Types of Machine Learning Focused on Hybrid Learning Algorithms

Akram Shojaei
Faculty of Electrical Engineering,
K. N. Toosi University of Technology
Tehran, Iran
Akram.Shojaeibagheini@email.kntu.ac.ir

Abstract

Machine learning has emerged as a powerful discipline that allows computers to learn autonomously from data without explicit programming instructions. This presentation provides an overview of machine learning, including its connection to artificial intelligence, and explores 14 types of learning techniques. The article covers supervised learning, where models establish relationships between inputs and target variables, and unsupervised learning, which identifies patterns within data without explicit guidance. Additionally, hybrid learning problems such as semi-supervised learning, self-supervised learning, and multi-instance learning are discussed, each addressing unique data scenarios. The concept of ensemble learning, which combines multiple machine learning models, is introduced, highlighting the benefits of aggregating predictions from weak learners. The article further explores the concept of hybrid machine learning (HML), where multiple algorithms are combined to achieve superior performance. Three approaches to HML are presented: architecture integration, information control, and model parameter optimization. Overall, this article provides a comprehensive understanding of machine learning techniques and the growing importance of hybrid algorithms in tackling complex learning problems.

What Is Machine Learning?

Machine learning is the discipline that involves the programming of computers to acquire the capability of learning autonomously from provided data.

It is a domain of study that empowers computers to learn and make predictions without the need for explicit programming instructions.

What Is Machine Learning?

Machine learning is an extensive area of research that encompasses and draws inspiration from various interconnected disciplines, including artificial intelligence.

Artificial intelligence:

- Involves computers functioning autonomously
- Computers respond to their surroundings
- Computers exhibit cognitive capabilities
- Computers make independent decisions

VS

Machine learning:

- Is a practical implementation of artificial intelligence
- Computers observe and analyze data
- Computers make predictions based on past patterns
- Computers employ pre-programmed algorithms

14 Types of Learning

Learning Problems
1. Supervised Learning
2. Unsupervised Learning
3. Reinforcement Learning

Hybrid Learning Problems
4. Semi-Supervised Learning
5. Self-Supervised Learning
6. Multi-Instance Learning

Statistical Inference
7. Inductive Learning
8. Deductive Inference
9. Transduction Learning

Learning Techniques
10. Multi-Task Learning
11. Active Learning
12. Online Learning
13. Transfer Learning
14. Ensemble Learning

Machine Learning Mastery. "Types of Learning in Machine Learning." https://machinelearningmastery.com/types-of-learning-in-machine-learning/

Learning Problems

1.Supervised Learning

Supervised Learning refers to a category of problems where a model is employed to establish a connection between input examples and the target variable. When the training data consists of input vectors and their corresponding target vectors, it is referred to as supervised learning.

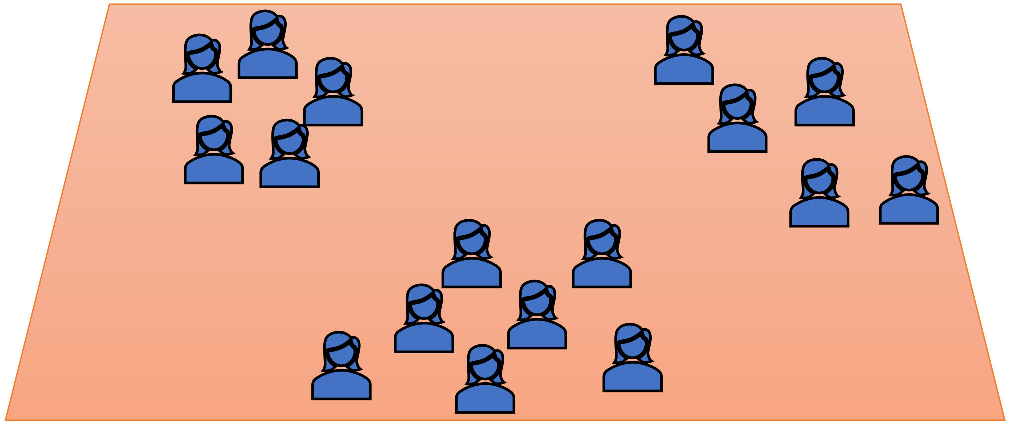
There are two primary types of supervised learning problems: classification, which involves predicting class labels, and regression, which involves predicting numerical labels.

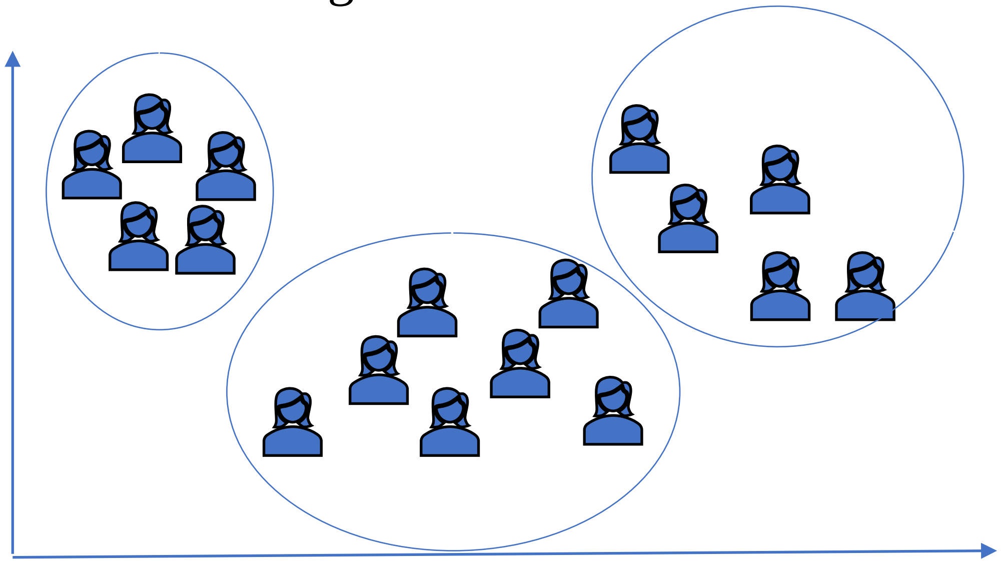
Learning Problems

2. Unsupervised Learning

Unsupervised Learning pertains to a category of problems where a model is employed to comprehend or extract connections within data without explicit guidance. Unlike supervised learning, unsupervised learning does not have an instructor or teacher to provide guidance, requiring the algorithm to make sense of the data independently.

There are two primary types of unsupervised learning problems: clustering, which involves identifying groups within the data, and density estimation, which involves summarizing the data distribution.





Hybrid Learning Problems

1.Semi-Supervised Learning

In many cases, labeling data can be a time-consuming and expensive task, resulting in a situation where a significant portion of the data remains unlabeled while only a small fraction is labeled. Certain algorithms are designed to handle such partially labeled data, and this approach is known as semi-supervised learning.

The objective of a semi-supervised learning model is to leverage all available data, including the unlabeled instances, in order to achieve effective learning, in contrast to supervised learning where only labeled data is utilized.

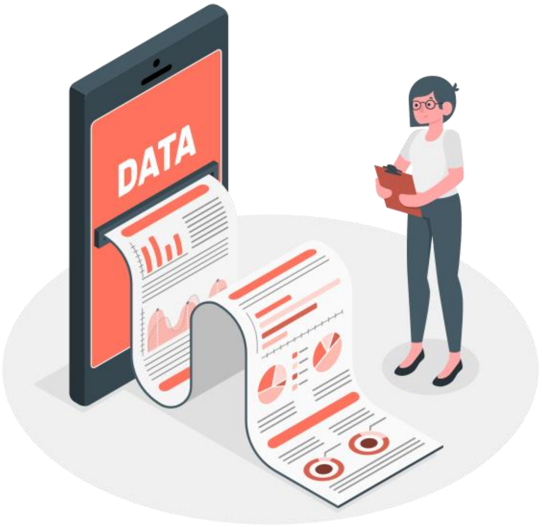


Image source: https://storyset.com/people">People illustrations by Storyset

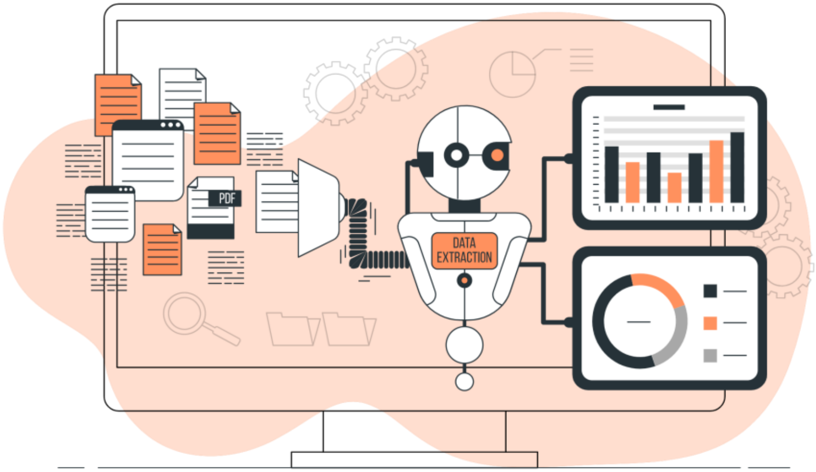
Hybrid Learning Problems

2.Self-Supervised Learning

An alternative method in machine learning entails generating a completely labeled dataset from an entirely unlabeled one. Once the entire dataset has been labeled, any supervised learning algorithm can be employed. This method is known as self-supervised learning. Self-supervised learning refers to the practice of framing an unsupervised learning problem as a supervised learning problem, enabling the utilization of supervised learning algorithms to address it.

Examples of self-supervised learning algorithms include:

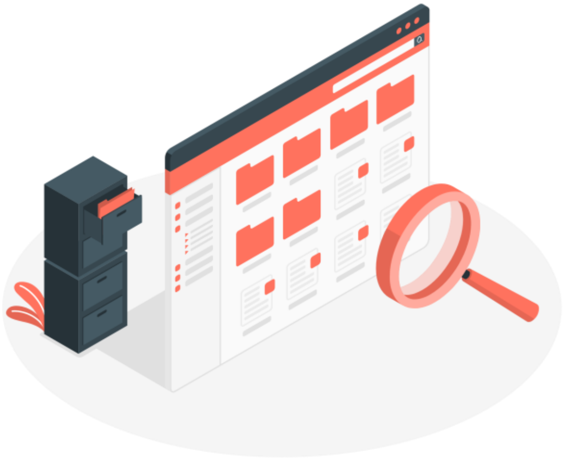
•Autoencoder
•Generative Adversarial Networks (GANs)



Hybrid Learning Problems

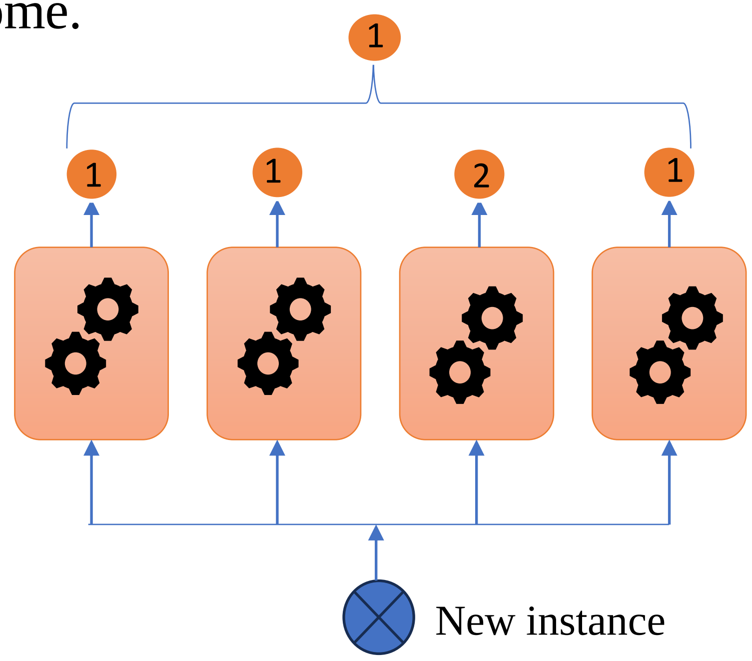
Multi-Instance Learning:
Multi-Instance Learning is a form of supervised learning wherein the individual instances lack labels, and instead, groups or bags of samples are labeled. In this approach, an entire collection of examples is labeled to indicate whether it contains an instance of a particular class or not, but the specific members within the collection do not have individual labels.

Initially, simple techniques like assigning class labels to individual instances and employing standard supervised learning algorithms often yield satisfactory results as an initial step.



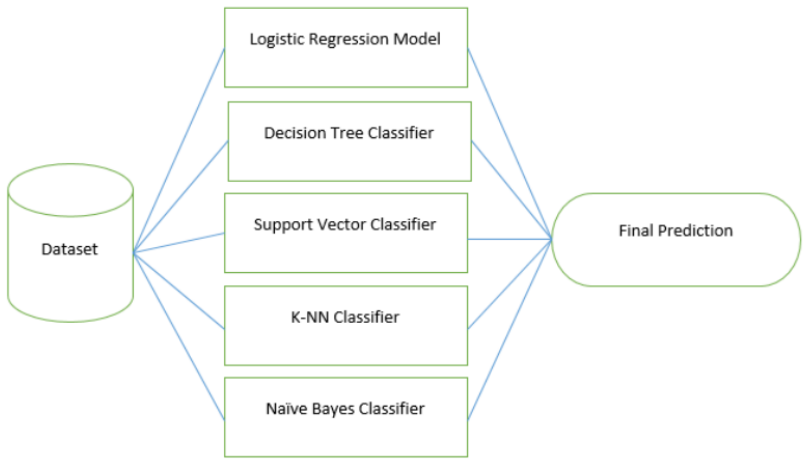
What is Ensemble Learning?

Ensemble Learning, also known as collective learning, is a machine learning concept that harnesses the collective strength of multiple machine learning models to tackle learning problems, including classification or regression tasks. In this approach, a set of homogeneous machine learning models, referred to as weak learners, are brought together. Each weak learner independently produces its own outcome, either on the entire training set or a subset of it, when applied to the problem. Subsequently, the results from each weak learner are merged or aggregated to obtain the final prediction or outcome.



Hybrid Ensemble Model

Example:
In this particular undertaking, a hybrid ensemble learning model is constructed by employing five distinct types of machine learning models as weak learners. These models include the Logistic Regression Model, Decision Tree, Support Vector Machine, K-Nearest Neighbor Model, and Naive Bayes Model. The term "hybrid" is employed in this context as opposed to other ensemble models, which typically employ a homogeneous set of weak learners. However, in this task, a heterogeneous collection of weak learners is utilized.



" A Hands-On Guide to Hybrid Ensemble Learning Models with Python Code." Analytics India Magazine. https://analyticsindiamag.com/a-hands-on-guide-to-hybrid-ensemble-learning-models-with-python-code/.

Hybrid Machine Learning (HML):

Hybrid Machine Learning (HML): Many of us have likely incorporated HML techniques into our designs without even realizing it. We may have utilized approaches that combine existing methodologies or integrated strategies from various domains.

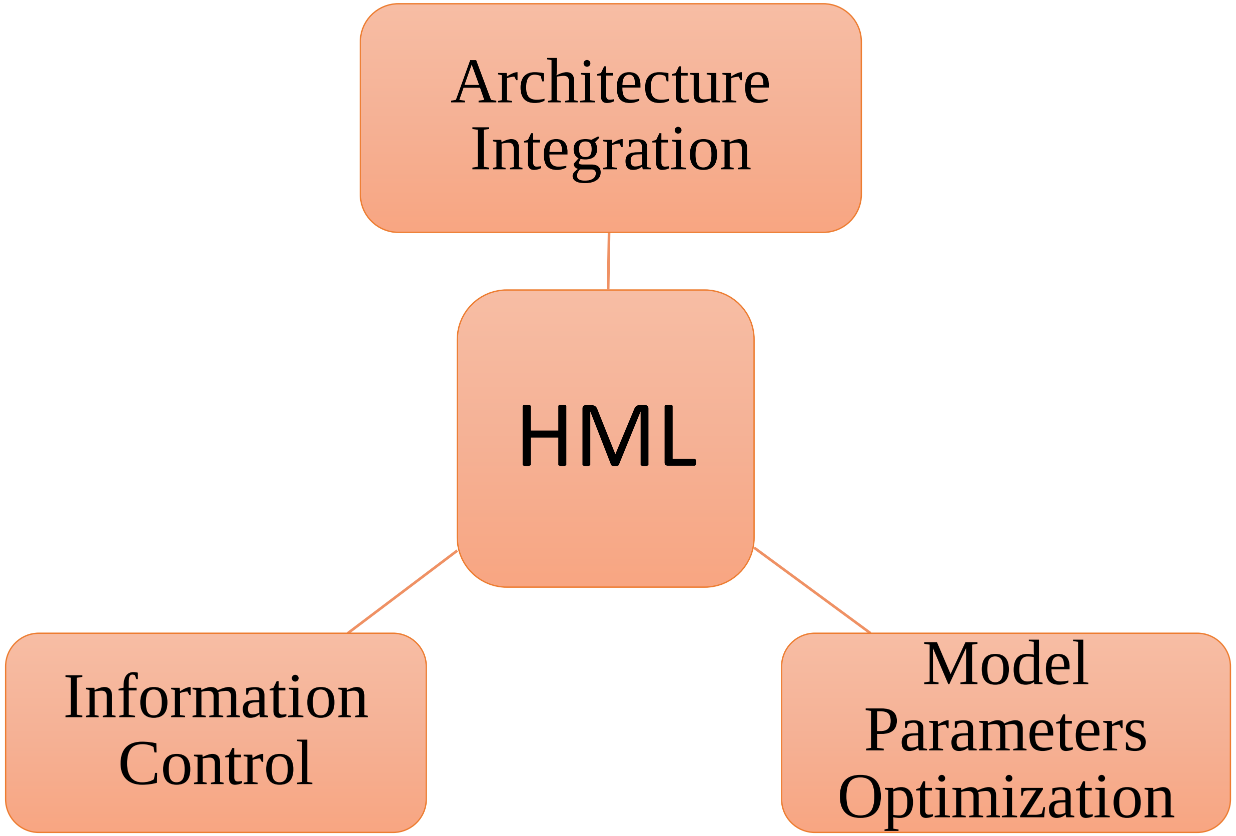
HML involves combining multiple algorithms to solve a problem effectively.

It entails merging two or more algorithms with the objective of achieving superior performance compared to using the individual algorithms alone.

In essence, HML refers to an algorithm composed of simpler algorithms

Hybrid Machine Learning (HML):

Hybrid Machine Learning (HML) can be approached from three different perspectives:



1. Architecture Integration

This particular approach to Hybrid Machine Learning (HML) involves integrating the architectures of two or more conventional algorithms, either entirely or partially, to create a more robust and self-sufficient algorithm. A commonly utilized model in this context is the Adaptive Neuro-Fuzzy Inference System (ANFIS). ANFIS has been widely employed as a standalone conventional machine learning strategy and combines the principles of fuzzy logic and artificial neural networks (ANN). The architecture of ANFIS consists of five layers, with the initial three layers derived from fuzzy logic principles and the remaining two layers from ANN.

2. Information Control

This form of Hybrid Machine Learning (HML) approach involves integrating information control processes or systems with traditional ML techniques to enhance the latter with the outcomes of the former. The following models serve as viable options for this type of hybrid learning strategy:

1.PCA-ANN Hybrid Model: When a Principal Component Analysis (PCA) module is employed to extract a submatrix of data that effectively captures the essence of the original data before applying a neural network, it can be referred to as a PCA-ANN hybrid model.

2.SVD-ELM Hybrid Model: If a Singular Value Decomposition (SVD) algorithm is used to reduce the dimensionality of a dataset prior to applying an Extreme Learning Machine (ELM) model, it can be termed an SVD-ELM hybrid model.

3. Model Parameters Optimization

It is well understood that each traditional ML technique employs a specific optimization or search algorithm.

If we consider the Particle Swarm Optimization (PSO) algorithm specifically, and use it to optimize the training parameters of an Artificial Neural Network (ANN) model, the resulting model can be referred to as a PSO-ANN hybrid model.

Similarly, other evolutionary optimization algorithms, such as Bee, Ant, Bat, and Fish Swarm, can be combined with traditional ML techniques to create corresponding hybrid models, where the optimization algorithms contribute to enhancing the performance of the base models.

Articles

1. A Hybrid Clustering Method Based on K-means Algorithm

This study introduces a novel clustering approach called Hierarchical Triangle K-means Clustering algorithm (HTK), which combines the strengths of K-means clustering and hierarchical clustering while leveraging the triangle inequality to expedite the clustering process.

The proposed HTK algorithm addresses the limitations of traditional K-means clustering, such as the requirement for initial cluster centers and the predetermined number of clusters. By integrating hierarchical clustering with K-means, it overcomes these challenges and significantly reduces the computation time associated with hierarchical clustering. The HTK algorithm produces clustering results that are comparable to the standard K-means algorithm, demonstrating its effectiveness.

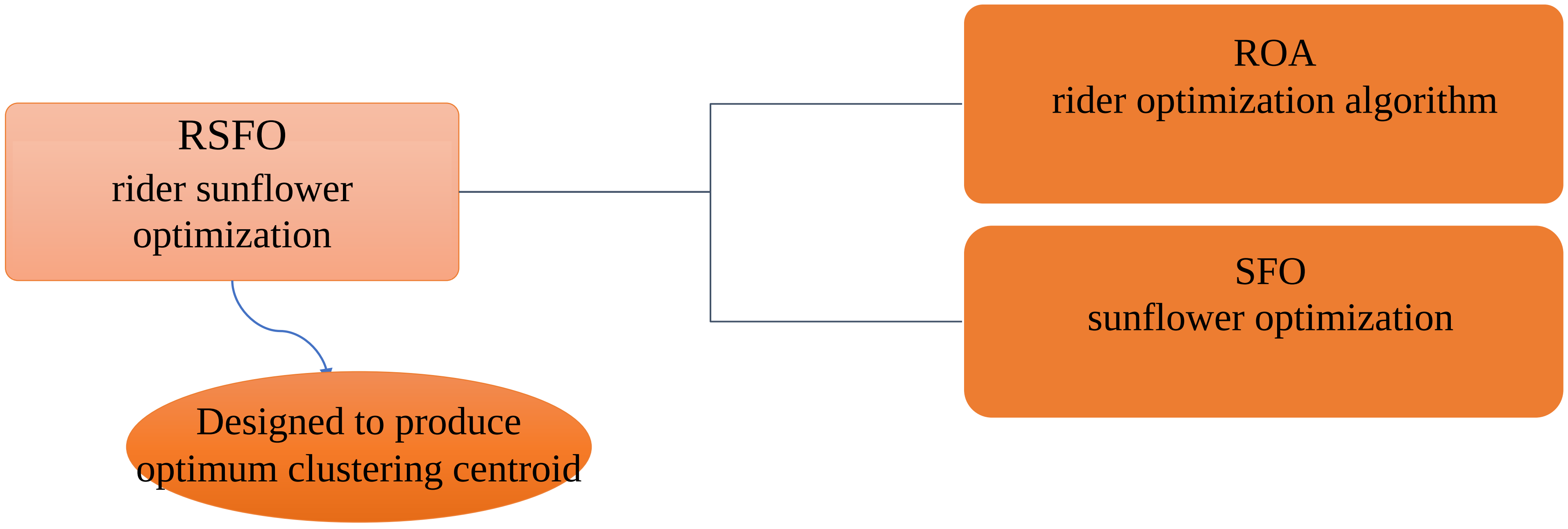
The HTK algorithm outperforms conventional methods in terms of both running time and accuracy. It enhances the clustering speed and optimizes the time complexity of the algorithm. Comparative evaluations with commonly used algorithms confirm that the HTK algorithm effectively reduces the clustering time, making it a valuable contribution to the field of clustering.

Hongwei Chen, "A hybrid clustering method based on k-means algorithm," Proc. SPIE 12350, 6th International Workshop on Advanced Algorithms and Control Engineering (IWAACE 2022), 123502D (20 October 2022).

Articles

2. Clustering by Hybrid K-Means-Based Rider Sunflower Optimization Algorithm for Medical Data

This hybrid algorithm combines the strengths of K-means clustering and the Rider Sunflower Optimization (RSFO) technique, while overcoming the limitations of each. It successfully avoids the issue of premature convergence associated with the K-means algorithm and mitigates the high computational cost typically associated with optimization techniques.

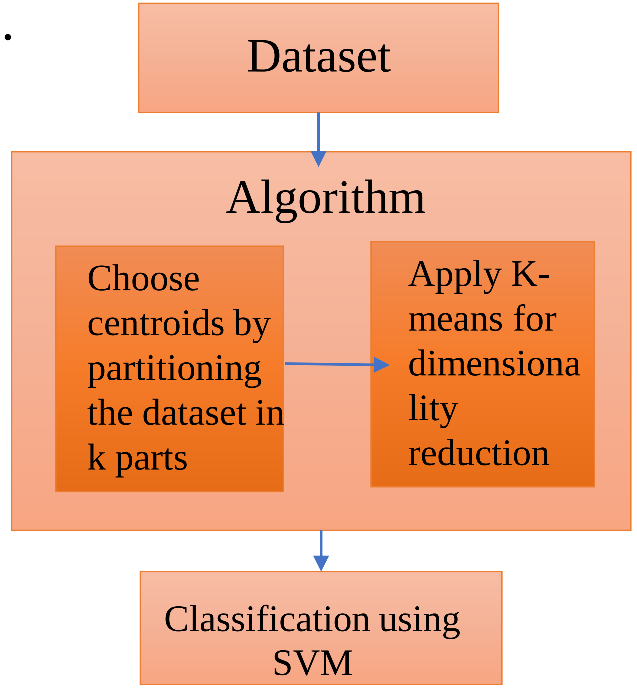


Articles

3. Disease Prediction using Hybrid K-means and Support Vector Machine

The presented hybrid approach combining K-means and Support Vector Machine (SVM) is valuable for multiple purposes, including selecting initial centroids and determining the number of clusters. It also enhances the overall efficiency of the K-means algorithm.

Moreover, the hybrid K-means algorithm plays a crucial role in reducing the dimensionality of the dataset before it is inputted into the Support Vector Machine classifier. This dimensionality reduction step helps optimize the performance of the SVM classifier in disease prediction tasks.



Conclusion

Machine learning is a dynamic field that encompasses various learning techniques, each suited to different data scenarios and problem domains. Supervised and unsupervised learning form the foundation, allowing computers to make predictions or identify patterns, respectively. Hybrid learning algorithms, such as semi-supervised learning, self-supervised learning, and multi-instance learning, bridge the gap between labeled and unlabeled data, expanding the possibilities of training models with limited labeled samples. Ensemble learning techniques aggregate the predictions of multiple models, capitalizing on their collective strengths and enhancing overall performance. Hybrid machine learning (HML) takes the fusion of algorithms a step further, combining multiple techniques to leverage their complementary advantages and achieve superior results. Whether through architecture integration, information control, or model parameter optimization, HML offers innovative approaches to tackle complex problems and optimize performance. As the field of machine learning continues to evolve, understanding and utilizing hybrid algorithms will become increasingly crucial for unlocking the full potential of data-driven decision-making.

References

1. Machine Learning Mastery. "Types of Learning in Machine Learning." https://machinelearningmastery.com/types-of-learning-in-machine-learning/.

2. " A Hands-On Guide to Hybrid Ensemble Learning Models with Python Code." Analytics India Magazine. https://analyticsindiamag.com/a-hands-on-guide-to-hybrid-ensemble-learning-models-with-python-code/.

3. Hongwei Chen, "A hybrid clustering method based on k-means algorithm," Proc. SPIE 12350, 6th International Workshop on Advanced Algorithms and Control Engineering (IWAACE 2022), 123502D (20 October 2022).

4. A. Jaya Mabel Rani, A. Pravin, "Clustering by Hybrid K-Means-Based Rider Sunflower Optimization Algorithm for Medical Data", *Advances in Fuzzy Systems*, vol. 2022, Article ID 7783196, 9 pages, 2022.

5. Kaur, Sandeep and Sheetal Kalra. “Disease prediction using hybrid K-means and support vector machine.” *2016 1st India International Conference on Information Processing (IICIP)* (2016): 1-6.