

УДК 615.471

**І.Е. Крашений<sup>1</sup>, А.О. Попов<sup>1</sup>**, канд. техн. наук, **Х. Рамірез<sup>2</sup>**, доктор філософії,  
**Х.М. Горріз<sup>2</sup>**, доктор філософії

<sup>1</sup>каф. фізичної та біомедичної електроніки,

Національний технічний університет України «Київський Політехнічний Інститут»,  
каб. 423, вул. Політехнічна, 16, Київ-56, 03056, Україна.

<sup>2</sup>каф. теорії сигналів, телематики та комунікацій, Університет Гранади,  
вул. Данієля Сауседо Аранди, Гранада, 18071, Іспанія.

## **Використання методів кластеризації в системах нечіткого виводу для діагностики хвороби Альцгеймера на основі ПЕТ-зображень**

*Робота присвячена використанню методів кластеризації в системах нечіткого виводу для класифікації ПЕТ-зображень з метою діагностики хвороби Альцгеймера. Оцінені характеристики кожного з трьох представлених кластеризаційних методів: Subtractive Clustering, C-means та Fuzzy Grid Partition. Надані рекомендації щодо використання методу Subtractive Clustering у системах нечіткого виводу для автоматичної діагностики хвороби Альцгеймера, як методу, що показав найкращі результати з AUC=0,8791. Бібл. 20, рис. 3, табл. 3.*

**Ключові слова:** хвороба Альцгеймера; слабоумство; нечітка логіка; позитронно-емісійна томографія.

### **Вступ**

Хвороба Альцгеймера (ХА) – це захворювання, при якому гинуть клітини головного мозку, що призводить до тяжких розладів пам'яті, інтелекту, інших когнітивних функцій, а також до серйозних проблем в емоційно-особистісній сфері і поведінці [1].

На сьогоднішній день хвороба Альцгеймера визнана однією з найбільш поширених причин старечого слабоумства (до 40% всіх випадків) [1,2]. У 2013 році статистичний звіт від Асоціації Альцгеймера показав, що у віковій групі старших 65 років частина хворих осіб складає 10%. У віковій групі від 85 років, частка становить приблизно 33%. Кількість хворих постійно зростає. За останні роки вона збільшилася з 24 мільйонів (2001) до майже 36 мільйонів (2010). За найбільш оптимістичним прогнозом, діагноз «хвороба Альцгеймера» буде поставлений до 2050 року більше 100 мільйонам хворих у всьому світі.

Головна причина зростання числа пацієнтів з хворобою Альцгеймера – збільшення тривалості життя в розвинених країнах і в більшості

країн, що розвиваються. За статистикою у людей пенсійного віку ризик захворіти хворобою Альцгеймера збільшується вдвічі через кожні п'ять років життя. При такій залежності, чим вище кількість літніх людей у віці від 65 років і чим довше живуть ці люди, тим масштабніше поширення хвороби Альцгеймера.

Від даної хвороби немає ліків. На сьогоднішній день існують тільки інгібітори ХА, метою застосування яких є полегшити її перебіг, що є тривалим процесом, що проходить впродовж років. Ці факти роблять ранню діагностику особливо важливою, оскільки своєчасна діагностика та лікування спроможні полегшити життя як хворому, так і його близьким та рідним.

Позитронно-емісійна томографія є одним із найбільш використовуваних методів об'єктивної діагностики ХА завдяки можливості оцінити церебральний кровоток, який зазнає суттєвих змін під час розвитку хвороби.

Діагностика на основі даних зображень є складною процедурою, оскільки потребує не тільки високоякісного обладнання, а й роботи високодосвідченого спеціаліста. Даний факт спонукає дослідників спрямовувати зусилля на розробку методів автоматизованої комп'ютерної діагностики.

Одним з можливих шляхів для побудови системи автоматизованої діагностики є використання підходу нечіткої логіки [3,4,5]. Даний підхід базується на використанні бази правил та функцій приналежності, на основі яких відбувається «поєднання» входу нечіткої системи із виходом. Але використання даного підходу суттєво ускладнено через необхідність створення бази правил вручну на основі експертних даних. Для малої кількості ознак даний недолік не є суттєвим, але зі зростанням їх кількості розробка ускладнюється.

Саме тому метою даної роботи є оглянути та порівняти існуючі методи автоматичного синтезу

системи нечіткого виводу для діагностики хвороби Альцгеймера.

**Системи нечіткого виводу**

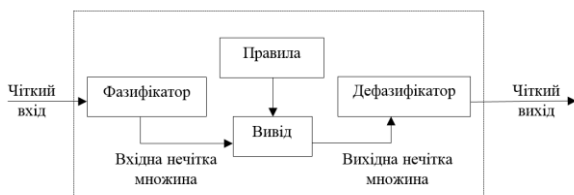
Нечітка логіка є узагальненням класичної логіки та теорії множин, що базується на понятті нечіткої множини як об'єкта з функцією приналежності елементу до множини, що набуває будь-які значення на інтервалі [0,1], а не тільки 0 або 1. Тобто, якщо описані вищі методи дозволяють чітко визначити хворий пацієнт чи здоровий, то використання апарату нечіткої логіки потенційно дає змогу оцінити «на скільки» хворий.

Поняття систем нечіткого виводу має центральне місце у нечіткій логіці та в теорії управління. Нечітким логічним виводом називається процес отримання нечітких висновків стосовно виду нечіткої множини, відповідного поточним значенням входів з використанням бази знань нечітких операцій. Основою нечіткого логічного виводу складає композиційне правило Заде, що формується наступним чином: якщо відомо нечітке відношення  $\tilde{R}$  між вхідною ( $x$ ) та вихідною ( $y$ ) змінними, то при нечіткому значенні вхідної змінної  $x = \tilde{A}$ , нечітке значення вихідної змінної визначається як:

$$y = \tilde{A} \circ \tilde{R}$$

де  $\circ$  – максимальна композиція.

Незалежно від алгоритму, за яким працює система нечіткого виводу, всі вони мають однукову внутрішню будову зображену на рис. 1 [6].



**Рис. 1. Загальна внутрішня будова систем нечіткого виводу**

В даній роботі пропонується використовувати один з найпростіших в реалізації алгоритм Сугено. Розглянемо його особливості більш детально і для цього введемо наступні позначення:  $x, y$  – вхідні змінні;  $x_0, y_0$  – чіткі значення  $x, y$ ;  $z$  – вихідна змінна;  $\mu_{A_1}, \mu_{A_2}, \mu_{B_1}, \mu_{B_2}, \mu_{C_1}, \mu_{C_2}$  – функції приналежності  $x, y$ ;

та правила:

ЯКЩО  $x \in \mu_{A_1}$  ТА  $y \in \mu_{B_1}$  ТОДІ  $z \in \mu_{C_1}$ ;

ЯКЩО  $x \in \mu_{A_2}$  ТА  $y \in \mu_{B_2}$  ТОДІ  $z \in \mu_{C_2}$ .

*Алгоритм Сугено.* В алгоритмі Сугено база знань будується наступним чином:

ЯКЩО  $x \in \mu_{A_1}$  ТА  $y \in \mu_{B_1}$  ТО  $z_1 = \alpha_1 x + \beta_1 y$ ,

ЯКЩО  $x \in \mu_{A_2}$  ТА  $y \in \mu_{B_2}$  ТО  $z_2 = \alpha_2 x + \beta_2 y$ .

Нечіткий вивід реалізується наступним чином:

Фазифікація: знаходяться ступені істинності для передумов кожного з правил:

$$\mu_{A_1}(x_0), \mu_{A_2}(x_0), \mu_{B_1}(y_0), \mu_{B_2}(y_0).$$

Вивід: знаходяться рівні відтину для передумов кожного з правил:

$$\alpha_1 = \mu_{A_1}(x_0) \cap \mu_{B_1}(y_0),$$

$$\alpha_2 = \mu_{A_2}(x_0) \cap \mu_{B_2}(y_0).$$

Також знаходяться індивідуальні виходи правил:

$$z_1^* = \alpha_1 x + \beta_1 y,$$

$$z_2^* = \alpha_2 x + \beta_2 y,$$

Визначення чіткого значення змінної виводу:

$$z_0 = \frac{\alpha_1 z_1^* + \alpha_2 z_2^*}{\alpha_1 + \alpha_2}.$$

**Методи синтезу систем нечіткого виводу**

Оскільки внутрішня структура (кількість правил, кількість входів та кількість змінних у правилах) системи нечіткого виводу суттєво залежить від кількості вхідних ознак, то при наявності великої кількості вхідних ознак стає важко створювати подібного роду систему вручну. Тому існують автоматичні методи створення та оптимізації нечітких систем.

*Fuzzy Grids partition.* Нечіткі сітки є одним із найбільш простих методів створення правил на основі тренувальної послідовності. Даний метод формує поділ за допомогою розділення вхідного простору на декілька нечітких шарів, кожен з яких є визначеним функцією приналежності для кожного виміру ознак. Для оптимізації розмір та розташування такої «сітки» зазвичай використовують метод градієнтного спуску. Головним недоліком такого генерування бази правил є експоненціальна залежність кількості правил від кількості входів або функцій приналежності [7].

S-means є одним із найбільш розповсюджених алгоритмів кластеризації, що використовується в побудові систем нечіткого виводу, завдяки невеликій кількості параметрів, що мають бути визначені. Для реалізації даного методу синтезу системи нечіткого виводу необхідно визначити кількість кластерів, рівень нечіткості,

критерій зупинки там матриця, що описує форму кластерів.

У загальному випадку задача кластеризації зводиться до оптимізації цільової функції:

$$J_m(U, V, X) = \sum_{i=1}^c \sum_{j=1}^N u_{ik}^m d_{ik}^2 = \sum_{i=1}^c \sum_{j=1}^N u_{ik}^m \|x_j - v_i\|^2,$$

де  $U = [u_{ik}] \in M_{fc}$  нечіткий поділ матриці  $Z$ ,  $V = [V_1, V_2, V_3, \dots, V_C]$ ,  $V_i \in R^n$  вектор «центрів» кластерів,  $D_{ikA} = \|x_k - V_i\|_A^2 = (x_k - V_i)^T A (x_k - V_i)$  – квадрат норми відстані, та  $m \in [1, \infty)$ , що визначає рівень нечіткості вихідних кластерів [8,9,10].

*Subtractive clustering algorithm* є іншим широко-використовуваним методом. Даний метод є модифікацією методу гір, що був описаний у [11]. Метод гір, в свою чергу, дуже подібний до алгоритму нечітких сіток, але реалізується завдяки поділу простору даних за допомогою розрахунку локальних потенціальних точок. Це досягається наступним чином: простір даних поділяється за допомогою сітки на кластери, надалі, визначається «важливість» (потенціал) кожного вузла сітки за допомогою обчислення кількості точок навколо.

В методі гір опорним точками (горами) є вузли сітки, в той час як у алгоритму від'ємної кластеризації такими точками є самі дані [11].

Потенціал «гори» визначається в даному методі як:

$$P_i = \sum_{j=1}^n \exp\left(-\frac{4\|x_i - x_j\|^2}{r_a^2}\right),$$

де  $x_i$  – нормалізований вектор даних, що визна-

чається як  $x_i = \frac{x_i^j - \min\{x^i\}}{\max\{x^i\} - \min\{x^i\}}$ ,  $n$  – загальна

кількість векторів даних,  $r_a$  – радіус гіперкулі у просторі даних, що описує величину кластеру.

Після визначення потенціалу всіх «гір» необхідно визначити «гору» з найбільшим потенціалом, яка в подальшому стане вершиною

першого кластера. В той час як значення потенціалу всіх інших «гір» мають бути оновленні відповідно до:

$$P_i^1 = P_i - P_1^* \exp\left(-\frac{4\|x_i - x_1^*\|^2}{r_b^2}\right),$$

де  $P_1^*$  – найбільша гора,  $x_1^*$  – координата найвищої гори, а  $r_b$  – радіус гіперкулі у просторі даних, що описує величину кластеру [11].

### Діагностична система нечіткого виводу для діагностики Хвороби Альцгеймера

Всі зображення були попередньо оброблені. Була виконана просторова нормалізація, а також реорганізація вокселів для відповідності всіх зображень одному єдиному стандартному шаблону. В якості ознак в даній роботі пропонується використовувати середній церебральний кровоток в локальних областях мозку у найбільш схильних до змін регіонах під час хвороби Альцгеймера.

Для розрахунку даних ознак був використаний спеціальний анатомічний атлас IBASPM[12]. IBASPM являє собою один із найбільш детальних та найбільш використовуваних атласів мозку людини. Даний атлас існують у кількох версіях, які відрізняються деталізацією: IBASPM116 (був використаний у даному дослідженні), IBASPM90, IBASPM84, IBASPM71, IBASPM69. Даний атлас також є статистичним та базується на використанні усереднених шаблонів мозку людини запропонованих Montreal Neurological Institute and Hospital (MNI). Дані шаблони отримані на основі 241 МРТ-зображення здорових людей. Ці зображення були вручну масштабовані та усереднені.

В даній роботі була використана частина бази даних томографічних зображень ADNI (adni.loni.usc.edu). Дана база містить в собі 249 FDG ПЕТ-зображень пацієнтів на різних стадіях хвороби Альцгеймера (табл. 1). Частина бази, що була використана в даній роботі містить в собі тільки зображення пацієнтів з діагностованою хворобою Альцгеймера (AD) та здорових контрольних пацієнтів (NC).

Таблиця 1. Демографічні дані пацієнтів в бази даних ADNI

Діагноз	Кількість	Вік	Стать (Ч/Ж)	MMSE
NC	68	75,97±5,0	43/25	29,06±1,08
MCI	111	74,85±7,4	76/35	26,68±2,16
AD	70	75,36±7,5	46/24	22,84±2,91

Для сортування регіонів за найбільшими змінами був використаний *t*-критерій [13,14] (1) для двох незалежних вибірок, в кожній з яких було розраховано середній кровоток.

$$t = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{\frac{(n_1 - 1)s_{X_1}^2 + (n_2 - 1)s_{X_2}^2}{n_1 + n_2 - 2}} \cdot \sqrt{\frac{2}{n}}}, \quad (1)$$

де  $n_1$  та  $n_2$  – кількість значень у виборці,  $s_{X_1}^2$  та  $s_{X_2}^2$  – незміщені оцінки дисперсії.

В даній роботі пропонується використовувати 24 області, зміни в яких являються найбільш суттєвими. Дані регіони приведені у табл. 2.

Таблиця 2. Відібрані за інформативністю області інтересу

№	Регіон ПЕТ	№	Регіон ПЕТ
1	Angular_R	13	Vermis_9
2	Angular_L	14	Paracentral_Lobule_L
3	Cerebelum_4_5_L	15	Vermis_8
4	Cerebelum_6_L	16	Parietal_Inf_L
5	Cingulum_Post_L	17	Cerebelum_8_R
6	Cerebelum_3_L	18	Postcentral_L
7	Cerebelum_4_5_R	19	Temporal_Inf_R
8	Cerebelum_3_R	20	Temporal_Mid_R
9	Cerebelum_6_R	21	Putamen_L
10	Vermis_3	22	Precentral_L
11	Postcentral_R	23	Precuneus_R
12	Vermis_1_2	24	Vermis_4_5

Загальна схема системи наведена на рис. 2.

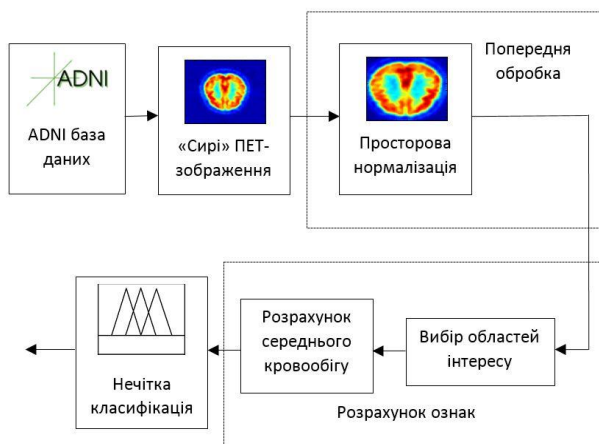


Рис. 2. Загальна схема діагностичної системи нечіткого виводу

### Експериментальне дослідження методів кластеризації в системах нечіткого виводу

В даній роботі були виконані експериментальні дослідження системи, представленої на рис. 2, з використанням три різні методи описані в дані роботі. Для валідації роботи класифікатора була використана методика leave-one-out [15]. В якості критеріїв порівняння якості даних методів кластеризації пропонується використовувати криві помилок (ROC-криві), площу під кривою помилок (AUC) та стандартні діагностичні міри, відомі у статистиці як класифікаційні функції. Дані міри можна обчислити на основі матриці невідповідностей [14,17,18,19,20].

### Результати

На рис. 3 зображені криві помилок для кожного методу кластеризації. Як видно, метод Fuzzy Grid Partition має найгірші результати серед трьох, в той час, як методи *c*-means та Subtractive Clustering мають приблизно однакові результати. Ці дані підтверджуються і величиною площі під кривою, що була розрахована. Так  $AUC_{FGP}=0,7395$ ,  $AUC_{CM}=0,8539$ ,  $AUC_{SC}=0,8791$ .

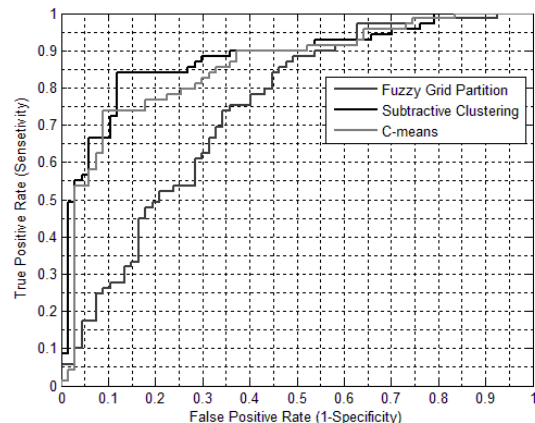


Рис. 3. Криві помилок для трьох методів нечіткої кластеризації

В даній роботі також були розраховані стандартні діагностичні міри (табл. 3) для трьох методів нечіткої кластеризації. Дані результати також підтверджують висновки, отримані на основі кривих помилок, а саме метод Subtractive Clustering та метод C-means дають приблизно однакові результати, суттєво вищі за метод Fuzzy Grid Partition.

Таблиця 3. Стандартні діагностичні міри розраховані для трьох методів нечіткої кластеризації

	Точність (accuracy)	Специфічність (specificity)	Чутливість (sensitivity)	Позитивна прогностична точність (positive predictive value)
Subtractive Clustering	0,9241	0,8965	0,8788	0,8406
C-means	0,8626	0,8743	0,8947	0,7391
Fuzzy Grid Partition	0,8574	0,8742	0,6892	0,7391

## Висновки

В роботі представлено можливе використання методів кластеризації в системах нечіткого виводу для діагностики хвороби Альцгеймера. Показано, що метод Fuzzy Grid Partition суттєво поступається за результатами методам c-means та Subtractive Clustering, які мають приблизно однакові результати. Алгоритм Subtractive Clustering за своїми результатами є більш застосовним у діагностиці хвороби Альцгеймера.

## Список використаних джерел

1. *Mayeux R.* Early Alzheimer's Disease / R. Mayeux. // *New England Journal of Medicine.* – 2010. – №362. – С. 2194–2201.
2. *Nowotny P.* Alzheimer Disease / P. Nowotny, J. Known, A. Goate. // *ENCYCLOPEDIA OF LIFE SCIENCES* Nature Publishing Group www.els.net. – 2001. – С. doi10.1038/npg.els.0000228.
3. *Zadeh L.* Fuzzy sets / L. Zadeh. // *Information and control.* – 1965. – №8. – С. 338–353.
4. *Zadeh L.* Fuzzy algorithms / L. Zadeh. // *Information and control.* – 1968. – №12. – С. 94–102.
5. *Zadeh L.* Fuzzy sets versus probability / L. Zadeh. // *Proceedings of the IEEE.* – 1980. – №68. – С. 421.
6. *Sharma D.* Designing and modeling fuzzy control Systems / D. Sharma. // *International Journal of Computer Applications.* – 2011. – №16. – С. 46–53.
7. *Hu Y.* Simple Fuzzy Grid Partition for Mining Multiple-level Fuzzy Sequential Patterns / Y.-C. Hu. // *Cybernetics and Systems.* – 2007. – №38. – С. 203–228.
8. *Bezdek J.* The fuzzy c-means clustering algorithm / J. Bezdek, R. Ehrlich, W. Full. // *Computers & Geosciences.* – 1984. – №10. – С. 191–203.
9. *Bezdek J.* Pattern recognition with fuzzy objective function algorithms / J. Bezdek. – New York: Springer, 1981. – 256 с.
10. *Cannon R.* Efficient implementation of the fuzzy c-means clustering algorithms / R. Cannon, J. Dave, J. Bezdek. // *IEEE Transactions on Pattern Analysis and Machine Intelligence.* – 1986. – №2. – С. 248–255.
11. *Yager R.* Approximate clustering via the mountain method / R. Yager, D. Filev. // *IEEE Transactions on Systems, Man, and Cybernetics.* – 1994. – №24. – С. 1279–1284.
12. *Alemán-Gómez Y.* IBASPM: Toolbox for automatic parcellation of brain structures. 12th Annual Meeting of the Organization for Human Brain Mapping / Y. Alemán-Gómez, L. Melie-García, P. Valdés-Hernandez. // *Neuroimage (CD-Rom in NeuroImage).* – 2006. – №27.
13. *Student.* THE PROBABLE ERROR OF A MEAN / Student. // *Biometrika.* – 1908. – №6. – С. 1–25.
14. *Mathematical statistics and data analysis* – Belmont: Duxbury Press, 2007. – 688 с. – (Available 2010 Titles Enhanced Web Assign).
15. *Arlot S.* A survey of cross-validation procedures for model selection. *Statistics Surveys* / S. Arlot, A. Celisse. // *Statistics Surveys.* – 2010. – №4. – С. 40–79.
16. *Geisser S.* Predictive inference: an introduction / Geisser. – New York: Chapman & Hall, 1993. – 240 с.
17. *Powers D.* Evaluation: from precision, recall and F-measure to ROC, informedness, markedness and correlation / Powers. // *Journal of Machine Learning Technologies.* – 2011. – №2. – С. 37–63.
18. *Goutte C.* A probabilistic interpretation of precision, recall and F-score, with implication for evaluation / C. Goutte, F. Gaussier. // *Advances in Information Retrieval.* – 2005. – №3408. – С. 345–359.
19. *Metz C.* Basic principles of ROC analysis / Metz. // *Seminars in Nuclear Medicine.* – 1978. – №11. – С. 283–298.
20. *Fawcett T.* An introduction to ROC analysis / Fawcett. // *Pattern Recognition Letters.* – 2006. – №27. – С. 861–874.

Поступила в редакцію 25 декабря 2015 г.

УДК 615.471

**И.Э. Крашенный<sup>1</sup>, А.А. Попов<sup>1</sup>**, канд.техн.наук, **Х. Рамирез<sup>2</sup>**, доктор философии,  
**Х.М. Горриз<sup>2</sup>**, доктор философии

<sup>1</sup>каф.физической и биомедицинской электроники,

Национальный технический университет Украины «Киевский Политехнический Институт»,  
каб. 423, ул. Политехническая, 16, Киев-56, 030566 Украина.

<sup>2</sup>каф. теорисигналов, телематики та комуникаций, Университет Гранады,  
вул. Даниэля Саусэдо Аранды, Гранада, 18071, Испания.

## Применение методов кластеризации для диагностики болезни Альцгеймера на основе ПЭТ-изображений

*Данная работа посвящена применению методов кластеризации в системах нечеткого вывода для классификации ПЭТ-изображений с целью диагностики болезни Альцгеймера. Оценены характеристики каждого из трех представленных методов кластеризации: Subtractive Clustering, C-means и Fuzzy Grid Partition. Представлены рекомендации касательно использования метода Subtractive Clustering в системах нечеткого вывода для автоматизированной диагностики болезни Альцгеймера, как метода, который показал наилучшие результаты с  $AUC=0,8791$ . Библ. 20, рис. 3, табл. 3.*

**Ключевые слова:** болезнь Альцгеймера; слабоумие; нечеткая логика; позитронно-эмиссионная томография.

UDC 615.471

**I. Krasheniy<sup>1</sup>, A. Popov<sup>1</sup>**, Ph.D, **J. Ramirez<sup>2</sup>**, Ph.D, **J.M. Gorriz<sup>2</sup>**, Ph.D

<sup>1</sup>Department of physical and biomedical electronics,

National technical university of Ukraine "Kyiv Polytechnic Institute",  
of. 423, Politekhnichna str., 16, Kyiv-56, 03056, Ukraine.

<sup>2</sup>Department of Signal Processing, Telematics and Communications, University of Granada,  
Daniel Sausedo Aranda str. Granada, 18071, Spain.

## Fuzzy clustering methods application for Alzheimer's diseases diagnosis based on PET images

*This work was dedicated to clustering methods application in fuzzy inference system for Alzheimer's disease diagnosis using PET-images. Three methods (Subtractive Clustering, C-means and Fuzzy Grid Partition) of clustering were discussed and their performance in Alzheimer's disease diagnosis were measured.*

*Recommendation of the future use of Subtractive Clustering algorithm in the computer-aided diagnosis system for Alzheimer's disease are given. The performance of this algorithm is  $AUC=0,8791$ . Ref. 20, fig. 3, tab. 3.*

**Keywords:** Alzheimer's disease; dementia; fuzzy logic; positron emission tomography.

### Reference

1. Mayeux, R. (2010). Early Alzheimer's disease. *New England Journal of Medicine* 362, pp. 2194–2201.
2. Nowotny, P., Kwon, J.M., and Goate, A. M. (2001). Alzheimer Disease. In *Encyclopedia of Life Sciences*, John Wiley & Sons, Ltd, ed. (Chichester, UK: John Wiley & Sons, Ltd),.
3. Zadeh, L. A. (1965). Fuzzy sets. *Information and Control* 8, 338–353.
4. Zadeh, L. A. (1968). Fuzzy algorithms. *Information and Control* 12, 94–102.
5. Zadeh, L. A. (1980). Fuzzy sets versus probability. *Proceedings of the IEEE* 68, 421–421.

6. *Sharma, D.* (2011). Designing and modeling fuzzy control Systems. *International Journal of Computer Applications* 16, 46–53.
7. *Hu, Y.-C.* (2007). SIMPLE FUZZY GRID PARTITION FOR MINING MULTIPLE-LEVEL FUZZY SEQUENTIAL PATTERNS. *Cybernetics and Systems* 38, 203–228.
8. *Bezdek, J. C., Ehrlich, R., and Full, W.* (1984). FCM: The fuzzy c-means clustering algorithm. *Computers & Geosciences* 10, 191–203.
9. *Bezdek, J. C.* (1981). *Pattern recognition with fuzzy objective function algorithms* (New York: Plenum Press).
10. *Cannon, R. L., Dave, J. V., and Bezdek, J. C.* (1986). Efficient implementation of the fuzzy c-means clustering algorithms. *Pattern Analysis and Machine Intelligence, IEEE Transactions on* pp. 248–255.
11. *Yager, R. R., and Filev, D. P.* (1994). Approximate clustering via the mountain method. *IEEE Transactions on Systems, Man, and Cybernetics* 24, pp. 1279–1284.
12. *Alemán-Gómez Y., Melie-García L., Valdés-Hernandez. P.* (2006) IBASPM: Toolbox for automatic parcellation of brain structures. The 12th Annual Meeting of the Organization for Human Brain Mapping, June 11-15, 2006, Florence, Italy. Available on CD-Rom in *NeuroImage*, Vol. 27, No.1.
13. *STUDENT* (1908). THE PROBABLE ERROR OF A MEAN. *Biometrika* 6, pp. 1–25.
14. *Rice, J. A.* (2007). *Mathematical statistics and data analysis* (Belmont, CA: Thomson/Brooks/Cole).
15. *Arlot, S., and Celisse, A.* (2010). A survey of cross-validation procedures for model selection. *Statistics Surveys* 4, pp. 40–79.
16. *Geisser, S.* (1993). *Predictive inference: an introduction* (New York: Chapman & Hall).
17. *Powers, D. M.* (2011). Evaluation: from precision, recall and F-measure to ROC, informedness, markedness and correlation.
18. *Goutte, C., and Gaussier, E.* (2005). A probabilistic interpretation of precision, recall and F-score, with implication for evaluation. In *Advances in Information Retrieval*, (Springer), pp. 345–359.
19. *Metz, C. E.* (1978). Basic principles of ROC analysis. In *Seminars in Nuclear Medicine*, (Elsevier), pp. 283–298.
20. *Fawcett, T.* (2006). An introduction to ROC analysis. *Pattern Recognition Letters* 27, pp. 861–874.