UNCONSTRAINED POSSIBILISTIC FUZZY C-MEANS ALGORITHM

A Thesis Submitted
In Partial Fulfillment of the Requirements for the Degree of

MASTER OF SCIENCE

in

Mathematics

by

Vedant Ajit Pansare
(Roll No.: 24/MSCMAT/04)

Under the Supervision of

Dr. Dhirendra Kumar
Department of Applied Mathematics
Delhi Technological University, Delhi, India

Department of Applied Mathematics

DELHI TECHNOLOGICAL UNIVERSITY

(Formerly Delhi College of Engineering)
Shahbad Daulatpur, Main Bawana Road, Delhi-110042, India

May, 2026

CANDIDATE'S DECLARATION

I, Vedant Ajit Pansare, hereby certify that the work presented in the thesis entitled *"Unconstrained Possibilistic Fuzzy C-Means Algorithm"*, submitted in partial fulfillment of the requirements for the award of the degree of Master of Science in Mathematics, Department of Applied Mathematics, Delhi Technological University, is an authentic record of my own work carried out under the supervision of Dr. Dhirendra Kumar.

The matter presented in this thesis has not been submitted by me for the award of any other degree of this or any other institute or university. All sources used in the preparation of this thesis have been appropriately cited and acknowledged.

Place: Delhi — Candidate's Signature

Date: — Vedant Ajit Pansare

CERTIFICATE BY THE SUPERVISOR

Certified that Vedant Ajit Pansare (Roll No.: 24/MSCMAT/04) has carried out the research work presented in this thesis entitled *"Unconstrained Possibilistic Fuzzy C-Means Algorithm"* for the award of the degree of Master of Science in Mathematics from the Department of Applied Mathematics, Delhi Technological University, Delhi, under my supervision.

The thesis embodies results of original work and studies carried out by the student himself, and the contents of the thesis do not form the basis for the award of any other degree to the candidate or to anybody else from this or any other university or institution.

Place: Delhi

Date:

Signature of Supervisor

Dr. Dhirendra Kumar

Department of Applied Mathematics

Delhi Technological University

ACKNOWLEDGEMENTS

I would like to thank my supervisor, Dr. Dhirendra Kumar, Department of Applied Mathematics, Delhi Technological University, for the time, guidance and patience he gave me during this dissertation. His remarks during the work made me look again at both the mathematical formulation and the computational steps of fuzzy clustering, and they helped me keep the study focused.

I am also grateful to the Department of Applied Mathematics, Delhi Technological University, where I had the academic support and working environment needed to complete this dissertation. I also acknowledge the authors of the research papers, datasets and open academic resources that formed the foundation for the literature review and experimental comparison reported in this thesis.

Finally, I thank my family and friends for their encouragement, patience and constant support during the preparation of this dissertation.

Vedant Ajit Pansare

ABSTRACT

Possibilistic Fuzzy C-Means (PFCM) jointly models fuzzy memberships alongside typicality values. It handles noisy data well and avoids the coincident-cluster trap, but it often converges to a suboptimal local minimum. This problem becomes particularly severe in high-dimensional or noisy data. MMPFCM rederives PFCM using the majorization-minimization method. It drops the explicit centroids and introduces a lower-rank surrogate variable that smooths out the optimization landscape. Separately, unconstrained fuzzy clustering relaxes the row-sum-to-one membership constraint by substituting the analytical membership solution into the objective, converting the constrained problem into an unconstrained one and opening the door to more flexible, more stable optimization. We unify both ideas – unconstrained optimization and the typicality mechanism - into a single framework. Our method, Unconstrained Possibilistic Fuzzy C-Means (UC-PFCM), substitutes the closed-form membership solution directly into the PFCM cost while keeping the typicality terms unchanged, and then minimises the resulting cost via gradient descent with momentum, updating centroids directly. The per-iteration cost stays the same as PFCM and MMPFCM. UC-PFCM converges to lower objective values than its competitors in almost every case on twelve UCI datasets and records the best ranks across four standard clustering metrics.

LIST OF PUBLICATIONS / MANUSCRIPTS

A manuscript based on the dissertation work has been prepared under the title *"Unconstrained Possibilistic Fuzzy C-Means Algorithm"* by Vedant Ajit Pansare and Dr. Dhirendra Kumar. The work reported in this thesis is aligned with that manuscript and extends it with a detailed dissertation-level introduction, literature review, mathematical preliminaries, methodology, experimental discussion, future scope and social impact.

CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF SYMBOLS, ABBREVIATIONS AND NOMENCLATURE

Symbol / Term	Meaning
X	Dataset or data matrix.
x_i	i-th data point, usually an element of $\mathbb{R}^d$.
n	Number of data points.
d	Number of features or data dimension.
c	Number of clusters.
v_j	Prototype or centroid of the j-th cluster.
$U = [u_{ij}]$	Fuzzy membership matrix.
u_{ij}	Membership degree of sample x_i in cluster j.
$T = [t_{ij}]$	Typicality matrix.
t_{ij}	Possibilistic typicality of sample x_i with respect to cluster j.
m	Fuzzy exponent or fuzzifier used in membership weighting.
η	Typicality exponent used in the possibilistic term.
γ_j	Scale parameter associated with cluster j in PFCM-type models.
α	Learning rate in gradient descent.
β	Momentum coefficient in momentum gradient descent.
d_{ij}^2	Squared Euclidean distance $\|x_i - v_j\|^2$.
FCM	Fuzzy C-Means.
PCM	Possibilistic C-Means.
PFCM	Possibilistic Fuzzy C-Means.
UC-FCM	Unconstrained Fuzzy C-Means.
UC-PFCM	Unconstrained Possibilistic Fuzzy C-Means.
MMPFCM	Majorization-minimization based PFCM.
NMI	Normalized Mutual Information.
ARI	Adjusted Rand Index.
PUR	Purity.
DBI	Davies-Bouldin Index.
XB	Xie-Beni Index.

CHAPTER 1

INTRODUCTION

1.1 Background

Clustering is one of the most frequently used methods in unsupervised machine learning. In a clustering problem, a dataset is available without class labels, and the objective is to discover meaningful groups from the data itself. The central intuition is simple: samples placed in the same cluster should be similar to each other, while samples placed in different clusters should be dissimilar. This idea appears in many application areas, including image segmentation, pattern recognition, document categorization, bioinformatics and exploratory data analysis [1–4].

The earliest prototype-based methods mostly follow a hard partitioning idea. In such methods, a point is placed in a single cluster, and K-Means is the standard example. This makes the output simple, but it is also a strong assumption. Many datasets do not separate themselves so neatly: two groups may overlap, a few observations may be noisy, and some samples may sit near the boundary between clusters. Assigning only one label in these cases can hide information that is actually useful for understanding the data.

Fuzzy clustering was introduced to keep this uncertainty visible. Instead of making a point belong completely to one group, it gives membership values for the different clusters. A large membership value means that the point is close to or strongly associated with that cluster; a small value means the association is weak. Fuzzy C-Means (FCM), formalized by [5, 6], became the most familiar fuzzy clustering method because its objective function is transparent, the update equations are easy to implement, and the method performs well in many practical settings.

The same membership rule, however, also causes a difficulty. In FCM the memberships of every point over all clusters must add up to one. That is reasonable for relative membership, but it is less suitable when a point is far from all clusters. Even an outlier is made to distribute a full unit of membership, and through this it can still pull the centres away from dense regions. This weakness is one of the reasons possibilistic clustering was proposed.

Possibilistic C-Means (PCM) was proposed by [7] to interpret membership-like values as typicalities rather than relative probabilities. A typicality value measures how compatible a sample is with a cluster prototype independently of other clusters. This is useful for noise handling because an outlier can have low typicality values for all clusters. However, PCM may suffer from coincident clusters, where two or more cluster prototypes collapse to the same region. Possibilistic Fuzzy C-Means (PFCM), proposed by [8], combines fuzzy memberships and possibilistic typicalities in one objective. It retains the cluster-separation ability of fuzzy memberships and the noise robustness of possibilistic typicalities.

PFCM improves the behaviour of both FCM and PCM, but its usual solution procedure still follows alternating updates. In this scheme, some variables are kept fixed while the others are updated. The method is convenient, yet the route followed by the iterations may still end at a weak local solution. For this reason, recent studies have paid more attention to the optimization step itself. [9] proposed UC-FCM by inserting the closed-form FCM membership solution back into the FCM objective, so that the centres can be optimized directly. In another direction, [10] studied PFCM through a majorization-minimization framework and presented the MMPFCM model.

The present thesis takes these two lines of work together. It proposes an Unconstrained Possibilistic Fuzzy C-Means algorithm, denoted by UC-PFCM. The basic aim is to keep the membership-free centre optimization of UC-FCM while also retaining the typicality term used in PFCM. In this way, the method keeps the noise-handling advantage of possibilistic clustering and at the same time works with an unconstrained prototype objective optimized by momentum gradient descent.

1.2 Motivation

The motivation for this work came from a fairly direct observation while studying FCM-type methods. A clustering algorithm is affected not only by the distance measure or by extra regularization terms, but also by the way in which the objective is actually minimized. FCM, PFCM and many related methods rely on alternating updates. These updates are fast and useful, but they also make the optimization follow a fixed pattern, which may not always lead to the best local solution.

The membership matrix is also large. For n samples and c clusters, it contains nc entries, and these entries must satisfy both the row-sum condition and non-negativity. When the algorithm switches back and forth between memberships and centres, it often converges in a small number of iterations, but a quick convergence does not always mean a better optimum. This concern becomes sharper in high-dimensional datasets, where the objective surface is usually more complicated.

The UC-FCM formulation shows a useful way around this issue. For fixed centres, the FCM membership values already have an analytical form. Therefore, one can place this expression back into the objective instead of optimizing the membership matrix separately. After this substitution, the objective is written only in terms of the prototypes and can be handled by gradient-based methods.

PFCM contributes the second ingredient needed here, namely typicality. A typicality value tells how representative a sample is for a particular cluster without forcing it to compete fully with the other clusters. This is the part that makes PFCM less sensitive to noisy observations. The question taken up in this dissertation is whether this typicality idea can be joined with the unconstrained centre optimization used in UC-FCM. The proposed UC-PFCM algorithm is developed for that purpose.

1.3 Problem Statement

The problem considered in this thesis is the following. For an unlabeled dataset $X = \{x_1, x_2, \ldots, x_n\} \subset \mathbb{R}^d$ and a chosen number of clusters c, the aim is to construct a fuzzy-possibilistic clustering method that gives reliable partitions, avoids an explicitly constrained membership matrix in the fuzzy part, and is less dependent on the local solution reached by standard alternating updates.

More specifically, the research problem has three mathematical components. First, the fuzzy membership component of PFCM must be reformulated so that the row-sum constrained membership matrix is not optimized as an independent variable. Second, the possibilistic typicality component must be preserved because it is responsible for reducing the effect of noisy and atypical points. Third, the final objective must be optimized using a practical method whose computational cost remains comparable to existing fuzzy clustering algorithms.

1.4 Research Gap

The literature shows a clear development path from hard clustering to fuzzy clustering, then to possibilistic and hybrid fuzzy-possibilistic clustering. FCM improves hard clustering by allowing partial memberships, but it is sensitive to noise due to the row-sum membership constraint. PCM reduces this issue by using typicality, but it may suffer from coincident clusters. PFCM combines the two and is generally more stable than either model alone.

However, PFCM is still commonly treated as an alternating optimization problem. The optimization mechanism itself is not fully redesigned. On the other hand, UC-FCM directly addresses the constrained optimization issue, but it does not include possibilistic typicalities. MMPFCM improves PFCM through a majorization-minimization framework, but it follows a different surrogate-based route rather than a direct gradient-based unconstrained centre optimization framework. Therefore, there is a gap between unconstrained fuzzy optimization and possibilistic fuzzy robustness.

This thesis fills that gap by constructing a UC-PFCM model. It removes the explicit constrained membership variable through analytical substitution, preserves the PFCM typicality terms, derives the gradient with respect to cluster centres, and applies momentum-based gradient descent. The research contribution is not merely the addition of a term to an existing algorithm; it is a restructuring of the optimization pathway for a fuzzy-possibilistic objective.

1.5 Objectives of the Study

The objectives of this dissertation are as follows:

1. To study the mathematical structure of FCM, PCM, PFCM, UC-FCM and recent PFCM variants.
2. To formulate an unconstrained fuzzy-possibilistic objective by combining the UC-FCM fuzzy component with the PFCM typicality mechanism.
3. To derive the gradient of the proposed objective with respect to cluster prototypes.
4. To design an iterative optimization algorithm based on momentum gradient descent and closed-form typicality updates.
5. To compare UC-PFCM with FCM, UC-FCM, PFCM and MMPFCM on benchmark datasets.
6. To evaluate the method using objective values, external clustering metrics, internal validity indices and statistical rank tests.

1.6 Major Contributions

The major contributions of this thesis are summarized below.

1. A new clustering objective, UC-PFCM, is formulated by embedding PFCM-style typicality into an unconstrained fuzzy clustering structure.
2. The constrained membership matrix is eliminated from the fuzzy part of the objective by using the analytical membership solution.
3. A closed-form gradient with respect to cluster centroids is derived, allowing the proposed model to be optimized through gradient descent with momentum.
4. The algorithm retains the same asymptotic per-iteration complexity as PFCM and MMPFCM, namely $\mathcal{O}(ncd)$.

5. Experiments on twelve benchmark datasets show that the proposed method achieves lower objective values and strong average ranks across multiple clustering measures.

1.7 Organization of the Thesis

The rest of the thesis is organized as follows. Chapter 2 presents the literature review, covering hard clustering, fuzzy clustering, possibilistic clustering, unconstrained fuzzy clustering and related evaluation methods. Chapter 3 gives the mathematical preliminaries required for the proposed method. Chapter 4 develops the UC-PFCM objective, derives the update mechanism and analyzes computational complexity. Chapter 5 reports the experimental setup, datasets, results and statistical analysis. Chapter 6 concludes the thesis and discusses future scope and social impact.

CHAPTER 2

LITERATURE REVIEW

2.1 Overview of Clustering Methods

Cluster analysis has a long history in pattern recognition and data analysis. The purpose of clustering is to discover groups without using predefined class labels. A clustering algorithm depends on a definition of similarity, a representation of clusters and a criterion for judging the quality of a partition. Different choices lead to different families of algorithms.

Partition-based methods, such as K-Means and FCM, represent each cluster by a prototype and assign data points according to their distances from prototypes. Density-based methods search for dense regions separated by sparse regions. Hierarchical methods build nested groupings of data points. Model-based methods assume that data are generated from an underlying probabilistic model. Among these families, prototype-based methods remain attractive because of their mathematical simplicity and computational efficiency.

The proposed work belongs to the prototype-based fuzzy clustering family. It focuses on how prototypes are optimized when the data may contain overlap or noise. This choice is motivated by the fact that many real-world datasets do not contain sharply separated clusters. In such data, fuzzy or possibilistic assignments provide more information than hard assignments.

2.2 Hard Clustering and K-Means

The K-Means algorithm partitions a dataset into c clusters by minimizing the within-cluster sum of squared distances. If v_j denotes the centre of the j-th cluster and C_j denotes the set of samples assigned to it, the objective is

$$J_{K\text{-Means}} = \sum_{j=1}^{c} \sum_{x_i \in C_j} \|x_i - v_j\|^2. \tag{2.1}$$

The method alternates between assigning each point to the nearest centre and recomputing centres as arithmetic means. The algorithm is simple, but its assignments are crisp. A point lying near a decision boundary receives the same type of final label as a point very close to a cluster centre. The uncertainty of assignment is lost.

This limitation is particularly important in data with overlapping clusters. For example, in image segmentation, boundary pixels may share properties of more than one region. In document clustering, a document can contain vocabulary from multiple topics. In such situations, a crisp label does not fully describe the data structure. Fuzzy clustering addresses this problem by replacing hard assignments with membership grades.

2.3 Fuzzy C-Means Clustering

FCM generalizes K-Means by assigning membership values $u_{ij} \in [0,1]$ to describe the degree to which sample x_i belongs to cluster j. The memberships for each sample must sum to one. The objective function is

$$J_{FCM}(U,V) = \sum_{i=1}^{n} \sum_{j=1}^{c} u_{ij}^{m} \|x_i - v_j\|^2, \quad \sum_{j=1}^{c} u_{ij} = 1, \tag{2.2}$$

where $m > 1$ is the fuzzifier. When m is close to one, the partition becomes sharper. Larger values of m produce fuzzier memberships.

The popularity of FCM comes from its clear objective and closed-form alternating updates. For fixed centres, memberships can be updated using the relative distances from each point to all centres. For fixed memberships, centres are weighted averages of the data. This procedure is easy to implement and often works well.

Nevertheless, FCM has a weakness in noisy environments. Because the memberships for each point sum to one, every point must contribute a complete unit of membership to the clustering structure. Even an outlier that is far from every cluster centre is forced to participate. Through this mechanism, distant observations can shift the prototypes away from the main data clouds and reduce the quality of the final partition. Many FCM variants have tried to address noise, initialization and high-dimensional data, but the row-sum membership condition itself remains part of the basic model.

2.4 Possibilistic C-Means

Possibilistic clustering looks at the assignment value in a different way. In PCM, the value attached to a point-cluster pair is treated as typicality rather than as relative membership. It tells how well the point fits a particular prototype on its own. Since these typicalities are not forced to add to one over the clusters, a point may be weakly typical for every cluster.

This difference is useful in noisy data. If a sample is far from all prototypes, it can receive small typicality values everywhere, and its influence on the centre update becomes limited. This gives PCM an advantage over FCM for outlier handling. At the same time, PCM has a well-known drawback: because typicality values are not competitive across clusters, two prototypes can sometimes settle in the same dense part of the data. This is the coincident-cluster problem.

The comparison between FCM and PCM therefore gives a clear trade-off. FCM keeps stronger competition between clusters, but it is more affected by noise. PCM handles atypical samples better, but it can weaken separation between prototypes. Hybrid methods such as PFCM were introduced to combine these two useful properties.

2.5 Possibilistic Fuzzy C-Means

PFCM puts fuzzy memberships and possibilistic typicalities into the same objective function. The membership part keeps the clusters competing with one another, and the typicality part down-weights samples that do not look representative of a cluster. A common form of the objective is

$$J_{PFCM} = \sum_{i=1}^{n} \sum_{j=1}^{c} \left(a u_{ij}^{m} + b t_{ij}^{\eta} \right) \|x_i - v_j\|^2 + \sum_{j=1}^{c} \gamma_j \sum_{i=1}^{n} (1 - t_{ij})^{\eta}, \tag{2.3}$$

Here u_{ij} denotes the fuzzy membership and t_{ij} denotes the typicality. The constants a and b balance the two terms, m and η control their fuzziness, and γ_j gives the scale for the j-th cluster.

PFCM is therefore a natural baseline for the present study, since it carries both the competitive part of FCM and the robustness idea of PCM. It can reduce the noise sensitivity of FCM and avoid the severe coincident-cluster behaviour of PCM. However, it is still solved by alternating updates. The update rules are local and efficient, but the algorithm may still settle into suboptimal local minima.

For this reason, recent work has revisited the optimization process of PFCM. Majorization-minimization methods build surrogate functions that are easier to minimize, while unconstrained methods remove membership constraints and use direct optimization. The proposed thesis follows the second direction while retaining the typicality mechanism.

2.6 Unconstrained Fuzzy Clustering

The unconstrained FCM approach introduced by [9] begins from a simple observation: for fixed cluster centres, the optimal FCM memberships have a closed-form expression. Therefore, instead of alternating between membership and centre updates, the membership formula can be substituted into the objective. The resulting objective depends only on the cluster centres.

This reformulation changes the nature of the optimization problem. The original FCM problem is constrained because of the row-sum and non-negativity requirements on the membership matrix. After substitution, the objective becomes unconstrained with respect to the prototypes. It can be solved by gradient descent or momentum gradient descent. The number of variables is also reduced from the membership matrix plus centre matrix to only the centre matrix.

UC-FCM has two important implications. First, it shows that the FCM objective can be studied from a different optimization viewpoint. Second, it creates a route for connecting fuzzy clustering with gradient-based learning systems. If a clustering loss is differentiable with respect to prototypes or representations, it can be embedded into larger models.

However, UC-FCM does not include a possibilistic typicality term. It improves the optimization of the fuzzy component but does not directly address the absolute compatibility of samples with clusters. This creates the opportunity for UC-PFCM.

2.7 Majorization-Minimization Based PFCM

[10] revisited PFCM using a majorization-minimization framework. The goal of such a framework is to replace a difficult objective at each iteration by a surrogate function that is easier to optimize and that guarantees a monotonic descent property. This surrogate-based route often makes the PFCM iterations more stable and can improve performance compared with the basic alternating scheme.

MMPFCM is important here because it is a recent PFCM-related baseline that changes the optimization strategy rather than merely adding a heuristic correction. For that reason, it is a useful comparison when checking whether the unconstrained prototype-based route of UC-PFCM is genuinely competitive.

2.8 Cluster Validity and Evaluation Measures

Clustering evaluation is not straightforward, especially because true labels are usually absent in real unsupervised problems. For benchmark data, however, class labels are available and external indices can be reported. Normalized Mutual Information (NMI) compares the information shared by the obtained labels and the true labels. Adjusted Rand Index (ARI) checks agreement between pairs of samples after correcting for chance [11]. Purity records the dominant true class inside each cluster, while the F^* score combines pairwise precision and recall.

Internal indices judge the partition using only the geometry of the data and the clusters. The Davies-Bouldin Index (DBI) compares the scatter within a cluster with its separation from the most similar neighbouring cluster [12]; smaller values are preferred. The Xie-Beni index is commonly used in fuzzy clustering and compares fuzzy compactness with the minimum centre separation [13]; again, a smaller value indicates a better partition.

No single index gives a complete judgement of clustering quality. For this reason, this thesis reports several external and internal measures. It also uses the Friedman test and average ranks to see whether the differences between algorithms remain consistent across the set of datasets [14].

2.9 Summary of Literature Review

The literature can be read as a gradual attempt to handle practical weaknesses in clustering. Hard clustering is simple but too rigid for overlapping data. FCM introduces memberships, yet remains sensitive to outliers because of its row-sum rule. PCM introduces typicality, but may create coincident prototypes. PFCM combines membership and typicality, although its alternating updates can still stop at weaker local minima. UC-FCM changes the optimization path by removing the constrained membership variable, but it does not include typicality. MMPFCM improves PFCM through a surrogate optimization framework.

The proposed UC-PFCM model is motivated by this gap. It combines the unconstrained fuzzy objective with PFCM-style typicality. This creates an algorithm that is mathematically connected to FCM, robust in the sense of possibilistic clustering, and optimized through a direct gradient-based framework.

CHAPTER 3

MATHEMATICAL PRELIMINARIES

3.1 Notation

Let $X = \{x_1, x_2, \ldots, x_n\} \subset \mathbb{R}^d$ denote a dataset containing n data points in d dimensions. The number of clusters is denoted by c. The j-th cluster prototype is denoted by $v_j \in \mathbb{R}^d$. The squared Euclidean distance between x_i and v_j is

$$d_{ij}^2 = \|x_i - v_j\|^2. \tag{3.1}$$

The fuzzy membership of point x_i in cluster j is u_{ij}, and the possibilistic typicality is t_{ij}. The membership matrix is $U = [u_{ij}]$ and the typicality matrix is $T = [t_{ij}]$.

3.2 Fuzzy C-Means Objective and Updates

The FCM objective is

$$J_{FCM}(U,V) = \sum_{i=1}^{n} \sum_{j=1}^{c} u_{ij}^m d_{ij}^2, \tag{3.2}$$

subject to

$$\sum_{j=1}^{c} u_{ij} = 1, \qquad 0 \le u_{ij} \le 1. \tag{3.3}$$

For fixed centres, the membership update is

$$u_{ij} = \frac{1}{\sum_{k=1}^{c} \left(\frac{\|x_i - v_j\|}{\|x_i - v_k\|} \right)^{2/(m-1)}}. \tag{3.4}$$

For fixed memberships, the prototype update is

$$v_j = \frac{\sum_{i=1}^{n} u_{ij}^m x_i}{\sum_{i=1}^{n} u_{ij}^m}. \tag{3.5}$$

These two updates are repeated until convergence.

3.3 PFCM Objective and Typicality Update

PFCM combines fuzzy and possibilistic terms. Its objective is

$$J_{PFCM}(U,T,V) = \sum_{i=1}^{n}\sum_{j=1}^{c}\left(au_{ij}^m + bt_{ij}^{\eta}\right) d_{ij}^2 + \sum_{j=1}^{c}\gamma_j \sum_{i=1}^{n}(1-t_{ij})^{\eta}. \tag{3.6}$$

Here $a > 0$ and $b > 0$ determine the relative importance of fuzzy memberships and typicalities. The parameter $\eta > 1$ controls the fuzziness of typicality values. The typicality update for fixed centres is

$$t_{ij} = \frac{1}{1 + \left(\frac{bd_{ij}^2}{\gamma_j}\right)^{1/(\eta-1)}}. \tag{3.7}$$

The prototype update is

$$v_j = \frac{\sum_{i=1}^{n}\left(au_{ij}^m + bt_{ij}^{\eta}\right) x_i}{\sum_{i=1}^{n}\left(au_{ij}^m + bt_{ij}^{\eta}\right)}. \tag{3.8}$$

The scale parameter γ_j is commonly set using the weighted within-cluster distance:

$$\gamma_j = K\frac{\sum_{i=1}^{n} u_{ij}^m d_{ij}^2}{\sum_{i=1}^{n} u_{ij}^m}, \tag{3.9}$$

where $K > 0$ is a constant.

3.4 Unconstrained FCM Reformulation

The UC-FCM objective is obtained by substituting the optimal FCM membership update into the FCM objective. For $m > 1$, define

$$S_i(V) = \sum_{j=1}^{c} d_{ij}^{2/(1-m)}. \tag{3.10}$$

Then the UC-FCM objective becomes

$$J_{UC\text{-}FCM}(V) = \sum_{i=1}^{n} \left(\sum_{j=1}^{c} d_{ij}^{2/(1-m)} \right)^{1-m} = \sum_{i=1}^{n} S_i(V)^{1-m}. \tag{3.11}$$

This objective depends only on the prototypes. The constrained membership matrix is not treated as an independent optimization variable.

The gradient of $J_{UC\text{-}FCM}$ with respect to v_j is

$$\frac{\partial J_{UC\text{-}FCM}}{\partial v_j} = 2 \sum_{i=1}^{n} S_i(V)^{-m} d_{ij}^{2m/(1-m)} (v_j - x_i). \tag{3.12}$$

This formula is central because it makes direct gradient-based prototype optimization possible.

3.5 Momentum Gradient Descent

Gradient descent changes the prototypes by moving opposite to the gradient. Writing $G_j^{(t)}$ for the gradient of the objective with respect to prototype v_j at iteration t, the ordinary update is

$$v_j^{(t+1)} = v_j^{(t)} - \alpha G_j^{(t)}, \tag{3.13}$$

where $\alpha > 0$ is the step size. In the momentum version, the new step also uses part of the previous step, so the update is written as

$$\Delta v_j^{(t)} = \beta \Delta v_j^{(t-1)} + \alpha G_j^{(t)}, \tag{3.14}$$

$$v_j^{(t+1)} = v_j^{(t)} - \Delta v_j^{(t)}, \tag{3.15}$$

where $0 \leq \beta < 1$ controls how much of the previous direction is carried forward. This usually reduces zig-zag movement and can make the search steadier on curved objective surfaces.

3.6 External Evaluation Measures

Let P denote the clustering produced by an algorithm and let Q denote the true partition. NMI reports how much information these two partitions share after normalization, so larger values mean stronger agreement. ARI works with pairs of samples and adjusts the Rand index for chance agreement; values closer to one indicate that the obtained partition matches the ground truth well.

Purity assigns each obtained cluster to the true class that appears most often inside it and then counts the matched samples. The F^* score uses the pairwise precision-recall viewpoint. These measures emphasize different aspects of a partition, so the experiments in this thesis report more than one external index.

3.7 Internal Validity Indices

The Davies-Bouldin Index compares each cluster with the other cluster that is most similar to it. If S_j denotes the scatter of cluster j and M_{jk} denotes the distance between prototypes j and k, one commonly used expression is

$$DBI = \frac{1}{c} \sum_{j=1}^{c} \max_{k \neq j} \frac{S_j + S_k}{M_{jk}}. \tag{3.16}$$

A smaller DBI value is preferred.

The Xie-Beni index is another internal index used frequently in fuzzy clustering. It relates the total fuzzy compactness of the partition to the smallest squared separation between two centres:

$$XB = \frac{\sum_{i=1}^{n} \sum_{j=1}^{c} u_{ij}^{m} \|x_i - v_j\|^2}{n \min_{j \neq k} \|v_j - v_k\|^2}. \tag{3.17}$$

A smaller XB value means that the clusters are relatively compact and their centres are well separated.

3.8 Statistical Ranking

When several algorithms are tested on several datasets, average performance alone can be misleading. Rank-based statistical tests are often used to compare algorithms across datasets. The Friedman test ranks the algorithms on each dataset and then tests whether the average ranks differ significantly. The statistic is

$$\tau_{\chi^2} = \frac{12N}{K(K+1)} \left(\sum_{i=1}^{K} r_i^2 - \frac{K(K+1)^2}{4} \right), \tag{3.18}$$

where N is the number of datasets, K is the number of algorithms, and r_i is the average rank of the i-th algorithm.

CHAPTER 4

PROPOSED METHODOLOGY

4.1 Overview of the Proposed Method

The proposed method is called Unconstrained Possibilistic Fuzzy C-Means, abbreviated as UC-PFCM. The method is designed by combining the unconstrained fuzzy objective of UC-FCM with the typicality terms of PFCM. The fuzzy membership part is not optimized as a separate constrained variable. Instead, the analytical membership solution is substituted into the fuzzy objective. The typicality part remains similar to PFCM so that the model preserves robustness to noisy or atypical samples.

The main idea is illustrated in Fig. 4.1. Classical PFCM alternates between membership update, typicality update and centre update. UC-FCM removes the membership variable and optimizes centres directly. UC-PFCM combines these two directions: it removes the constrained membership variable from the fuzzy component, keeps the typicality update, and updates prototypes by momentum gradient descent.

Data and Initial Prototypes
⇓
Compute Pairwise Distances $d_{ij} = \|x_i - v_j\|$
⇓
Evaluate UC-FCM Fuzzy Term + PFCM Typicality Term
⇓
Update Prototypes Using Momentum Gradient Descent
⇓
Update Typicalities Using Closed-Form Rule
⇓
Nearest-Prototype Final Assignment

Figure 4.1: Workflow of the proposed UC-PFCM algorithm.

4.2 UC-PFCM Objective Function

The UC-PFCM objective is defined as

$$\begin{aligned} J_{UC\text{-}PFCM}(V,T) = a \sum_{i=1}^{n} \left(\sum_{j=1}^{c} d_{ij}^{2/(1-m)} \right)^{1-m} \\ + b \sum_{i=1}^{n} \sum_{j=1}^{c} t_{ij}^{\eta} d_{ij}^{2} + \sum_{j=1}^{c} \gamma_j \sum_{i=1}^{n} (1 - t_{ij})^{\eta}. \end{aligned} \tag{4.1}$$

The first term is the unconstrained fuzzy term. It is inherited from UC-FCM and is responsible for fuzzy cluster competition without explicitly storing a constrained membership matrix. The second term is the typicality-weighted distortion term. It reduces the influence of data points that are not typical of a cluster. The third term regularizes typicality values and avoids the trivial solution in which all typicalities become zero.

This objective has two important features. First, it avoids the explicit row-sum membership constraint. Second, it keeps typicality values, which are crucial for noise robustness. Therefore, UC-PFCM is not simply UC-FCM with another distance term; it is a hybrid unconstrained fuzzy-possibilistic model.

4.3 Gradient with Respect to Prototypes

Let

$$S_i(V) = \sum_{k=1}^{c} \|x_i - v_k\|^{2/(1-m)}. \tag{4.2}$$

The derivative of the unconstrained fuzzy component with respect to v_j is

$$2a\sum_{i=1}^{n} S_i(V)^{-m}\|x_i - v_j\|^{2m/(1-m)}(v_j - x_i). \tag{4.3}$$

The derivative of the typicality distortion term with respect to v_j is

$$2b\sum_{i=1}^{n} t_{ij}^{\eta}(v_j - x_i). \tag{4.4}$$

The regularization term involving $(1-t_{ij})^{\eta}$ does not directly depend on v_j when typicalities are treated as fixed during the prototype update step. Therefore, the full gradient is

$$\frac{\partial J_{UC\text{-}PFCM}}{\partial v_j} = 2a\sum_{i=1}^{n} S_i(V)^{-m}\|x_i - v_j\|^{2m/(1-m)}(v_j - x_i) + 2b\sum_{i=1}^{n} t_{ij}^{\eta}(v_j - x_i). \tag{4.5}$$

This expression also has an intuitive meaning. The first summation gives the movement suggested by the unconstrained fuzzy part. The second summation gives the movement suggested by the typicality-weighted part. Thus, points with high typicality have more say in the centre update, while points that look atypical contribute less.

4.4 Typicality Update

When the prototypes are fixed, the typicality values are updated by

$$t_{ij} = \frac{1}{1 + \left(\frac{bd_{ij}^2}{\gamma_j}\right)^{1/(\eta-1)}}. \tag{4.6}$$

The formula behaves as expected. A point close to prototype v_j has a small d_{ij}^2 and therefore receives a large t_{ij}. A distant point receives a small typicality. In this sense, t_{ij} works like a reliability weight for the pair (x_i, v_j).

The scale parameter is set as

$$\gamma_j = K \frac{\sum_{i=1}^{n} u_{ij}^m d_{ij}^2}{\sum_{i=1}^{n} u_{ij}^m}, \tag{4.7}$$

where K is taken as one in the experiments. Although the explicit membership matrix is not optimized in the objective, the membership values can still be computed from the closed-form FCM rule when needed for initialization and scale estimation.

4.5 Prototype Update

The prototype update uses momentum gradient descent. Let

$$G_j^{(t)} = \left. \frac{\partial J_{UC\text{-}PFCM}}{\partial v_j} \right|_{V=V^{(t)}, T=T^{(t)}}. \tag{4.8}$$

Then

$$\Delta v_j^{(t)} = \beta \Delta v_j^{(t-1)} + \alpha G_j^{(t)}, \tag{4.9}$$

$$v_j^{(t+1)} = v_j^{(t)} - \Delta v_j^{(t)}. \tag{4.10}$$

The learning rate controls the step size. The momentum coefficient controls how much of the previous update direction is retained. In this work, the learning rate is selected from a small candidate set and the momentum coefficient is set to 0.6, following the practical setting used in the experimental design.

4.6 Final Cluster Assignment

After convergence, each sample is assigned to the nearest cluster centre. The indicator value f_{ij} is defined as

$$f_{ij} = \begin{cases} 1, & j = \arg\min_k d_{ik}, \\ 0, & j \neq \arg\min_k d_{ik}. \end{cases} \tag{4.11}$$

This final crisp assignment is used only for evaluating clustering performance against ground-truth labels. The optimization process itself uses fuzzy and possibilistic information.

4.7 Algorithm

Algorithm 1 Unconstrained Possibilistic Fuzzy C-Means (UC-PFCM)

Input: Data matrix X, number of clusters c, parameters a, b, m, η, K, learning rate α, momentum β.
Initialize memberships $U^{(0)}$ randomly and compute initial prototypes $V^{(0)}$ using the FCM centroid update.
Compute γ_j using Eq. (4.7) and initialize typicalities using Eq. (4.6).
Set $\Delta v_j^{(0)} = 0$ for all clusters.
repeat
    Compute distances $d_{ij} = \|x_i - v_j\|$.
    Evaluate the gradient G_j using Eq. (4.5).
    Update prototypes by momentum gradient descent.
    Update typicalities using Eq. (4.6).
    Compute the objective value $J_{UC\text{-}PFCM}$.
until $|J^{(t+1)} - J^{(t)}| < 10^{-5}$ or maximum iterations reached.
Assign each sample to the nearest prototype.
Output: Cluster labels, prototypes V, typicalities T.

4.8 Computational Complexity

Let n be the number of samples, d the number of features, c the number of clusters and t the number of iterations. The main operations are distance computation, typicality update and gradient computation. Computing all distances requires $\mathcal{O}(ncd)$ operations.

Updating all typicalities requires visiting each sample-cluster pair and is therefore $\mathcal{O}(nc)$ after distances are available; including distance information, it is dominated by $\mathcal{O}(ncd)$. Computing the gradient for all prototypes also requires summations over samples, clusters and dimensions, so it is $\mathcal{O}(ncd)$.

Therefore, each iteration has complexity

$$\mathcal{O}(ncd), \tag{4.12}$$

and the total complexity over t iterations is

$$\mathcal{O}(tncd). \tag{4.13}$$

This is the same asymptotic order as standard PFCM and MMPFCM. Hence, the proposed model changes the optimization structure without increasing the asymptotic per-iteration cost.

4.9 Methodological Significance

The significance of UC-PFCM lies in how it combines two complementary ideas. The unconstrained fuzzy term reduces the number of constrained variables and gives a direct prototype-based objective. The possibilistic term introduces typicality, which reduces the effect of atypical samples. Momentum gradient descent then provides a simple practical mechanism for optimizing the resulting objective.

The method is also modular. Other distance functions, initialization strategies or representation-learning components can be connected to the same framework if a differentiable distance is available. This makes UC-PFCM useful not only as a standalone clustering algorithm but also as a possible clustering loss in larger unsupervised learning systems.

CHAPTER 5

EXPERIMENTAL RESULTS AND DISCUSSION

5.1 Experimental Setup

The proposed UC-PFCM algorithm is compared with four baseline methods: FCM, UC-FCM, PFCM and MMPFCM. These baselines were selected because they represent the major stages relevant to the thesis: classical fuzzy clustering, unconstrained fuzzy clustering, possibilistic fuzzy clustering and recent majorization-minimization based PFCM.

All datasets were normalized to the interval $[0,1]$ before clustering. Each algorithm was run 20 times with random initialization. Results are reported as mean values with standard deviations where applicable. The convergence criterion for UC-PFCM is

$$|J_{UC\text{-}PFCM}^{(t+1)} - J_{UC\text{-}PFCM}^{(t)}| < 10^{-5}. \tag{5.1}$$

The experiments were conducted on a Windows 11 system with AMD Ryzen 5 5600H processor and 16 GB RAM.

5.2 Datasets

Twelve benchmark datasets were used. They include low-dimensional datasets such as Iris and Balance, medium-dimensional datasets such as Vehicle and Segment, and high-dimensional image datasets such as ORL, Yale and COIL20. This diversity is important because a clustering algorithm should not be judged only on one type of data.

Table 5.1: Summary of datasets used in the experiments.

Dataset	Samples	Features	Classes
Iris	150	4	3
Glass	214	9	6
Urban	675	147	9
Vehicle	846	18	4
ORL	400	1024	40
Yale32	165	1024	15
Yale64	165	4096	15
Balance	625	4	3
COIL20	1440	1024	20
Segment	2310	19	7
SCADI	206	205	7
Isolet5	1559	617	26

5.3 Objective Function Comparison

Table 5.2 reports the mean, minimum and maximum final objective values over 20 runs for PFCM, MMPFCM and UC-PFCM. UC-PFCM obtains the lowest mean objective value on every dataset. This is an important result because all three methods are fuzzy-possibilistic in spirit, but they follow different optimization routes.

Table 5.2: Final objective values over 20 runs. The best mean value in each dataset is highlighted in bold.

Dataset	Method	Mean	Min	Max
Iris	PFCM	23.347	20.963	28.768
	MMPFCM	22.148	18.858	26.473
	UC-PFCM	**22.096**	**17.781**	**25.756**
Glass	PFCM	122.483	93.527	193.280
	MMPFCM	121.948	86.126	196.410
	UC-PFCM	**114.726**	**83.385**	**140.160**
Urban	PFCM	5020.292	4266.400	6165.100
	MMPFCM	4785.992	4238.200	5673.500
	UC-PFCM	**4738.511**	**4137.900**	**5498.300**
Vehicle	PFCM	951.109	739.320	1131.200
	MMPFCM	943.449	773.350	1129.400
	UC-PFCM	**895.194**	**718.020**	**1043.700**
ORL	PFCM	268161.534	248160	288160
	MMPFCM	232584.404	219970	245200
	UC-PFCM	**229296.627**	**217750**	**245150**
Yale32	PFCM	65379.305	65379.000	65379.000
	MMPFCM	63619.546	62692.000	64547.000
	UC-PFCM	**61452.428**	**61452.000**	**61452.000**
Yale64	PFCM	217507.588	215440	226130
	MMPFCM	215003.058	204930	219580
	UC-PFCM	**207069.906**	**203140**	**209210**
Balance	PFCM	678.467	477.140	814.200
	MMPFCM	649.653	459.160	808.620
	UC-PFCM	**648.108**	**443.540**	**772.980**
COIL20	PFCM	665860.251	595070	730200
	MMPFCM	649264.773	588500	708310
	UC-PFCM	**643498.632**	**556790**	**689450**
Segment	PFCM	2825.033	2530.600	3363.100
	MMPFCM	2820.541	2387.200	3307.300
	UC-PFCM	**2750.179**	**2365.900**	**3093.400**
SCADI	PFCM	4276.292	3624.900	5127.600
	MMPFCM	4229.470	4163.600	4720.900
	UC-PFCM	**4192.539**	**3137.600**	**4458.800**
Isolet5	PFCM	663137.121	654510	671760
	MMPFCM	658257.917	653390	663130
	UC-PFCM	**635716.706**	**613590**	**657840**

Across the twelve datasets, UC-PFCM gives the smallest mean objective value. It also gives the smallest minimum value in the repeated runs. This indicates that, under the same experimental design, the proposed unconstrained gradient-based route is able to locate more favourable local minima than standard PFCM and MMPFCM.

5.4 Runtime Comparison

Table 5.3 reports the average runtime in seconds. On most datasets the runtime of UC-PFCM is within a practical range. It is slower on some high-dimensional datasets, which is understandable because gradient-based centre updates may need more iterations than closed-form alternating updates.

Table 5.3: Average running time in seconds.

Dataset	FCM	UC-FCM	PFCM	MMPFCM	UC-PFCM
Iris	0.002	0.012	0.009	0.008	0.010
Glass	0.011	0.043	0.015	0.023	0.072
Urban	0.048	0.130	0.079	0.040	0.111
Vehicle	0.064	0.139	0.109	0.079	0.078
ORL	0.798	3.963	0.908	0.283	0.894
Yale32	0.133	1.884	0.157	0.091	0.418
Yale64	0.369	6.289	0.555	0.534	1.530
Balance	0.001	0.011	0.008	0.012	0.014
COIL20	3.563	10.196	4.794	1.645	2.809
Segment	1.622	7.057	1.881	1.447	1.820
SCADI	0.030	0.088	0.037	0.023	0.120
Isolet5	3.560	17.705	0.983	4.476	5.491

The runtime results also make clear that UC-PFCM is not proposed as the fastest method in every case. Its main advantage is the quality of the solution reached, as seen from objective values and external metrics, while the asymptotic per-iteration order remains the same. On datasets such as Vehicle, ORL, COIL20 and Segment, the actual running time is still reasonable for experimental use.

5.5 External Clustering Performance

Tables 5.4 and 5.5 report external clustering performance. The metrics are NMI, ARI, purity and F^*. Higher values are better.

Table 5.4: Clustering performance in terms of NMI and ARI. Values are mean (standard deviation).

Dataset	Metric	FCM	UC-FCM	PFCM	MMPFCM	UC-PFCM
Iris	NMI	0.737 (0.018)	0.716 (0.090)	0.737 (0.016)	0.759 (0.024)	**0.763 (0.007)**
	ARI	0.714 (0.024)	0.695 (0.125)	0.718 (0.023)	0.754 (0.034)	**0.762 (0.008)**
Glass	NMI	0.328 (0.022)	0.309 (0.047)	0.311 (0.025)	0.317 (0.038)	**0.329 (0.017)**
	ARI	0.180 (0.014)	0.169 (0.037)	0.168 (0.027)	0.162 (0.023)	0.180 (0.012)
Urban	NMI	0.615 (0.021)	0.587 (0.092)	0.626 (0.026)	0.631 (0.022)	**0.645 (0.006)**
	ARI	0.425 (0.023)	**0.452 (0.144)**	0.411 (0.026)	0.437 (0.029)	0.447 (0.009)
Vehicle	NMI	0.098 (0.006)	0.101 (0.032)	0.105 (0.006)	0.114 (0.004)	**0.142 (0.002)**
	ARI	0.074 (0.006)	0.065 (0.032)	0.080 (0.005)	0.084 (0.005)	**0.090 (0.002)**
ORL	NMI	0.780 (0.009)	0.781 (0.009)	0.762 (0.007)	0.771 (0.004)	**0.794 (0.003)**
	ARI	0.450 (0.022)	0.437 (0.005)	0.412 (0.009)	0.412 (0.010)	**0.470 (0.068)**
Yale32	NMI	0.523 (0.016)	0.531 (0.080)	0.525 (0.018)	0.533 (0.014)	**0.543 (0.010)**
	ARI	0.266 (0.011)	0.280 (0.093)	0.251 (0.005)	0.266 (0.013)	**0.281 (0.009)**
Yale64	NMI	0.651 (0.016)	**0.660 (0.113)**	0.625 (0.010)	0.635 (0.021)	0.656 (0.035)
	ARI	**0.426 (0.034)**	0.404 (0.118)	0.379 (0.033)	0.384 (0.017)	0.423 (0.059)
Balance	NMI	0.109 (0.057)	0.119 (0.098)	0.117 (0.074)	0.143 (0.068)	**0.168 (0.038)**
	ARI	0.127 (0.064)	0.136 (0.114)	0.133 (0.074)	0.161 (0.076)	**0.195 (0.040)**
Segment	NMI	0.603 (0.010)	0.605 (0.048)	0.607 (0.009)	0.616 (0.012)	**0.617 (0.014)**
	ARI	0.491 (0.028)	0.505 (0.026)	0.499 (0.021)	0.505 (0.016)	**0.524 (0.015)**
SCADI	NMI	0.628 (0.028)	0.636 (0.028)	0.602 (0.066)	0.612 (0.040)	**0.669 (0.023)**
	ARI	0.458 (0.065)	0.467 (0.045)	0.517 (0.116)	0.413 (0.082)	**0.662 (0.033)**
COIL20	NMI	0.772 (0.011)	**0.784 (0.011)**	0.722 (0.043)	0.767 (0.006)	0.776 (0.009)
	ARI	0.584 (0.021)	**0.622 (0.025)**	0.508 (0.082)	0.586 (0.015)	0.604 (0.008)
Isolet5	NMI	0.666 (0.015)	0.708 (0.044)	0.707 (0.008)	0.708 (0.034)	**0.710 (0.000)**
	ARI	0.443 (0.019)	0.430 (0.011)	**0.448 (0.019)**	0.440 (0.041)	0.447 (0.000)

Table 5.5: Clustering performance in terms of Purity and F^*. Values are mean (standard deviation).

Dataset	Metric	FCM	UC-FCM	PFCM	MMPFCM	UC-PFCM
Iris	PUR	0.885 (0.011)	0.870 (0.077)	0.888 (0.011)	0.905 (0.016)	**0.909 (0.004)**
	F^*	0.884 (0.014)	0.869 (0.080)	0.887 (0.078)	0.905 (0.016)	**0.909 (0.011)**
Glass	PUR	0.546 (0.006)	0.513 (0.039)	0.536 (0.033)	0.540 (0.031)	**0.551 (0.016)**
	F^*	0.483 (0.020)	0.473 (0.025)	0.468 (0.035)	0.468 (0.017)	**0.489 (0.009)**
Urban	PUR	0.653 (0.026)	0.633 (0.092)	0.653 (0.025)	0.669 (0.028)	**0.691 (0.008)**
	F^*	0.649 (0.031)	0.641 (0.065)	0.642 (0.034)	0.662 (0.031)	**0.689 (0.014)**
Vehicle	PUR	0.395 (0.006)	0.371 (0.034)	0.398 (0.008)	**0.404 (0.009)**	0.402 (0.006)
	F^*	0.416 (0.014)	0.429 (0.028)	0.424 (0.011)	0.427 (0.014)	**0.446 (0.010)**
ORL	PUR	0.633 (0.033)	0.617 (0.005)	0.609 (0.057)	0.629 (0.019)	**0.640 (0.002)**
	F^*	0.625 (0.022)	0.611 (0.008)	0.588 (0.008)	0.625 (0.017)	**0.627 (0.050)**
Yale32	PUR	0.473 (0.013)	0.491 (0.078)	0.473 (0.021)	0.485 (0.037)	**0.497 (0.030)**
	F^*	0.474 (0.015)	0.534 (0.100)	0.505 (0.017)	0.501 (0.022)	**0.541 (0.019)**
Yale64	PUR	0.580 (0.023)	0.655 (0.132)	0.576 (0.026)	0.591 (0.021)	**0.675 (0.035)**
	F^*	0.668 (0.031)	**0.679 (0.142)**	0.645 (0.024)	0.652 (0.017)	0.669 (0.030)
Balance	PUR	0.642 (0.052)	0.642 (0.074)	0.649 (0.060)	0.659 (0.105)	**0.677 (0.075)**
	F^*	0.563 (0.060)	0.556 (0.093)	0.567 (0.057)	0.601 (0.080)	**0.625 (0.075)**
Segment	PUR	0.664 (0.019)	0.675 (0.010)	0.672 (0.012)	0.669 (0.010)	**0.689 (0.006)**
	F^*	0.659 (0.024)	0.673 (0.010)	0.669 (0.021)	0.670 (0.015)	**0.693 (0.012)**
SCADI	PUR	0.839 (0.028)	0.843 (0.032)	0.814 (0.061)	0.819 (0.033)	**0.857 (0.020)**
	F^*	0.655 (0.042)	0.663 (0.024)	0.606 (0.060)	0.633 (0.022)	**0.666 (0.020)**
COIL20	PUR	0.680 (0.019)	**0.717 (0.023)**	0.642 (0.067)	0.677 (0.016)	0.714 (0.024)
	F^*	0.683 (0.026)	**0.714 (0.020)**	0.647 (0.069)	0.676 (0.011)	0.699 (0.012)
Isolet5	PUR	0.543 (0.007)	0.561 (0.031)	0.561 (0.016)	0.564 (0.025)	**0.685 (0.001)**
	F^*	0.561 (0.003)	0.563 (0.022)	0.582 (0.016)	0.580 (0.042)	**0.685 (0.000)**

The results show that UC-PFCM performs especially well on NMI and purity. It obtains the best NMI on ten out of twelve datasets and the best purity on ten out of twelve datasets. For ARI and F^*, it also performs strongly and has the best average rank. These results are consistent with the lower objective values shown earlier.

5.6 Internal Validity Measures

The internal metrics DBI and XB are reported in Table 5.6. Lower values are better for both metrics. UC-PFCM obtains the best DBI value on most datasets and the best XB value on all datasets in the reported comparison.

Table 5.6: DBI and XB indices for PFCM, MMPFCM and UC-PFCM. A smaller DBI value is preferred.

Dataset	DBI $\downarrow$			XB $\downarrow$		
	PFCM	MMPFCM	UC-PFCM	PFCM	MMPFCM	UC-PFCM
Iris	1.867 (0.105)	0.998 (0.580)	**0.890 (0.184)**	2.019 (0.749)	0.609 (1.297)	**0.340 (0.213)**
Glass	2.173 (0.501)	1.965 (0.467)	**1.015 (0.124)**	2.543 (1.096)	1.875 (1.096)	**0.654 (0.831)**
Urban	2.105 (0.290)	1.971 (0.049)	**1.789 (0.074)**	2.023 (0.208)	1.952 (0.022)	**1.737 (0.061)**
Vehicle	2.694 (1.329)	2.878 (0.940)	**2.136 (0.923)**	3.498 (0.511)	6.593 (7.120)	**3.176 (4.877)**
ORL	2.057 (0.192)	**1.749 (0.110)**	1.980 (1.263)	3.588 (0.192)	2.540 (1.745)	**1.467 (2.223)**
Yale32	2.587 (0.162)	2.211 (0.236)	**1.923 (0.213)**	3.703 (0.326)	3.980 (2.539)	**2.406 (1.565)**
Yale64	1.794 (0.177)	1.719 (0.102)	**1.479 (0.044)**	2.506 (0.926)	1.276 (0.172)	**1.062 (0.085)**
Balance	2.098 (0.016)	1.835 (0.695)	**1.531 (0.228)**	1.299 (0.110)	1.201 (0.457)	**1.014 (0.243)**
COIL20	4.500 (0.164)	1.681 (0.082)	**1.508 (0.033)**	2.878 (0.727)	1.722 (0.138)	**1.336 (0.021)**
Segment	1.484 (0.174)	1.397 (0.110)	**1.308 (0.120)**	1.345 (0.496)	1.151 (0.425)	**1.054 (0.117)**
SCADI	1.681 (0.402)	1.421 (0.126)	**1.388 (0.171)**	1.523 (0.685)	0.970 (0.377)	**0.937 (0.200)**
Isolet5	2.606 (0.243)	2.551 (0.239)	**2.173 (0.228)**	2.739 (1.374)	3.264 (1.289)	**2.543 (1.168)**

The DBI and XB values support the same general conclusion as the external metrics. UC-PFCM tends to produce partitions that are not only closer to ground-truth labels but also internally compact and separated.

5.7 Statistical Significance

The Friedman test is used to evaluate whether differences among algorithms are statistically meaningful across datasets. Here $N = 12$ datasets and $K = 5$ algorithms. The critical value at significance level $\alpha = 0.05$ with $K - 1 = 4$ degrees of freedom is approximately 9.488.

Table 5.7: Friedman test results at $\alpha = 0.05$.

Metric	τ_{χ^2}	p-value
NMI	16.9	0.0020
ARI	14.6	0.0056
PUR	17.9	0.0013
F^*	17.1	0.0019

Since all test statistics exceed the critical value and all p-values are below 0.05, the null hypothesis of equal performance is rejected. This indicates that the algorithms do not perform equivalently across the tested datasets.

Table 5.8: Average ranks over the datasets; a smaller rank indicates better performance.

Algorithm	NMI	ARI	PUR	F^*
FCM	3.50	3.38	3.83	3.83
UC-FCM	3.04	3.33	3.04	3.21
PFCM	4.00	3.58	4.00	3.75
MMPFCM	2.96	3.25	2.42	2.67
UC-PFCM	**1.50**	**1.46**	**1.71**	**1.54**

The rank table favours UC-PFCM in all four external measures. This is useful because it shows that the improvement is not limited to one or two favourable datasets, but appears across the complete benchmark set.

5.8 Convergence Behaviour

The convergence pattern of UC-PFCM can be understood from the two terms in its objective. The unconstrained fuzzy term keeps the prototypes competing through their

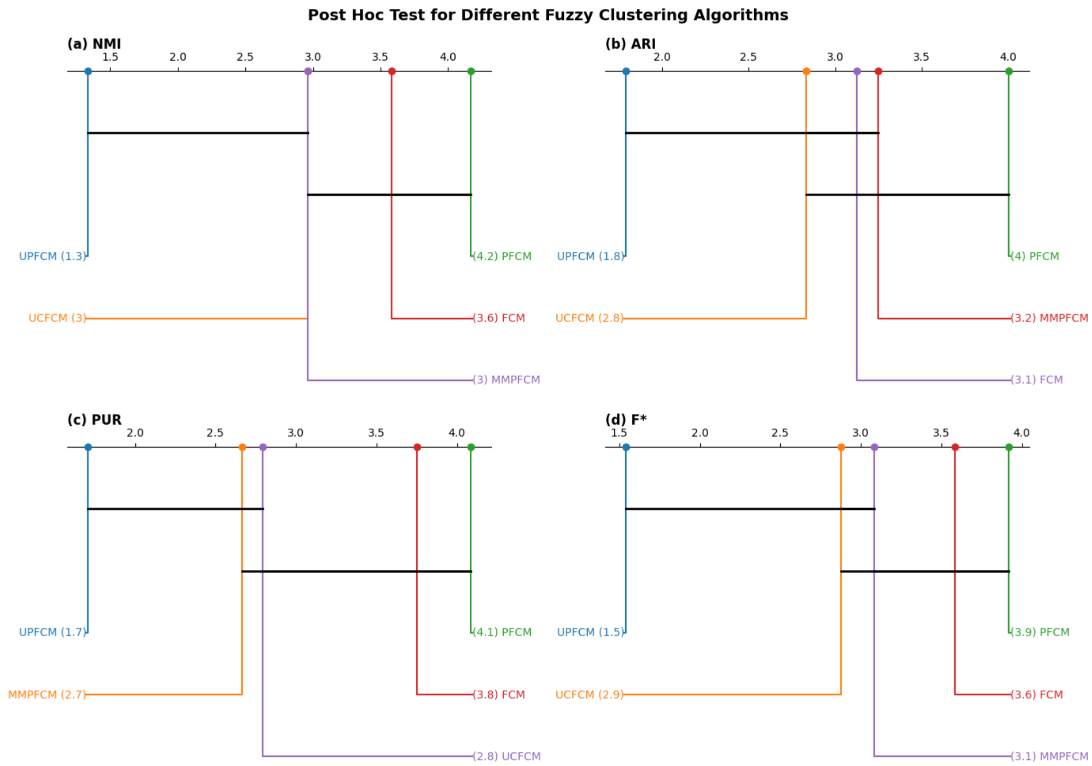

Figure 5.1: Critical difference (CD) diagrams based on the Nemenyi post-hoc test for (a) NMI, (b) ARI, (c) PUR, and (d) F^*.

distances to every sample. The typicality term then reduces the weight of samples that do not fit a prototype well. In many runs, this combination gives a more controlled path than plain alternating PFCM.

PFCM often stops quickly because its main updates have closed-form expressions. Still, reaching a stopping criterion quickly is not the same as reaching a better solution. UC-PFCM may require more time in some cases, but the objective comparison shows that it often settles at a lower value. For clustering, this distinction matters because a slightly longer run can be worthwhile if the final partition is more stable.

5.9 Discussion

The experimental results demonstrate that the proposed method is effective from several viewpoints. First, the objective values are consistently lower for UC-PFCM. Second, the average ranks across NMI, ARI, purity and F^* are best for UC-PFCM. Third, the internal metrics DBI and XB also show strong performance.

The result is especially meaningful because the proposed model does not increase the asymptotic per-iteration complexity. The improvement comes from a change in optimization structure rather than from an expensive additional procedure. This makes UC-PFCM a practical extension of PFCM.

There are also limitations. UC-PFCM depends on learning-rate selection. If the learning rate is too large, convergence may be unstable. If it is too small, convergence may be slow. The method also assumes that the number of clusters is known in advance, as do most FCM-type methods. Finally, the present experiments are based on benchmark datasets. Tests on larger and more domain-specific datasets would give additional evidence about how the method behaves in practical applications.

5.10 Research Gap Validation Through Results

The empirical results can also be read as a validation of the research gap stated in Chapter 1. If the only weakness of PFCM were the absence of typicality, then adding typicality to FCM would be sufficient. However, the experiments show that typicality alone is not the full answer: PFCM already includes it, yet it does not reach the lowest objective values. The route taken during optimization is also important. In the same way, unconstrained optimization alone is not enough, because UC-FCM is strong on some datasets but does not give the best ranks across all measures. This supports the central argument of the thesis: unconstrained optimization and possibilistic typicality need to be combined rather than treated as separate improvements.

The objective-value comparison is particularly important in this regard. UC-PFCM obtains a lower mean objective on every dataset in the PFCM-family comparison. Since the same datasets, same number of clusters and same repeated-run structure are used, the improvement cannot be explained merely by a different evaluation metric. It indicates that the proposed objective and optimization pathway are able to move the prototypes toward more favorable local minima. In clustering, this is a meaningful improvement because the quality of final labels is strongly connected to the quality of the local minimum reached by the algorithm.

The external metrics provide a second level of support. NMI, ARI, purity and F^* do not measure exactly the same aspect of clustering. NMI is information-theoretic, ARI is pair-count based, purity is cluster-dominance based and F^* combines precision-recall style reasoning. UC-PFCM does not win every single cell of every table, but it achieves the best average rank across all four metrics. This is more convincing than isolated wins because it shows that the proposed method is not overfitted to one evaluation criterion.

The internal indices give another check on the results. DBI and XB do not use class labels; they look only at the compactness and separation of the obtained clusters. The values obtained by UC-PFCM therefore suggest that the partitions are geometrically reasonable, not merely better matched to the benchmark labels. This is important for unsupervised learning, where labels are usually not available.

5.11 Limitations of the Present Study

The results are encouraging, but the study still has limitations. The first is that the number of clusters is taken as known. This is common in FCM-type experiments because benchmark datasets provide the class count. In a real unsupervised problem, that information may not be available, so UC-PFCM would need to be used together with a separate validity-based method for choosing c.

Second, the present version uses Euclidean distance. This is a standard choice in FCM research and works well for many numerical datasets, but it may be too simple when the features are correlated, sparse or arranged on a nonlinear structure. Kernel versions or metric-learning versions of UC-PFCM could be explored for such cases.

Third, the learning rate needs attention. Momentum gradient descent gives more flexibility than a fixed closed-form update, but the step size affects the run strongly. A very small value slows the method, while a large value can make the updates unstable. In this work the learning rate is chosen from a fixed candidate set; future implementations can use adaptive or line-search based choices.

Fourth, the experiments use benchmark datasets. Such datasets are useful because they make comparison possible, but they cannot represent every kind of real data. Additional experiments on large noisy datasets, medical images, text embeddings and streaming data would give a broader view of the method's behaviour.

These limitations mainly define the boundary of the present dissertation. They show where the method can be extended rather than contradicting the contribution. The work establishes UC-PFCM as a mathematically motivated clustering framework, and later studies can adapt it more specifically to different data domains.

5.12 Key Findings

The main findings from the experiments are:

1. UC-PFCM achieves the lowest mean objective value on all twelve datasets in

the PFCM-family comparison.

2. UC-PFCM obtains the best average rank for NMI, ARI, purity and F^*.
3. UC-PFCM gives strong DBI and XB values, indicating compactness and separation.
4. The proposed method remains computationally practical despite using gradient-based optimization.
5. The combination of unconstrained fuzzy optimization and possibilistic typicality is empirically more effective than using either idea alone.

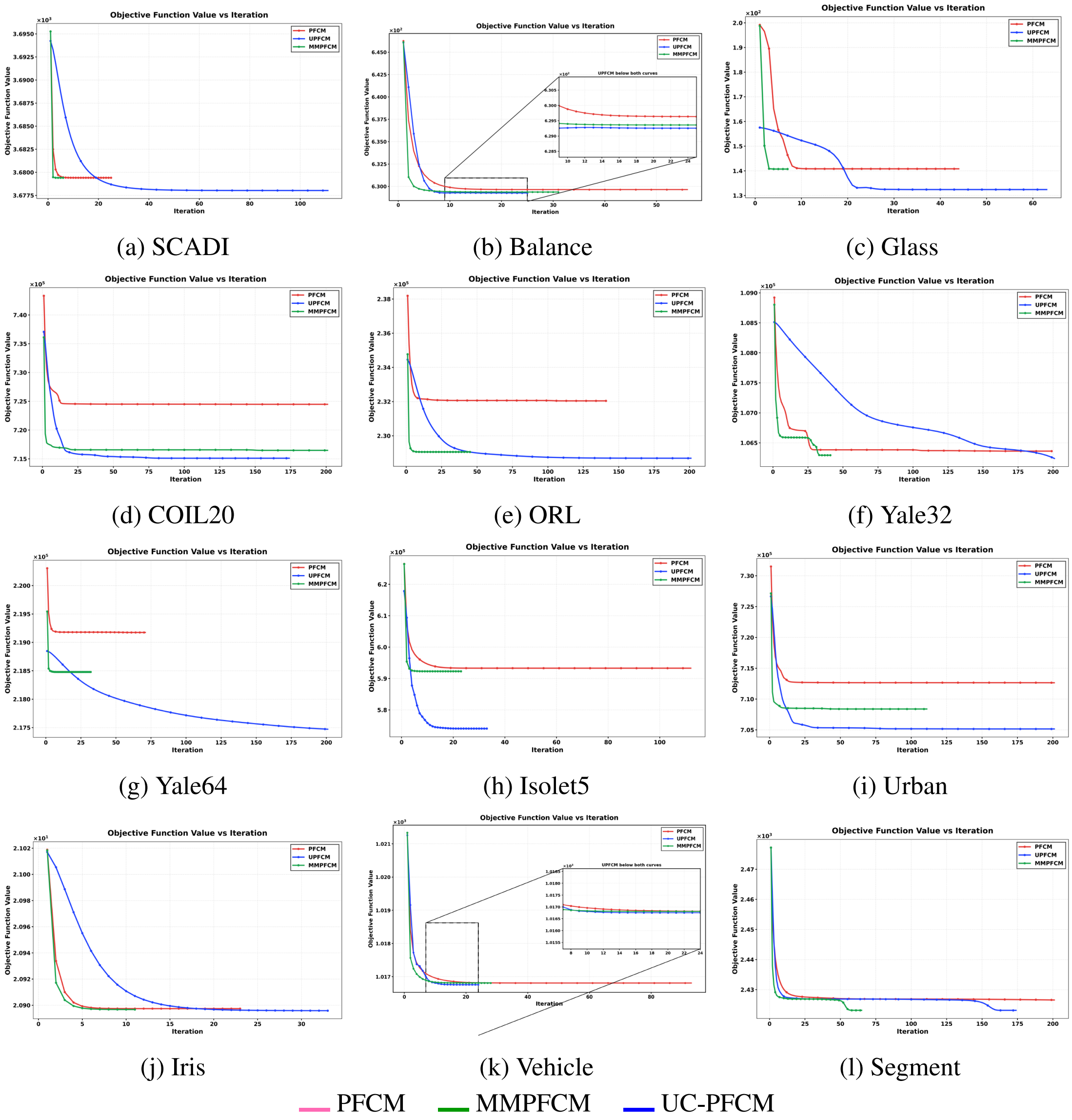

Figure 5.2: Convergence comparison of PFCM, MMPFCM, and UC-PFCM across all datasets. Lower objective values indicate better convergence.

CHAPTER 6

CONCLUSION, FUTURE SCOPE AND SOCIAL IMPACT

6.1 Conclusion

This thesis proposed an Unconstrained Possibilistic Fuzzy C-Means algorithm for unsupervised clustering. The method was motivated by two limitations in the existing literature. Classical FCM is constrained by the row-sum membership condition and is sensitive to noise. PFCM improves robustness by combining memberships and typicalities, but it is still usually optimized through alternating minimization. UC-FCM removes the explicit membership constraint, but it does not include possibilistic typicality.

The proposed UC-PFCM method combines the advantages of these approaches. The fuzzy membership component is reformulated as an unconstrained objective over cluster prototypes by substituting the analytical membership solution. The possibilistic typicality component is retained to reduce the influence of noisy and atypical samples. The resulting objective is optimized using momentum gradient descent, while typicalities are updated using a closed-form expression.

The experimental study on twelve benchmark datasets showed that UC-PFCM obtains lower objective values than PFCM and MMPFCM across all datasets. It also achieves the best average ranks across NMI, ARI, purity and F^* score. Internal validity indices such as DBI and XB further support the effectiveness of the proposed method. Overall, the results show that unconstrained prototype optimization and possibilistic typicality form a strong combination for fuzzy clustering.

6.2 Future Scope

Several directions can extend the present work.

1. **Adaptive learning-rate strategies:** The present implementation chooses the learning rate from a fixed set. Later work can try line-search methods or adaptive optimizers so that the step size changes with the progress of the algorithm.
2. **Automatic cluster-number estimation:** The present model, like FCM and PFCM, takes the number of clusters as input. A future version may combine UC-PFCM with a validity index or a penalty term so that c can be selected automatically.
3. **Kernel and deep extensions:** The method can be studied in kernel spaces, and it may also be used as a clustering loss after a learned representation in deep models.
4. **Large-scale datasets:** For very large datasets, mini-batch versions can be developed and tested so that the full distance matrix is not needed at every step.
5. **Image segmentation applications:** Since fuzzy and possibilistic methods are already common in image segmentation, UC-PFCM can be modified with spatial information and tested on medical images.
6. **Robust distance measures:** Replacing Euclidean distance with more robust distances may help when the data contain heavy-tailed noise or many outliers.

6.3 Social Impact

Clustering is often used near the beginning of a data analysis pipeline, before detailed modelling or manual inspection. Improvements in clustering can therefore help in areas such as medical imaging, remote sensing, document organization and pattern recognition. In medical image analysis, a more robust unsupervised segmentation method may reduce some manual effort for experts. In information systems, clustering can make large unlabelled collections easier to browse and organize.

At the same time, clustering results need careful interpretation. An unsupervised group is not automatically a meaningful, fair or actionable category. In sensitive areas, the output should be checked by domain experts and should not be used as a final decision by itself. The proposed method improves the mathematical and empirical side of clustering, but responsible use, clear reporting and domain-specific validation remain necessary.

6.4 Final Remarks

The main contribution of this thesis is to connect unconstrained fuzzy prototype optimization with possibilistic typicality in one model. The experiments indicate that this combination is both effective and computationally feasible. The work therefore suggests that, in fuzzy clustering, reformulating the optimization problem can be as important as adding a new distance or regularization term.

REFERENCES

[1] Guy Barrett Coleman and Harry C. Andrews. Image segmentation by clustering. *Proceedings of the IEEE*, 67(5):773–785, 1979.

[2] Edwin Diday and Jean-Claude Simon. Clustering analysis. In *Digital Pattern Recognition*, pages 47–94. Springer, Berlin, 1976.

[3] Andrea Baraldi and Palma Blonda. A survey of fuzzy clustering algorithms for pattern recognition. I. *IEEE Transactions on Systems, Man, and Cybernetics, Part B*, 29(6):778–785, 1999.

[4] Jian-Ping Mei, Yangtao Wang, Lihui Chen, and Chunyan Miao. Large scale document categorization with fuzzy clustering. *IEEE Transactions on Fuzzy Systems*, 25(5):1239–1251, 2017.

[5] James C. Bezdek. *Pattern Recognition with Fuzzy Objective Function Algorithms*. Plenum Press, New York, 1981.

[6] James C. Bezdek, Robert Ehrlich, and William Full. FCM: The fuzzy C-means clustering algorithm. *Computers & Geosciences*, 10(2–3):191–203, 1984.

[7] Raghu Krishnapuram and James M. Keller. A possibilistic approach to clustering. *IEEE Transactions on Fuzzy Systems*, 1(2):98–110, 1993.

[8] Nikhil R. Pal, Kuhu Pal, James M. Keller, and James C. Bezdek. A possibilistic fuzzy c-means clustering algorithm. *IEEE Transactions on Fuzzy Systems*, 13(4): 517–530, 2005.

[9] Feiping Nie, Runxin Zhang, Weizhong Yu, and Xuelong Li. Unconstrained fuzzy C-means algorithm. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 47(5):3440–3451, 2025.

[10] Yuxue Chen and Shuisheng Zhou. Revisiting possibilistic fuzzy c-means clustering using the majorization-minimization method. *Entropy*, 26(8):670, 2024.

[11] Lawrence Hubert and Phipps Arabie. Comparing partitions. *Journal of Classification*, 2(1):193–218, 1985.

[12] David L. Davies and Donald W. Bouldin. A cluster separation measure. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, PAMI-1(2):224–227, 1979.

[13] Xuanli Lisa Xie and Gerardo Beni. A validity measure for fuzzy clustering. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 13(8):841–847, 1991.

[14] Janez Demšar. Statistical comparisons of classifiers over multiple data sets. *Journal of Machine Learning Research*, 7:1–30, 2006.

CHAPTER A

ADDITIONAL NOTES ON IMPLEMENTATION

For consistency, the experiments were carried out with the following implementation choices:

1. Before applying the algorithms, every dataset was scaled to the interval $[0,1]$.
2. Each algorithm was run 20 times with random initialization.
3. For UC-PFCM, the stopping tolerance was fixed at 10^{-5} using the change in objective value.
4. The momentum coefficient was kept at 0.6.
5. The learning rate was chosen from $\{0.001, 0.005, 0.01, 0.05, 0.1\}$.
6. Final cluster labels were assigned by the nearest-prototype rule.

When UC-PFCM is applied to a new dataset, the parameters should be checked rather than copied blindly. The learning rate and the scale parameter, in particular, can affect both the speed of convergence and the stability of the run.

FICTA-2026:-Notification of Acceptance for Paper-ID:227 - and Guidelines for Registration & CRC Submission

1 message

FICTA-2026 <ficta2026@easychair.org>
To: Vedant Pansare <vedant3.142pansare@gmail.com>

Sat, 9 May, 2026 at 9:31 pm

Dear Authors,

Paper-ID: 227

Title: Unconstrained Possibilistic Fuzzy C-Means Algorithm

Congratulations!

On behalf of editorial board of FICTA-2026, we are happy to inform you that your above referenced paper has been ACCEPTED (based on the review reports) for presentation in FICTA-2026 conference in hybrid mode to be organized by AI and Data Science (AI & DS) Research Group, London Metropolitan University, London, United Kingdom during 08-09 June 2026.

Please refer to the reviewers' comments attached below which are intended to help you to improve your paper for final publication. You are advised to address the comments of the reviewers in your final paper suitably. Kindly take careful note of the editorial guidelines/comments in preparing the final version of your paper (CRC).

An accepted and presented paper will be published in the conference proceedings (Springer SIST Series-indexed in SCOPUS, EI Compendex, INSPEC, WTI Frankfurt eG, zbMATH, Japanese Science and Technology Agency (JST), SCImago, DBLP) only if the final version is accompanied by the registration fee payment information for at least one of the authors.

The detailed guidelines for registration and preparation of final CRC are given in the attachment.

The registration and CRC submission process should be completed by or before: 13 May 2026.

For any queries related to the registration visit the conference website or email the undersigned. Kindly mention your Paper-ID (as mentioned above) in each communication for easy tracking at our end and serving you better in answering your queries.

Thanking You

Warm Regards

General Chair

FICTA-2026, London, UK

Conference Website: https://ficta.co.uk/

E-Mail: info@ficta.co.uk

SUBMISSION: 227
TITLE: Unconstrained Possibilistic Fuzzy C-Means Algorithm

----------------------- REVIEW 1 ---------------------

SUBMISSION: 227
TITLE: Unconstrained Possibilistic Fuzzy C-Means Algorithm

----------- Overall evaluation -----------
SCORE: 2 (accept)
----- TEXT:

- Dataset has been used in the manuscript; however, no relevant citation is given.
- You have implemented effort in drawing the attention of the reader with well-organized manner.

----------------------- REVIEW 2 ---------------------

SUBMISSION: 227
TITLE: Unconstrained Possibilistic Fuzzy C-Means Algorithm

----------- Overall evaluation -----------
SCORE: 2 (accept)
----- TEXT:

- The analysis is insightful, and the conclusions are supported by strong evidence.
- Overall, the research paper is well-written and demonstrates a strong understanding of the subject matter.

Author Guidelines for Registration and Final CRC Submission.pdf
81 KB

INTERNATIONAL CONFERENCE

ON

Recent Advances in Mathematics and Mathematical Sciences

Organized by

Department of Mathematics

H. N. B. Garhwal University, Dr B. G. R. Campus Pauri, Uttarakhand-246001, India

Collaboration with

Vijñāna Parishad of India

(27th Annual and 8th International Conference)

Dr. U. C. Gairola
Convener

Mobile No. (+91) 9760231159
E-mail: mathsbgr@gmail.com

Ref. RAMMS/2026/ASP028 Date: 15/05/2026

Dear Vedant Ajit Pansare,

Thanks for the abstract submission to the **27th Annual & 8th International Conference** of **Vijñāna Parishad of India (VPI)** on **Recent Advances in Mathematics and Mathematical Sciences** (RAMMS-2026).

It is our pleasure to inform that your Abstract entitled "Unconstrained Possibilistic Fuzzy C-Means Algorithm" has been accepted for the presentation. You are cordially invited to present your paper orally at RAMMS-2026 to be held during during **June 4-6, 2026** at HNB Garhwal University, BGR Campus, Pauri. We would like to request for the mandatory registration through the following link (ignore, if you have done already).

Registration Link: **https://forms.gle/tQhpXbMZcQhLZQVK6**

Thank you for your cooperation. We look forward!

With best regards

Team
(Organizing Committee)
RAMMS-2026

Vedant Pansare

Final_MSc_Thesis_VedantPansare(1)

Assignment 1

Document Details

Submission ID

trn:oid:::27535:139761142

Submission Date

May 20, 2026, 10:14 PM GMT+5:30

Download Date

May 20, 2026, 10:21 PM GMT+5:30

File Name

Final_MSc_Thesis_VedantPansare(1).pdf

File Size

1.2 MB

54 Pages

11,522 Words

59,610 Characters

6% Overall Similarity

The combined total of all matches, including overlapping sources, for each database.

Filtered from the Report

- Bibliography
- Small Matches (less than 10 words)

Match Groups

46 Not Cited or Quoted 6%
Matches with neither in-text citation nor quotation marks

0 Missing Quotations 0%
Matches that are still very similar to source material

0 Missing Citation 0%
Matches that have quotation marks, but no in-text citation

0 Cited and Quoted 0%
Matches with in-text citation present, but no quotation marks

Top Sources

5% Internet sources
2% Publications
4% Submitted works (Student Papers)

Final_MSc_Thesis_VedantPansare(1).pdf

Indian Institute of Technology Jodhpur

Document Details

Submission ID

trn:oid:::29334:139676955

Submission Date

May 20, 2026, 10:40 AM GMT+5:30

Download Date

May 20, 2026, 10:44 AM GMT+5:30

File Name

Final_MSc_Thesis_VedantPansare(1).pdf

File Size

1.2 MB

54 Pages

11,522 Words

59,610 Characters

*% detected as AI

AI detection includes the possibility of false positives. Although some text in this submission is likely AI generated, scores below the 20% threshold are not surfaced because they have a higher likelihood of false positives.

Caution: Review required.

It is essential to understand the limitations of AI detection before making decisions about a student's work. We encourage you to learn more about Turnitin's AI detection capabilities before using the tool.

Disclaimer

Our AI writing assessment is designed to help educators identify text that might be prepared by a generative AI tool. Our AI writing assessment may not always be accurate (i.e., our AI models may produce either false positive results or false negative results), so it should not be used as the sole basis for adverse actions against a student. It takes further scrutiny and human judgment in conjunction with an organization's application of its specific academic policies to determine whether any academic misconduct has occurred.

CHAPTER B

PLAGIARISM VERIFICATION FORM

DELHI TECHNOLOGICAL UNIVERSITY

(Formerly Delhi College of Engineering)

Shahbad Daulatpur, Main Bawana Road, Delhi-110042

PLAGIARISM VERIFICATION

Title of the Thesis: *Unconstrained Possibilistic Fuzzy C-Means Algorithm*

Total Pages: ________________ Name of the Scholar: Vedant Ajit Pansare

Supervisor: Dr. Dhirendra Kumar

Department: Department of Applied Mathematics

This form records the similarity-check details for the thesis. The process and outcome are given below:

Software used: ___________________________ Similarity Index: _____________

Total Word Count: ___________________

Date: ___________________

Candidate's Signature Signature of Supervisor

DEPARTMENT OF APPLIED MATHEMATICS

DELHI TECHNOLOGICAL UNIVERSITY

(Formerly Delhi College of Engineering)

Shahbad Daulatpur, Main Bawana Road, Delhi-110042, India

CERTIFICATE OF FINAL THESIS SUBMISSION

1. Name: Vedant Ajit Pansare
2. Roll No.: 24/MSCMAT/04
3. Thesis title: *Unconstrained Possibilistic Fuzzy C-Means Algorithm*
4. Degree for which the thesis is submitted: M.Sc. Mathematics
5. Faculty of the University to which the thesis is submitted: Dr. Dhirendra Kumar

6. Thesis Preparation Guide was referred to for preparing the thesis.	YES ☐	NO ☐
7. Specifications regarding thesis format have been closely followed.	YES ☐	NO ☐
8. The contents have been organized based on the guidelines.	YES ☐	NO ☐
9. The thesis has been prepared without resorting to plagiarism.	YES ☐	NO ☐
10. All sources used have been cited appropriately.	YES ☐	NO ☐
11. The thesis has not been submitted elsewhere for a degree.	YES ☐	NO ☐
12. All the corrections have been incorporated.	YES ☐	NO ☐
13. Submitted 2 hard bound copies plus one CD.	YES ☐	NO ☐

Signature of Supervisor
Dr. Dhirendra Kumar

Signature of Candidate
Vedant Ajit Pansare
24/MSCMAT/04