Tuijin Jishu/Journal of Propulsion Technology
ISSN: 1001-4055
Vol. 45 No. 1 (2024)

Integrating Disaster Data Clustering with Neural Networks for Comprehensive Analysis

[1] Dr. Pradeep Kumar N S, [2] B C Divakara, [3] Keerthy N [4] Dr. Girish H
[5] Mr Suhas S K, [6] Mrs. K Revathi

[1]*Professor, Department of ECE, SEA College of Engineering and Technology, Bangalore*

[2]*Assistant Professor, Department of ECE, Global Academy of Technology, Bangalore*

[3]*Assistant Professor, Department of ECE, Global Academy of Technology, Bengaluru*

[4]*Professor, Department of ECE, Cambridge Institute of Technology, Bengaluru*

[5]*Assistant Professor, Department of ECE, SEA College of Engineering and Technology, Bangalore*

[6]*Assistant Professor, Department of ECE SEA College of Engineering and Technology, Bangalore*

***Abstract*:** The rise in both natural and human-induced disasters has caused considerable harm to the inhabitants of our planet. These disasters, ranging from tsunamis and floods to earthquakes and forest fires, not only wreak havoc on property but also tragically claim human lives. With the increasing accessibility of large datasets and the growing prominence of parallel computing architectures, clustering algorithms are once again taking center stage. Spectral Clustering, in particular, has proven to outperform many traditional clustering methods. It transforms the clustering problem into a graph-partitioning challenge by treating each data point as a graph node. What sets Spectral Clustering apart is its ability to handle a broader range of clustering problems without imposing specific data assumptions. Despite its ease of implementation and computational efficiency, it can be time-consuming for dense datasets due to matrix construction and eigenvalue calculations. BIRCH (Balanced Iterative Reducing and Clustering using Hierarchies), an unsupervised data mining algorithm, is designed for hierarchical clustering in large datasets. A notable advantage of BIRCH is its adaptive and incremental clustering approach for multi-dimensional metric data points, aiming for optimal clustering quality within available resources. In the domain of disaster data analysis, neural networks play a pivotal role. An implemented neural network architecture for the forest fire dataset demonstrated impressive accuracy (97%) with minimal loss. This underscores the significant contribution of neural networks in addressing challenges related to disaster data analysis

Keywords: *Neural Networks, Data Clustering, BIRCH Clustering, Spectral Clustering.*

1. Introduction

The Disaster management stands as a critical discipline dedicated to lessening the impact of both natural and human-induced disasters on communities and environments [1]. Within the realm of disaster management, a significant obstacle lies in deciphering the inherent patterns and structures embedded in disaster-related data, encompassing details about previous incidents, geographical information, weather patterns, and more [2]. Spectral Clustering emerges as a sophisticated data analysis technique with the potential to enhance disaster management endeavors. This method, rooted in spectral graph theory, excels in categorizing data points into clusters based on their similarities [3]. Diverging from conventional clustering algorithms, Spectral Clustering doesn't assume predefined shapes or distributions of clusters, making it especially adept at handling intricate and irregular disaster-related datasets [4]. The approach of Spectral Clustering involves interpreting the data as a graph, where data points serve as nodes, and pairwise similarities between these points are represented as edges. This graph-based representation facilitates the discovery of meaningful clusters within the data [5]. The versatility of Spectral Clustering allows it to adapt to various forms of disaster-related data, be it spatial, time-series, or multi-dimensional, demonstrating its utility in analyzing the diverse datasets commonly encountered in disaster management.

In the sphere of disaster management, Spectral Clustering plays a pivotal role in pinpointing communities or clusters of data points sharing common characteristics or exhibiting similar patterns. This capability proves crucial for comprehending the spatial distribution of disaster events or the underlying behaviors leading to such incidents. Moreover, Spectral Clustering serves as a valuable tool for dimensionality reduction, essential for visualizing and interpreting high-dimensional disaster data, simplifying complexity while retaining meaningful structural information [6]. The integration of Spectral Clustering with Geographic Information System (GIS) tools and geographical data enables a detailed analysis of the spatial aspects of disaster events. This integration empowers disaster management professionals to make well-informed decisions regarding resource allocation, evacuation planning, and risk assessment. Under the umbrella of data-driven insights, the application of Spectral Clustering to disaster-related datasets yields valuable understanding of underlying patterns and relationships. This knowledge, in turn, informs strategies for disaster preparedness, response, and recovery efforts [7].

.**BIRCH Clustering:**

The BIRCH, an acronym for Balanced Iterative Reducing and Clustering using Hierarchies, represents a hierarchical clustering algorithm meticulously crafted for the efficient analysis of vast datasets. It emerged as a solution to the intricate challenges posed by clustering extensive data volumes while maintaining a delicate balance between memory usage and computational efficiency. BIRCH proves its mettle in domains such as data mining, pattern recognition, and exploratory data analysis.Within the realm of BIRCH clustering, fundamental concepts and advantages unfold. The algorithm constructs a hierarchical cluster structure where each level of the hierarchy encapsulates clusters of varying granularity. This hierarchical paradigm facilitates a nuanced exploration of data, ranging from broad clusters to more detailed subclusters.

A notable strength of BIRCH lies in its capability for incremental clustering. This means that it can adeptly process and cluster incoming data points without necessitating a reprocessing of the entire dataset. Such adaptability renders BIRCH suitable for scenarios involving streaming data or continuous influxes of new data.BIRCH operates within the feature space, a multi-dimensional realm defined by the attributes of data points. Employing a tree-like structure known as the "Clustering Feature Tree" (CF Tree), BIRCH efficiently organizes and represents clusters within this feature space. The CF Tree serves as the nucleus of BIRCH, storing statistical information about clusters and their data points, such as cluster size, mean, and variance, facilitating swift identification of potential cluster candidates. Maintaining a balanced tree structure is a distinctive feature of BIRCH, ensuring that the number of subclusters at each hierarchy level remains roughly constant. This equilibrium contributes to manageable memory requirements and computational complexity, even for extensive datasets. BIRCH's scalability shines through, enabling it to efficiently handle large datasets. Its incremental clustering prowess proves particularly advantageous in scenarios involving data streams or dynamic datasets. By condensing a concise summary of data distribution through the CF Tree, BIRCH minimizes memory demands, rendering it suitable for systems constrained by limited memory resources. Renowned for its speed, BIRCH relies on simple mathematical calculations, facilitating rapid processing of data points. This attribute makes BIRCH an ideal choice for applications requiring real-time or near-real-time clustering. The hierarchical architecture of BIRCH offers users the ability to traverse data at various levels of granularity, facilitating the discovery of meaningful patterns or clusters within the dataset.

1.1 Neural Networks for Deep Learning

In the realm of disaster management and response, neural networks have emerged as a formidable tool, wielding the capacity to process and scrutinize extensive datasets from diverse sources. This empowers more astute decision-making, facilitates the implementation of early warning systems, and refines disaster response strategies. Delving into the application of neural networks in deep learning for disaster management, let's explore the following facets: Prediction of Natural Disasters: Neural networks exhibit proficiency in analyzing historical weather data, satellite imagery, seismic activity, and environmental factors, contributing to the prediction of natural disasters such as hurricanes, tsunamis, and earthquakes. This predictive capability offers crucial lead time for evacuation and preparedness. Fire and Flood Prediction: Neural networks prove instrumental in processing data from sensors, weather stations, and remote sensing satellites to predict and monitor wildfires, floods, and landslides. Convolutional Neural Networks (CNNs) excel in assessing the extent of damage caused by disasters through the analysis of aerial and satellite images, guiding resource allocation and response prioritization.

Search and Rescue Operations: Drones equipped with neural networks become invaluable assets in search and rescue missions, identifying survivors in disaster-stricken areas through the analysis of image or video data. Real-time Information Gathering: Natural Language Processing (NLP) models come into play as they process social media posts and news articles in real-time, providing critical information about disaster events, identifying affected areas, and gauging public sentiment. This real-time data proves indispensable for coordinating emergency responses. Supportive Chatbots and Virtual Assistants: Chatbots and virtual assistants powered by NLP offer information and support to affected individuals during and after disasters, aiding in the dissemination of crucial information and resources.

Anomaly Detection with IoT Sensors: Neural networks analyze data from Internet of Things (IoT) sensors in disaster-prone areas, detecting anomalies, monitoring environmental conditions, and triggering alerts when unusual patterns emerge. Optimization of Resource Allocation: Neural networks play a pivotal role in optimizing the allocation of resources, including food, water, medical supplies, and manpower during disaster response. This ensures efficient distribution to affected areas. Risk Assessment and Infrastructure Planning: Neural networks assist in evaluating the vulnerability and risk levels of different regions, enabling urban planners and policymakers to make informed decisions about disaster-resistant infrastructure. Simulation of Disaster Scenarios: Neural networks simulate various disaster scenarios to predict the impact on infrastructure, transportation, and public health, contributing to preparedness and mitigation efforts. Communication Network Optimization: Neural networks optimize communication networks to ensure their functionality during disasters, facilitating effective coordination among first responders. Post-Disaster Recovery Analysis: Neural networks continue to play a role in post-disaster recovery by analyzing data related to rebuilding efforts, resource distribution, and long-term impact assessment..

2. Literature Survey

In the paper titled "Early Forest Fire Region Segmentation Based on Deep Learning" authored by Guangyi Wang, Youmin Zhang, and collaborators, the focus is on overcoming the challenge of detecting small forest fire areas that evade identification through traditional methods. The authors propose an innovative forest fire monitoring framework founded on convolutional neural networks (CNNs). To gauge the efficacy and precision of this framework in early forest fire detection, numerous experiments were conducted using a self-generated forest fire dataset and actual forest fire monitoring videos. The outcomes of these experiments illustrate the framework's adeptness in functioning effectively under diverse challenging conditions related to fire and illumination, ultimately demonstrating its prowess in the early detection of forest fires. The study's discoveries hold substantial promise for enhancing forest fire monitoring and response initiatives.

The paper titled "Classification Algorithms Analysis in the Forest Fire Detection Problem" by Mikhail D. Molovtsev, Irina S. Sineva, and their colleagues, delves into the application of diverse machine learning algorithms for detecting forest fires and non-fire conditions. Utilizing data from the national Park of Montesino in the North of Porto Galia, the study concentrates on four machine learning algorithms: Support Vector Machine (SVM), Random Forest, Logistic Regression, and Decision Tree. These algorithms undergo binary classification to differentiate between fire and non-fire instances. The paper evaluates their performance by constructing an error matrix for each classifier, which serves as the foundation for calculating fundamental metrics that assess the efficiency and accuracy of the classification methods. Summary of the Paper "Attribute Profiles in Earthquake Damage Identification from Very High-Resolution Post-Event Image":

In the paper titled "Attribute Profiles in Earthquake Damage Identification from Very High-Resolution Post-Event Image," crafted by Enes O˘guzhan Alatas¸, G¨uls¸en Tas¸kın, and others, the challenge of accurately assessing earthquake damage through very high-resolution (VHR) satellite images takes center stage. Departing from traditional damage assessment methods reliant on spectral information, the paper underscores the significance of incorporating contextual relationships between pixels to elevate classification accuracy. Introducing the use of Attribute Profiles (APs) and Multi Attribute Profiles (MAPs), the study creates a multi-dimensional representation of an image by sequentially applying various attribute filters. This innovative approach generates complex features capturing specific earthquake-induced damage patterns. The authors, for the first time, apply APs and MAPs to extract additional contextual features from a VHR satellite image of the City of Bam in Iran, captured eight days after an earthquake. In assessing the performance of these morphological attribute features, a comparison with Haralick's features (HFs) using the k-nearest neighbors (k-NN) classifier is conducted. Preliminary findings suggest that APs and MAPs outperform HFs in detecting earthquake damage with greater accuracy..

3. Overview

The surge in both natural and human-induced disasters has inflicted substantial harm on our planet's inhabitants, ranging from property damage to tragic loss of human lives. Tsunamis, floods, earthquakes, and forest fires are among the devastating events that continue to pose threats globally. With the increasing accessibility of large datasets and the growing prominence of parallel computing architectures, clustering algorithms have once again taken the spotlight. Spectral Clustering, in particular, has emerged as a standout performer, transforming the clustering problem into a graph-partitioning challenge by treating each data point as a graph node. What distinguishes Spectral Clustering is its remarkable ability to handle a diverse range of clustering problems without imposing specific data assumptions. Despite its ease of implementation and computational efficiency, dense datasets may pose a time challenge due to the intricacies of matrix construction and eigenvalue calculations. Enter BIRCH (Balanced Iterative Reducing and Clustering using Hierarchies), an unsupervised data mining algorithm tailored for hierarchical clustering in large datasets. BIRCH stands out for its adaptive and incremental clustering approach for multi-dimensional metric data points, striving for optimal clustering quality within available resources. In the realm of disaster data analysis, neural networks assume a crucial role. An implemented neural network architecture, particularly for the forest fire dataset, showcased impressive accuracy (97%) with minimal loss. This underscores the significant contribution of neural networks in effectively addressing challenges associated with disaster data analysis.

The evolving landscape of both natural and human-induced disasters has underscored the need for a comprehensive approach to understanding and managing the associated data. "Integrating Disaster Data Clustering with Neural Networks for Comprehensive Analysis" explores a synergistic approach to address this challenge. The rise in disasters, from tsunamis and floods to earthquakes and forest fires, has prompted a closer examination of large datasets and the utilization of advanced computational architectures. This paper delves into the fusion of two powerful analytical techniques: disaster data clustering and neural networks. The exploration begins with an acknowledgment of the devastating impact of disasters, not only on property but tragically on human lives. As the accessibility of large datasets increases and parallel computing architectures become more prominent, clustering algorithms once again take center stage. Spectral Clustering is highlighted as a noteworthy method, particularly for its ability to handle a broad range of clustering problems without imposing stringent data assumptions. However, the potential time constraints for dense datasets are recognized due to matrix construction and eigenvalue calculations. The paper introduces BIRCH (Balanced Iterative Reducing and Clustering using Hierarchies), an unsupervised data mining algorithm specifically designed for hierarchical clustering in large datasets. Notable for its adaptive and incremental clustering approach, BIRCH aims to achieve optimal clustering quality within available resources.

In the context of disaster data analysis, the role of neural networks is emphasized. The implementation of a neural network architecture for a forest fire dataset is showcased, demonstrating impressive accuracy (97%) with minimal loss. This illustrates the substantial contribution of neural networks in addressing the challenges inherent in disaster data analysis. The overarching goal of the paper is to present a holistic and integrated approach, combining the strengths of disaster data clustering and neural networks. By merging these analytical techniques, the aim is to provide a comprehensive framework for understanding, processing, and responding to the complexities of disaster-related datasets. The potential implications of this integrated approach are discussed, with a focus on enhancing the efficiency and effectiveness of disaster management and response strategies.

4. Methodology

The Algorithm and Methodology: Steps

Spectral Clustering, an increasingly favored clustering algorithm outperforming many traditional counterparts, adopts a unique approach by treating each data point as a graph-node. This transforms the clustering problem into a graph-partitioning challenge, with a typical implementation involving three fundamental steps.

Constructing the Similarity Graph:

In this initial step, the Similarity Graph takes shape as an adjacency matrix denoted by A.

Two primary methods for building the graph include:

Epsilon-neighbourhood Graph: This method involves fixing a parameter epsilon in advance. Subsequently, each point connects to all points within its epsilon-radius. In cases where distances between any two points are similar in scale, edge weights, representing the distance, may not be stored as they do not provide additional information. This results in an undirected and unweighted graph.

K-Nearest Neighbours: Here, a pre-fixed parameter k guides the process. An edge is directed from vertex u to v only if v is among the k-nearest neighbors of u. Note that this leads to a weighted and directed graph.

To convert it to an undirected form, two approaches are commonly used:

Direct an edge from u to v and from v to u if either v is among the k-nearest neighbors of u OR u is among the k-nearest neighbors of v.

Direct an edge from u to v and from v to u if v is among the k-nearest neighbors of u AND u is among the k-nearest neighbors of v.

Fully-Connected Graph: This type of graph connects each point with an undirected edge weighted by the distance between the two points. The Gaussian similarity metric is typically used to calculate distances, as this approach models local neighborhood relationships.

The second step involves projecting the data into a lower-dimensional space. This reduction in dimensionality accounts for situations where members of the same cluster may be widely dispersed in the original dimensional space. The reduction is achieved by computing the Graph Laplacian Matrix. To calculate this matrix, the degree

$$d_i = \sum_{j=1|(i,j)\epsilon E}^{n} w_{ij}$$

of each node must first be defined. The degree of the ith node is determined by...

Note that w_{ij} is the edge between the nodes I and j as defined in the adjacency matrix above.

The degree matrix is defined as follows:

$$D_{ij} = \begin{cases} d_i, i = j \\ 0, i \neq j \end{cases}$$

Thus the Graph Laplacian Matrix is defined as: L=D-A

Afterward, this matrix is normalized to improve mathematical efficiency. To reduce dimensions, the next step involves computing eigenvalues and their corresponding eigenvectors. If the desired number of clusters is denoted as k, the top k eigenvalues and their corresponding eigenvectors are extracted and organized into a matrix, with the eigenvectors serving as the columns.

Moving on to the clustering phase: The primary objective here is to cluster the dimensionally reduced data using a conventional clustering method, typically the K-Means Clustering algorithm. Initially, each node is assigned a row from the normalized Graph Laplacian Matrix. Subsequently, the clustered data is processed using a traditional technique. The node identifiers are preserved to facilitate the transformation of the clustering results.

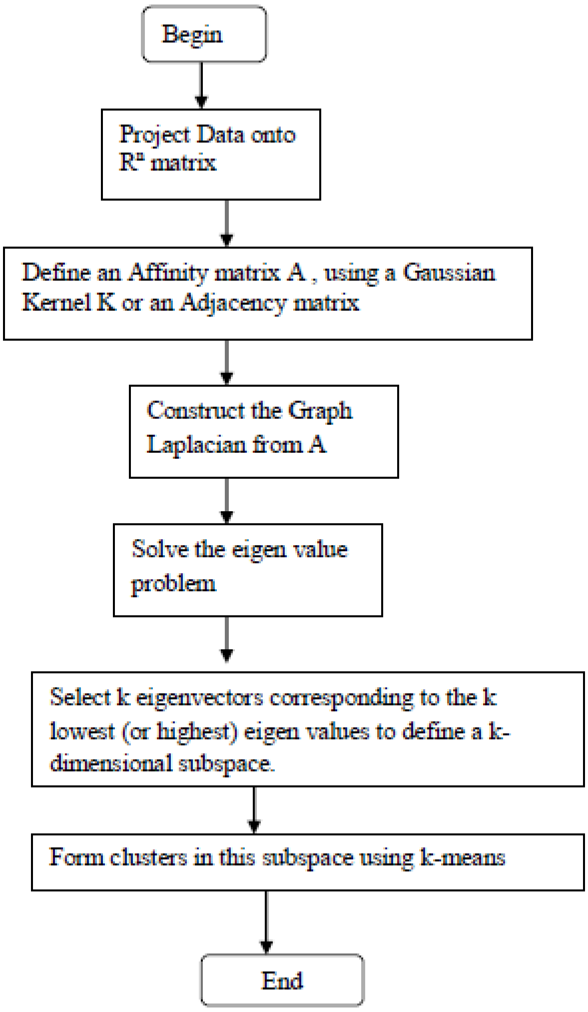


Fig 1: Flowchart of Spectral Clustering

4.2 BIRCH Clustering

Algorithm and Methodology: Steps:

Phase 1:

The algorithm initiates with an initial threshold value and proceeds to scan the data, inserting points into the tree. Should the algorithm run out of memory before completing the scan, it adjusts the threshold value upwards. Subsequently, a new, smaller CF-tree is constructed by re-inserting the leaf entries from the previous CF-tree into the new one. Once all the old leaf entries have been successfully re-inserted, the scanning and insertion into the new CF-tree resume from where they were paused. Optimal selection of the initial threshold value significantly minimizes the need for rebuilding the tree. However, setting the initial threshold too high may result in a less detailed CF-tree compared to what could be achieved with the available memory.

Alternatively, it is possible to allocate a fixed amount of disk space to manage outliers. Outliers are leaf entries characterized by low density, which, in the context of overall clustering, are deemed unimportant.

During CF-tree rebuilding, the size of the new CF-tree is reduced in two ways. First, the threshold value is increased, allowing each leaf entry to accommodate more data points. Second, certain leaf entries are designated as potential outliers and are stored on disk. An old leaf entry qualifies as a potential outlier if it contains significantly fewer data points than the average. A change in the threshold value or data distribution may render these potential outliers no longer outliers. Consequently, these potential outliers are reviewed to determine if they can be reintegrated into the tree without increasing its size.

Phase 2:

Moving on to Phase 2: Recognizing that specific clustering algorithms exhibit optimal performance within a particular range of object numbers, crowded subclusters are amalgamated into larger ones, resulting in an overall reduction in the size of the CF-tree.

Phase 3:

In Phase 3: Virtually any clustering algorithm can be adapted to classify Clustering Features instead of individual data points.

Phase 4:

Phase 4 marks a departure from the previous stages where the tree may have undergone multiple rebuilds but the original data was scanned only once. Phase 4 entails additional passes over the data to rectify inaccuracies arising from the application of the clustering algorithm to a coarse data summary. Additionally, Phase 4 provides the option to discard outliers.

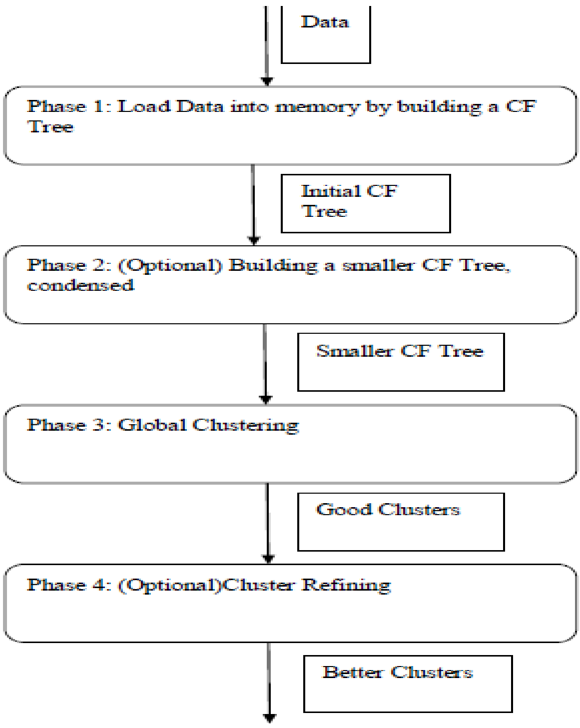


Fig 2 : Flow chart of BIRCH Clustering

4.3 Neural Networks for deep learning

Neural networks are typically structured into layers, each comprising interconnected 'nodes' housing an 'activation function'. Information flows through the network, starting from the 'input layer,' proceeding through one or more 'hidden layers' for processing, and ultimately reaching the 'output layer' for generating a response. In neural networks, organization is achieved through layers. These layers consist of interconnected 'nodes,' each equipped with an 'activation function.' Data patterns are introduced into the network through the 'input layer,' which then propagates information through one or more 'hidden layers'

Tuijin Jishu/Journal of Propulsion Technology
ISSN: 1001-4055
Vol. 45 No. 1 (2024)

---

responsible for complex processing. Finally, the result emerges from the 'output layer.'

The fundamental structure of neural networks involves layers, with each layer consisting of interconnected 'nodes' possessing their respective 'activation functions.' Information is initially input through the 'input layer,' and it subsequently undergoes processing across one or more 'hidden layers.' The final output is produced by the 'output layer.'

One of the most well-known neural network architectures is the feed-forward multilayer neural network. It comprises an 'input layer,' one or more 'hidden layers,' and a solitary 'output layer.' These layers can vary in the number of neurons they contain, and each layer establishes full connectivity with its neighboring layers.

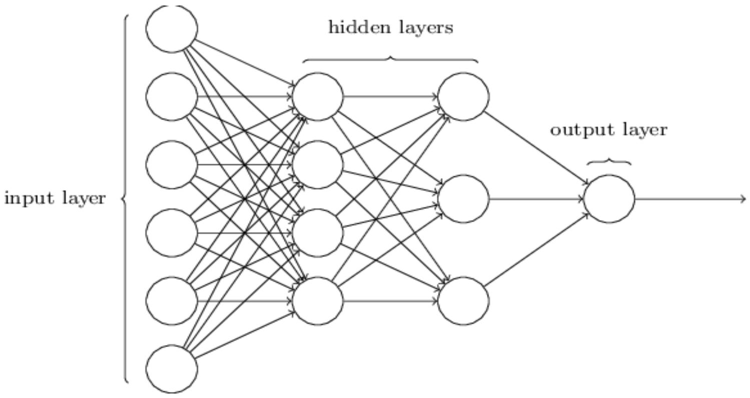


Fig 3 Neural Network Architecture

A neural network structure is employed to analyze the forest fire dataset with the aim of achieving high accuracy. To obtain clusters for the datasets below using BIRCH and Spectral Clustering Data Mining algorithms Spectral Clustering Input dataset:

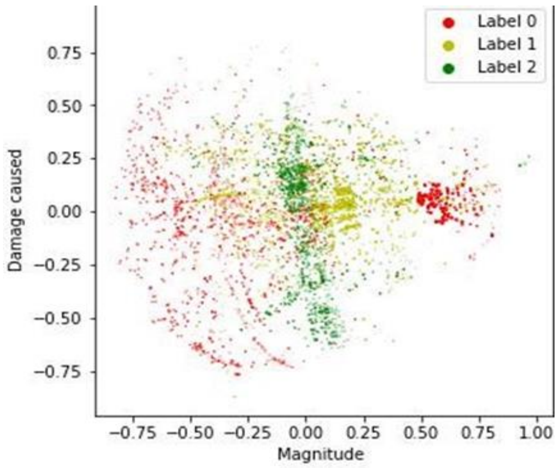
Earthquake dataset Link:https://www.kaggle.com/glorykim/earthquake-building-damage-data.

```
X = pd.read_csv(io.BytesIO(uploaded['earthquake.csv']))
print(X)

       Unnamed: 0 struct_typ  occ_type  ...  magnitude  distance  meandamage
0               0        URM         1  ...          5  8.758655     0.13637
1               1         W1         2  ...          5  8.750071     0.13302
2               2        URM         1  ...          5  8.761617     0.18070
3               3        URM         1  ...          5  8.767489     0.14398
4               4        URM         1  ...          5  8.764526     0.17911
...           ...        ...       ...  ...        ...       ...         ...
11087       11766         S1         3  ...          5  8.783250     0.09335
11088       11767         S1        10  ...          5  8.720425     0.07537
11089       11768         C4         1  ...          5  8.882139     0.21389
11090       11769         S1         3  ...          5  8.909023     0.10686
11091       11770         S1         3  ...          5  8.724296     0.11963

[11092 rows x 8 columns]
```

Fig4: Description of dataset



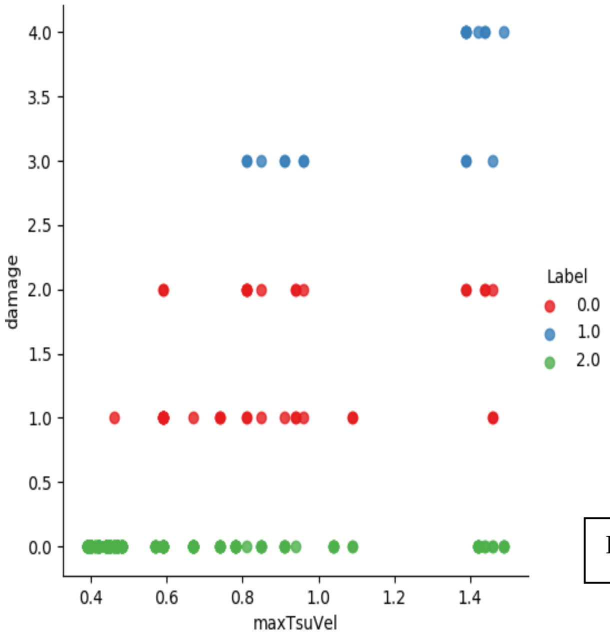
Link:https://www.kaggle.com/huseyincot/inlandsea-tsunami-damage

Fig5: Spectral Clustering clusters

```
     berthing_place  gauge_point  depth  ...  maxTsuH  maxTsuVel  damage
0              bp01            1  17.68  ...     0.39       0.59       0
1              bp01            1  17.68  ...     0.39       0.59       0
2              bp01            1  17.68  ...     0.39       0.59       0
3              bp01            1  17.68  ...     0.39       0.59       0
4              bp01            1  17.68  ...     0.39       0.59       0
..              ...          ...    ...  ...      ...        ...     ...
300            bp16           36   1.24  ...     1.49       1.37       0
301            bp16           36   1.24  ...     1.49       1.37       0
302            bp16           36   1.24  ...     1.49       1.37       0
303            bp16           36   1.24  ...     1.49       1.37       0
304            bp16           36   1.24  ...     1.49       1.37       4

[305 rows x 9 columns]
```

Fig 6: Description of dataset



Label 0:Severe Damage:To Take action
Label 1:Less Damage caused Label 2:Moderate Damage

Fig 7: Spectral Clustering clusters

Spectral Clustering Input dataset: FOREST FIRE dataset

Link to dataset: https://www.kaggle.com/sumitm004/forest-fire-area)

The clusters are computed taking 13 attributes of the dataset for both the clustering algorithms. The clusters are formed on the basis of similarity between the attributes

```
import pandas as pd
import io

X = pd.read_csv(io.BytesIO(uploaded['forestfires_1.csv']))
print(X)

     X  Y  month  day  FFMC  DMC   DC  ISI  temp  RH  wind  rain   area
0    7  5      3    6  86.0   26   94    5   8.0  51     7   0.0   0.00
1    7  4     10    3  91.0   35  669    7  18.0  33     1   0.0   0.00
2    7  4     10    7  91.0   44  687    7  15.0  33     1   0.0   0.00
3    8  6      3    6  92.0   33   78    9   8.0  97     4   0.2   0.00
4    8  6      3    1  89.0   51  102   10  11.0  99     2   0.0   0.00
..  .. ..    ...  ...   ...  ...  ...  ...   ...  ..   ...   ...    ...
512  4  3      8    1  81.6   57  666    2  27.8  32     3   0.0   6.44
513  2  4      8    1  81.6   57  666    2  21.9  71     6   0.0  54.29
514  7  4      8    1  81.6   57  666    2  21.2  70     7   0.0  11.16
515  1  4      8    7  94.4  146  615   11  25.6  42     4   0.0   0.00
516  6  3     11    3  79.5    3  107    1  11.8  31     5   0.0   0.00

[517 rows x 13 columns]
```

Fig 8:Description of dataset

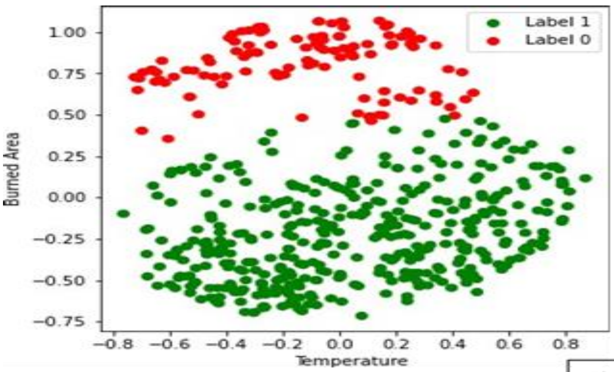
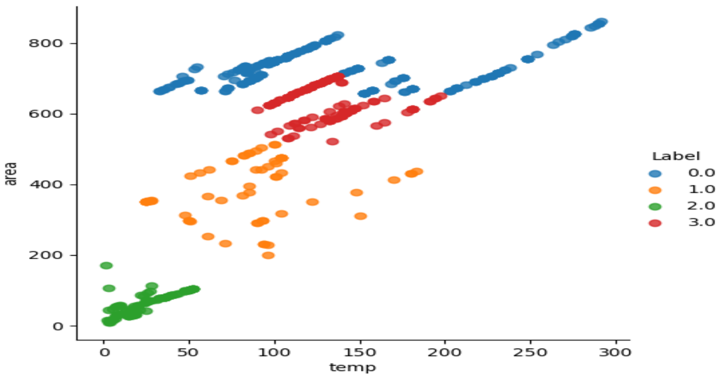


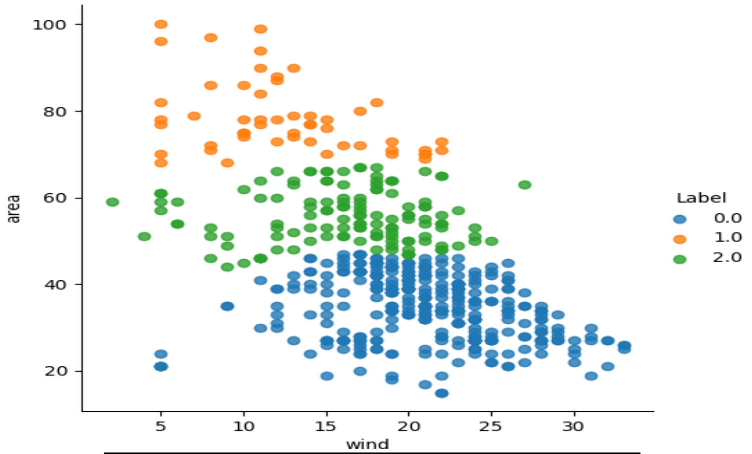
Fig 9:Spectral Clustering clusters

BIRCH Clustering Results

Link to dataset: https://www.kaggle.com/sumitm004/forest-fire-area)

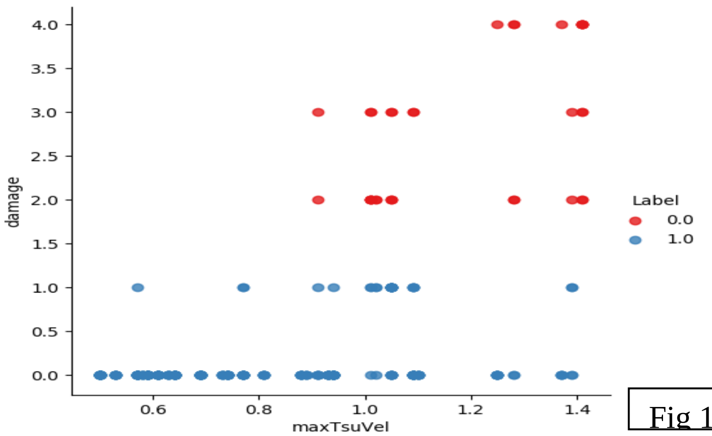


As temperature and RH(Relative Humidity)increases forest firesrise at specific area
Label 0, Label 3 :Severe Areaburnt
Label1 :Less burnt area Label 2 : Very Less Areaburnt



As the wind speed increasesforest fire burnt area too increases.
Label 0 :Severe Area burntLabel1 :Less burnt area Label 2 : Very Less Area burnt

Fig 10,11: BIRCH Clustering clusters



As the Velocity increases ofTsunami damage is severe Label 0: Severe damage Label 1:Less damage

Fig 12: BIRCH Clustering clusters

BIRCH Clustering Input dataset: Earthquake dataset

Link:https://www.kaggle.com/glorykim/earthquake-building-damage-data

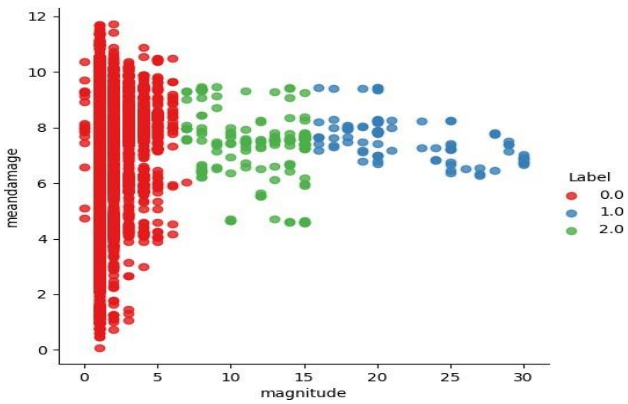


Fig 13:BIRCH Clustering clusters

Label 0:Severe Damage:ToTake action

Label 1:Moderate Damage,Beprepared

Label 2:Less Damage caused

Neural Network results(forest fire dataset): Input:Forestfire dataset

Method:Neural Network model Output: Accuracy and loss calculation

```
361/361 [==============================] - 0s 72us/step - loss: 0.1112 - accuracy: 0.9584
Epoch 187/200
361/361 [==============================] - 0s 72us/step - loss: 0.0727 - accuracy: 0.9640
Epoch 188/200
361/361 [==============================] - 0s 69us/step - loss: 0.0770 - accuracy: 0.9751
Epoch 189/200
361/361 [==============================] - 0s 72us/step - loss: 0.0773 - accuracy: 0.9723
Epoch 190/200
361/361 [==============================] - 0s 72us/step - loss: 0.0680 - accuracy: 0.9806
Epoch 191/200
361/361 [==============================] - 0s 72us/step - loss: 0.0836 - accuracy: 0.9640
Epoch 192/200
361/361 [==============================] - 0s 72us/step - loss: 0.0636 - accuracy: 0.9834
Epoch 193/200
361/361 [==============================] - 0s 72us/step - loss: 0.0885 - accuracy: 0.9723
Epoch 194/200
361/361 [==============================] - 0s 69us/step - loss: 0.0766 - accuracy: 0.9695
Epoch 195/200
361/361 [==============================] - 0s 72us/step - loss: 0.0800 - accuracy: 0.9640
Epoch 196/200
361/361 [==============================] - 0s 75us/step - loss: 0.0592 - accuracy: 0.9723
Epoch 197/200
361/361 [==============================] - 0s 75us/step - loss: 0.0681 - accuracy: 0.9751
Epoch 198/200
361/361 [==============================] - 0s 72us/step - loss: 0.0681 - accuracy: 0.9751
Epoch 199/200
361/361 [==============================] - 0s 75us/step - loss: 0.0701 - accuracy: 0.9723
Epoch 200/200
361/361 [==============================] - 0s 83us/step - loss: 0.0792 - accuracy: 0.9723
```

Fig 14: Output of Neural Network for Forest Fire dataset

5. Conclusive

A comprehensive examination of BIRCH and Spectral Clustering techniques is conducted. The algorithm's performance is assessed across various datasets, and the resulting clusters are computed. BIRCH is recognized for its locality, as it makes clustering decisions without the need to scan all data points or consider pre-existing clusters. It leverages the non-uniform distribution of data points and assigns varying importance to each data point while optimizing memory usage for the creation of sub-clusters. Additionally, it operates incrementally, negating the need for the entire dataset in advance. Spectral clustering, rooted in graph theory, excels at identifying node communities within graphs based on their interconnections. Notably, this approach extends its flexibility to non-graph data clustering. The application of Neural Networks to the forest fire dataset yields impressive accuracy results.

References

[1] G. Eason, B. Noble, and I. N. Sneddon, "On certain integrals of Lipschitz-Hankel type involving products of Bessel functions," Phil. Trans. Roy. Soc. London, vol. A247, pp. 529–551, April 1955. (references) M. Sardogan, A. Tuncer and Y. Ozen, "Plant Leaf Disease Detection and Classification Based on CNN with LVQ Algorithm," 2018 3rd International Conference on Computer Science and Engineering (UBMK), 2018, pp. 382-385.

[2] D. Bini, D. Pamela and S. Prince, "Machine Vision and Machine Learning for Intelligent Agrobots: A review" 2020 5th International Conference on Devices, Circuits and Systems (ICDCS), 2020, pp. 12- 16.

[3] G. M. Sharif Ullah Al- Mamun et al., "Performance Analysis of Multipurpose AGROBOT," 2019 IEEE International WIE Conference on Electrical and Computer Engineering (WIECON-ECE), 2019, pp. 1-4.

[4] H Durmuş, E. O. Güneş, M. Kırcı and B. B. Üstündağ, "The design of general purpose autonomous agricultural mobile-robot: "AGROBOT"," 2015 Fourth International Conference on Agro- Geoinformatics (Agro- geoinformatics),2015,pp.49-53.

[5] N G. Kurale and M. V. Vaidya, "Classification of Leaf Disease Using Texture Feature and Neural Network Classifier," 2018 International Conference on Inventive Research in Computing Applications

(ICIRCA),2018,pp. 1-6.

[6] H. Girish, T. G. Manjunat and A. C. Vikramathithan, "Detection and Alerting Animals in Forest using Artificial Intelligence and IoT," 2022 IEEE Fourth International Conference on Advances in Electronics, Computers and Communications (ICAECC), 2022, pp. 1-5, doi: 10.1109/ICAECC54045.2022.9716679.

[7] A. Devipriya, H. Girish, V. Srinivas, N. Adi Reddy and D. Rajkiran, "RTL Design and Logic Synthesis of Traffic Light Controller for 45nm Technology," 2022 3rd International Conference for Emerging Technology (INCET), 2022, pp. 1-5, doi: 10.1109/INCET54531.2022.9824833.

[8] S. Vaddadi, V. Srinivas, N. A. Reddy, G. H, R. D and A. Devipriya, "Factory Inventory Automation using Industry 4.0 Technologies," 2022 IEEE IAS Global Conference on Emerging Technologies (GlobConET), 2022, pp. 734-738, doi: 10.1109/GlobConET53749.2022.9872416.

[9] T G Manjunath , A C Vikramathithan , H Girish, “Analysis of Total Harmonic Distortion and implementation of Inverter Fault Diagnosis using Artificial Neural Network”, Journal of Physics: Conference Series, Volume 2161, 1st International Conference on Artificial Intelligence, Computational Electronics and Communication System (AICECS 2021) 28-30 October 2021, Manipal, India. https://iopscience.iop.org/issue/1742-6596/2161/1

[10] GIRISH H , “Internet of Things Based Heart Beat Rate Monitoring System”, © September 2022 | IJIRT | Volume 6 Issue 4 | ISSN: 2349-6002 IJIRT 156592 INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH IN TECHNOLOGY 227

[11] Dr Girish H, Dr. Mangala Gowri , Dr. Keshava Murthy , Chetan Naik J, Autonomous Car Using Deep Learning and Open CV, PAGE NO: 107-115, VOLUME 8 ISSUE 10 2022, Gradiva review journal, ISSN NO : 0363-8057, DOI:10.37897.GRJ.2022.V8I8.22.50332

[12] Girish H. "Intelligent Traffic Tracking System Using Wi-Fi ." International Journal for Scientific Research and Development 8.12 (2021): 86-90.

[13] Girish H, Shashikumar D R, “Emission monitoring system using IOT”, Dogo Rangsang Research Journal, UGC Care Group I Journal, ISSN: 347-7180, Vol – 8 Issue-14, No.6, 2020.

[14] Girish H, Shashikumar D R, “Display and misson computer software loading”, International journal of engineering research & Technology (IJERT), ISSN: 2278-0181, Vol – 10 Issue-08, No.13, August 2020.

[15] Girish H, Shashikumar D R, “A Novel Optimization Framework for Controlling Stabilization Issue in Design Principle of FinFET based SRAM”, International Journal of Electrical and Computer Engineering (IJECE) Vol. 9, No. 5 October 2019, pp. 4027~4034. ISSN: 2088-8708, DOI: 10.11591/ijece.v9i5.pp.4027-4034

[16] Girish H, Shashikumar D R, “PAOD: a predictive approach for optimization of design in FinFET/SRAM”, International Journal of Electrical and Computer Engineering (IJECE) Vol. 9, No. 2, April 2019, pp. 960~966. ISSN: 2088-8708, DOI: 10.11591/ijece.v9i2.pp.960-966

[17] Girish H, Shashikumar D R, “SOPA: Search Optimization Based Predictive Approach for Design Optimization in FinFET/SRAM”, © Springer International Publishing AG, part of Springer Nature 2019 Silhavy (Ed.): CSOC 2018, AISC 764, pp. 21–29, 2019. https://doi.org/10.1007/978-3-319-91189-2_3.

[18] Girish H, Shashikumar D R, “Cost-Effective Computational Modelling of Fault Tolerant Optimization of FinFET-based SRAM Cells”, © Springer International Publishing AG 2017 R. Silhavy et al. (eds.), Cybernetics and Mathematics Applications in Intelligent Systems, Advances in Intelligent Systems and Computing 574, DOI 10.1007/978-3-319-57264-2_1.

[19] Girish H, Shashikumar D R, “A Survey on the Performance Analysis of FinFET SRAM Cells for Different Technologies”, International Journal of Engineering and Advanced Technology (IJEAT) ISSN: 2249 – 8958, Volume-4 Issue-6, 2016.

[20] Girish H, Shashikumar D R, “Insights of Performance Enhancement Techniques on FinFET-based SRAM Cells”, Communications on Applied Electronics (CAE) – ISSN: 2394-4714, Foundation of Computer Science FCS, New York, USA .Volume 5 – No.6, July 2016 – www.caeaccess.org

[21] Girish H, Shashikumar D R, “DESIGN OF FINFET”, International Journal of Engineering Research ISSN: 2319-6890) (online), 2347-5013(print) Volume No.5 Issue: Special 5, pp: 992-1128, doi: 10.17950/ijer/v5i5/013 2016.

[22] Dr. MANGALA GOWRI S G, Dr. SHASHIDHAR T M, Dr. SUNITHA R, SHYLAJA V, Dr. GIRISH H, “DESIGN OF 3-BIT CMOS WALLACE MULTIPLIER” , Seybold report, ISSN 1533-9211, Volume 18, Page No: 1260-1271 DOI: 10.5281/zenodo.8300481

[23] Dr. Mangala Gowri S G, Dr. Girish H, Ramesh N, Dr. Nataraj Vijaypur, “IOT based plant monitoring system and smart irrigation using new features” ResMilitaris, vol.13, n°2, January Issue 2023 https://resmilitaris.net/menu-script/index.php/resmilitaris/article/view/3312/2608

[24] Girish H.2023, Smart Theft Securityvehicular System Using Iot. Int J Recent Sci Res. 14(02), pp. 2881-2884. DOI: http://dx.doi.org/10.24327/ijrsr.2023.1402.0591

[25] Dr. Mangala Gowri S G, Dr. Girish H, Dr. Santosh Dattatray Bhopale, Weed Detection based on Neural Networks, International Journal of All Research Education and Scientific Methods (IJARESM), ISSN: 2455-6211 Volume 11, Issue 3, March-2023, Impact Factor: 7.429, Available online at: www.ijaresm.com DOI: https://doi.org/11.56025/IJARESM.2023.11323351

[26] Shashidhara, K.S., Girish, H., Parameshwara, M.C., Rai, B.K., Dakulagi, V. (2023). A Novel Approach for Identification of Healthy and Unhealthy Leaves Using Scale Invariant Feature Transform and Shading Histogram-PCA Techniques. In: Shetty, N.R., Patnaik, L.M., Prasad, N.H. (eds) Emerging Research in Computing, Information, Communication and Applications. Lecture Notes in Electrical Engineering, vol 928. Springer, Singapore. https://doi.org/10.1007/978-981-19-5482-5_47