

Метод полной интуиционистской нечеткой кластеризации С-средних

Нгуен Тхи Тху Зунг

Аннотация—В настоящее время, важную роль в области интеллектуального анализа данных играют методы кластеризации данных. Однако требуется разработка новых методов, позволяющих повысить эффективность обработки, увеличить результативности кластеризации. На сегодняшний день методы классификации совершенствуются, получены положительные результаты при применении интуиционистской нечеткой логики. Однако, данная техника является новой и требует разработки новых методов и алгоритмов для повышения достоверности результатов. В данной работе разработан метод интуиционистской нечеткой кластеризации С-средних на основе модификации целевой функции, учитывающей интуиционистский индекс нечеткости. Предложенный метод апробирован при решении задачи сегментации изображений. Результаты сегментации изображений оценены и сравнены с другими методами. Результат показывает, что предлагаемый метод дает лучшие результаты, чем существующие методы.

Ключевые слова— Кластеризация, нечеткая логика, интуиционистская нечеткая логика, сегментация изображений, интуиционистская нечеткая кластеризация, интуиционистский индекс нечеткости.

I. ВВЕДЕНИЕ

В последние годы информационные технологии претерпели значительные преобразования, оказав огромное влияние на развитие общества. Информационный взрыв приводит к возникновению гигантских объемов данных [1], [2]. Это побуждает людей эффективно эксплуатировать данные и перерабатывать информацию в полезные знания для удовлетворения потребностей социальной жизни. Однако реальность показывает, что всестороннее привлечение информации из огромного количества многомерных данных не может быть гарантированно эффективным при использовании обычных ручных методов. Поэтому разработка алгоритмов, которые вписываются в реальность, более востребована. Ведущим методом, внесшим большой вклад в процесс добычи данных и обработки больших данных, является метод кластеризации данных [3], [4].

Основными вопросами, которым уделяется большое внимание в кластеризации, является улучшение качества кластеризации. Первым алгоритмом кластеризации была жесткая кластеризация К-средних (KM) [5]. Алгоритм KM позволяет каждому элементу

полностью принадлежать кластеру. Это привело к ограничениям в реалистичном моделировании [6]. Поэтому появился алгоритм кластеризации нечетких С-средних (FCM) [7], [8]. FCM учитывает мягкие границы между кластерами и позволяет каждой кластеризуемой выборке данных принадлежать к разным кластерам в определенной степени [9], [10], [11]. Однако реальность в очередной раз показывает, что в случае несовершенства неточная и неполная информация приводит к неопределенности (нерешительности) [12], [13], [14], [15], [16], [17]. Поэтому родился интуиционистский нечеткий алгоритм кластеризации, в этом, каждому нечеткому элементу присваивается как значение принадлежности, так и значение непринадлежности с определенной степенью нерешительности (интуиционистский индекс нечеткости) [18], [19], [20]. Сначала в [21] предложен алгоритм кластеризации для интуиционистских нечетких множеств (IFS), основанный на концепции матрицы ассоциаций и эквивалентной матрицы ассоциаций. Затем в [22] предлагается новая интуиционистская нечеткая кластеризация С-средних с использованием новой меры расстояния для линейно и нелинейно разделяемых данных. Кроме того, в [23] предложена сегментация изображений с использованием пространственной интуиционистской нечеткой кластеризации С-средних. Далее, в [24] предложен алгоритм интуиционистской нечеткой энтропийной кластеризации для сегментации инфракрасных изображений. В работе [25] разработан новый интуиционистский нечеткий генератор множеств с применением к кластеризации. Здесь авторы предложили обобщение порождающей функции Ягера, в которой настроены все три компонента IFS. После этого в [26] возник модифицированный интуиционистский нечеткий алгоритм С-средних, учитывающий степень колебаний. Кроме того, в [27] используется алгоритм IFCM для двумерных данных о клиентах в помещении, продолжительности дел и количестве посещенных мест. Действительно, в [28] представлен обобщенный интуиционистский нечеткий алгоритм кластеризации С-средних, использующий адаптивную технику интуиционистской фазификации любого набора данных реального мира с использованием меры расстояния на основе AIFS. Исследования показывают, что применение интуиционистской теории нечеткости в методах кластеризации дает лучшие результаты, чем обычные методы, поскольку интуиционистская нечеткая кластеризация направлена на добавление коэффициента колебания точек к кластерам [29], [30], [31].

Статья получена 20 апреля 2024.

Нгуен Тхи Тху Зунг, аспирантка, Россия, Санкт-Петербург, Санкт-Петербургский Политехнический университет Петра Великого (e-mail: thudung.mta.tb@gmail.com; ORCID: <https://orcid.org/0000-0001-9206-5968>).

Меняется цель моделирования алгоритма кластеризации, что приводит к необходимости соответствующего обновления целевой функции. Поскольку интуиционистский нечеткий алгоритм кластеризации учитывает коэффициент нерешительности, это делает необходимым учет коэффициента нерешительности в целевой функции. Поэтому в этом направлении и зародилось данное исследование, улучшающее эффективность кластеризации путем модификации целевой функции алгоритма кластеризации. Кроме того, на сегодняшний день методы классификации получены положительные результаты при применении интуиционистской нечеткой логики. Однако, данная техника является новой и требует разработки новых методов и алгоритмов для повышения достоверности результатов. Существующая интуиционистская нечеткая кластеризация С-средних:

- Не учитывает интуиционистский индекс нечеткости в целевой функции;
- Алгоритм зависит от результатов FCM;
- Интуиционистский индекс нечеткости полностью включен в степень принадлежности.

Поэтому в данной работе разработан метод интуиционистской нечеткой кластеризации С-средних на основе модификации целевой функции, учитывающей интуиционистский индекс нечеткости.

Статья построена на основе 4 разделов. В разделе 1 представлено введение, в разделе 2 показаны математические основы интуиционистского метода нечеткой кластеризации С-средних и алгоритм этого метода. В разделе 3 показан эксперимент предлагаемого метода и даны оценки результатов, а также проведено сравнение результатов с другими методами. Последняя часть содержит заключение.

II. ТЕОРЕТИЧЕСКИЕ ОСНОВЫ

A. Математическая основа метода интуиционистской нечеткой кластеризации С-средних

Предположим, в s -мерном пространстве \mathbf{R}^s , имеет n элементов $\mathbf{X} = \{X_1, X_2, \dots, X_i, \dots, X_n\}$ c кластеров $X_i = (x_{i1}, x_{i2}, \dots, x_{ik}, \dots, x_{is})$, $i = \overline{1, n}$, $k = \overline{1, s}$, разделенных на c кластеры с центрами кластеров с матрицей $\hat{\mathbf{V}} \in \mathbf{V}_{cn}$, состоящей из центров кластеров $\hat{\mathbf{V}} = \{\hat{V}_1, \hat{V}_2, \dots, \hat{V}_j, \dots, \hat{V}_c\}$ с $\hat{V}_j = (\hat{v}_{j1}, \hat{v}_{j2}, \dots, \hat{v}_{jk}, \dots, \hat{v}_{js})$. Тогда кластеры, определяемые разделенными точками принадлежности, представляются следующим образом: $\hat{V}_j = \{(x_i, \hat{\mu}_{ij}(x_i), \delta_{ij}(x_i), \hat{\pi}_{ij}(x_i))\}$; $j = \overline{1, c}$. Здесь принадлежность i -го элемента к j -му кластеру определяется следующим образом:

$$\begin{cases} \mu_{ij} = \hat{\mu}_{ij} + \gamma \times \hat{\pi}_{ij} \\ \delta_{ij} = \delta_{ij} + (1 - \gamma) \times \hat{\pi}_{ij} \end{cases} \quad (1.1)$$

Где $\gamma \geq 0$ - коэффициент преобразования для X_i в кластере \hat{V}_j . Полная принадлежность i -го элемента к c кластерам определяется как $\sum_{j=1}^c \hat{\mu}_{ij} = 1$. Матрица принадлежности $\hat{\mathbf{U}} = [\hat{\mu}_{ij}] \in \mathbf{V}_{cn}$, $i = \overline{1, n}$, $j = \overline{1, c}$. Пусть множество $M_{ifc} = \{\hat{\mathbf{U}} \in \mathbf{V}_{cn}\}$ - множество интуиционистских нечетких с-разбиений X .

Степень неопределенности первоначально вычисляется вычисляются из интуиционистского нечеткого дополнения Ягера, как:

$$\hat{\pi}_{ij} = 1 - \hat{\mu}_{ij} - (1 - \hat{\mu}_{ij}^\alpha)^{1/\alpha} \quad (1.2)$$

Где $\alpha > 0$. В предлагаемом алгоритме интуиционистской нечеткой кластеризации целевая функция $J_m: (M_{ifc} \times \mathbf{R}^{cs}) \rightarrow \mathbf{R}$, которая должна быть минимизирована, имеет следующий вид:

$$J_m = \sum_{j=1}^c \sum_{i=1}^n (\hat{\mu}_{ij} + \gamma \hat{\pi}_{ij})^m d^2(X_i, \hat{V}_j) \quad (1.3)$$

Где m - коэффициент нечеткости. В предлагаемом методе используется евклидова мера расстояния, т. е.

$$d^2(X_i, \hat{V}_j) = (x_{i1} - \hat{v}_{j1})^2 + \dots + (x_{ik} - \hat{v}_{jk})^2 + \dots + (x_{is} - \hat{v}_{js})^2 \quad (1.4)$$

Использован метод множителей Лагранжа для минимизации целевой функции при условии ограничения $\sum_{j=1}^c \hat{\mu}_{ij} = 1$. Лагранжиан $L(m, \lambda)$ строится на основе целевой функции и ограничения следующим образом:

$$L(m, \lambda) = \sum_{i=1}^n \sum_{j=1}^c (\hat{\mu}_{ij} + \gamma \hat{\pi}_{ij})^m d^2(X_i, \hat{V}_j) - \sum_{i=1}^n \lambda_i (\sum_{j=1}^c \hat{\mu}_{ij} - 1) \quad (1.5)$$

Дифференцируем уравнение (1.5) относительно множителей Лагранжа λ_i и приравняем их нулю следующим образом:

$$\begin{aligned} \frac{\partial J_m}{\partial \hat{\mu}_{ij}} &= 0 \\ \Rightarrow m(\hat{\mu}_{ij} + \gamma \hat{\pi}_{ij})^{m-1} &\left(1 - \gamma + \gamma \hat{\mu}_{ij}^{\alpha-1} (1 - \hat{\mu}_{ij}^\alpha)^{\frac{1}{\alpha}-1}\right) d^2(X_i, \hat{V}_j) - \lambda \end{aligned}$$

$$\Rightarrow \hat{\mu}_{ij} = \left(\frac{\lambda}{m}\right)^{\frac{1}{m-1}} \left[\left(1 - \gamma + \gamma \hat{\mu}_{ij}^{(\alpha-1)} (1 - \hat{\mu}_{ij}^\alpha)^{\frac{1}{\alpha}-1}\right) d^2 \right]^{\frac{-1}{m-1}} - \gamma \hat{\pi}_{ij} \quad (1.6)$$

Аналогично производные условия Лагранжа (1.5) приравниваются к нулю по параметру принадлежности $\hat{\mu}_{ij}$:

$$\frac{\partial J_m}{\partial \lambda} = \sum_{j=1}^c \hat{\mu}_{ij} - 1$$

$$\frac{\partial J_m}{\partial \lambda} = 0 \Rightarrow \sum_{j=1}^c \hat{\mu}_{ij} = 1 \quad (1.7)$$

От (1.6) и (1.7) получим

$$\left(\frac{\lambda}{m}\right)^{\frac{1}{m-1}} \sum_{j=1}^c \left[\left(1 - \gamma + \gamma \hat{\mu}_{ij}^{(\alpha-1)} (1 - \hat{\mu}_{ij}^\alpha)^{\frac{1}{\alpha}-1}\right) d^2 \right]^{\frac{-1}{m-1}} - \gamma \sum_{j=1}^c \hat{\pi}_{ij} = 1$$

В результате получается итерационная формула для значения принадлежности $\hat{\mu}_{ij}$ следующим образом:

$$\hat{\mu}_{ij} = \frac{1}{\sum_{j=1}^c \frac{z d^{\frac{m-1}{\alpha}}}{z d^{\frac{m-1}{\alpha}}}} (1 + \gamma \sum_{j=1}^c \hat{\pi}_{ij}) - \gamma \hat{\pi}_{ij} \quad (1.8)$$

где $z = \left[1 - \gamma + \gamma \hat{\mu}_{ij}^{(\alpha-1)} (1 - \hat{\mu}_{ij}^\alpha)^{\frac{1}{\alpha}-1}\right]^{\frac{-1}{m-1}}$

Тогда, центр кластера $\hat{V}_j = (\hat{v}_{j1}, \hat{v}_{j2}, \dots, \hat{v}_{jk}, \dots, \hat{v}_{jd})$ вычислена как:

$$\hat{V}_j = \frac{\sum_j^N [\hat{\mu}_{ij} + \gamma - \gamma \hat{\mu}_{ij} - \gamma (1 - \hat{\mu}_{ij}^\alpha)^{1/\alpha}]^m X_i}{\sum_j^N [\hat{\mu}_{ij} + \gamma - \gamma \hat{\mu}_{ij} - \gamma (1 - \hat{\mu}_{ij}^\alpha)^{1/\alpha}]^m} \quad (1.9)$$

В. Алгоритма и этапы реализации предложенного интуиционистского нечеткого метода кластеризации С-средних

Этапы реализации предложенного интуиционистского нечеткого метода кластеризации С-средних состоят в следующем:

Шаг 1: Вводные данные. Сначала входные данные с n элементами в пространстве p размерностей.

Шаг 2: Определение количества кластеров c . На этом этапе для расчета количества кластеров в методе кластеризации будет использован статистический алгоритм интуиционистского нечеткого разрыва.



Схема 1. Процесс реализации предложенного интуитивного метода нечеткой кластеризации.

Шаг 3: Настройки начальных параметров. На этом шаге можно задать: значение непрозрачности m , степень принадлежности значений матрицы $U = \{\mu_j(x_i)\}, i = \overline{1, n}, j = \overline{1, c}$ случайным образом, и значение коэффициента преобразования γ .

Шаг 4: Определение центра кластера. Вычисление центров кластеров $V_j^* = (\mu_{V_j^*}, \nu_{V_j^*}, \pi_{V_j^*}), j = \overline{1, c}$ с помощью формул (1.9).

Шаг 5: Определение матрицы принадлежности. Рассчитывается матрица принадлежности $U = \{\mu_j(x_i)\}, i = \overline{1, n}, j = \overline{1, c}$ с помощью формул (1.8) со значениями принадлежности элементов к кластерам по центрам вычисленных кластеров.

Шаг 6: подсчитать значение целевой функции. Целевая функция J_m вычисляется по формуле (1.3), а расстояние - по формуле (1.4). Если $J_m \leq \epsilon$, то остановить алгоритм и принять результаты. Если $J_m > \epsilon$ то вернуться к шагу 4.

Шаг 7: Получение результатов.

Алгоритм остановится, когда функция J_m достигнет минимального значения.

Шаг 8: Получение результатов.

Алгоритм предлагаемого интуиционистского метода нечеткой кластеризации представлен на схеме 1.

III. ЭКСПЕРИМЕНТЫ И АНАЛИЗ

Предложенный метод интуиционистской нечеткой кластеризации С-средних применяется в задаче сегментации изображений. Результаты сегментации изображений сравниваются с методом нечеткой кластеризации С-средних и традиционным интуиционистским методом нечеткой кластеризации С-средних.

А. Вход

В данной работе для эксперимента используются 6 различных изображения, все изображения имеют соответствующие эталоны для сравнения и оценки эффективности кластеризации.

В. Реализация

Рассмотрим каждый пиксель как элемент, значение которого (R, G, B) является значением цвета пикселя, тогда задача определена: Количество элементов - это количество пикселей в изображении; Каждое значение каждого элемента имеет 3 измерения (R, G, B); Количество кластеров - это количество областей изображения, которые мы хотим разделить.

С. Результаты и оценки

Результаты кластеризации представлены в рисунке 1.

Сравнение результатов с мерами точности кластеризации: В этой работе используются метрики точности пикселей для оценки точности кластеризации методов, включая в таблице 1:

Таблица 1. Меры точности кластеризации

Мера	Значение / Ожидаемое значение
ASA – Средняя точность сегментации $ASA = \sum_{i=1}^c \frac{ X_i \cap Y_i }{\sum_{j=1}^c X_j }$	- Измеряет долю истинного сегмента, которая правильно помечена. - Максимальное значение ASA
DSC – Коэффициент Дайса $DSC = \frac{2 X_i \cap Y_i }{ X_i + Y_i }$	- Определяет меру «похожести» двух областей. - Максимальное значение DSC.
JS – Коэффициент Жаккара $JS = \frac{ X_i \cap Y_i }{ X_i \cup Y_i }$	- Измеряет сходство путем оценки доли общих элементов между двумя наборами. - Максимальное значение JS.
PC – Коэффициент разделения $PC = \frac{1}{n} \sum_{i=1}^c \sum_{j=1}^n \mu_{ij}^{*m}$	- Разбиение с меньшей нечеткостью означает лучшую эффективность. - Максимальное значение PC.
PE – Энтропия разделения $PE = -\frac{1}{N} \sum_{i=1}^c \sum_{j=1}^n \mu_{ij}^* \log \mu_{ij}^*$	- Разбиение с меньшей нечеткостью означает лучшую эффективность. - Минимальное значение PE.
XB –	- При хорошей

<p>Индекс Xie-Beni</p> $V_{XB} = \frac{\sum_{i=1}^c \sum_{j=1}^n \mu_{ij}^{*2} \ V_i - X_j\ ^2}{n \cdot \max \ V_i - X_j\ ^2}$	<p>кластеризации образцы уплотняются в пределах одного кластера, а образцы разделяются между разными кластерами. - Минимальное значение XB.</p>
--------------------------------------------------------------------------------------------------------------------------------	-----------------------------------------------------------------------------------------------------------------------------------------------------

Результаты расчетов по метрикам, приведенным в таблице 2, для FCM, IFCM и предложенного метода, которые экспериментально использовались для сегментации 3 изображений на рисунке 1, приведены в таблице 2.

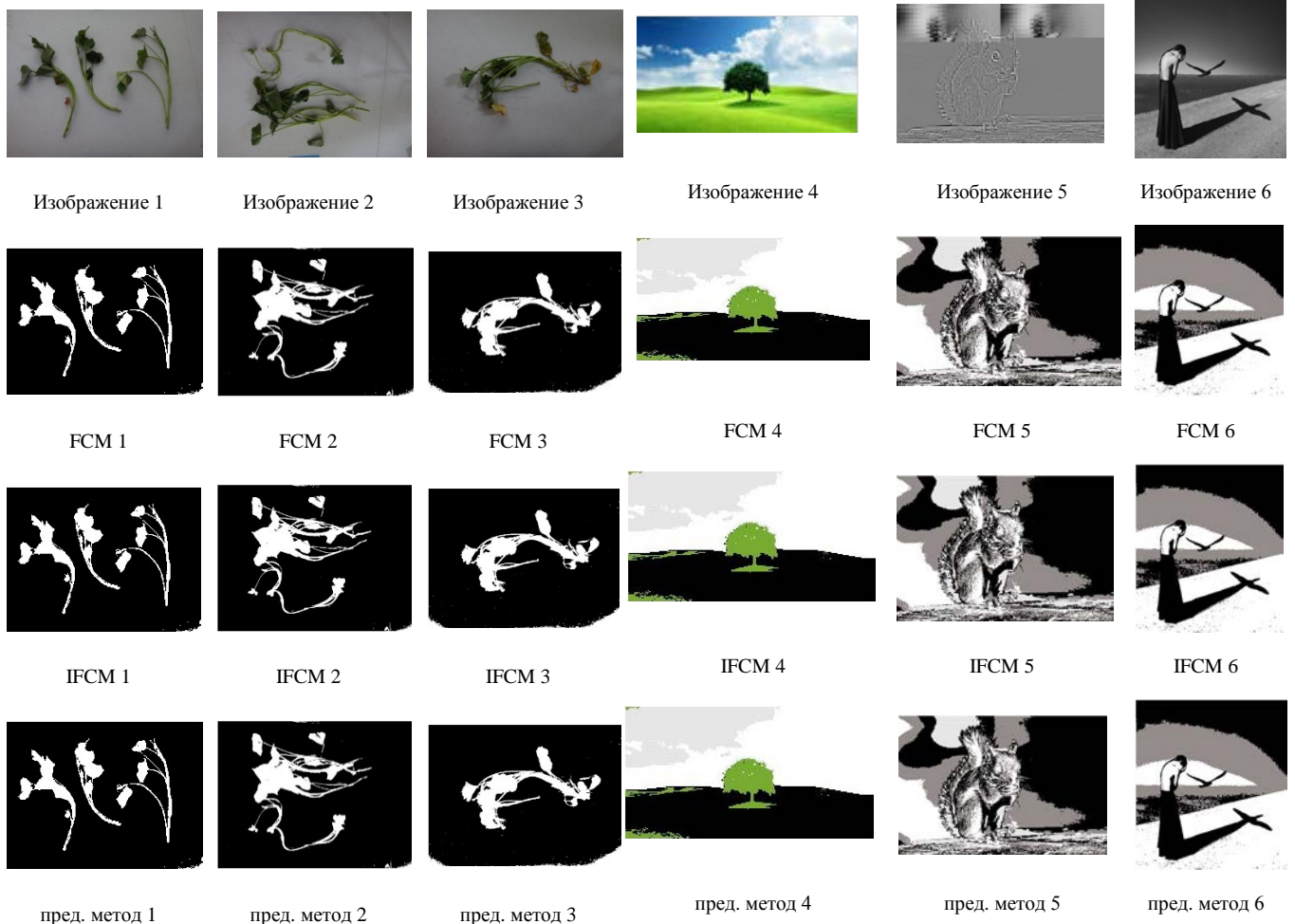


Рис 1. Сегментации изображений

Таблица 2. Сравнение результатов сегментации изображений с помощью метрик

Мера $m = 2;$ $\alpha = 0.95$ $\gamma = 1.$	Изображение 1			Изображение 2			Изображение 3			Средние эффективнее (%)	
	FCM	IFCM	Пред. м	FCM	IFCM	Пред. м	FCM	IFCM	Пред. м	чем FCM	чем IFCM
ASA	0.9705	0.9729	0.9789	0.9613	0.9624	0.9735	0.9204	0.9212	0.9381	1,35	1,21
DSC	0.8886	0.8817	0.8905	0.8825	0.8796	0.8833	0.6879	0.6649	0.6956	0.48	2.02
JS	0.7996	0.7884	0.8025	0.7897	0.7850	0.7911	0.5243	0.4980	0.5333	0.76	3.22
PC	0.9224	0.9202	0.9458	0.9253	0.9236	0.9459	0.8801	0.8750	0.9100	2.72	3.06
PE	0.0630	0.0646	0.0448	0.0601	0.0614	0.0440	0.0909	0.0943	0.0695	26.44	28.44
V_{XB}	0.0335	0.0352	0.0335	0.0305	0.0319	0.0309	0.0590	0.0651	0.0576	0.31	6.50

Выводы по результатам:

По результатам сравнения с использованием метрик в таблице видно, что **предлагаемый метод дает лучшие результаты, чем метод FCM и метод IFCM** (среднее значение измерений ASA, DSC, JS, PC, PE, XB) на 5.3%, чем FCM, и на 7.4%, чем IFCM конкретно:

ASA (Средняя точность сегментации): Предлагаемый метод на **1,35% лучше, чем FCM**, и на **1,21% лучше, чем IFCM**.

DSC (Коэффициент Дайса - определяет меру «похожести» двух областей): Предлагаемый метод на **0,48% лучше, чем FCM**, и на **2,02% лучше, чем IFCM**.

JS (Коэффициент Жаккара - Измеряет сходство путем оценки доли общих элементов между двумя наборами): Предлагаемый метод на **0,76% лучше, чем FCM**, и на **3,22% лучше, чем IFCM**.

PC (Коэффициент разделения - Разбиение с меньшей нечеткостью означает лучшую эффективность):

Предлагаемый метод на **2,72% лучше, чем FCM**, и на **3,06% лучше, чем IFCM**.

PE (Энтропия разделения - Разбиение с меньшей нечеткостью означает лучшую эффективность): Предложенный метод на **26,44% лучше, чем FCM** и на **28,44% лучше, чем IFCM**.

XB (Индекс Xie-Beni - При хорошей кластеризации образцы уплотняются в пределах одного кластера, а образцы разделяются между разными кластерами): Предложенный метод на **0,31% лучше, чем FCM**, и на **6,5% лучше, чем IFCM**.

IV. ЗАКЛЮЧЕНИЕ

В данной работе разработан метод интуиционистской нечеткой кластеризации С-средних решает недостатки существующих методов нечеткой кластеризации С-средних путем учета интуиционистского индекса нечеткости в объективной функции, так как интуиционистский нечеткий алгоритм кластеризации рассматривает нерешительности. В этом направлении улучшена производительность кластеризации путем модификации объективной функции алгоритма кластеризации. Предложенный метод целесообразно применять в обработке изображений, решении задач кластеризации многомерных данных и проводить интуитивную фазификацию данных, применяя в моделях прогнозирования нечетких временных рядов и интуитивных нечетких системах вывода. Предложенная интуиционистская нечеткая модель кластеризации С-средних применяется в обнаружении, сегментировании областей изображения объектов на изображении. При сравнении результатов видно, что предлагаемый метод дает лучшие результаты, чем метод FCM и метод IFCM, при этом предлагаемый метод лучше (среднее значение измерений ASA, DSC, JS, PC, PE, XB) на 5.3%, чем FCM, и на 7.4%, чем IFCM.

БЛАГОДАРНОСТИ

Выражаем благодарность учёным и исследователям в рассматриваемой области, предложившим эффективное направление прогнозирования временных рядов, которое является основой для разработки новых улучшений, предложенных в этой работе.

БИБЛИОГРАФИЯ

- [1] Т. Т. З. Нгуен and Л. В. Черненькая, "Системный анализ в управлении развитием территориальных комплексов вьетнама," in *В сборнике: Системный анализ в проектировании и управлении. сборник научных трудов XXV Международной научной и учебно-практической конференции*, СПб, 2021, pp. 346–352.
- [2] Т. Т. З. Нгуен and Л. В. Черненькая, "Модель анализа факторов на основе нечеткой кластеризации с-средних," in *Журнал Известия Тульского государственного университета – Технические науки (ТулГУ, г. Тула)*, no. 1, Тула: Издательство ТулГУ, 2023, pp. 329–336.
- [3] Т. Т. З. Нгуен and Л. В. Черненькая, "Эвристическая нечеткая модель прогнозирования высокого порядка с хедж-алгеброй," in *Системный анализ в проектировании и управлении*, 2023.
- [4] Т. Т. З. Нгуен and Л. В. Черненькая, "Новая гауссовская система нечёткого логического вывода типа Такаги-Сугено-Канга с весом главного компонента," *International Journal of Open Information Technologies*, vol. 12, no. 5, pp. 116–124, 2024.
- [5] S. S. Khan and A. Ahmad, "Cluster center initialization algorithm for K-means clustering," *Pattern Recognit Lett*, vol. 25, no. 11, pp. 1293–1302, 2004, doi: 10.1016/j.patrec.2004.04.007.
- [6] Нгуен Тхи Тху Зунг and Л. В. Черненькая, "Дискретизация в моделях прогнозирования нечетких временных рядов," *Журнал Известия Тульского государственного университета – Технические науки (ТулГУ, г. Тула)*, vol. 8, no. Системный анализ, Управление и обработка информации, pp. 296–304, 2023, doi: 10.24412/2071-6168-2023-8-296-297.
- [7] James C. Bezdek, *Pattern Recognition with Fuzzy Objective Function Algorithms*.
- [8] J. C. Bezdek, "A Convergence Theorem for the Fuzzy ISODATA Clustering Algorithms," *IEEE Trans Pattern Anal Mach Intell*, vol. PAMI-2, no. 1, pp. 1–8, 1980, doi: 10.1109/TPAMI.1980.4766964.
- [9] Т. Т. З. Нгуен and Л. В. Черненькая, "Модель Прогнозирования Эвристических Нечетких Временных Рядов Высокого Порядка, Основанная На Хедж-Алгебраическом Подходе Часть 3," *Журнал Известия Тульского государственного университета – Технические науки (ТулГУ, г. Тула)*, vol. 9, no. Системный анализ, Управление и обработка информации, 2023.
- [10] Нгуен Тхи Тху Зунг and Л. В. Черненькая, "Модели для анализа развития экономики нескольких районов вьетнама на основе математических методов многомерной статистики," in *В сборнике: Импульс организационных инноваций. Сборник конкурсных работ 1 межвузовского конкурса студентов, магистрантов и аспирантов*, Москва, 2020, pp. 301–324.
- [11] V. Garcia Domingo and J. M. Nuñez Do Rio Carles Ventura Royo, "Fuzzy C-means and clustering algorithms: a comparative study," 2019.
- [12] Нгуен Тхи Тху Зунг and Васильевна Черненькая Людмила, "Модель Прогнозирования Эвристических Нечетких Временных Рядов Высокого Порядка, Основанная На Хедж-Алгебраическом Подходе Часть 1," *Журнал Известия Тульского государственного университета – Технические науки (ТулГУ, г. Тула)*, vol. 9, no. Системный анализ, Управление и обработка информации, 2023.
- [13] A. Jurio, D. Paternain, H. Bustince, C. Guerra, and G. Beliakov, "A construction method of Atanassov's Intuitionistic fuzzy set for image processing," in *5th IEEE International Conference Intelligent Systems*, Institute of Electrical and Electronics Engineers, Ed., University of Westminster, London, UK: 2020 IEEE International Conference on Intelligent Systems, Jul. 2010, pp. 337–342.
- [14] K. Atanassov, "Type-1 Fuzzy Sets and Intuitionistic Fuzzy Sets," *Algorithms*, vol. 10, no. 3, p. 106, 2017, doi: 10.3390/a10030106.
- [15] K. T. Atanassov, "Operators over interval valued intuitionistic fuzzy sets," *Fuzzy Sets Syst*, vol. 64, no. 2, pp. 159–174, 1994, doi: 10.1016/0165-0114(94)90331-X.
- [16] K. T. Atanassov, "Two theorems for intuitionistic fuzzy sets," *Fuzzy Sets Syst*, vol. 110, no. 2, pp. 267–269, 2000, doi: 10.1016/S0165-0114(99)00112-8.
- [17] K. T. Atanassov, "Review and new results on intuitionistic fuzzy sets," *International Journal Bioautomation*, vol. 20, pp. S17–S26, 2016.
- [18] K. T. Atanassov, "Intuitionistic fuzzy sets," *Fuzzy Sets Syst*, vol. 20, pp. 87–96, 1986.
- [19] K. T. Atanassov, "More on intuitionistic fuzzy sets," *Fuzzy set and Systems*, vol. 33, pp. 37–45, 1989.
- [20] K. Atanassov and G. Gargov, "Interval valued intuitionistic fuzzy sets," *Fuzzy Sets Syst*, vol. 31, no. 3, pp. 343–349, 1989, doi: 10.1016/0165-0114(89)90205-4.
- [21] Z. Xu and J. Wu, "Intuitionistic fuzzy C-means clustering algorithms," *Journal of Systems Engineering and Electronics*, vol. 21, no. 4, pp. 580–590, 2010, doi: 10.3969/j.issn.1004-4132.2010.04.009.
- [22] P. Kaur, A. K. Soni, and A. Gosain, "Novel intuitionistic fuzzy c-means clustering for linearly and nonlinearly separable data," *WSEAS Transactions on Computers*, vol. 11, no. 3, pp. 65–76, 2012.
- [23] B. K. Tripathy, A. Basu, and S. Govel, "Image segmentation using spatial intuitionistic fuzzy C means clustering," *2014 IEEE International Conference on Computational Intelligence and Computing Research, IEEE ICCIC 2014*, no. September 2015, pp. 10–12, 2014, doi: 10.1109/ICCIC.2014.7238446.
- [24] X. Zhou, R. Zhao, F. Yu, and H. Tian, "Intuitionistic fuzzy entropy clustering algorithm for infrared image segmentation," *Journal of Intelligent and Fuzzy Systems*, vol. 30, no. 3, pp. 1831–1840, 2016, doi: 10.3233/IFS-151894.
- [25] M. Kaushal, R. Solanki, Q. M. Danish Lohani, and P. K. Muhuri, "A novel intuitionistic fuzzy set generator with application to clustering," *IEEE International Conference on Fuzzy Systems*, vol. 2018-July, pp. 1–8, 2018, doi: 10.1109/FUZZ-IEEE.2018.8491602.

- [26] H. Verma, A. Gupta, and D. Kumar, "A modified intuitionistic fuzzy c-means algorithm incorporating hesitation degree," *Pattern Recognit Lett*, vol. 122, pp. 45–52, 2019, doi: 10.1016/j.patrec.2019.02.017.
- [27] O. Dogan, B. Oztaysi, and C. Fernandez-Llatas, "Segmentation of indoor customer paths using intuitionistic fuzzy clustering: Process mining visualization," *Journal of Intelligent and Fuzzy Systems*, vol. 38, no. 1, pp. 675–684, 2020, doi: 10.3233/JIFS-179440.
- [28] M. Kaushal and Q. M. D. Lohani, "Generalized intuitionistic fuzzy c-means clustering algorithm using an adaptive intuitionistic fuzzification technique," *Granular Computing*, vol. 7, no. 1, pp. 183–195, 2022, doi: 10.1007/s41066-021-00259-1.
- [29] Z. Xu and R. R. Yager, "Some geometric aggregation operators based on intuitionistic fuzzy sets," *Int J Gen Syst*, vol. 35, no. 4, pp. 417–433, Aug. 2006, doi: 10.1080/03081070600574353.
- [30] M. Dhyani, G. S. Kushwaha, and S. Kumar, "A novel intuitionistic fuzzy inference system for sentiment analysis," *International Journal of Information Technology (Singapore)*, vol. 14, no. 6, pp. 3193–3200, Oct. 2022, doi: 10.1007/s41870-022-01014-8.
- [31] O. Castillo and P. Melin, "A new method for fuzzy inference in intuitionistic fuzzy systems," *Annual Conference of the North American Fuzzy Information Processing Society - NAFIPS*, vol. 2003-Janua, pp. 20–25, 2003, doi: 10.1109/NAFIPS.2003.1226749.

A Complete Intuitionistic Fuzzy C-Means Clustering Method

Nguyen Thi Thu Dung

Abstract—Currently, data clustering methods play an important role in the field of data mining. However, the development of new methods is required to increase the efficiency of processing and increase the effectiveness of clustering. Today, classification methods are being improved, and positive results have been obtained when using intuitionistic fuzzy logic. However, this technique is new and requires the development of new methods and algorithms to increase the reliability of the results. In this paper, a method of intuitionistic fuzzy C-means clustering is developed based on a modification of the objective function that takes into account the intuitionistic fuzzy index. The proposed method was tested in solving the problem of image segmentation. Image segmentation results are evaluated and compared with other methods, showing that the proposed method produces better results than existing methods.

Keywords— Clustering, fuzzy logic, intuitionistic fuzzy logic, image segmentation, intuitionistic fuzzy clustering, hesitance.

REFERENCES

- [1] T. T. Z. Nguen and L. V. Chernen'kaja, "Sistemnyj analiz v upravlenii razvitiem territorial'nyh kompleksov v'etnama," in V sbornike: Sistemnyj analiz v proektirovanii i upravlenii. sbornik nauchnyh trudov XXV Mezhdunarodnoj nauchnoj i uchebno-prakticheskoy konferencii, SPB, 2021, pp. 346–352.
- [2] T. T. Z. Nguen and L. V. Chernen'kaja, "Model' analiza faktorov na osnove nechetkoj klasterizacii s-srednih," in Zhurnal Izvestija Tul'skogo gosudarstvennogo universiteta – Tehnicheskie nauki (TulGU, g. Tula), no. 1, Tula: Izdatel'stve TulGu, 2023, pp. 329–336.
- [3] T. T. Z. Nguen and L. V. Chernen'kaja, "Jevristicheskaja nechetkaja model' prognozirovanija vysokogo porjadka s hedzh-algebroj," in Sistemnyj analiz v proektirovanii i upravlenii, 2023.
- [4] T. T. Z. Nguen and L. V. Chernen'kaja, "Novaja gaussovskaja sistema nechjotkogo logicheskogo vyvoda tipa Takagi-Sugeno-Kanga s vesom glavnogo komponenta," International Journal of Open Information Technologies, vol. 12, no. 5, pp. 116–124, 2024.
- [5] S. S. Khan and A. Ahmad, "Cluster center initialization algorithm for K-means clustering," Pattern Recognit Lett, vol. 25, no. 11, pp. 1293–1302, 2004, doi: 10.1016/j.patrec.2004.04.007.
- [6] Nguyen Thi Thu Zung and L. V. Chernen'kaja, "Diskretizacija v modeljah prognozirovanija nechetkih vremennyh rjadov," Zhurnal Izvestija Tul'skogo gosudarstvennogo universiteta – Tehnicheskie nauki (TulGU, g. Tula), vol. 8, no. Sistemnyj analiz, Upravlenie i obrabotka informacii, pp. 296–304, 2023, doi: 10.24412/2071-6168-2023-8-296-297.
- [7] James C. Bezdek, Pattern Recognition with Fuzzy Objective Function Algorithms.
- [8] J. C. Bezdek, "A Convergence Theorem for the Fuzzy ISODATA Clustering Algorithms," IEEE Trans Pattern Anal Mach Intell, vol. PAMI-2, no. 1, pp. 1–8, 1980, doi: 10.1109/TPAMI.1980.4766964.
- [9] T. T. Z. Nguen and L. V. Chernen'kaja, "Model' Prognozirovanija Jevristicheskikh Nechetkih Vremennyh Rjadov Vysokogo Porjadka, Osnovannaja Na Hedzh-Algebraicheskom Podhode Chast' 3," Zhurnal Izvestija Tul'skogo gosudarstvennogo universiteta – Tehnicheskie nauki (TulGU, g. Tula), vol. 9, no. Sistemnyj analiz, Upravlenie i obrabotka informacii, 2023.
- [10] Nguyen Thi Thu Zung and L. V. Chernen'kaja, "Modeli dlja analiza razvitiya jekonomiki neskol'kih rajonov v'etnama na osnove matematiceskikh metodov mnogomernoj statistiki," in V sbornike: Impul's organizacionnyh innovacij. Sbornik konkursnyh rabot I mezhvuzovskogo konkursa studentov, magistrantov i aspirantov, Moskva, 2020, pp. 301–324.
- [11] V. Garcia Domingo and J. M. Nuñez Do Rio Carles Ventura Royo, "Fuzzy C-means and clustering algorithms: a comparative study," 2019.
- [12] Nguen Thi Thu Zung and Vasil'evna Chernen'kaja Ljudmila, "Model' Prognozirovanija Jevristicheskikh Nechetkih Vremennyh Rjadov Vysokogo Porjadka, Osnovannaja Na Hedzh-Algebraicheskom Podhode Chast' 1," Zhurnal Izvestija Tul'skogo gosudarstvennogo universiteta – Tehnicheskie nauki (TulGU, g. Tula), vol. 9, no. Sistemnyj analiz, Upravlenie i obrabotka informacii, 2023.
- [13] A. Jurio, D. Paternain, H. Bustince, C. Guerra, and G. Beliakov, "A construction method of Atanassov's Intuitionistic fuzzy set for image processing," in 5th IEEE International Conference Intelligent Systems, Institute of Electrical and Electronics Engineers, Ed., University of Westminster, London, UK: 2020 IEEE International Conference on Intelligent Systems , Jul. 2010, pp. 337–342.
- [14] K. Atanassov, "Type-1 Fuzzy Sets and Intuitionistic Fuzzy Sets," Algorithms, vol. 10, no. 3, p. 106, 2017, doi: 10.3390/a10030106.
- [15] K. T. Atanassov, "Operators over interval valued intuitionistic fuzzy sets," Fuzzy Sets Syst, vol. 64, no. 2, pp. 159–174, 1994, doi: 10.1016/0165-0114(94)90331-X.
- [16] K. T. Atanassov, "Two theorems for intuitionistic fuzzy sets," Fuzzy Sets Syst, vol. 110, no. 2, pp. 267–269, 2000, doi: 10.1016/S0165-0114(99)00112-8.
- [17] K. T. Atanassov, "Review and new results on intuitionistic fuzzy sets," International Journal Bioautomation, vol. 20, pp. S17–S26, 2016.
- [18] K. T. Atanassov, "Intuitionistic fuzzy sets," Fuzzy Sets Syst, vol. 20, pp. 87–96, 1986.
- [19] K. T. Atanassov, "More on intuitionistic fuzzy sets," Fuzzy set and Systems, vol. 33, pp. 37–45, 1989.
- [20] K. Atanassov and G. Gargov, "Interval valued intuitionistic fuzzy sets," Fuzzy Sets Syst, vol. 31, no. 3, pp. 343–349, 1989, doi: 10.1016/0165-0114(89)90205-4.
- [21] Z. Xu and J. Wu, "Intuitionistic fuzzy C-means clustering algorithms," Journal of Systems Engineering and Electronics, vol. 21, no. 4, pp. 580–590, 2010, doi: 10.3969/j.issn.1004-4132.2010.04.009.
- [22] P. Kaur, A. K. Soni, and A. Gosain, "Novel intuitionistic fuzzy c-means clustering for linearly and nonlinearly separable data," WSEAS Transactions on Computers, vol. 11, no. 3, pp. 65–76, 2012.
- [23] B. K. Tripathy, A. Basu, and S. Govel, "Image segmentation using spatial intuitionistic fuzzy C means clustering," 2014 IEEE International Conference on Computational Intelligence and Computing Research, IEEE ICCIC 2014, no. September 2015, pp. 10–12, 2014, doi: 10.1109/ICCIC.2014.7238446.
- [24] X. Zhou, R. Zhao, F. Yu, and H. Tian, "Intuitionistic fuzzy entropy clustering algorithm for infrared image segmentation," Journal of Intelligent and Fuzzy Systems, vol. 30, no. 3, pp. 1831–1840, 2016, doi: 10.3233/IFS-151894.
- [25] M. Kaushal, R. Solanki, Q. M. Danish Lohani, and P. K. Muhuri, "A novel intuitionistic fuzzy set generator with application to clustering," IEEE International Conference on Fuzzy Systems, vol. 2018-July, pp. 1–8, 2018, doi: 10.1109/FUZZ-IEEE.2018.8491602.
- [26] H. Verma, A. Gupta, and D. Kumar, "A modified intuitionistic fuzzy c-means algorithm incorporating hesitation degree," Pattern Recognit Lett, vol. 122, pp. 45–52, 2019, doi: 10.1016/j.patrec.2019.02.017.
- [27] O. Dogan, B. Oztaysi, and C. Fernandez-Llatas, "Segmentation of indoor customer paths using intuitionistic fuzzy clustering: Process mining visualization," Journal of Intelligent and Fuzzy Systems, vol. 38, no. 1, pp. 675–684, 2020, doi: 10.3233/JIFS-179440.
- [28] M. Kaushal and Q. M. D. Lohani, "Generalized intuitionistic fuzzy c-means clustering algorithm using an adaptive intuitionistic fuzzification technique," Granular Computing, vol. 7, no. 1, pp. 183–195, 2022, doi: 10.1007/s41066-021-00259-1.
- [29] Z. Xu and R. R. Yager, "Some geometric aggregation operators based on intuitionistic fuzzy sets," Int J Gen Syst, vol. 35, no. 4, pp. 417–433, Aug. 2006, doi: 10.1080/03081070600574353.

[30] M. Dhyani, G. S. Kushwaha, and S. Kumar, "A novel intuitionistic fuzzy inference system for sentiment analysis," *International Journal of Information Technology (Singapore)*, vol. 14, no. 6, pp. 3193–3200, Oct. 2022, doi: 10.1007/s41870-022-01014-8.

[31] O. Castillo and P. Melin, "A new method for fuzzy inference in intuitionistic fuzzy systems," *Annual Conference of the North American Fuzzy Information Processing Society - NAFIPS*, vol. 2003-Janua, pp. 20–25, 2003, doi: 10.1109/NAFIPS.2003.1226749.