Transactions on Fuzzy Sets and Systems (TFSS)
URL: https://sanad.iau.ir/journal/tfss/
Online ISSN: 2821-0131
Vol.4, No.2, (2025), pp-pp
DOI: https://doi.org/10.71602/tfss.2025.1183012

Article Type: Original Research Article
(Journal Pre-proof)

Some Granular Computing Based Machine Learning Algorithms

Vijay R. Tiwari*

(Special Issue on "Granular Computing" Dedicated to Prof. Witold Pedrycz.)

Abstract. Granular computing has emerged as a new computational method that is beneficial when dealing with large amounts of data. In recent years, several machine learning models based on the granular framework have been developed, outperforming traditional machine learning models. This article reviews some newly developed techniques in terms of granular framework settings.

AMS Subject Classification 2020: 68T01; 68T09; 68T37
Keywords and Phrases: Big data, Information granules, GrC, ML.

1 Introduction

During Lin's sabbatical leave in 1996-97, Zadeh proposed the concept of "Granular Mathematics" as a potential area of research for him. To further refine the focus, Lin later suggested the term "Granular Computing" [1], which initially referred to the computational aspect of granular mathematics. Granular computing involves breaking data of massive size into information granules. These granules are then used to solve complex real-world computational problems, utilizing techniques such as machine learning, rough sets, fuzzy sets and logic, data mining, and clustering.

Machine learning (ML) algorithms when combined with Granular Computing (GrC) techniques can be instrumental in achieving various complex machine learning tasks such as GrC-based feature selection, GrC-based classification algorithm, concept-cognitive learning with GrC, cost-sensitive active learning via GrC, and multi granulation learning in cognition [2]. Ensemble learning is a powerful machine learning technique that involves combining two or more ML algorithms to create a more effective model [3]. The GrC-based approach has the potential to significantly enhance traditional ensemble learning methods. In [4], a novel ensemble learning method based on a granular framework is proposed. This approach involves transforming data and breaking down multi-class classification tasks into multiple binary classification tasks. The classifier is trained on various feature sets using different algorithms, and the results are then combined, which is a key characteristic of ensemble learning.

GrC techniques are also used in rule mining. Rule mining is characterized by the formation of concepts and identification of concepts [5]. A concept is described by its intension and extension. Intension refers to the set of properties that apply to an object when a concept is applied to it. On the other hand, extension

∗**Corresponding Author:** Vijay R. Tiwari, Email: vijay.tiwari@jaihindcollege.edu.in, ORCID: 0000-0001-9747-3389
Received: 4 September 2024; **Revised**: 22 October 2024; **Accepted**: 25 October 2024; **Available Online**: 2 December 2024; **Published Online**: 7 November 2025.

How to cite: Tiwari VR. Some granular computing based machine learning algorithms. *Transactions on Fuzzy Sets and Systems.* 2025; 4(2): pp-pp. DOI: https://doi.org/10.71602/tfss.2025.1183012

refers to the set of objects that are instances of a concept and are subsets of the universe. In data mining, granulation is dividing the universe into subsets which are called granules. Partition and covering are some commonly used granulation techniques in data mining. Partition results in the collection of disjoint subsets of the universe whereas covering results in the collection of overlapping subsets of the universe. Yao [6] proposed a data mining model based on the GrC approach, which is based on the construction of an information table.

In this article, we will explore the concepts of "big data," "machine learning," and "granular computing" in Section 2. In Section 3, we will explore newly developed algorithms based on the granular setting. Additionally, Section 4 will focus on several real-life applications that use the granular framework. Finally, this article will conclude with remarks in Section 5.

2 Definition of Terms

In this section, we will review fundamental definitions and explore various concepts related to granular computing.

2.1 Big Data

As the name suggests, "Big Data" refers to a large amount of information. The term "Big data" was first documented in the year 1997, by Cox and Ellsworth, in their paper, where they described the challenges of using the vast amount of data generated by supercomputers. According to the Oxford Dictionary, big data is defined as "Data of very large size that presents significant logistical challenges in its manipulation and management." Over the time various descriptions of big data have emerged. Robert Allison [7] defines it as, an "Amalgamation of information from a variety of sources."

Laney [8], introduced the 3V's of big data: "high Volume, high Velocity, and/or high Variety information assets that require new forms of processing to enable enhanced decision-making, insight discovery, and process optimization." We recall the 5V's of big data [9]:

- Volume: This refers to the amount of space required to store the data. A large volume of data can impact the computational cost of machine learning tasks.
- Velocity: This refers to the speed of data transformation, including collection and processing.
- Variety: This refers to the type of data which can be structured or unstructured. In machine learning, one encounters either discrete or continuous data. The format of the data can impact the interpretation ability of machine learning models.
- Veracity: This refers to the degree to which data can be trusted. Veracity is evaluated by the accuracy of the information provided by the data.
- Value: This refers to the value addition a data does towards a process.

Variability and visualisation add to these 5V's making it to 7V's which are crucial in building a strategy for machine learning models for big data. An efficient visualisation method can enhance the presentation of data and hence, useful in its analytics.

2.2 Machine Learning

Machine learning is a branch of artificial intelligence that aims to replicate human thinking and decision-making abilities. It allows machines to learn from data and improve their performance over time. The process of machine learning typically involves two stages: training and testing.

Training Stage

In the training stage, the machine learns through various algorithms. Machine learning can broadly be classified into three types: supervised learning, unsupervised learning and reinforcement learning. In supervised learning input is labelled data and output is supervised output. Supervised machine learning algorithms are employed in solving regression and classification problems. In unsupervised learning algorithms, input data need not be labelled and the machine acts on data without any guidance. In unsupervised learning machine discovers patterns and relationships in the data without any explicit guidance. Unsupervised machine learning algorithms are employed in solving clustering and association problems. Reinforcement learning is a type of machine learning that learns through a feedback mechanism based on rewards and punishments for certain behaviours. This approach is commonly used in decision-making and control problems.

Testing Stage

After a machine is trained with an algorithm, its performance is determined by testing it using appropriate testing techniques. Training and testing involve data partitioning. The most widely used data partitioning techniques used for testing machine learning algorithms are:

- Cross-Validation: The cross-validation technique works by dividing the data set into n disjoint subsets and performing n iterations. In each iteration, one subset is used for testing while the remaining $n-1$ subsets are used for training. However, it should be noted that cross-validation can be computationally expensive.

- Data-Split: In the data-split method, the data set is divided into two groups. Approximately 70% of the data is used to train the machine, while the remaining 30% is used for testing.

An appropriate sampling technique is a must for sampling training and test data to build an unbiased and accurate model.

2.3 Granular Computing (GrC)

Machine learning tasks become increasingly complex with the large size of data. Traditional machine learning is not sufficient for working with big data. Therefore, it is crucial to develop techniques that can effectively partition the data into smaller parts and work on these smaller parts. Data partitioning can greatly improve the performance of machine learning algorithms. One efficient approach to performing machine learning for big data is through granular computing which involves two important stages: granulation and organization. GrC aims to build efficient machine-learning models that utilize granules to process huge amounts of data, information, and knowledge. Here a granule refers to the smallest unit of information.

In Merriam-Websters dictionary [10], a granule is defined as "a small particle; especially, one of numerous particles forming a larger unit".

Keet [11], describes granules as "collections of entities grouped together based on similarity, functional or physical adjacency, or indistinguishability".

A mathematical notion of a granule given by Lin [12] is: "A granule can be a crisp or fuzzy subset, a function, an algorithm, a random variable(measurable function), a generalized function, etc."

Hu and Shi [13], define a granule as "a clump of points drawn together by indistinguishability, similarity, and functionality." According to Hu and Shi GrC is defined as:

Definition 2.1. ([13], §2) *Granular computing (GrC) is characterized by three aspects:* (1) *a leveled granular system based on the tolerance relation of distance function,* (2) *a layered granular computing,* (3) *an optimisation problem about above layered granular computing.*

They treat a set of granules as a convex region and utilize a continuous function to simplify the complex boundaries of the sample distribution. This function is beneficial for feature transformation and can also aid in dimension reduction.

Zadeh [14] proposed three main issues with GrC namely., granulation, organization, and causation. Granulation is the decomposing of the whole into parts, organisation is the integration of parts into a whole and causation is association of causes with effects.

The process of granulation follows a top-down approach and the process of organisation follows a bottom-up approach. In top down approach, an overview model is formed first then each part of the model is refined into details whereas in the bottom-up approach details of each part of the model are formed first which are then linked to form larger components. This large component is used to form the required model.

3 GrC-ML Algorithms

The performance of traditional machine learning methods is often inadequate when applied to big data. The 5V's of big data has a significant impact on the four dimensions of machine learning namely., "accuracy", "efficiency", "interpretability", and "stability" when using traditional approaches. These learning algorithms can produce errors such as imbalances between bias-variance tradeoffs and often cause overfitting. To address these issues, GrC-based machine learning can be applied. In the context of GrC-based learning, traditional methods are treated as algorithms of a single granularity. In [15], a multi-granularity learning approach is proposed, which includes techniques such as semi-supervised learning, semi-heuristic learning, multi-task learning, generative learning, and semi-random data partitioning. A brief overview of these methods will be discussed in the subsequent subsections.

3.1 Semi-Supervised Learning

Supervised learning requires a labelled data set for training and testing purposes. However, when dealing with a large amount of data, the task of labelling can become tedious and challenging. To address this issue, semi-supervised learning can be used. This involves dividing the big data into smaller and larger portions. The smaller portion is labelled by experts (humans), while the larger portion is labelled using machine learning algorithms. In ([15], Chapter 2), two classification algorithms are proposed. These algorithms utilize KNN and SVM to label a large data set.

3.1.1 KNN Based Labelling

K-Nearest Neighbour or KNN is a type of supervised machine learning algorithm used for solving classification and prediction problems. In KNN, the value of K is an integer and it represents the number of neighbours selected for the task ([16], §4). When KNN is applied to semi-supervised learning, the classifier first learns from a small portion of labelled data and then uses this knowledge to classify the remaining unlabelled instances. The classification is determined by the majority label type of the K nearest neighbours. After classification, the newly labelled instance is added to the training set and the classifier learns again from this updated data set. To minimize the risk of incorrect labelling of subsequent unlabelled instances, Liu and Cocea ([15], Chapter 2) suggest labelling one instance at a time. The granular framework of semi-supervised learning based on KNN labelling involves two levels of granularity ([15], p.25) and it is illustrated in Figure 1.

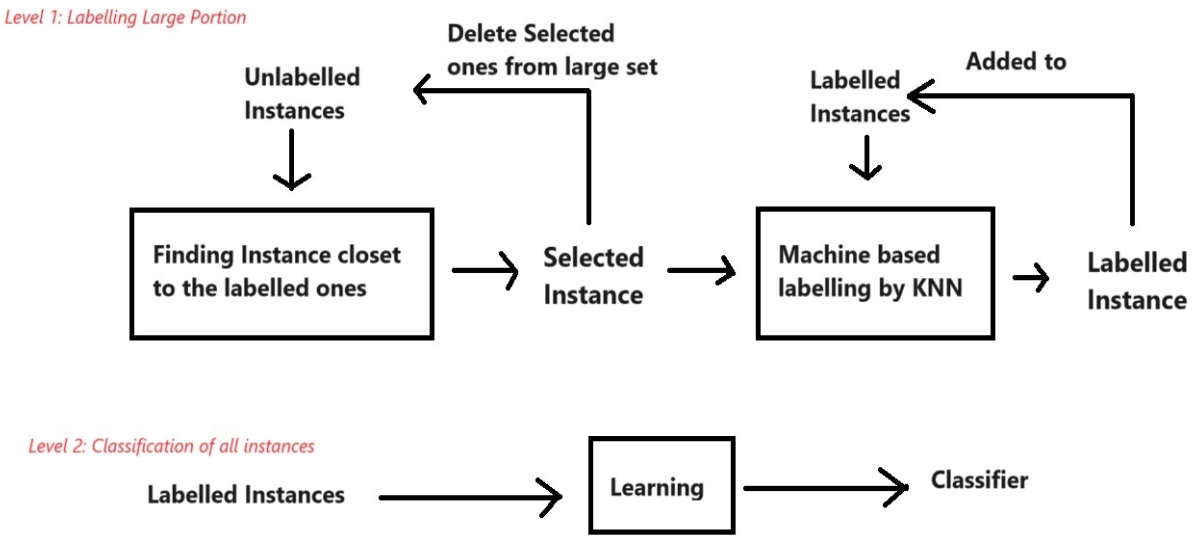

Figure 1: Semi-supervised learning through KNN ([15], p.25)

3.1.2 SVM Based Labelling

Support Vector Machine or SVM creates an optimal boundary for separating labelled instances from unlabelled instances. In semi-supervised learning based on SVM, the classifier learns from the small portion of labelled instances. Once the learning is done, it is used for labelling the large portion of unlabelled instances. The process starts by labelling instances farthest from the boundary. The newly labelled instances are then added to the training set and the whole process is repeated till all unlabelled instances are labelled. To minimize the risk of incorrect labelling of subsequent unlabelled instances Liu and Cocea ([15], Chapter 2) suggest labelling one instance at a time. The granular framework of semi-supervised learning based on SVM labelling involves two levels of granularity ([15], p.26) and it is shown in Figure 2.

Semi-supervised learning has various applications, including text classification, natural language processing, and handwritten digit recognition.

3.2 Semi Heuristic Learning

In traditional heuristic learning, a strategy based on statistical techniques is used for the learning process. When traditional heuristic learning is applied to data with high variability, it can result in a high bias-variance trade-off. In semi-heuristic learning, instances are classified through voting, using techniques such as bagging and boosting, which are types of ensemble machine learning. Bagging (Bootstrap Aggregating) is an ensemble learning technique that reduces the chances of overfitting by decreasing the variance. Boosting is another ensemble learning technique that improves accuracy by combining multiple weak classifiers to create a strong classifier. Bagging uses majority voting for classification while boosting uses weighted voting. In terms of the granular framework, bagging has two levels of granularity where in level 1, the model uses an algorithm to learn from training datasets and in level 2, the fitness of the trained model is learned ([15], Chapter 4).

In ([15], Chapter 4), a nature-inspired semi-heuristic learning approach is proposed that has been shown to outperform traditional heuristic learning methods. This semi-heuristic learning technique utilizes probabilistic

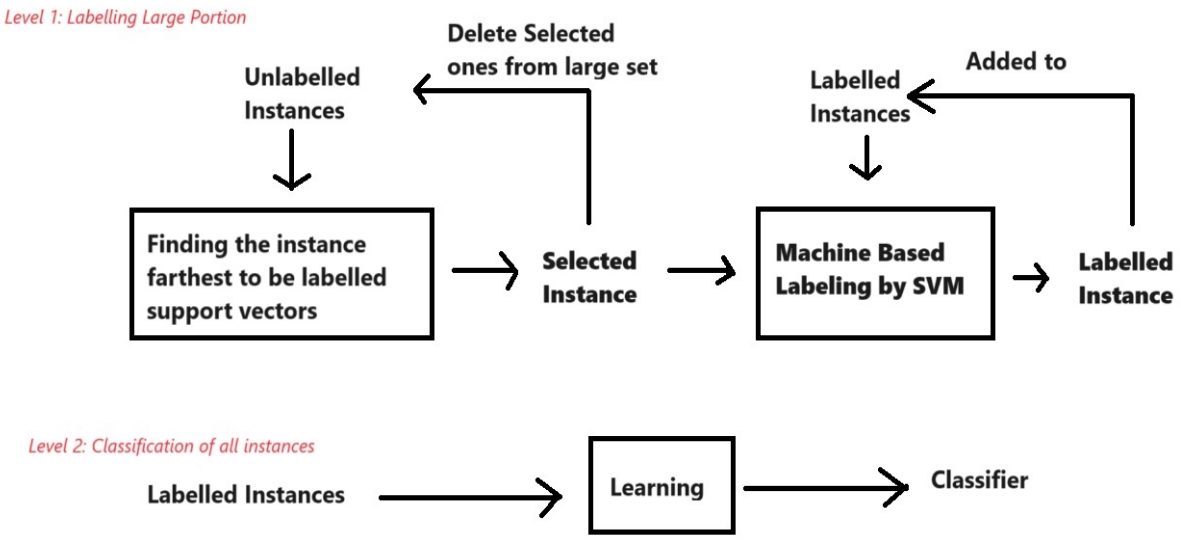

Figure 2: Semi-supervised learning through SVM ([15], p.26)

voting [17], specifically in the context of instance classification, to reduce bias. The process of probabilistic voting involves calculating the weight and percentage weight of each class. Once the weights are determined, a class is randomly selected using a method inspired by the Roulette Wheel. Each class is then assigned an interval based on its weight, referred to as a probabilistic interval. This interval is viewed as a "probabilistic granule" and the frequency of the class serves as an information granule for class selection when assigning a class to a test instance. The weight of the class also acts as an information granule when assigning a class to an unseen instance.

In the context of GrC, granulation is the process of selecting several samples from the training set with replacement, while organisation involves combining models that have been trained on different samples into an ensemble model.

3.3 Multi-Task Learning

A single-task learning is not sufficient to learn from multiple subsets of big data. Since each subset is divided into a training subset and a test subset for a particular learning task, thus it becomes necessary to convert single-task learning to multiple-task learning. Another issue in traditional single-task learning is that single-task learning is discriminative in nature. Thus single-task learning discriminates one class from a group of mutually exclusive classes. In real situations, classes need not be mutually exclusive. In fact, classes can be correlated to each other. Therefore single-task learning for big data results in a poor classifier. While using single-task learning one has to redo the classifier learning every time when the class labels of data sets are updated. This leads to poor maintainability, usability, extensibility and flexibility of the classifier [18].

Multi-task learning is a solution to the challenges faced in single-task learning when dealing with large amounts of data. In single-task learning, each class is treated independently and given equal weight when assigning class labels to new instances. However, in multi-task learning, the classifier can easily adapt to accommodate new class labels. For instance, in ([15], Chapter 5), Liu and Cocea describe a fuzzy approach to multi-task learning. This approach utilizes a fuzzy rule-based system (FRBS) to avoid bias from deterministic

rules. In [19], Liu and Haig use the Tsukamoto system, which is a type of FRBS, for classifying instances. To better understand the functionality of the Tsukamoto system, we refer to the example in [19]. This system has two variables: x_1 with linguistic terms "Tall" and "Short," and x_2 with linguistic terms "Large" and "Small." The output variable, y, corresponds to the membership degree of the fuzzy set, which is equal to the rule firing strength. It has two outputs: positive and negative. Therefore, there are four rules in total.

R1: If $x_1 =$ 'tall' and $x_2 =$ 'large' then $y =$ 'positive'.

R2: If $x_1 =$ 'tall' and $x_2 =$ 'small' then $y =$ 'positive'.

R3: If $x_1 =$ 'short' and $x_2 =$ 'large' then $y =$ 'negative'.

R4: If $x_1 =$ 'short' and $x_2 =$ 'small' then $y =$ 'negative'.

Each numerical input attribute is converted into multiple linguistic attributes using a trapezoidal fuzzy membership function. For instance, the values $a = 1.3$, $b = 1.8$, $c = 1.8$, and $d = 1.8$ are utilized to construct the fuzzy membership function for the linguistic term "Tall," while the values $a = 2$, $b = 8$, $c = 8$, and $d = 8$ are used for the term "Short." The next step is to classify the instance $(x_1 = 1.425, x_2 = 6.5)$. By applying the trapezoidal fuzzy membership function, we can determine the fuzzy membership values for the input instances as $f_{Tall}(x_1) = 0.25$, $f_{Short}(x_1) = 0.75$, $f_{Large}(x_2) = 0.75$, and $f_{Small}(x_2) = 0.75$. Fuzzification of the four rules results in the following:

R1: $f_{Tall}(x_1) = 0.25$, $f_{Large}(x_2) = 0.75$.

R2: $f_{Tall}(x_1) = 0.25$, $f_{Small}(x_2) = 0.75$.

R3: $f_{Short}(x_1) = 0.75$, $f_{Large}(x_2) = 0.75$,

R4: $f_{Short}(x_1) = 0.75$, $f_{Small}(x_2) = 0.75$.

After the process of fuzzification, the minimum membership degree of the antecedents is determined. This is followed by the steps of application, implication, aggregation, and defuzzification. For a better understanding of these steps, please refer to ([15], [19]). The application of rules involves the use of conjunction ($\wedge$), which takes the minimum of two values, representing the minimum of the membership degrees of the antecedents. Hence,

$$f_{Tall}(x_1) \wedge f_{Large}(x_2) = 0.25, \text{ and } f_{Tall}(x_1) \wedge f_{Small}(x_2) = 0.25.$$

$$f_{Short}(x_1) \wedge f_{Large}(x_2) = 0.75 \text{ and } f_{Short}(x_1) \wedge f_{Small}(x_2) = 0.75.$$

The next step is implication $(\rightarrow)$:

$$f_{R1\rightarrow Positive}(x_1, x_2) = 0.25 \text{ and } f_{R2\rightarrow Positive}(x_1, x_2) = 0.25$$

$$f_{R3\rightarrow Negative}(x_1, x_2) = 0.75 \text{ and } f_{R4\rightarrow Negative}(x_1, x_2) = 0.25$$

After implication aggregation is applied as:

$$f_{Positive}(x_1, x_2) = f_{R1\rightarrow Positive}(x_1, x_2) \wedge f_{R2\rightarrow Positive}(x_1, x_2) = 0.25$$

$$f_{Negative}(x_1, x_2) = f_{R3\rightarrow Negative}(x_1, x_2) \wedge f_{R4\rightarrow Negative}(x_1, x_2) = 0.75.$$

The last step is defuzzification and we observe

$$f_{Negative}(x_1, x_2) > f_{Positive}(x_1, x_2).$$

Therefore the class label of the instance $(x_1 = 1.425, x_2 = 6.5)$ is negative. If the total fuzzy membership degrees for an instance add up to one for a specific class and zero for all other classes, then the instance is classified under that particular class. However, if the total fuzzy membership degrees do not add up to one, the instance may belong to multiple classes. Furthermore, if the total fuzzy membership degrees for a group of classes add up to one, then these classes are considered mutually exclusive. However, if the total fuzzy membership degrees exceed one, then they are not mutually exclusive.

The granular framework of multi-task learning has three levels of granularity. In level 1, learning of fuzzy membership function happens, level 2 involves learning towards identifying fuzzy membership degree and level 3 involves learning of relation identification among classes ([15], Chapter 5).

In terms of GrC, granulation involves transforming the class attributes into a number of binary attributes and dividing the training set into subsets, with each subset representing a specific class. Organisation is achieved by linking classes based on factors such as mutual exclusion, correlation, and independence.

3.4 Semi-Data Partitioning

Data partitioning is a crucial step in training and testing machine learning algorithms. There are various methods available for partitioning data, with data splitting being the most commonly used approach. In data splitting, the entire dataset is divided into a 70:30 ratio, with 70% of the data used for training the machine and the remaining 30% used for testing the learned algorithm. The two main approaches for data splitting are fixed and random. However, the fixed method can result in differences between the training and test sets, leading to issues with sample representation. As a result, this method is often avoided. While traditional random partitioning works well for small datasets, it may not be sufficient for larger datasets. When applied to large datasets, random partitioning can lead to problems such as "class imbalance" and "sample representativeness."

Liu and Cocea [20], proposed a modified version of the random partition method to address the challenges of applying traditional methods to big data. This modified method, called "semi-random partitioning," is based on a stratified sampling approach. In stratified sampling, the data set is divided into subgroups based on similarity, which helps to reduce errors caused by data variability. Additionally, this method addresses class imbalance issues in the training and test sets by assigning instances to a class based on the sampling probability of each class. However, one limitation of this method is that it may struggle to maintain balance between the training and test sets.

The granular framework of semi-random partitioning involves three levels of granularity. In level 1, the data set is partitioned randomly into a training set and a test set. In level 2, the original dataset is divided into a number of subsets where each subset contains instances with a particular class label. After this, subsets are partitioned randomly into training subsets and test subsets. The training and test subsets are then merged into a training set and test set respectively. The second level, as discussed in ([15], Chapter 6), is a crucial step that addresses the issue of class imbalance. This level ensures that a fixed percentage of instances (70%) are included in both the training and test sets. Level 3 is dividing subsets of level 2 into a number of sub-subsets which contain instances of subclass of a class taken in level 2. After this data from each subsets are partitioned randomly into training sub-subsets and test sub-subsets. These sub-subsets are merged together to obtain a training set and test set. The third level, also discussed in ([15], Chapter 6), eliminates dissimilarities between the training and test instances. At this level, instances of the same type are grouped together before being partitioned into the training and test sets. This ensures that instances from each group are present in both sets.

In terms of GrC, granulation involves dividing a data set into multiple subsets. Each subset is then further divided into a training subset and a test subset, which are necessary for level two of granularity. This process is repeated by dividing the data subsets into smaller sub-subsets, with each sub-subset also being split into a training sub-subset and a test sub-subset for level three of granularity. The organisation aspect involves

integrating the training and test subsets (or sub-subsets) into a complete training and test set.

3.5 Rule Learning

The Most commonly used rule learning methods in traditional machine learning are based on techniques such as "Divide and Conquer" and "Separate and Conquer". Divide and conquer works on a top-down approach which often leads to overfitting, hence causing poor interpretability of decision tree. Traditional methods work on "attribute-value pair selection" which sometimes produces low-quality rule sets. ID3 and C4.5 algorithms follow divide and conquer towards rule learning and the PRISM algorithm follows the separate and conquer technique for rule learning.

PRISM produces a complex rule-based classifier as it learns a rule set for each class. Experiments in [21], show that rule learning for each class is not required in many problems and rule learning for one target class is sufficient. Thus Liu and Cocea [21], propose a modified version of PRISM which is called 'PISMSTC' where STC stands for "single target class." PRISMSTC produces a simpler rule-based classifier and is hence effective in terms of granular framework as it reduces computational complexity without losing accuracy.

In ([15], Chapter 7), Liu and Cocea describe a modified version of the "separate and conquer" technique for handling big data. This approach operates on three levels of granularity, allowing for more efficient selection of attribute-value pairs. At level 1, a heuristic is used to learn the rule set, while also evaluating the overall quality of the rule set. Moving on to level 2, each individual rule is learned by combining multiple heuristics, and the quality of each rule is evaluated. In a previous study, a collaborative approach was proposed for level 2 learning [22]. Finally, at level 3, each term in the rule is learned using a combination of heuristics, and the impact of each rule term on optimizing the full rule set is evaluated.

In terms of GrC, granulation involves separating training instances required for attribute-value pair selection and for learning of rules. Organisation involves the integration of all rules learned at different iterations into a whole rule set.

4 Applications of GrC

Sentiment analysis utilizes natural language processing to determine the tone of a digital text through the use of machine learning algorithms. One commonly used method for extracting features from textual data is the Bag-of-Words (BOW) approach. However, this traditional method may not be suitable for larger documents as it can result in complex and high-dimensional data. To address this issue, Liu and Cocea [15] propose a multi-granular approach where the instances of textual data are broken down into smaller sub-instances for local feature extraction. This approach is effective in handling the complexity of learning and classifying high-dimensional data. In ([15], Chapter 8), a fuzzy text classification approach is utilized, which has been shown to perform slightly better than traditional methods such as "Naive Bayes" and "C4.5."

Recently, Behzadidoost et al. [23] proposed a GrC-based deep learning model for text classification that boasts enhanced accuracy. This is due to the fact that deep learning models utilize automatic feature extraction within deep layers, making them more accurate than traditional machine learning models. In this method, numerical and textual granules are formed and used for training alongside raw training texts. The proposed model also employs a novel algorithm for constructing large textual and numerical granules, which are then fed into a variety of deep-learning models. This model utilizes granular computing for data augmentation.

In [24], Muhammad et al. utilize a granular computing framework for the detection of cardiovascular disease (CVD). In this framework, the attributes are information granules and z numbers are used for classification.

5 Concluding Remarks

The field of machine learning using a granular framework is currently a popular area of study. From large learning models to deep learning models, the concept of granular computation has proven to be beneficial in solving real-world problems. Numerous GrC-based models have been created and evaluated in recent years. It is likely that improved versions of the models discussed in this article will emerge as a new area of research. As technology continues to advance, the demand for GrC-based models will only continue to grow and find its way into several important fields.

Acknowledgements: The author would like to express gratitude to Dr. Sovan Samanta for their encouragement and support. Additionally, the author would like to thank the Department of Mathematics at Jai Hind College for providing the necessary infrastructure for conducting this study. The author would like to thank anonymous referee for their efforts towards scrutiny of this manuscript."

Conflict of Interest: The author declare no conflict of interest.

References

[1] Zadeh LA. Some reflections on soft computing, granular computing and their roles in the conception, design and utilization of information/intelligent systems. *Soft Computing.* 1998; 2(1): 23-25. DOI: https://doi.org/10.1007/s005000050030

[2] Chen D, Xu W, Li J. Granular computing in machine learning. *Granular Computing.* 2019; 4: 299-300. DOI: https://doi.org/10.1007/s41066-018-00146-2

[3] Mienye ID, Sun Y. A survey of ensemble learning: Concepts, algorithms, applications, and prospects. *IEEE Access.* 2022; 10: 99129-99149. DOI: https://doi.org/10.1109/ACCESS.2022.3207287

[4] Liu H, Zhang L. Advancing ensemble learning performance through data transformation and classifiers fusion in granular computing context. *Expert Systems with Applications.* 2019; 131: 20-29. DOI: https://doi.org/10.1016/j.eswa.2019.04.051

[5] Yao JT, Yao YY. Granular computing approach to machine learning. In: Wang L, Halgamuge SK, Yao X. (eds.) *Fsdk 2002: Proceedings of the 1st International Conference on Fuzzy Systems and Knowledge Discovery: Computational Intelligence for the E-Age, FSDK 2002, 18-22 November 2002, Singapore.* Orchid Country Club; 2002. p.732-736. https://www.semanticscholar.org/paper/A-Granular-Computing-Approach-to-Machine-Learning-Yao-Yao/0186028b4aaf265d8590dcb0890f900ee4095b0a

[6] Yao YY. On modeling data mining with granular computing. In: *25th Annual International Computer Software and Applications Conference. COMPSAC 2001, 08-12 October 2001, Chicago, IL, USA*. IEEE; 2001. p.638-643. DOI: https://doi.org/10.1109/CMPSAC.2001.960680

[7] Allison R. *Visualizing Data with SAS®: Selected Topics.* Cary NC USA: SAS Institute Inc; 2017. https://support.sas.com/content/dam/SAS/support/en/books/free-books/vds.pdf

[8] Laney D. *3D Data Management: Controlling Data Volume, Velocity, and Variety (). META Group* 2001. https://www.bibsonomy.org/bibtex/742811cb00b303261f79a98e9b80bf49 [Accessed 1st December 2024].

[9] Ishwarappa, Anuradha J. A brief introduction on Big Data 5Vs characteristics and Hadoop technology. *Procedia Computer Science.* 2015; 48: 319-324. DOI: https://doi.org/10.1016/j.procs.2015.04.188

[10] Webster M. *Granule: Merriam-Webster.com Dictionary.* https://www.merriam-webster.com/dictionary/granule [Accessed 26th October 2024].

[11] Keet CM. Abstraction. In: Dubitzky W, Wolkenhauer O, Cho KH, Yokota H. (eds.) *Encyclopedia of Systems Biology.* New York: Springer; 2013. p.3-4. DOI: https://doi.org/10.1007/978-1-4419-9863-7

[12] Lin TY. Granular Computing: Practices, Theories, and Future Directions. In: Meyers R. (eds.) *Encyclopedia of Complexity and Systems Science.* Springer, New York, NY; 2009. p.4339-4355. DOI: https://doi.org/10.1007/978-0-387-30440-3_256

[13] Hu H, Shi Z. Machine learning as granular computing. In: *2009 IEEE International Conference on Granular Computing, Nanchang, China.* IEEE; 2009. p.229-234. DOI: https://doi.org/10.1109/GRC.2009.5255125

[14] Zadeh LA. Toward a theory of fuzzy information granulation and its centrality in human reasoning and fuzzy logic. *Fuzzy Sets and Systems.* 1997; 90(2): 111-127. DOI: https://doi.org/10.1016/S0165-0114(97)00077-8

[15] Liu H, Cocea M. *Granular Computing Based Machine Learning: A Big Data Processing Approach.* Springer International Publishing AG; 2018. https://link.springer.com/book/10.1007/978-3-319-70058-8

[16] Tiwari VR. Developments in KD Tree and KNN searches. *International Journal of Computer Applications.* 2023; 185(17): 17-23. DOI: https://doi.org/10.5120/ijca2023922879

[17] Liu H, Gegov A, Cocea M. Nature and biology inspired approach of classification towards reduction of bias in machine learning. In: *2016 International Conference on Machine Learning and Cybernetics (ICMLC), 10-13 July 2016, Jeju, Korea (South).* IEEE; 2016. p.588-593. DOI: https://doi.org/10.1109/ICMLC.2016.7872953

[18] Liu H, Cocea M, Mohasseb A, Bader M. Transformation of discriminative single-task classification into generative multi-task classification in machine learning context. In: *2017 Ninth International Conference on Advanced Computational Intelligence (ICACI), 4–6 February 2017, Doha, Qatar.* IEEE; 2017. p.66-73. DOI: http://doi.org/10.1109/ICACI.2017.7974487

[19] Liu H, Cocea M. Fuzzy information granulation towards interpretable sentiment analysis. *Granular Computing.* 2017; 2: 289-302. DOI: https://doi.org/10.1007/s41066-017-0043-8

[20] Liu H, Cocea M. Semi-random partitioning of data into training and test sets in granular computing context. *Granular Computing.* 2017; 2: 357-386. DOI: https://doi.org/10.1007/s41066-017-0049-2

[21] Liu H, Cocea M. Granular computing-based approach of rule learning for binary classification. *Granular Computing.* 2019; 4: 275-283. DOI: https://doi.org/10.1007/s41066-018-0097-2

[22] Liu H, Gegov A, Cocea M. Collaborative rule generation: An ensemble learning approach. *Journal of Intelligent and Fuzzy Systems.* 2016; 30(4): 2277-2287. DOI: https://doi.org/10.3233/IFS-151997

[23] Behzadidoost R, Mahan F, Izadkhah H. Granular computing-based deep learning for text classification. *Information Sciences.* 2024; 119746. DOI: https://doi.org/10.1016/j.ins.2023.119746

[24] Khalif KMNK, Muhammad N, Aziz MKBM, Irawan MI, Iqbal M, Setiawan MN. Advancing machine learning for identifying cardiovascular disease via granular computing. *IAES International Journal of Artificial Intelligence.* 2024; 13(2): 2433-2440. DOI: https://doi.org/10.11591/ijai.v13.i2.pp2433-2440

Vijay R. Tiwari
Department of Mathematics
Assistant Professor of Jai Hind College, Jai Hind College
Churchgate, Mumbai 400020, India
E-mail: vijay.tiwari@jaihindcollege.edu.in