Table 1 States created by fuzzy k-medoids with 9 number of states using 60 features chosen by hiererchical clustering.

States	Aminoacids
1	R, N, D, C, Q, M, P, W, Y
2	C, I, M, F, P, W, Y
3	R, N, D, C, G, P, S, T, Y
4	A, N, C, G, M, P, S
5	I, L, M, W
6	C, G, I, M, V
7	R, N, K
8	D, E, P
9	N, C, H, W, Y

Sequence	A	I	M	A	L	K	M	R
States	4	2	1	4	5	7	1	1
		5	2				2	3
		6	4				4	7
			5				5	
			6				6	

Table 2 Corresponding states of the aminoacids in the sequence AIMALKMR

$$
\begin{array}{c}
 \\
St1 \\
St2 \\
St3 \\
St4 \\
St5 \\
St6 \\
St7 \\
St8 \\
St9
\end{array}
\begin{array}{c}
\begin{array}{ccccccccc}
St1 & St2 & St3 & St4 & St5 & St6 & St7 & St8 & St9
\end{array} \\
\left(\begin{array}{ccccccccc}
1 & 0 & 1 & 1 & 0 & 0 & 1 & 0 & 0 \\
2 & 1 & 1 & 1 & 1 & 1 & 1 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
1 & 1 & 1 & 1 & 2 & 1 & 1 & 0 & 0 \\
2 & 1 & 1 & 2 & 1 & 1 & 2 & 0 & 0 \\
2 & 1 & 1 & 2 & 1 & 1 & 1 & 0 & 0 \\
1 & 1 & 0 & 1 & 1 & 1 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0
\end{array}\right)
\end{array}
$$

Fig. 2 Counts of the transitions between states produced from sequence AIMALKMR.

$$
\begin{array}{c} \\ St1 \\ St2 \\ St3 \\ St4 \\ St5 \\ St6 \\ St7 \\ St8 \\ St9 \end{array}
\begin{array}{c}
\begin{array}{ccccccccc} St1 & St2 & St3 & St4 & St5 & St6 & St7 & St8 & St9 \end{array} \\
\left(\begin{array}{ccccccccc}
1/4 & 0 & 1/4 & 1/4 & 0 & 0 & 1/4 & 0 & 0 \\
2/8 & 1/8 & 1/8 & 1/8 & 1/8 & 1/8 & 1/8 & 0 & 0 \\
1/9 & 1/9 & 1/9 & 1/9 & 1/9 & 1/9 & 1/9 & 1/9 & 1/9 \\
1/8 & 1/8 & 1/8 & 1/8 & 2/8 & 1/8 & 1/8 & 0 & 0 \\
2/10 & 1/10 & 1/10 & 2/10 & 1/10 & 1/10 & 2/10 & 0 & 0 \\
2/9 & 1/9 & 1/9 & 2/9 & 1/9 & 1/9 & 1/9 & 0 & 0 \\
1/5 & 1/5 & 0 & 1/5 & 1/5 & 1/5 & 0 & 0 & 0 \\
1/9 & 1/9 & 1/9 & 1/9 & 1/9 & 1/9 & 1/9 & 1/9 & 1/9 \\
1/9 & 1/9 & 1/9 & 1/9 & 1/9 & 1/9 & 1/9 & 1/9 & 1/9
\end{array} \right)
\end{array}
$$

Fig. 3 Transition matrix produced from sequence AIMALKMR.

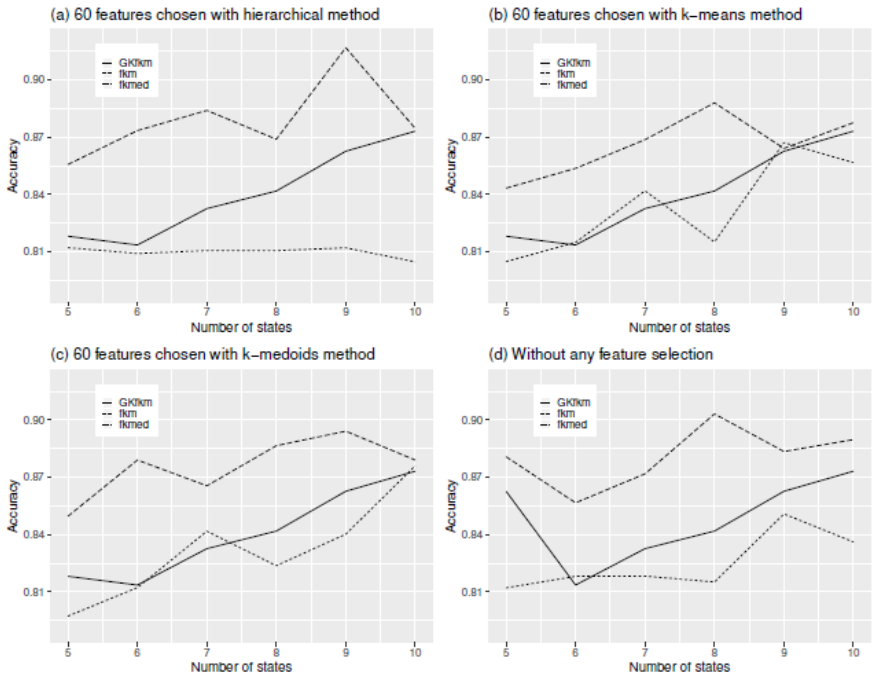

Fig. 4 Effect of the number of states on the accuracy values

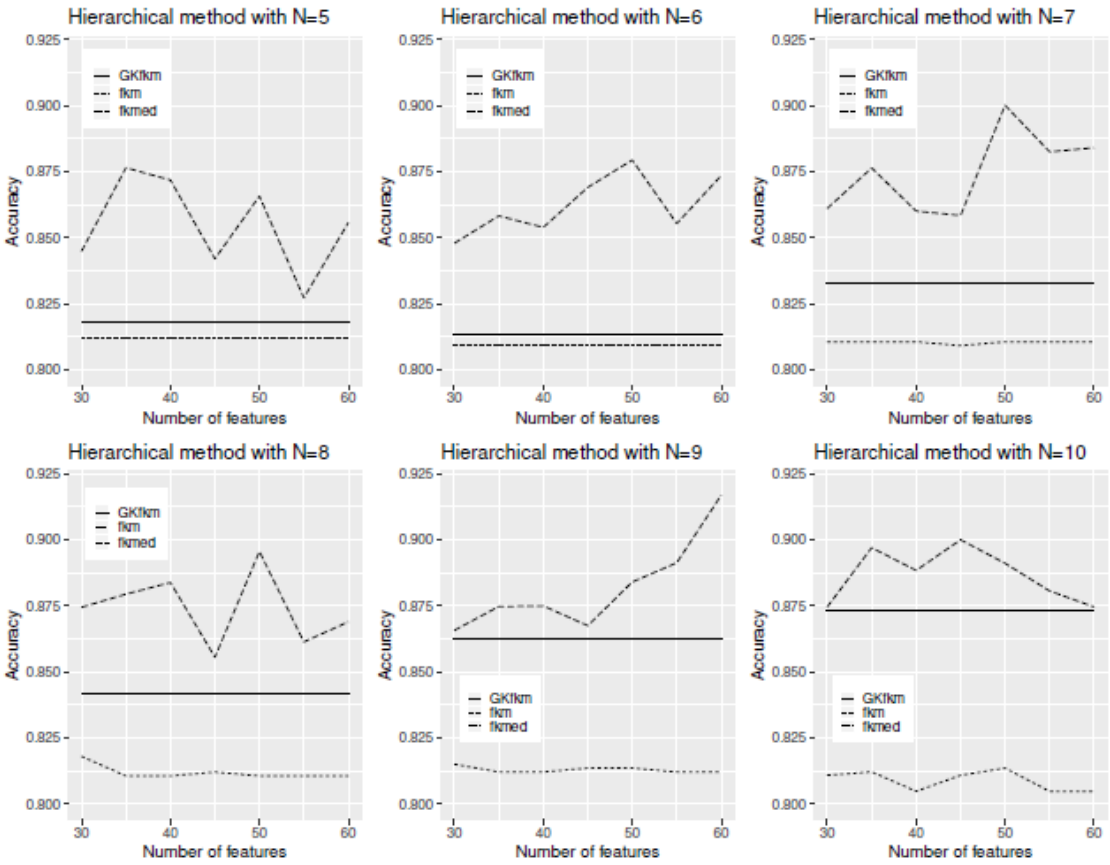

Fig. 5 Accuracy values for hierarchical feature selection

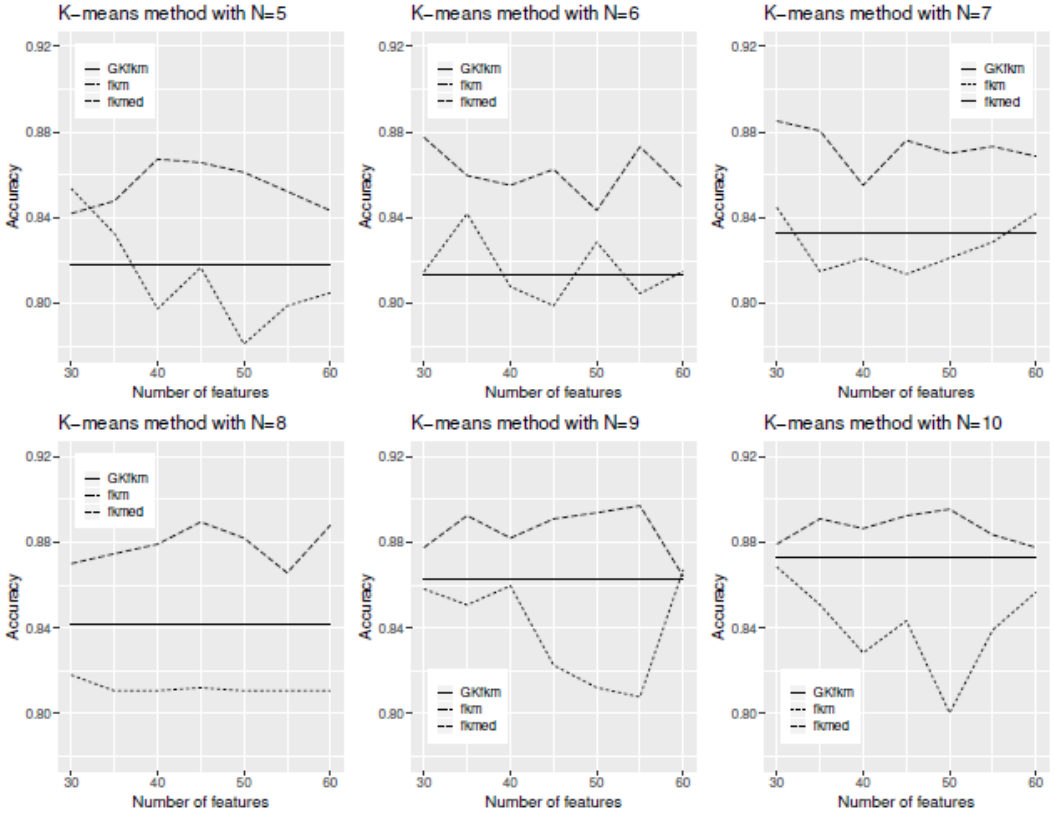

Fig. 6 Accuracy values for k-means feature selection

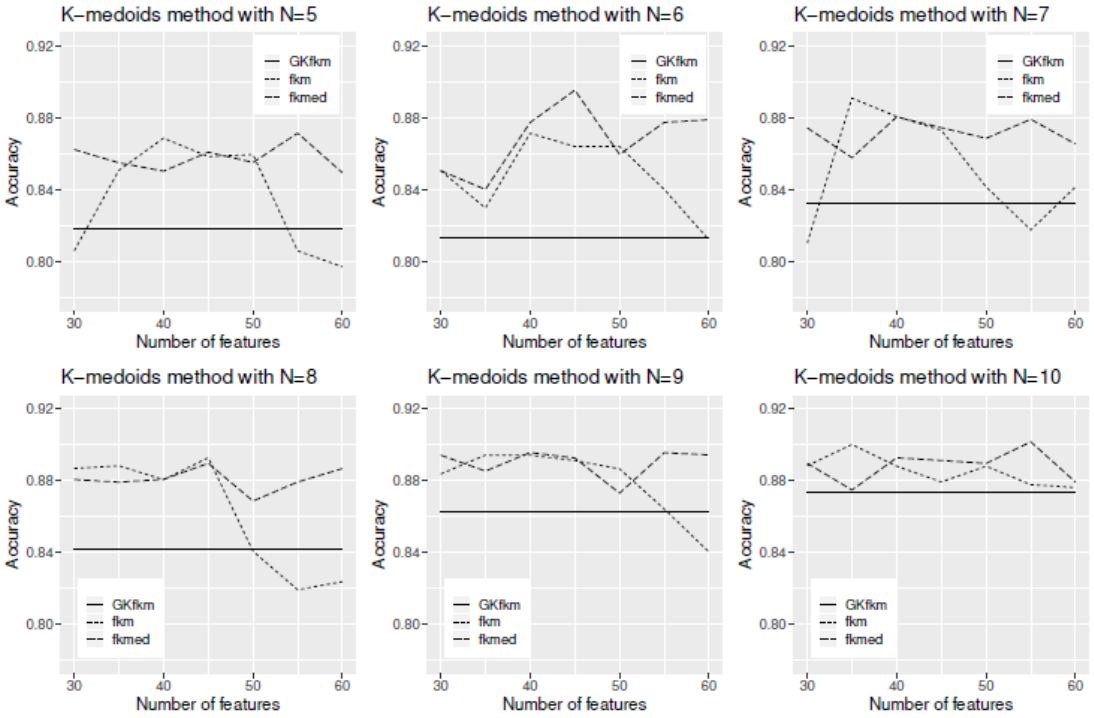

Fig. 7 Accuracy values for k-medoids feature selection

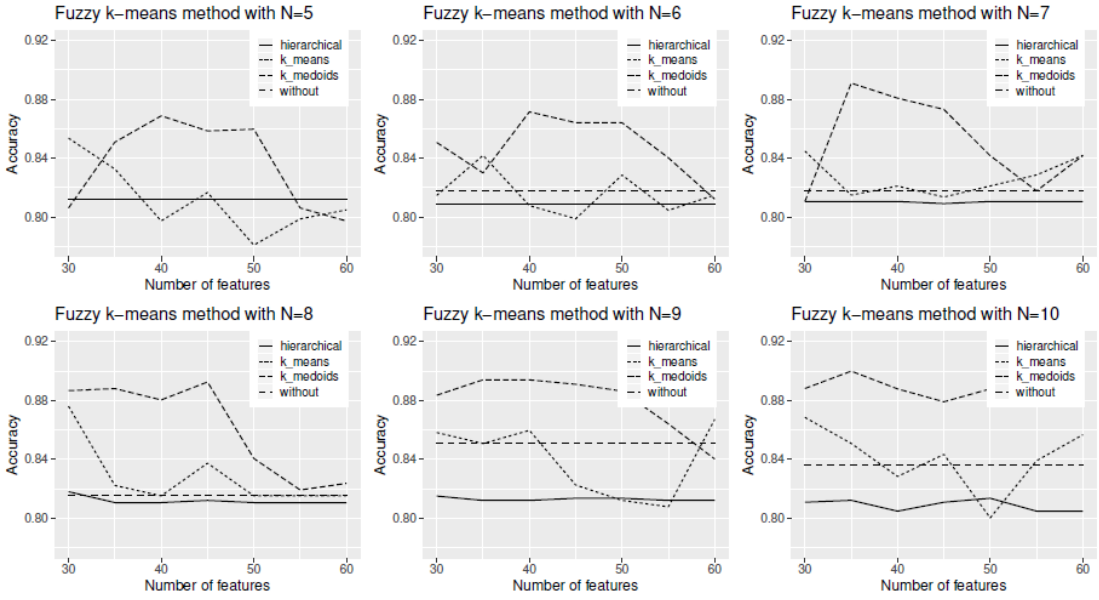

Fig. 8 Accuracy values for fuzzy k-means state selection

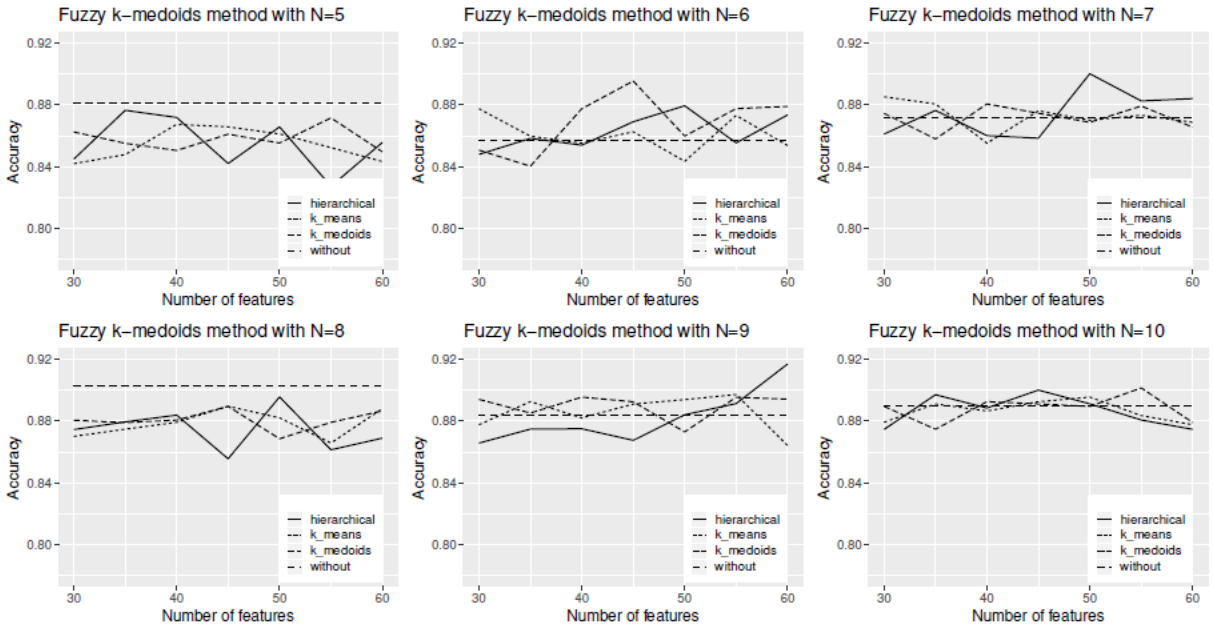

Fig. 9 Accuracy values for fuzzy k-medoids state selection

Model	Precision	Recall	F-score	MCC	Accuracy
Proposed model training values	0.908	0.945	0.924	0.837	0.917
Multi Property model training values	0.886	0.917	0.900	0.777	0.888
Proposed model test values	0.864	0.950	0.905	0.789	0.893
Multi Property model test values	0.875	0.875	0.875	0.732	0.867

Table 3 Comparison of the proposed model with multiple property grouping state selection method