is not straightforward, although the current cyberspace enables the production of a huge amount of data with very high fre- quency. Thus, collecting useful data for the target machine learning-based applications, e.g., smart city applications, and their management is important to further analysis. Therefore, a more in-depth investigation of data collection methods is needed while working on the real-world data. Moreover, the historical data may contain many ambiguous values, missing values, outliers, and meaningless data. The machine learning algorithms, discussed in Sect “Machine Learning Tasks and Algorithms” highly impact on data qual- ity, and availability for training, and consequently on the resultant model. Thus, to accurately clean and pre-process the diverse data collected from diverse sources is a chal- lenging task. Therefore, effectively modifying or enhance existing pre-processing methods, or proposing new data preparation techniques are required to effectively use the learning algorithms in the associated application domain.

To analyze the data and extract insights, there exist many machine learning algorithms, summarized in Sect. “Machine Learning Tasks and Algorithms”. Thus, selecting a proper learning algorithm that is suitable for the target application is challenging. The reason is that the outcome of different learning algorithms may vary depending on the data charac- teristics [106]. Selecting a wrong learning algorithm would result in producing unexpected outcomes that may lead to loss of effort, as well as the model’s effectiveness and accu- racy. In terms of model building, the techniques discussed in Sect. “Machine Learning Tasks and Algorithms” can directly be used to solve many real-world issues in diverse domains, such as cybersecurity, smart cities and healthcare summarized in Sect. “Applications of Machine Learning”. However, the hybrid learning model, e.g., the ensemble of methods, modifying or enhancement of the existing learning techniques, or designing new learning methods, could be a potential future work in the area.

Thus, the ultimate success of a machine learning-based solution and corresponding applications mainly depends on both the data and the learning algorithms. If the data are bad to learn, such as non-representative, poor-quality, irrelevant features, or insufficient quantity for training, then the machine learning models may become useless or will produce lower accuracy. Therefore, effectively processing the data and handling the diverse learning algorithms are important, for a machine learning-based solution and eventu- ally building intelligent applications.

Conclusion

In this paper, we have conducted a comprehensive overview of machine learning algorithms for intelligent data analysis and applications. According to our goal, we have briefly dis- cussed how various types of machine learning methods can

be used for making solutions to various real-world issues. A successful machine learning model depends on both the data and the performance of the learning algorithms. The sophisticated learning algorithms then need to be trained through the collected real-world data and knowledge related to the target application before the system can assist with intelligent decision-making. We also discussed several popular application areas based on machine learning tech- niques to highlight their applicability in various real-world issues. Finally, we have summarized and discussed the chal- lenges faced and the potential research opportunities and future directions in the area. Therefore, the challenges that are identified create promising research opportunities in the field which must be addressed with effective solutions in various application areas. Overall, we believe that our study on machine learning-based solutions opens up a promising direction and can be used as a reference guide for potential research and applications for both academia and industry professionals as well as for decision-makers, from a techni- cal point of view.

References

1. . Canadian Institute of Cybersecurity, University of New Brunswick, ISCX Dataset. http://www.unb.ca/cic/datasets/index.html/ (Accessed on 20 October 2019).
2. . CIC-DDoS2019 [online]. Available: https://www.unb.ca/cic/datasets/ddos-2019.html/ (Accessed on 28 March 2020).

Copyright © Author(s). This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY).

3. . World Health Organization: WHO. http://www.who.int/.
4. . Google Trends. In https://trends.google.com/trends/, 2019.
5. . Adnan, N., Nordin, S.M., Rahman, I., & Noor, A. The effects of knowledge transfer on farmers' decision-making toward sustainable agriculture practices. World Journal of Science, Technology, and Sustainable Development, 2018.
6. . Agrawal, R., Gehrke, J., Gunopulos, D., & Raghavan, P. Automatic subspace clustering of high-dimensional data for data mining applications. In Proceedings of the 1998 ACM SIGMOD International Conference on Management of Data, 1998; 94–105.
7. . Agrawal, R., Imieliński, T., & Swami, A. Mining association rules between sets of items in large databases. In ACM SIGMOD Record. ACM, 1993; 22: 207–216.
8. . Agrawal, R., Gehrke, J., Gunopulos, D., & Raghavan, P. Fast algorithms for mining association rules. In Proceedings of the International Joint Conference on Very Large Data Bases, Santiago, Chile, 1994; 1215: 487–499.
9. . Aha, D.W., Kibler, D., & Albert, M. Instance-based learning algorithms. Machine Learning, 1991; 6(1): 37–66.
10. . Alakus, T.B., & Turkoglu, I. Comparison of deep learning approaches to predict COVID-19 infection. Chaos, Solitons & Fractals, 2020; 140.
11. . Amit, Y., & Geman, D. Shape quantization and recognition with randomized trees. Neural Computation, 1997; 9(7): 1545–1588.
12. . Ankerst, M., Breunig, M.M., Kriegel, H.-P., & Sander, J. OPTICS: Ordering Points to Identify the Clustering Structure. ACM SIGMOD Record, 1999; 28(2): 49–60.
13. . Anzai, Y. Pattern Recognition and Machine Learning. Elsevier, 2012.
14. . Ardabili, S.F., Mosavi, A., Ghamisi, P., Ferdinand, F., Varkonyi-Koczy, A.R., Reuter, U., Rabczuk, T., Atkinson, P.M. COVID-19 outbreak prediction with machine learning. Algorithms, 2020; 13(10): 249.
15. . Baldi, P. Autoencoders, Unsupervised Learning, and Deep Architectures. In Proceedings of the ICML Workshop on Unsupervised and Transfer Learning, 2012; 37–49.
16. . Balducci, F., Impedovo, D., & Pirlo, G. Machine learning applications on agricultural datasets for smart farm enhancement. Machines, 2018; 6(3): 38.
17. . Boukerche, A., & Wang, J. Machine learning-based traffic prediction models for intelligent transportation systems. Computer Networks, 2020; 181.
18. . Breiman, L. Bagging Predictors. Machine Learning, 1996; 24(2): 123–140.
19. . Breiman, L. Random Forests. Machine Learning, 2001; 45(1): 5–32.
20. . Breiman, L., Friedman, J., Stone, C.J., & Olshen, R.A. Classification and Regression Trees. CRC Press, 1984.
21. . Cao, L. Data Science: A Comprehensive Overview. ACM Computing Surveys (CSUR), 2017; 50(3): 43.
22. . Carpenter, G.A., & Grossberg, S. A Massively Parallel Architecture for a Self-Organizing Neural Pattern Recognition Machine. Computer Vision, Graphics, and Image Processing, 1987; 37(1): 54–115.
23. . Chiu, C.-C., Sainath, T.N., Wu, Y., Prabhavalkar, R., Nguyen, P., Chen, Z., Kannan, A., Weiss, R.J., Rao, K., & Gonina, E., et al. State-of-the-art speech recognition with sequence-to-sequence models. In 2018 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), 2018; pages 4774–4778. IEEE.
24. . Chollet, F. Xception: Deep Learning with Depthwise Separable Convolutions. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2017; pages 1251–1258.
25. . Cobuloglu, H., & Büyüktahtakın, I.E. A Stochastic Multi-Criteria Decision Analysis for Sustainable Biomass Crop Selection. Expert Systems with Applications, 2015; 42(15–16): 6065–68.
26. . Das, A., Ng, W.-K., & Woon, Y.-K. Rapid Association Rule Mining. In Proceedings of the Tenth International Conference on Information and Knowledge Management, pages 474–481. ACM, 2001.
27. . de Amorim, R.C. Constrained Clustering with Minkowski Weighted K-Means. In 2012 IEEE 13th International Symposium on Computational Intelligence and Informatics (CINTI), pages 13–17. IEEE, 2012.
28. . Dey, A.K. Understanding and Using Context. Personal and Ubiquitous Computing, 2001; 5(1): 4–7.
29. . Eagle, N., & Pentland, A.S. Reality Mining: Sensing Complex Social Systems. Personal and Ubiquitous Computing, 2006; 10(4): 255–268.
30. . Essien, A., Petrounias, I., Sampaio, P., & Sampaio, S. Improving Urban Traffic Speed Prediction Using Data Source Fusion and Deep Learning. In 2019 IEEE International Conference on Big Data and Smart Computing (BigComp). IEEE, 2019: 1–8.
31. . Essien A, Petrounias I, Sampaio P, Sampaio S. A deep-learning model for urban traffic flow prediction with traffic events mined from Twitter. In: World Wide Web, 2020: 1–24.
32. . Ester M, Kriegel H-P, Sander J, Xiaowei X, et al. A density-based algorithm for discovering clusters in large spatial databases with noise. KDD. 1996;96:226–31.
33. . Fatima M, Pasha M, et al. Survey of machine learning algorithms for disease diagnostic. J Intell Learn Syst Appl. 2017;9(01):1.
34. . Flach PA, Lachiche N. Confirmation-guided discovery of first-order rules with Tertius. Mach Learn. 2001;42(1–2):61–95.
35. . Freund Y, Schapire RE, et al. Experiments with a new boosting algorithm. In: ICML, Citeseer. 1996;96:148–156.
36. . Fujiyoshi H, Hirakawa T, Yamashita T. Deep learning-based image recognition for autonomous driving. IATSS Res. 2019;43(4):244–52.
37. . Fukunaga K, Hostetler L. The estimation of the gradient of a density function, with applications in pattern

Copyright © Author(s). This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY).

recognition. IEEE Trans Inform Theory. 1975;21(1):32–40.
38. . Goodfellow I, Bengio Y, Courville A, Bengio Y. Deep learning. Cambridge: MIT Press; 2016.
39. . Goodfellow I, Pouget-Abadie J, Mirza M, Xu B, Warde-Farley D, Ozair S, Courville A, Bengio Y. Generative adversarial nets. In: Advances in neural information processing systems. 2014: 2672–2680.
40. . Guerrero-Ibáñez J, Zeadally S, Contreras-Castillo J. Sensor technologies for intelligent transportation systems. Sensors. 2018;18(4):1212.
41. . Han J, Pei J, Kamber M. Data mining: concepts and techniques. Amsterdam: Elsevier; 2011.
42. . Han J, Pei J, Yin Y. Mining frequent patterns without candidate generation. In: ACM SIGMOD Record, ACM. 2000;29:1–12.
43. . Harmon SA, Sanford TH, Sheng X, Turkbey EB, Roth H, Ziyue X, Yang D, Myronenko A, Anderson V, Amalou A, et al. Artificial intelligence for the detection of COVID-19 pneumonia on chest CT using multinational datasets. Nat Commun. 2020;11(1):1–7.
44. . He K, Zhang X, Ren S, Sun J. Spatial pyramid pooling in deep convolutional networks for visual recognition. IEEE Trans Pattern Anal Mach Intell. 2015;37(9):1904–16.
45. . He K, Zhang X, Ren S, Sun J. Deep residual learning for image recognition. In: Proceedings of the IEEE conference on computer vision and pattern recognition. 2016: 770–778.
46. . Hinton GE. A practical guide to training restricted Boltzmann machines. In: Neural Networks: Tricks of the Trade. Springer. 2012;599–619.
47. . Holte RC. Very simple classification rules perform well on most commonly used datasets. Mach Learn. 1993;11(1):63–90.
48. . Hotelling H. Analysis of a complex of statistical variables into principal components. J Educ Psychol. 1933;24(6):417–441.
49. . Houtsma M, Swami A. Set-oriented mining for association rules in relational databases. In: Data Engineering, Proceedings of the Eleventh International Conference on, IEEE. 1995:25–33.
50. . Jamshidi M, Lalbakhsh A, Talla J, Peroutka Z, Hadjilooei F, Lalbakhsh P, Jamshidi M, La Spada L, Mirmozafari M, Dehghani M, et al. Artificial intelligence and COVID-19: Deep learning approaches for diagnosis and treatment. IEEE Access. 2020;8:109581–95.
51. . John GH, Langley P. Estimating continuous distributions in Bayesian classifiers. In: Proceedings of the Eleventh Conference on Uncertainty in Artificial Intelligence, Morgan Kaufmann Publishers Inc. 1995;338–345.
52. . Kaelbling LP, Littman ML, Moore AW. Reinforcement learning: a survey. J Artif Intell Res. 1996;4:237–85.
53. . Kamble SS, Gunasekaran A, Gawankar SA. Sustainable industry 4.0 framework: A systematic literature review identifying current trends and future perspectives. Process Saf Environ Prot. 2018;117:408—25.
54. . Kamble SS, Gunasekaran A, Gawankar SA. Achieving sustainable performance in a data-driven agriculture supply chain: A review for research and applications. Int J Prod Econ. 2020;219:179–94.
55. . Kaufman L, Rousseeuw PJ. Finding groups in data: An introduction to cluster analysis. John Wiley & Sons; 2009.
56. . Keerthi SS, Shevade SK, Bhattacharyya C, Radha Krishna MK. Improvements to Platt's SMO algorithm for SVM classifier design. Neural Comput. 2001;13(3):637–656.
57. . Khadse V, Mahalle PN, Biraris SV. An empirical comparison of supervised machine learning algorithms for Internet of Things data. In: 2018 Fourth International Conference on Computing Communication Control and Automation (ICCUBEA), IEEE. 2018:1–6.
58. . Kohonen T. The self-organizing map. Proc IEEE. 1990;78(9):1464—80.
59. . Koroniotis N, Moustafa N, Sitnikova E, Turnbull B. Towards the development of realistic botnet dataset in the Internet of Things for network forensic analytics: Bot-IoT dataset. Fut Gen Comput Syst. 2019;100:779–96.
60. . Krizhevsky A, Sutskever I, Hinton GE. ImageNet classification with deep convolutional neural networks. In: Advances in neural information processing systems. 2012:1097–1105.
61. . Kushwaha S, Bahl S, Bagha AK, Parmar KS, Javaid M, Haleem A, Singh RP. Significant applications of machine learning for COVID-19 pandemic. J Ind Integr Manag. 2020;5(4):413–432.
62. . Lade P, Ghosh R, Srinivasan S. Manufacturing analytics and Industrial Internet of Things. IEEE Intell Syst. 2017;32(3):74–89.
63. . Lalmuanawma S, Hussain J, Chhakchhuak L. Applications of machine learning and artificial intelligence for COVID-19 (SARS-CoV-2) pandemic: A review. Chaos Sol Fract. 2020:110059.
64. . LeCessie S, Van Houwelingen JC. Ridge estimators in logistic regression. J R Stat Soc Ser C (Appl Stat). 1992;41(1):191–201.
65. . LeCun Y, Bottou L, Bengio Y, Haffner P. Gradient-based learning applied to document recognition. Proc IEEE. 1998;86(11):2278–2324.
66. . Liu H, Motoda H. Feature extraction, construction, and selection: A data mining perspective. Springer Science & Business Media; 1998.
67. . López G, Quesada L, Guerrero LA. Alexa vs. Siri vs. Cortana vs. Google Assistant: A comparison of speech-based natural user interfaces. In: International Conference on Applied Human Factors and Ergonomics, Springer. 2017:241–250.
68. . Liu B, Hsu W, Ma Y. Integrating classification and association rule mining. In: Proceedings of the Fourth International Conference on Knowledge Discovery and Data Mining. 1998.
69. . MacQueen J. Some methods for classification and analysis of multivariate observations. In: Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability. 1967;1:281–297.
70. . Mahdavinejad MS, Rezvan M, Barekatain M, Adibi P, Barnaghi P, Sheth AP. Machine learning for Internet

Copyright © Author(s). This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY).

of Things data analysis: A survey. Digit Commun Netw. 2018;4(3):161–175.
71. . A. Marchand and P. Marx, "Automated Product Recommendations with Preference-Based Explanations," Journal of Retailing, vol. 96, no. 3, pp. 328-343, 2020.
72. . A. McCallum, "Information Extraction: Distilling Structured Data from Unstructured Text," Queue, vol. 3, no. 9, pp. 48-57, 2005.
73. . A. Mehrotra, R. Hendley, and M. Musolesi, "Prefminer: Mining User's Preferences for Intelligent Mobile Notification Management," in Proceedings of the International Joint Conference on Pervasive and Ubiquitous Computing, Heidelberg, Germany, Sept. 12-16, 2016, pp. 1223-1234, New York, NY, USA: ACM.
74. . Y. Mohamadou, A. Halidou, and P. T. Kapen, "A Review of Mathematical Modeling, Artificial Intelligence, and Datasets Used in the Study, Prediction, and Management of COVID-19," Applied Intelligence, vol. 50, no. 11, pp. 3913-3925, 2020.
75. . M. Mohammed, M. B. Khan, and B. E. Bashier Mohammed, Machine Learning: Algorithms and Applications, Boca Raton, FL, USA: CRC Press, 2016.
76. . N. Moustafa and J. Slay, "UNSW-NB15: A Comprehensive Data Set for Network Intrusion Detection Systems," in 2015 Military Communications and Information Systems Conference (MilCIS), Canberra, Australia, 2015, pp. 1-6, IEEE.
77. . M. Nilashi, O. B. Ibrahim, H. Ahmadi, and L. Shahmoradi, "An Analytical Method for Diseases Prediction Using Machine Learning Techniques," Computers & Chemical Engineering, vol. 106, pp. 212-223, 2017.
78. . Y. O. Yujin, S. Park, and J. C. Ye, "Deep Learning COVID-19 Features on CXR Using Limited Training Datasets," IEEE Transactions on Medical Imaging, vol. 39, no. 8, pp. 2688-2700, 2020.
79. . D. W. Otter, J. R. Medina, and J. K. Kalita, "A Survey of the Usages of Deep Learning for Natural Language Processing," IEEE Transactions on Neural Networks and Learning Systems, 2020.
80. . H.-S. Park and C.-H. Jun, "A Simple and Fast Algorithm for K-Medoids Clustering," Expert Systems with Applications, vol. 36, no. 2, pp. 3336-3341, 2009.
81. . K. Pearson, "On Lines and Planes of Closest Fit to Systems of Points in Space," Philosophical Magazine, vol. 2, no. 11, pp. 559-572, 1901.
82. . F. Pedregosa et al., "Scikit-Learn: Machine Learning in Python," Journal of Machine Learning Research, vol. 12, pp. 2825-2830, 2011.
83. . S. Perveen, M. Shahbaz, K. Keshavjee, and A. Guergachi, "Metabolic Syndrome and Development of Diabetes Mellitus: Predictive Modeling Based on Machine Learning Techniques," IEEE Access, vol. 7, pp. 1365-1375, 2018.
84. . P. Santi, D. Ram, C. Rob, and E. Nathan, "Behavior-Based Adaptive Call Predictor," ACM Transactions on Autonomous and Adaptive Systems, vol. 6, no. 3, pp. 21:1-21:28, 2011.
85. . A. S. Polydoros and L. Nalpantidis, "Survey of Model-Based Reinforcement Learning: Applications on Robotics," Journal of Intelligent and Robotic Systems, vol. 86, no. 2, pp. 153-173, 2017.
86. . M. L. Puterman, Markov Decision Processes: Discrete Stochastic Dynamic Programming, Hoboken, NJ, USA: John Wiley & Sons, 2014.
87. . J. R. Quinlan, "Induction of Decision Trees," Machine Learning, vol. 1, pp. 81-106, 1986.
88. . J. R. Quinlan, C4.5: Programs for Machine Learning, San Francisco, CA, USA: Morgan Kaufmann, 1993.
89. . C. Rasmussen, "The Infinite Gaussian Mixture Model," in Advances in Neural Information Processing Systems, vol. 12, pp. 554-560, 1999.
90. . K. Ravi and V. Ravi, "A Survey on Opinion Mining and Sentiment Analysis: Tasks, Approaches and Applications," Knowledge-Based Systems, vol. 89, pp. 14-46, 2015.
91. . L. Rokach, "A Survey of Clustering Algorithms," in Data Mining and Knowledge Discovery Handbook, Boston, MA, USA: Springer, pp. 269-298, 2010.
92. . S. Safdar, S. Zafar, N. Zafar, and N. F. Khan, "Machine Learning-Based Decision Support Systems for Heart Disease Diagnosis: A Review," Artificial Intelligence Review, vol. 50, no. 4, pp. 597-623, 2018.
93. . I. H. Sarker, "Context-Aware Rule Learning from Smartphone Data: Survey, Challenges and Future Directions," Journal of Big Data, vol. 6, no. 1, pp. 1-25, 2019.
94. . I. H. Sarker, "A Machine Learning-Based Robust Prediction Model for Real-Life Mobile Phone Data," Internet of Things, vol. 5, pp. 180-193, 2019.
95. . I. H. Sarker, "AI-Driven Cybersecurity: An Overview, Security Intelligence Modeling and Research Directions," SN Computer Science, vol. 2, no. 3, 2021.
96. . I. H. Sarker, "Deep Cybersecurity: A Comprehensive Overview from Neural Network and Deep Learning Perspective," SN Computer Science, vol. 2, no. 4, 2021.
97. . I. H. Sarker, Y. B. Abushark, F. Alsolami, and A. Khan, "IntrudTree: A Machine Learning-Based Cybersecurity Intrusion Detection Model," Symmetry, vol. 12, no. 5, pp. 754-768, 2020.
98. . I.H. Sarker, Y.B. Abushark, and A. Khan, "ContextPCA: Predicting Context-Aware Smartphone Apps Usage Based on Machine Learning Techniques," Symmetry, vol. 12, no. 4, p. 499, 2020.
99. . I.H. Sarker, H. Alqahtani, F. Alsolami, A. Khan, Y.B. Abushark, and M.K. Siddiqui, "Context Pre-Modeling: An Empirical Analysis for Classification-Based User-Centric Context-Aware Predictive Modeling," J. Big Data, vol. 7, no. 1, pp. 1-23, 2020.
100. . I.H. Sarker, A. Alan, H. Jun, A.I. Khan, Y.B. Abushark, and S. Khaled, "BehavDT: A Behavioral Decision Tree Learning to Build User-Centric Context-Aware Predictive Model," Mob. Netw. Appl., pp. 1-11, 2019.
101. . I.H. Sarker, A. Colman, M.A. Kabir, and J. Han, "Phone Call Log as a Context Source to Modeling Individual User Behavior," in Proc. 2016 ACM Int. Joint Conf. Pervasive Ubiquitous Comput. (Ubicomp): Adjunct, Germany, 2016, pp. 630–634.
102. . I.H. Sarker, A. Colman, M.A. Kabir, and J. Han, "Individualized Time-Series Segmentation for Mining

Copyright © Author(s). This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY).

Mobile Phone User Behavior," Comput. J., vol. 61, no. 3, pp. 349–368, 2018.
103. . I.H. Sarker, M.M. Hoque, M.K. Uddin, and A. Tawfeeq, "Mobile Data Science and Intelligent Apps: Concepts, AI-Based Modeling and Research Directions," Mob. Netw. Appl., pp. 1-19, 2020.
104. . I.H. Sarker and A.S.M. Kayes, "ABC-RuleMiner: User Behavioral Rule-Based Machine Learning Method for Context-Aware Intelligent Services," J. Netw. Comput. Appl., p. 102762, 2020.
105. . I.H. Sarker, A.S.M. Kayes, S. Badsha, H. Alqahtani, P. Watters, and A. Ng, "Cybersecurity Data Science: An Overview from Machine Learning Perspective," J. Big Data, vol. 7, no. 1, pp. 1–29, 2020.
106. . I.H. Sarker, P. Watters, and A.S.M. Kayes, "Effectiveness Analysis of Machine Learning Classification Models for Predicting Personalized Context-Aware Smartphone Usage," J. Big Data, vol. 6, no. 1, pp. 1–28, 2019.
107. . I.H. Sarker and K. Salah, "AppsPred: Predicting Context-Aware Smartphone Apps Using Random Forest Learning," Internet Things, vol. 8, 2019.
108. . T. Scheffer, "Finding Association Rules That Trade Support Optimally Against Confidence," Intell. Data Anal., vol. 9, no. 4, pp. 381–395, 2005.
109. . R. Sharma, S.S. Kamble, A. Gunasekaran, V. Kumar, and A. Kumar, "A Systematic Literature Review on Machine Learning Applications for Sustainable Agriculture Supply Chain Performance," Comput. Oper. Res., vol. 119, 2020.
110. . S. Shengli and C.X. Ling, "Hybrid Cost-Sensitive Decision Tree, Knowledge Discovery in Databases," in Proc. PKDD 2005, 9th European Conf. Principles Practice Knowl. Discovery Databases, vol. 3721, 2005.
111. . C. Shorten, T.M. Khoshgoftaar, and B. Furht, "Deep Learning Applications for COVID-19," J. Big Data, vol. 8, no. 1, pp. 1–54, 2021.
112. . G. Sökhan and N. Y., "Data Analysis in Health and Big Data: A Machine Learning Medical Diagnosis Model Based on Patients' Complaints," Commun. Stat. Theory Methods, pp. 1–10, 2019.
113. . D. Silver, A. Huang, C.J. Maddison, A. Guez, L. Sifre, G. Van Den Driessche, J. Schrittwieser, I. Antonoglou, V. Panneershelvam, M. Lanctot, et al., "Mastering the Game of Go with Deep Neural Networks and Tree Search," Nature, vol. 529, no. 7587, pp. 484–489, 2016.
114. . B. Ślusarczyk, "Industry 4.0: Are We Ready?" Polish J. Manag. Stud., vol. 17, 2018.
115. . P.H.A. Sneath, "The Application of Computers to Taxonomy," J. Gen. Microbiol., vol. 17, no. 1, 1957.
116. . T. Sørensen, "Method of Establishing Groups of Equal Amplitude in Plant Sociology Based on Similarity of Species," Biol. Skr., vol. 5, 1948.
117. . V. Srinivasan, S. Moghaddam, and A. Mukherji, "MobileMiner: Mining Your Frequent Patterns on Your Phone," in Proc. Int. Joint Conf. Pervasive Ubiquitous Comput., Seattle, WA, USA, 13-17 September, 2014, pp. 389–400.
118. . C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, and A. Rabinovich, "Going Deeper with Convolutions," in Proc. IEEE Conf. Comput. Vis. Pattern Recognit., 2015, pp. 1–9.
119. . M. Tavallaee, E. Bagheri, W. Lu, and A.A. Ghorbani, "A Detailed Analysis of the KDD Cup 99 Data Set," in IEEE Symp. Comput. Intell. Secur. Defense Appl., 2009, pp. 1–6.
120. . M. Tsagkias, T.K. Tracy, S. Surya, V. M., and M. de Rijke, "Challenges and Research Opportunities in E-Commerce Search and Recommendations," ACM SIGIR Forum, vol. 54, 2021, pp. 1–23.
121. . K. Wagstaff, C. Cardie, S. Rogers, and S. Schrödl, "Constrained K-Means Clustering with Background Knowledge," in Proc. ICML, 2001, pp. 577–584.
122. . W. Wang, J. Yang, and R. Muntz, "STING: A Statistical Information Grid Approach to Spatial Data Mining," VLDB, 1997, pp. 186–195.
123. . P. Wei, Y. Li, Z. Zhang, H. Tao, Z. Li, and D. Liu, "An Optimization Method for Intrusion Detection Classification Model Based on Deep Belief Network," IEEE Access, vol. 7, pp. 87593–87605, 2019.
124. . K. Weiss, T.M. Khoshgoftaar, and D.D. Wang, "A Survey of Transfer Learning," J. Big Data, vol. 3, no. 1, p. 9, 2016.
125. . I.H. Witten and E. Frank, Data Mining: Practical Machine Learning Tools and Techniques, Morgan Kaufmann, 2005.
126. . I.H. Witten, E. Frank, L.E. Trigg, M.A. Hall, G. Holmes, and S.J. Cunningham, Weka: Practical Machine Learning Tools and Techniques with Java Implementations, 1999.
127. . C.-C. Wu, L.-L. Yen, Y.-H. Li, and X.-Y. Yang, "Decision Tree Induction with a Constrained Number of Leaf Nodes," Appl. Intell., vol. 45, no. 3, pp. 673–685, 2016.
128. . X. Wu, V. Kumar, J.R. Quinlan, J. Ghosh, Q. Yang, H. Motoda, G.J. McLachlan, A. Ng, B. Liu, S.Y. Philip, et al., "Top 10 Algorithms in Data Mining," Knowl. Inf. Syst., vol. 14, no. 1, pp. 1–37, 2008.
129. . Y. Xin, L. Kong, Z. Liu, Y. Chen, Y. Li, H. Zhu, M. Gao, H. Hou, and C. Wang, "Machine Learning and Deep Learning Methods for Cybersecurity," IEEE Access, vol. 6, pp. 35365–35381, 2018.
130. . D. Xu and Y. Yingjie, "A Comprehensive Survey of Clustering Algorithms," Ann. Data Sci., vol. 2, no. 2, pp. 165–193, 2015.
131. . M.J. Zaki, "Scalable Algorithms for Association Mining," IEEE Trans. Knowl. Data Eng., vol. 12, no. 3, pp. 372–390, 2000.
132. . A. Zanella, N. Bui, A. Castellani, L. Vangelista, and M. Zorzi, "Internet of Things for Smart Cities," IEEE Internet Things J., vol. 1, no. 1, pp. 22–32, 2014.
133. . Q. Zhao and S.S. Bhowmick, "Association Rule Mining: A Survey," Singapore: Nanyang Technological University, 2003.
134. . T. Zheng, W. Xie, L. Xu, X. He, Y. Zhang, M. You, G. Yang, and Y. Chen, "A Machine Learning-Based Framework to Identify Type 2 Diabetes Through Electronic Health Records," Int. J. Med. Inform., vol. 97, pp. 120–127, 2017.

ISSN 2714-7444 (online), https://acopen.umsida.ac.id, published by Universitas Muhammadiyah Sidoarjo
Copyright © Author(s). This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY).

135. . Y. Zheng, S. Rajasegarar, and C. Leckie, "Parking Availability Prediction for Sensor-Enabled Car Parks in Smart Cities," in Intelligent Sensors, Sensor Networks and Information Processing (ISSNIP), 2015 IEEE Tenth International Conf., 2015, pp. 1–6.
136. . H. Zhu, H. Cao, E. Chen, H. Xiong, and J. Tian, "Exploiting Enriched Contextual Information for Mobile App Classification," in Proc. 21st ACM Int. Conf. Inf. Knowledge Manag., 2012, pp. 1617–1621.
137. . H. Zhu, E. Chen, H. Xiong, K. Y., H. Cao, and J. Tian, "Mining Mobile User Preferences for Personalized Context-Aware Recommendation," ACM Trans. Intell. Syst. Technol., vol. 5, no. 4, p. 58, 2014.
138. . Z. Zikang, Y. Yong, G. Guofeng, and Z. Xinyu, "Sentiment Analysis of Agricultural Product E-Commerce Review Data Based on Deep Learning," in 2020 Int. Conf. Internet Things Intell. Appl. (ITIA), IEEE, 2020, pp. 1–7.
139. . S. Zulkernain, P. Madiraju, and S.I. Ahamed, "A Context-Aware Interruption Management System for Mobile Devices," in Mobile Wireless Middleware, Operating Systems, and Applications, Springer, 2010, pp. 221–234.
140. . S. Zulkernain, P. Madiraju, S. Ahamed, and K. Stamm, "A Mobile Intelligent Interruption Management System," J. UCS, vol. 16, no. 15, pp. 2060–2080, 2010.

ISSN 2714-7444 (online), https://acopen.umsida.ac.id, published by Universitas Muhammadiyah Sidoarjo
Copyright © Author(s). This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY).