 [www.onlinedoctranslator.com - naisreP ot hsilgnE morf detalsnarT](https://www.onlinedoctranslator.com/en/?utm_source=onlinedoctranslator&utm_medium=pdf&utm_campaign=attribution)

[ﺳﯿﺴﺘﻢﻫﺎﯼ ﺧﺒﺮﻩ ﺑﺎ ﺑﺮﻧﺎﻣﻪ ﻫﺎﯼ ﮐﺎﺭﺑﺮﺩﯼ 195 (2022) 116573](https://doi.org/10.1016/j.eswa.2022.116573)

[ﻓﻬﺮﺳﺖﻣﻄﺎﻟﺐ](http://www.elsevier.com/locate/eswa) ﻣﻮﺟﻮﺩ ﺩﺭScienceDirect

ﺳﯿﺴﺘﻢﻫﺎﯼ ﺧﺒﺮﻩ ﺑﺎ ﺑﺮﻧﺎﻣﻪ ﻫﺎﯼ ﮐﺎﺭﺑﺮﺩﯼ

[www.elsevier.com](http://www.elsevier.com/)[/locate/eswa:ﻣﺠﻠﻪ ﺻﻔﺤﻪﺍﺻﻠﯽ](http://www.elsevier.com/locate/eswa)

ﯾﮏﺍﻟﮕﻮﺭﯾﺘﻢ ﺧﻮﺷﻪ ﺑﻨﺪﯼ ﻗﻮﯼ ﺑﺮ ﺍﺳﺎﺱ ﺷﻨﺎﺳﺎﯾﯽ ﻧﻘﺎﻁ ﺍﺻﻠﯽ ﻭ ﺗﺨﻤﯿﻦ ﭼﮕﺎﻟﯽ ﻫﺴﺘﻪ

KNN

ﮊﻭﮊﻭ، ﮔﺎﻧﮕﻮﺍﻥ ﺳﯽ∗، ﻫﺎﺉﻮﺩﻭﻧﮓ ﺳﺎﻥ، ﮐﺎﯼ ﮐﻮ، ﻭﺍﯾﭽﻨﮓ ﻫﻮ

ﺁﺯﻣﺎﯾﺸﮕﺎﻩﮐﻠﯿﺪ ﺍﯾﺎﻟﺘﯽ ﻋﺎﯾﻖ ﺑﺮﻕ ﻭ ﺗﺠﻬﯿﺰﺍﺕ ﺑﺮﻕ، ﺁﺯﻣﺎﯾﺸﮕﺎﻩ ﮐﻠﯿﺪ *Shaanxi* ﺩﺍﻧﺸﮑﺪﻩ ﻣﻬﻨﺪﺳﯽ ﺑﺮﻕ ﺷﺒﮑﻪ ﻫﻮﺷﻤﻨﺪ، ﺩﺍﻧﺸﮕﺎﻩ ﺷﯿﺎﻥ ﺟﯿﺎﺉﻮﺗﻨﮓ، ﺷﯿﺎﻥ، ﺍﺳﺘﺎﻥ ﺷﺎﻧﺸﯽ، 710049، ﭼﯿﻦ

ﭼﮑﯿﺪﻩ

ﺛﺎﺑﺖﺷﺪﻩ ﺍﺳﺖ ﮐﻪ ﺧﻮﺷﻪ ﺑﻨﺪﯼ ﭘﯿﮏ ﭼﮕﺎﻟﯽ (DPC) ﯾﮏ ﺭﻭﺵ ﺧﻮﺷﻪ ﺑﻨﺪﯼ ﻣﺆﺛﺮ ﺍﺳﺖ ﻭ ﺩﺭ ﺑﺴﯿﺎﺭﯼ ﺍﺯ ﺯﻣﯿﻨﻪ ﻫﺎﯼ ﻋﻠﻤﯽ ﮐﺎﺭﺑﺮﺩ ﺩﺍﺭﺩ. ﻣﯽ ﺗﻮﺍﻧﺪ ﭘﯿﮏ ﭼﮕﺎﻟﯽ ﺭﺍ ﺩﺭ ﻫﺮ ﺧﻮﺷﻪ ﺗﺸﺨﯿﺺ ﺩﻫﺪ ﻭ ﺑﻘﯿﻪﻧﻘﺎﻁ ﺭﺍ ﺩﺭ ﯾﮏ ﻣﺮﺣﻠﻪ ﺍﺧﺘﺼﺎﺹ ﺩﻫﺪ. ﺑﺎ ﺍﯾﻦ ﺣﺎﻝ، ﺭﻭﺵ ﺍﺻﻠﯽ ﺍﺯ ﭼﻨﺪﯾﻦ ﺍﺷﮑﺎﻝ ﺭﻧﺞ ﻣﯽ ﺑﺮﺩ. ﺍﻭﻝ ﺍﺯ ﻫﻤﻪ، ﻧﻮﯾﺴﻨﺪﮔﺎﻥ ﺩﻭ ﻣﻌﯿﺎﺭ ﭼﮕﺎﻟﯽ ﺭﺍ ﺑﺮﺍﯼ ﺗﺨﻤﯿﻦ ﭼﮕﺎﻟﯽ ﻣﺠﻤﻮﻋﻪ ﺩﺍﺩﻩ ﻫﺎﯼ ﺑﺰﺭﮒ ﻭ ﮐﻮﭼﮏﺍﺭﺍﺉﻪ ﮐﺮﺩﻧﺪ. ﺑﺎ ﺍﯾﻦ ﺣﺎﻝ، ﺍﻧﺘﺨﺎﺏ ﺩﻟﺨﻮﺍﻩ ﻣﺘﺮﯾﮏ ﭼﮕﺎﻟﯽ ﺑﺮ ﺍﺛﺮﺑﺨﺸﯽ DPC ﺗﺄﺛﯿﺮ ﺧﻮﺍﻫﺪ ﮔﺬﺍﺷﺖ. ﻋﻼﻭﻩ ﺑﺮ ﺍﯾﻦ، ﺷﻨﺎﺳﺎﯾﯽ ﻣﺮﺍﮐﺰ ﺧﻮﺷﻪ ﺍﯼ ﻣﺴﺘﻠﺰﻡ ﺗﺨﺼﯿﺺ ﺁﺳﺘﺎﻧﻪ ﺑﻪ ﺻﻮﺭﺕ ﺩﺳﺘﯽ ﺑﺮﺍﯼﺩﻭ ﮐﻤﯿﺖ ﺍﺿﺎﻓﯽ ﺍﺳﺖ ﮐﻪ ﺍﺳﺘﺤﮑﺎﻡ DPC ﺭﺍ ﮐﺎﻫﺶ ﻣﯽ ﺩﻫﺪ. ﻋﻼﻭﻩ ﺑﺮ ﺍﯾﻦ، ﺭﻭﺵ ﮔﺎﻡ ﺑﻪ ﮔﺎﻡ DPC ﺩﺭ ﺗﺨﺼﯿﺺ ﻧﻘﺎﻁ ﺑﺎﻗﯽ ﻣﺎﻧﺪﻩ ﭘﺲ ﺍﺯ ﯾﺎﻓﺘﻦ ﻣﺮﺍﮐﺰ ﺧﻮﺷﻪ ﻣﻤﮑﻦ ﺍﺳﺖ ﺑﺎﻋﺚ ﺍﻧﺘﺸﺎﺭ ﺧﻄﺎﻫﺎﺷﻮﺩ. ﺑﺮﺍﯼ ﺭﻓﻊ ﺍﯾﻦ ﻣﻌﺎﯾﺐ، ﻣﺎ ﯾﮏ ﺭﻭﺵ ﺟﺪﯾﺪ ﺧﻮﺷﻪ ﺑﻨﺪﯼ ﻣﺒﺘﻨﯽ ﺑﺮ KNN ﺑﻪ ﻧﺎﻡ ICKDC ﺭﺍ ﭘﯿﺸﻨﻬﺎﺩ ﮐﺮﺩﯾﻢ. ﻫﺪﻑ ﭼﺎﺭﭼﻮﺏ ﺟﺪﯾﺪ ﻧﺸﺎﻥ ﺩﺍﺩﻥ ﯾﮏ ﺧﻮﺷﻪ ﺑﺎ ﻣﺠﻤﻮﻋﻪ ﺍﯼ ﺍﺯ ﻧﻘﺎﻁ ﺍﺻﻠﯽ ﺑﻪﺟﺎﯼ ﯾﮏ ﺷﯽ ﻭﺍﺣﺪ ﺑﺮﺍﯼ ﻣﺮﮐﺰ ﻭ ﻃﺒﻘﻪ ﺑﻨﺪﯼ ﺑﻘﯿﻪ ﺍﺷﯿﺎء ﺑﻪ ﻃﻮﺭ ﻣﻮﺛﺮ ﺑﺮ ﺍﺳﺎﺱ ﺍﻃﻼﻋﺎﺕ ﺗﻮﺯﯾﻊ ﺍﺳﺖ. ﺑﻪ ﻃﻮﺭ ﮐﻠﯽ، ﻓﺮﺁﯾﻨﺪ ﺗﺨﻤﯿﻦ ﭼﮕﺎﻟﯽ ﻣﺤﻠﯽ ﻣﺒﺘﻨﯽ ﺑﺮ KNN، ﺷﻨﺎﺳﺎﯾﯽ ﻭ ﺍﺩﻏﺎﻡ ﻧﻘﺎﻁ ﺍﺻﻠﯽ،ﻭ ﻃﺒﻘﻪ ﺑﻨﺪﯼ ﺍﺷﯿﺎء ﻏﯿﺮ ﻫﺴﺘﻪ ﺍﯼ ﺑﻪ ICKDC ﺍﺟﺎﺯﻩ ﻣﯽ ﺩﻫﺪ ﺗﺎ ﺳﺎﺧﺘﺎﺭ ﺧﻮﺷﻪ ﻫﺎ ﺭﺍ ﺁﺷﮑﺎﺭ ﮐﻨﺪ. ﺁﺯﻣﺎﯾﺶ ﻫﺎ ﺑﺮ ﺭﻭﯼ ﻣﺠﻤﻮﻋﻪ ﻫﺎﯼ ﺩﺍﺩﻩ ﻣﺼﻨﻮﻋﯽ ﻭ ﺩﻧﯿﺎﯼ ﻭﺍﻗﻌﯽ ﺑﺮﺗﺮﯼ ICKDC ﺭﺍ ﻧﺴﺒﺖ ﺑﻪﭼﻨﺪﯾﻦ ﺍﻟﮕﻮﺭﯾﺘﻢ ﺧﻮﺷﻪ ﺑﻨﺪﯼ ﭘﺮﮐﺎﺭﺑﺮﺩ ﻭ ﺍﻧﻮﺍﻉ ﺟﺪﯾﺪ DPC ﻧﺸﺎﻥ ﻣﯽ ﺩﻫﺪ. ﻫﺪﻑ ﭼﺎﺭﭼﻮﺏ ﺟﺪﯾﺪ ﻧﺸﺎﻥ ﺩﺍﺩﻥ ﯾﮏ ﺧﻮﺷﻪ ﺑﺎ ﻣﺠﻤﻮﻋﻪ ﺍﯼ ﺍﺯ ﻧﻘﺎﻁ ﺍﺻﻠﯽ ﺑﻪ ﺟﺎﯼ ﯾﮏ ﺷﯽ ﻭﺍﺣﺪ ﺑﺮﺍﯼ ﻣﺮﮐﺰ ﻭ ﻃﺒﻘﻪ ﺑﻨﺪﯼﺑﻘﯿﻪ ﺍﺷﯿﺎء ﺑﻪ ﻃﻮﺭ ﻣﻮﺛﺮ ﺑﺮ ﺍﺳﺎﺱ ﺍﻃﻼﻋﺎﺕ ﺗﻮﺯﯾﻊ ﺍﺳﺖ. ﺑﻪ ﻃﻮﺭ ﮐﻠﯽ، ﻓﺮﺁﯾﻨﺪ ﺗﺨﻤﯿﻦ ﭼﮕﺎﻟﯽ ﻣﺤﻠﯽ ﻣﺒﺘﻨﯽ ﺑﺮ KNN، ﺷﻨﺎﺳﺎﯾﯽ ﻭ ﺍﺩﻏﺎﻡ ﻧﻘﺎﻁ ﺍﺻﻠﯽ، ﻭ ﻃﺒﻘﻪ ﺑﻨﺪﯼ ﺍﺷﯿﺎء ﻏﯿﺮ ﻫﺴﺘﻪ ﺍﯼ ﺑﻪ

ICKDC ﺍﺟﺎﺯﻩ ﻣﯽ ﺩﻫﺪ ﺗﺎ ﺳﺎﺧﺘﺎﺭ ﺧﻮﺷﻪ ﻫﺎ ﺭﺍ ﺁﺷﮑﺎﺭ ﮐﻨﺪ. ﺁﺯﻣﺎﯾﺶ ﻫﺎ ﺑﺮ ﺭﻭﯼ ﻣﺠﻤﻮﻋﻪ ﻫﺎﯼ ﺩﺍﺩﻩ ﻣﺼﻨﻮﻋﯽ ﻭ ﺩﻧﯿﺎﯼ ﻭﺍﻗﻌﯽ ﺑﺮﺗﺮﯼ ICKDC ﺭﺍ ﻧﺴﺒﺖ ﺑﻪ ﭼﻨﺪﯾﻦ ﺍﻟﮕﻮﺭﯾﺘﻢ ﺧﻮﺷﻪ ﺑﻨﺪﯼ ﭘﺮﮐﺎﺭﺑﺮﺩ ﻭ ﺍﻧﻮﺍﻉﺟﺪﯾﺪ DPC ﻧﺸﺎﻥ ﻣﯽ ﺩﻫﺪ. ﻫﺪﻑ ﭼﺎﺭﭼﻮﺏ ﺟﺪﯾﺪ ﻧﺸﺎﻥ ﺩﺍﺩﻥ ﯾﮏ ﺧﻮﺷﻪ ﺑﺎ ﻣﺠﻤﻮﻋﻪ ﺍﯼ ﺍﺯ ﻧﻘﺎﻁ ﺍﺻﻠﯽ ﺑﻪ ﺟﺎﯼ ﯾﮏ ﺷﯽ ﻭﺍﺣﺪ ﺑﺮﺍﯼ ﻣﺮﮐﺰ ﻭ ﻃﺒﻘﻪ ﺑﻨﺪﯼ ﺑﻘﯿﻪ ﺍﺷﯿﺎء ﺑﻪ ﻃﻮﺭ ﻣﻮﺛﺮ ﺑﺮ ﺍﺳﺎﺱ ﺍﻃﻼﻋﺎﺕﺗﻮﺯﯾﻊ ﺍﺳﺖ. ﺑﻪ ﻃﻮﺭ ﮐﻠﯽ، ﻓﺮﺁﯾﻨﺪ ﺗﺨﻤﯿﻦ ﭼﮕﺎﻟﯽ ﻣﺤﻠﯽ ﻣﺒﺘﻨﯽ ﺑﺮ KNN، ﺷﻨﺎﺳﺎﯾﯽ ﻭ ﺍﺩﻏﺎﻡ ﻧﻘﺎﻁ ﺍﺻﻠﯽ، ﻭ ﻃﺒﻘﻪ ﺑﻨﺪﯼ ﺍﺷﯿﺎء ﻏﯿﺮ ﻫﺴﺘﻪ ﺍﯼ ﺑﻪ ICKDC ﺍﺟﺎﺯﻩ ﻣﯽ ﺩﻫﺪ ﺗﺎ ﺳﺎﺧﺘﺎﺭ ﺧﻮﺷﻪ

ﻫﺎﺭﺍ ﺁﺷﮑﺎﺭ ﮐﻨﺪ. ﺁﺯﻣﺎﯾﺶ ﻫﺎ ﺑﺮ ﺭﻭﯼ ﻣﺠﻤﻮﻋﻪ ﻫﺎﯼ ﺩﺍﺩﻩ ﻣﺼﻨﻮﻋﯽ ﻭ ﺩﻧﯿﺎﯼ ﻭﺍﻗﻌﯽ ﺑﺮﺗﺮﯼ ICKDC ﺭﺍ ﻧﺴﺒﺖ ﺑﻪ ﭼﻨﺪﯾﻦ ﺍﻟﮕﻮﺭﯾﺘﻢ ﺧﻮﺷﻪ ﺑﻨﺪﯼ ﭘﺮﮐﺎﺭﺑﺮﺩ ﻭ ﺍﻧﻮﺍﻉ ﺟﺪﯾﺪ DPC ﻧﺸﺎﻥ ﻣﯽ ﺩﻫﺪ.

ARTICLEINFO

ﮐﻠﯿﺪﻭﺍﮊﻩ ﻫﺎ:

ﺧﻮﺷﻪﺑﻨﺪﯼ

ﭼﮕﺎﻟﯽﺑﻪ ﺍﻭﺝ ﻣﯽ ﺭﺳﺪ

ﻧﻘﺎﻁﺍﺻﻠﯽ ﺗﺨﻤﯿﻦﭼﮕﺎﻟﯽ ﻫﺴﺘﻪKNN

ﻧﻤﻮﻧﻪ ﻫﺎﺩﺭ ﺧﻮﺷﻪ ﻫﺎﯼ K ﻗﺮﺍﺭ ﻣﯽ ﮔﯿﺮﻧﺪ، ﺩﺭ ﺣﺎﻟﯽ ﮐﻪ ﺧﻮﺷﻪ ﺑﻨﺪﯼ ﺳﻠﺴﻠﻪ ﻣﺮﺍﺗﺒﯽ ﺑﻪ ﺩﻧﺒﺎﻝ ﺗﺸﮑﯿﻞﯾﮏ ﺳﺎﺧﺘﺎﺭ ﺩﺭﺧﺖ ﻣﺎﻧﻨﺪ ﺍﺯ ﻧﻤﻮﻧﻪ ﻫﺎﯼ ﺩﺳﺘﻪ ﺑﻨﺪﯼ ﻭ ﭘﺎﺭﺗﯿﺸﻦ ﺩﺭ ﻫﺮ ﮔﺮﻩ ﺩﺭﺧﺖ

ﺍﺳﺖ. ﯾﮑﯽﺍﺯ ﺷﺎﺧﻪ ﻫﺎﯼ ﻣﻬﻢ ﺧﻮﺷﻪ ﺑﻨﺪﯼ ﭘﺎﺭﺗﯿﺸﻦ ﺑﻨﺪﯼ، ﺭﻭﺵ ﻣﺒﺘﻨﯽ ﺑﺮ ﭼﮕﺎﻟﯽ ﺍﺳﺖ ﮐﻪ

ﻣﯽ ﺗﻮﺍﻧﺪﺍﺷﮑﺎﻝ ﻧﺎﻣﻨﻈﻢ ﻭ ﻏﯿﺮ ﮐﺮﻭﯼ ﻧﻤﻮﻧﻪ ﻫﺎﯼ ﺩﺍﺩﻩ ﺭﺍ ﻃﺒﻘﻪ ﺑﻨﺪﯼ ﮐﻨﺪ. ﺑﻪ ﻃﻮﺭ ﮐﻠﯽ، [ﺗﻮﺯﯾﻊﺍﺷﯿﺎء ﺩﺍﺩﻩ ﺭﺍ ﺑﺎ ﺍﺳﺘﻔﺎﺩﻩ ﺍﺯ](#_bookmark32) ﺍﺳﺘﺮﺍﺗﮋﯼ ﻫﺎﯼ ﺗﺨﻤﯿﻦ ﭼﮕﺎﻟﯽ ﻣﺨﺘﻠﻒ ﺩﺭ ﻣﺮﺣﻠﻪ ﺍﻭﻝ ﺗﻮﺻﯿﻒﻣﯽ ﮐﻨﺪ، ﺳﭙﺲ ﻣﺮﺍﮐﺰ ﺧﻮﺷﻪ ﺭﺍ ﺑﻪ ﻋﻨﻮﺍﻥ ﻣﻨﺎﻃﻘﯽ ﺑﺎ ﺗﺮﺍﮐﻢ ﻣﺤﻠﯽ [ﺑﺎﻻﺗﺮ ﺷﻨﺎﺳﺎﯾﯽ](#_bookmark32)

ﻣﯽﮐﻨﺪ ﻭ ﺧﻮﺷﻪ ﻧﻬﺎﯾﯽ ﺭﺍ ﺗﺤﺖ ﻃﺮﺡ ﻫﺎﯼ ﭘﺎﺭﺗﯿﺸﻦ ﺑﻨﺪﯼ ﻣﺨﺘﻠﻒ ﺗﺸﮑﯿﻞ ﻣﯽ ﺩﻫﺪ.

DBSCAN )ﺧﻮﺷﻪ ﺑﻨﺪﯼ ﻓﻀﺎﯾﯽ ﻣﺒﺘﻨﯽ ﺑﺮ ﭼﮕﺎﻟﯽ ﺑﺮﻧﺎﻣﻪ ﻫﺎﯼ ﮐﺎﺭﺑﺮﺩﯼ ﺑﺎ ﻧﻮﯾﺰ( )ﺍﺳﺘﺮ، [ﮐﺮﯾﮕﻞ،ﺳﺎﻧﺪﺭ،](#_bookmark60) ﺧﻮ، ﻭ ﻫﻤﮑﺎﺭﺍﻥ.،(1996 ﯾﮑﯽ ﺍﺯ ﻣﻌﺮﻑ ﺗﺮﯾﻦ ﺭﻭﺵ ﻫﺎﯼ ﺧﻮﺷﻪ ﺑﻨﺪﯼ ﻣﺒﺘﻨﯽ ﺑﺮﭼﮕﺎﻟﯽ ﺍﺳﺖ. ﺩﺭ DBSCAN، ﺍﺷﯿﺎء ﺩﺍﺧﻞ ﻣﻨﻄﻘﻪ ﺑﺎ ﺷﻌﺎﻉ ﺍﮔﺮ ﺗﻌﺪﺍﺩ ﺁﻧﻬﺎ [ﺍﺯ](#_bookmark60) ﺣﺪﺍﻗﻞ [ﺍﻣﺘﯿﺎﺯ](#_bookmark60) [ﺍﺯ](#_bookmark40)[ﭘﯿﺶ](#_bookmark31) [ﺗﻌﺮﯾﻒ ﺷﺪﻩ](#_bookmark27) ﺑﯿﺸﺘﺮ ﺑﺎﺷﺪ، ﺧﻮﺷﻪ ﺑﻨﺪﯼ ﻣﯽ [ﺷﻮﻧﺪ.](#_bookmark35)(*minPts*[. ﺳﭙﺲ ﻣﻨﺎﻃﻘﯽ ﮐﻪ](#_bookmark35) [ﺍﺷﯿﺎءﻣﺸﺎﺑﻬﯽ ﺩﺍﺭﻧﺪ](#_bookmark67) ﺑﻪ ﻫﻤﺎﻥ ﺧﻮﺷﻪ ﺍﺧﺘﺼﺎﺹ ﺩﺍﺩﻩ [ﻣﯽ ﺷﻮﻧﺪ.](#_bookmark50) ﺑﺎ [ﺍﯾﻦ](#_bookmark40) ﺣﺎﻝ، [ﻋﻤﻠﮑﺮﺩ](#_bookmark40) DB[-](#_bookmark61)SCAN [ﺑﻪ ﺷﺪﺕ ﺑﺮ ﺩﻭ ﭘﺎﺭﺍﻣﺘﺮ ﻣﺸﺨﺺ](#_bookmark61) ﺷﺪﻩ ﻣﺘﮑﯽ ﺍﺳﺖ، ﺩﺭ [ﺣﺎﻟﯽ](#_bookmark67) ﮐﻪ [ﻓﺮﺁﯾﻨﺪ ﺗﻨﻈﯿﻢ](#_bookmark67)

ﻣﯽﺗﻮﺍﻧﺪ ﻏﯿﺮ ﺿﺮﻭﺭﯼ ﺑﺎﺷﺪ. ﺑﺮﺍﯼ ﻏﻠﺒﻪ ﺑﺮ ﺩﺷﻮﺍﺭﯼ ﺩﺭ [ﺍﻧﺘﺨﺎﺏ](#_bookmark63)

.1ﻣﻘﺪﻣﻪ

ﻃﺒﻘﻪﺑﻨﺪﯼ ﺑﺪﻭﻥ ﻧﻈﺎﺭﺕ، ﮐﻪ ﺑﻪ ﺁﻥ ﺧﻮﺷﻪ ﺑﻨﺪﯼ ﯾﺎ ﺗﺠﺰﯾﻪ ﻭ ﺗﺤﻠﯿﻞ ﺩﺍﺩﻩ ﻫﺎﯼ ﺍﮐﺘﺸﺎﻓﯽ

ﻧﯿﺰﮔﻔﺘﻪ ﻣﯽ ﺷﻮﺩ، ﻧﻘﺸﯽ ﺿﺮﻭﺭﯼ ﺩﺭ ﺍﺳﺘﺨﺮﺍﺝ ﻭ ﺩﺭﮎ ﺩﺍﻧﺶ ﺑﺎﻟﻘﻮﻩ ﺩﺭﻭﻥ ﺩﺍﺩﻩ ﻫﺎ ﺍﯾﻔﺎ ﻣﯽ ﮐﻨﺪ.ﺗﺎﺭﯾﺨﭽﻪ ﺍﯼ ﻃﻮﻻﻧﯽ ﺩﺍﺭﺩ ﮐﻪ ﺑﻪ ﺯﻣﺎﻥ ﺍﺭﺳﻄﻮ ﺑﺮﻣﯽ ﮔﺮﺩﺩﺗﺎﺭﯾﺨﭽﻪ ﺣﯿﻮﺍﻧﺎﺕﮐﻪ ﺩﺭ ﺁﻥ ﺍﻭ ﺣﯿﻮﺍﻧﺎﺕﺭﺍ ﺑﺮ ﺍﺳﺎﺱ ﭘﯿﭽﯿﺪﮔﯽ ﺳﺎﺧﺘﺎﺭ ﺁﻧﻬﺎ ﻃﺒﻘﻪ ﺑﻨﺪﯼ ﮐﺮﺩ )ﺑﻮﺗﻨﮑﻮ، ﭼﺎﺉﻮﺍﻟﯿﺖ ﻭﺍﻧﮕﺴﻪ ﻭ ﭘﺎﺭﺩﺍﻟﻮﺱ،.(2009 ﺩﺭ ﺗﺤﻠﯿﻞ ﺧﻮﺷﻪ ﺍﯼ، ﻣﺠﻤﻮﻋﻪ ﺍﯼ ﺍﺯ ﻧﻤﻮﻧﻪ ﻫﺎ ﺑﺮ ﺍﺳﺎﺱ

ﺷﺒﺎﻫﺖ ﻫﺎﯾﺸﺎﻥﺑﻪ ﭼﻨﺪ ﺧﻮﺷﻪ ﺗﻘﺴﯿﻢ ﻣﯽ ﺷﻮﻧﺪ ﮐﻪ ﻣﻌﻤﻮﻻ ًﺑﺎ ﻓﻮﺍﺻﻞ ﺍﻗﻠﯿﺪﺳﯽ ﺍﻧﺪﺍﺯﻩ ﮔﯿﺮﯼﻣﯽ ﺷﻮﻧﺪ، ﺑﻪ ﻃﻮﺭﯼ ﮐﻪ ﻧﻤﻮﻧﻪ ﻫﺎﯼ ﺩﺭﻭﻥ ﯾﮏ ﺧﻮﺷﻪ ﻣﺸﺎﺑﻪ ﯾﮑﺪﯾﮕﺮ ﻫﺴﺘﻨﺪ ﻭ ﺑﺎ

ﻧﻤﻮﻧﻪ ﻫﺎﯼﻣﻮﺟﻮﺩ ﺩﺭ ﺧﻮﺷﻪ ﻫﺎﯼ ﺩﯾﮕﺮ ﻣﺘﻔﺎﻭﺕ ﻫﺴﺘﻨﺪ.ﺟﯿﻦ، ﻣﻮﺭﺗﯽ ﻭ ﻓﻠﯿﻦ،.(1999

[ﺗﺠﺰﯾﻪﻭ](#_bookmark39) ﺗﺤﻠﯿﻞ ﺧﻮﺷﻪ ﺍﯼ ﺷﺎﻣﻞ ﺗﺤﻘﯿﻘﺎﺗﯽ ﺍﺳﺖ ﮐﻪ ﺷﺎﻣﻞ ﻃﯿﻒ ﮔﺴﺘﺮﺩﻩ ﺍﯼ ﺍﺯ ﺟﻮﺍﻣﻊ

ﻣﺎﻧﻨﺪﭘﺮﺩﺍﺯﺵ ﺗﺼﻮﯾﺮ )ﺑﻮﻧﺪﯼ ﻭ ﻫﻤﮑﺎﺭﺍﻥ،;2017 [ﺩﺍﻧﺎﭼﺎﻧﺪﺭﺍ ﻭ ﭼﺎﻧﻮ،](#_bookmark39)([2017، ﺗﺸﺨﯿﺺ](#_bookmark39)

ﺟﺎﻣﻌﻪ)ﺁﺑﻪ،;2017ﻓﺮﯾﺮﺍ ﻭ ﮊﺍﺉﻮ،;2016ﻣﺎ ﻭ ﻫﻤﮑﺎﺭﺍﻥ،2016، ﺑﯿﻮ ﺍﻧﻔﻮﺭﻣﺎﺗﯿﮏ )ﺯﻭ، ﻟﯿﻦ، ﺟﯿﺎﻧﮓ،ﻟﯿﻮ ﻭ ﺯﻧﮓ،(2018، ﺍﻣﻨﯿﺖ ﺳﺎﯾﺒﺮﯼ )ﺑﻮﺳﻪ، ﺟﻨﺞ، ﻫﺎﻟﺮ ﻭ ﺳﺒﺴﺘﯿﻦ، ;2014ﻻﺧﻨﻮ ﻭ ﻫﻤﮑﺎﺭﺍﻥ،(2017ﻭ ﻏﯿﺮﻩ. ﺩﺭ ﺍﯾﻦ ﻣﯿﺎﻥ ﺭﻭﺵ ﻫﺎﯼ ﻣﺨﺘﻠﻒ ﺧﻮﺷﻪ ﺑﻨﺪﯼ ﭘﯿﺸﻨﻬﺎﺩ ﺷﺪﻩ ﺍﺳﺖ.ﯾﮏ ﺭﻭﺵ ﺗﻘﺮﯾﺒﯽ ﺍﻣﺎ ﺑﻪ ﻃﻮﺭ ﮔﺴﺘﺮﺩﻩ ﭘﺬﯾﺮﻓﺘﻪ ﺷﺪﻩ، ﻃﺒﻘﻪ ﺑﻨﺪﯼ ﺁﻧﻬﺎ ﺑﻪ ﺭﻭﺵ ﻫﺎﯼ ﭘﺎﺭﺗﯿﺸﻦﺑﻨﺪﯼ ﻭ ﺭﻭﺵ ﻫﺎﯼ ﺳﻠﺴﻠﻪ ﻣﺮﺍﺗﺒﯽ ﺍﺳﺖWunsch. & Xu،.(2005 ﺧﻮﺷﻪ ﺑﻨﺪﯼ

ﭘﺎﺭﺗﯿﺸﻦﺑﻨﺪﯼ ﻣﺴﺘﻘﯿﻤﺎ ًﻃﺒﻘﻪ ﺑﻨﺪﯼ ﻣﯽ ﮐﻨﺪ

∗ ﻧﻮﯾﺴﻨﺪﻩﻣﺘﻨﺎﻇﺮ.

)K. Qu(، qukaihenshuai@163.com،(ﺳﺎﻥ .ﺍﭺ)haodongshd@163.com )G. Si(,sigangquan@mail.xjtu.edu.cn )Z. Zhou(،zz\_790539463@stu.xjtu.edu.cn :ﺁﺩﺭﺱﺍﯾﻤﯿﻞ

)W. Hou(.weicheng123@stu.xjtu.edu.cn

https://doi.org/10.1016/j.eswa.2022.116573

ﺩﺭﯾﺎﻓﺖﺷﺪﻩ ﺩﺭ 10 ﺁﮔﻮﺳﺖ 2020؛ ﺩﺭﯾﺎﻓﺖ ﺩﺭ ﻓﺮﻡ ﺍﺻﻼﺡ ﺷﺪﻩ 28 ﺩﺳﺎﻣﺒﺮ 2021؛ ﭘﺬﯾﺮﺵ ﺩﺭ 17 ﮊﺍﻧﻮﯾﻪ 2022 ﺩﺭ ﺩﺳﺘﺮﺱ ﺁﻧﻼﯾﻦ 3

ﻓﻮﺭﯾﻪ2022

.ﮐﻠﯿﻪﺣﻘﻮﻕ ﻣﺤﻔﻮﻅ ﺍﺳﺖ Ltd. Elsevier 2022 0957-4174/©

ﺩﺭﺍﯾﻦ ﻣﻘﺎﻟﻪ، ﯾﮏ ﺭﻭﯾﮑﺮﺩ ﺟﺪﯾﺪ ICKDC ﺑﺮﺍﯼ ﺣﺬﻑ ﻋﯿﻮﺏ DPC ﭘﯿﺸﻨﻬﺎﺩ ﺷﺪﻩ ﺍﺳﺖ. ﺑﻪﻃﻮﺭ ﮐﻠﯽ، ﯾﮏ ﻣﺘﺮﯾﮏ ﺗﺨﻤﯿﻦ ﭼﮕﺎﻟﯽ [ﯾﮑﻨﻮﺍﺧﺖ ﺟﺪﯾﺪ ﺑﺮ ﺍﺳﺎﺱ ﺍﯾﺪﻩ KNN ﺑﺮﺍﯼ](#_bookmark28) ﺗﺮﮐﯿﺐ ﺍﻃﻼﻋﺎﺕﺗﻮﺯﯾﻊ ﺍﺯ ﻧﻤﻮﻧﻪ ﻫﺎﯼ ﻣﺠﺎﻭﺭ ﻣﺤﻠﯽ ﺍﺳﺘﻔﺎﺩﻩ ﻣﯽ ﺷﻮﺩ. ﺑﻨﺎﺑﺮﺍﯾﻦ ﺍﺯ ﻣﺸﮑﻼﺕ DPC ﺩﺭ

ﺗﻌﯿﯿﻦﺷﮑﻞ ﻣﻌﯿﺎﺭﻫﺎﯼ ﺗﺨﻤﯿﻦ ﭼﮕﺎﻟﯽ ﺍﺟﺘﻨﺎﺏ ﻣﯽ ﺷﻮﺩ. ﺳﭙﺲ، ﯾﮏ »ﻓﺮﺍﯾﻨﺪ ﻋﻤﻮﻣﯽ ﺍﻧﺘﺨﺎﺑﺎﺕ«ﺑﺮﺍﯼ ﺗﻌﺮﯾﻒ ﻧﻘﺎﻁ ﺍﺻﻠﯽ ﻃﺮﺍﺣﯽ ﻣﯽ ﺷﻮﺩ ﻭ ﺑﻪ ﺩﻧﺒﺎﻝ ﺁﻥ ﯾﮏ ﻣﺮﺣﻠﻪ ﺍﺩﻏﺎﻡ ﺑﺮﺍﯼ ﺷﻨﺎﺳﺎﯾﯽﺳﺎﺧﺘﺎﺭ ﺍﺻﻠﯽ ﻫﺮ ﺧﻮﺷﻪ ﺍﻧﺠﺎﻡ ﻣﯽ ﺷﻮﺩ. ﺗﻌﺪﺍﺩ ﺍﯾﺪﻩ ﺁﻝ ﺧﻮﺷﻪ ﻫﺎ ﻧﯿﺰ ﻣﯽ ﺗﻮﺍﻧﺪ ﺑﻪ [ﻃﻮﺭﺧﻮﺩﮐﺎﺭ](#_bookmark30) ﺑﺮﺍﯼ ﺍﯾﻦ ﻣﻨﻈﻮﺭ [ﺍﺳﺘﺨﺮﺍﺝ](#_bookmark55) ﺷﻮﺩ. [ﺑﻘﯿﻪ ﻧﻘﺎﻁ، ﺑﻪ ﻭﯾﮋﻩ ﻧﻘﺎﻁ ﻧﺰﺩﯾﮏ ﺑﻪ](#_bookmark55) ﻟﺒﻪ ﺧﻮﺷﻪ ﻫﺎ،ﻣﯽ ﺗﻮﺍﻧﻨﺪ ﺑﻪ ﺭﺍﺣﺘﯽ [ﺑﺎ](#_bookmark38) ﺍﻃﻼﻋﺎﺕ [ﺗﻮﺯﯾﻊ ﻃﺒﻘﻪ ﺑﻨﺪﯼ](#_bookmark38) ﺷﻮﻧﺪ. ﺍﺳﺘﺮﺍﺗﮋﯼ ﺍﻧﺘﺸﺎﺭ [ﺑﺮﭼﺴﺐ](#_bookmark30)

[ﭘﯿﺸﻨﻬﺎﺩﯼﺧﻄﺮ](#_bookmark58) [ﻭﺍﮐﻨﺶ](#_bookmark46) ﺯﻧﺠﯿﺮﻩ [ﺍﯼ ﺩﺭ DPC ﺭﺍ ﮐﺎﻫﺶ](#_bookmark46) ﻣﯽ ﺩﻫﺪ ﻭ ﮐﯿﻔﯿﺖ ﻧﺘﺎﯾﺞ ﺧﻮﺷﻪ

ﺑﻨﺪﯼﺭﺍ ﺑﻬﺒﻮﺩ ﻣﯽ ﺑﺨﺸﺪ.ﻋﮑﺲ. .1 ﻓﺮﺁﯾﻨﺪ ﺷﻨﺎﺳﺎﯾﯽ ﻭ ﺍﺩﻏﺎﻡ ﻣﻨﺎﻃﻖ ﺍﺻﻠﯽ ﺑﻪ ICKDC

ﺍﺟﺎﺯﻩﻣﯽ ﺩﻫﺪ ﺗﺎ ﺗﻮﺯﯾﻊ ﭼﮕﺎﻟﯽ ﺧﻮﺷﻪ ﻫﺎ ﻭ ﺩﺭ ﻧﺘﯿﺠﻪ ﻋﻤﻠﮑﺮﺩ ﺑﻬﺘﺮ ﺭﺍ ﺗﺸﺨﯿﺺ ﺩﻫﺪ.

ﺑﻘﯿﻪﺍﯾﻦ ﻣﻘﺎﻟﻪ ﺑﻪ ﺷﺮﺡ ﺯﯾﺮ [ﺳﺎﺯﻣﺎﻧﺪﻫﯽ](#_bookmark62) ﺷﺪﻩ [ﺍﺳﺖ. ﺩﺭ ﺑﺨﺶ2، ﻣﺎ ﺑﻪ ﻃﻮﺭ ﺧﻼﺻﻪ](#_bookmark62)

ﺭﻭﺵDPC ﺍﺻﻠﯽ ﻭ ﺍﺳﺘﺮﺍﺗﮋﯼ ﻫﺎﯼ ﺗﺨﻤﯿﻦ ﭼﮕﺎﻟﯽ ﻣﺒﺘﻨﯽ ﺑﺮ KNN ﺭﺍ ﻣﺮﻭﺭ ﻣﯽ ﮐﻨﯿﻢ. ﺩﺭ ﺑﺨﺶ3، ﭘﯿﺎﺩﻩ ﺳﺎﺯﯼ ICKDC ﺭﺍ ﺷﺮﺡ ﻣﯽ ﺩﻫﯿﻢ ﻭ ﺗﺠﺰﯾﻪ ﻭ ﺗﺤﻠﯿﻞ ﻣﻔﺼﻠﯽ ﺍﺭﺍﺉﻪ ﻣﯽ

ﺩﻫﯿﻢ.ﺩﺭ ﺑﺨﺶ4، ﻣﺎ ﺭﻭﺵ ﺟﺪﯾﺪ ﺭﺍ ﺑﺎ ﺍﺳﺘﺮﺍﺗﮋﯼ [ﻫﺎﯼ ﺧﻮﺷﻪ ﺑﻨﺪﯼ ﭘﯿﺸﺮﻓﺘﻪ ﺑﺎ](#_bookmark53) ﺍﺳﺘﻔﺎﺩﻩ ﺍﺯ ﻣﺠﻤﻮﻋﻪﺩﺍﺩﻩ ﻫﺎﯼ ﻣﺼﻨﻮﻋﯽ ﻭ ﺩﻧﯿﺎﯼ ﻭﺍﻗﻌﯽ ﻣﻘﺎﯾﺴﻪ ﻣﯽ ﮐﻨﯿﻢ. ﺩﺭ ﺑﺨﺶ5، ﻣﺎ ﻧﺘﯿﺠﻪ ﮔﯿﺮﯼ

ﻣﯽﮐﻨﯿﻢ.

.2ﺁﺛﺎﺭ ﻣﺮﺗﺒﻂ

ﭘﺎﺭﺍﻣﺘﺮﻫﺎﯼﺑﻬﯿﻨﻪ DBSCAN، ﯾﮏ ﻧﻮﻉ ﻣﻌﺮﻭﻑ ﺑﻪ ﻧﺎﻡ OP-TICS )ﺁﻧﮑﺮﺳﺖ، ﺑﺮﻭﻧﯿﮓ، ﮐﺮﯾﮕﻞ ﻭﺳﺎﻧﺪﺭ،(1999 ﭘﯿﺸﻨﻬﺎﺩ ﻣﯽ ﺷﻮﺩ. ﺍﯾﻦ ﯾﮏ ﺗﺮﺗﯿﺐ ﺧﻮﺷﻪ ﺍﯼ ﺭﺍ ﺍﺭﺍﺉﻪ ﻣﯽ ﺩﻫﺪ ﮐﻪ ﺣﺎﻭﯼ ﻧﺘﺎﯾﺞﺧﻮﺷﻪ ﺑﻨﺪﯼ ﺑﻪ ﺩﺳﺖ ﺁﻣﺪﻩ ﺍﺯ ﻃﯿﻒ ﮔﺴﺘﺮﺩﻩ ﺍﯼ ﺍﺯ ﺗﻨﻈﯿﻤﺎﺕ ﭘﺎﺭﺍﻣﺘﺮ ﺍﺳﺖ. ﺩﺭ ﻧﺘﯿﺠﻪ،OPTICS ﺩﺭ ﮐﺸﻒ ﺧﻮﺷﻪ ﻫﺎﯾﯽ ﺑﺎ ﭼﮕﺎﻟﯽ ﻣﺘﻔﺎﻭﺕ ﻗﻮﯼ ﺗﺮ ﺍﺳﺖ. ﺗﻌﺪﺍﺩﯼ ﺍﺯ ﺍﻧﻮﺍﻉ

ﺩﯾﮕﺮDBSCAN ﻧﯿﺰ ﭘﯿﺸﻨﻬﺎﺩ ﺷﺪﻩ ﺍﺳﺖ ﮐﻪ ﺗﻮﺳﻂ DB-CLASD ﻧﺸﺎﻥ ﺩﺍﺩﻩ ﺷﺪﻩ ﺍﺳﺖ.

ﺷﺎﻩ،ﺑﻨﺴﺪﺍﺩﯾﺎ ﻭ ﮔﺎﻧﺎﺗﺮﺍ،)2012 ST-DBSCAN ،ﺑﯿﺮﺍﻧﺖ ﻭ ﮐﻮﺕ،)2007 DM-DBSCAN , ﺍﻟﺒﺎﺗﺎﻭ ﺁﺷﻮﺭ،(2013 ﻭ ﻏﯿﺮﻩ DENCLUE CLUstEring) ﻣﺒﺘﻨﯽ ﺑﺮ (DENsity )ﻫﯿﻨﻬﺒﻮﺭﮒ ﻭ ﮔﺎﺑﺮﯾﻞ،;2007ﻫﯿﻨﻬﺒﻮﺭﮒ ﻭ ﮐﯿﻢ،(2003 ﯾﮑﯽ ﺩﯾﮕﺮ ﺍﺯ ﺭﻭﺵ ﻫﺎﯼ ﺧﻮﺷﻪ ﺑﻨﺪﯼ ﻣﺒﺘﻨﯽ ﺑﺮ ﭼﮕﺎﻟﯽﺑﺎ ﺍﺳﺘﻔﺎﺩﻩ ﺍﺯ ﺗﺨﻤﯿﻦ ﮔﺮ ﭼﮕﺎﻟﯽ ﺍﺳﺖ. ﺩﺭ DENCLUE، ﺣﺪﺍﮐﺜﺮ ﻣﺤﻠﯽ ﺗﺨﻤﯿﻦ ﺗﻮﺯﯾﻊ ﭼﮕﺎﻟﯽﺑﻪ ﻋﻨﻮﺍﻥ ﯾﮏ ﺟﺬﺏ ﮐﻨﻨﺪﻩ ﭼﮕﺎﻟﯽ ﺗﻌﺮﯾﻒ ﻣﯽ ﺷﻮﺩ. ﺳﭙﺲ ﻫﺮ ﻧﻘﻄﻪ ﺩﺭ ﺟﻬﺖ

ﺑﯿﺸﺘﺮﯾﻦﺍﻓﺰﺍﯾﺶ ﺷﯿﺐ ﭼﮕﺎﻟﯽ ﺑﺎ ﺟﺬﺏ ﮐﻨﻨﺪﻩ ﻣﺮﺗﺒﻂ ﻣﯽ ﺷﻮﺩ ﺗﺎ ﺧﻮﺷﻪ ﻫﺎ ﺭﺍ ﺗﺸﮑﯿﻞ

ﺩﻫﻨﺪ.ﺑﺮﺭﺳﯽ ﻫﺎﯼ ﺟﺎﻣﻊ ﺧﻮﺷﻪ ﺑﻨﺪﯼ ﻣﺒﺘﻨﯽ ﺑﺮ ﭼﮕﺎﻟﯽ ﺭﺍ ﻣﯽ ﺗﻮﺍﻥ ﺩﺭ ﺍﯾﻦ ﻣﻘﺎﻟﻪ ﯾﺎﻓﺖﻫﺎﻥ،

ﭘﯽﻭ ﮐﺎﻣﺒﺮ(2011) ﻭ [ﮐﺮﯾﮕﻞ،](#_bookmark3) ﮐﺮﻭﮔﺮ، ﺳﺎﻧﺪﺭ ﻭ ﺯﯾﻤﮏ.(2011)

ﯾﮏﺍﻟﮕﻮﺭﯾﺘﻢ ﺑﺮﺍﯼ ﺍﺟﺮﺍﯼ ﺧﻮﺷﻪ ﺑﻨﺪﯼ ﻣﺒﺘﻨﯽ ﺑﺮ ﭼﮕﺎﻟﯽ ﺑﻪ ﻧﺎﻡ DPC )ﺧﻮﺷﻪ ﺑﻨﺪﯼ ﺑﺎ

ﺟﺴﺘﺠﻮﯼﺳﺮﯾﻊ ﻭ ﯾﺎﻓﺘﻦ ﭘﯿﮏ ﻫﺎﯼ ﭼﮕﺎﻟﯽ( ﺩﺭ ﺳﺎﻝ 2014 [ﭘﯿﺸﻨﻬﺎﺩ](#_bookmark6) ﺷﺪ.ﺭﻭﺩﺭﯾﮕﺰ ﻭ ﻻﯾﻮ، 2014،ﻣﺤﺎﺳﺒﻪ ﻣﯽ ﺷﻮﺩ i ﺑﺮ ﺍﺳﺎﺱ ﺍﯾﻦ [ﻓﺮﺽ](#_bookmark9) ﺍﯾﺠﺎﺩ ﻣﯽ ﺷﻮﺩ ﮐﻪ ﻣﺮﺍﮐﺰ ﺧﻮﺷﻪ ﺑﻬﯿﻨﻪ ﺩﻭ ﻭﯾﮋﮔﯽﺯﯾﺮ ﺭﺍ ﺑﺮﺁﻭﺭﺩﻩ ﻣﯽ ﮐﻨﻨﺪ: (1) ﺁﻧﻬﺎ ﺑﺎﻻﺗﺮﯾﻦ ﭼﮕﺎﻟﯽ ﻣﺤﻠﯽ ﺭﺍ ﺩﺭ ﺧﻮﺷﻪ ﻫﺎﯼ ﻣﺮﺑﻮﻃﻪ ﺩﺍﺭﻧﺪ(2). ﺁﻧﻬﺎ ﺩﺭ ﻓﺎﺻﻠﻪ ﻧﺴﺒﺘﺎ ﺯﯾﺎﺩﯼ ﺍﺯ ﻧﻘﺎﻁ ﺑﺎ ﭼﮕﺎﻟﯽ ﺑﺎﻻﺗﺮ [ﻗﺮﺍﺭ](#_bookmark26) ﺩﺍﺭﻧﺪ. ﺩﺭ ﻧﺘﯿﺠﻪ، ﺩﻭ ﻣﻘﺪﺍﺭﻣﺘﻨﺎﻇﺮ ﺍﺯ ﻧﻘﻄﻪ ﺩﺍﺩﻩ DPC (.ﯾﻌﻨﯽ، ، ﺗﺮﺍﮐﻢ ﻣﺤﻠﯽ ﻭ ، ﻓﺎﺻﻠﻪ ﺗﺎ ﻧﺰﺩﯾﮑﺘﺮﯾﻦ ﻧﻘﻄﻪ ﺩﺍﺩﻩﺑﺎ ﭼﮕﺎﻟﯽ ﺑﺎﻻﺗﺮ. ﺑﺎ ﺗﻮﺟﻪ ﺑﻪ ﻣﺠﻤﻮﻋﻪ ﺩﺍﺩﻩ = ] [ × ﮐﻪ ﺷﺎﻣﻞ*n*ﻧﻘﺎﻁ ﺩﺍﺩﻩ ﺑﺎﺩﺍﺑﻌﺎﺩ ﻫﺮ

ﺛﺎﺑﺖﺷﺪﻩ ﺍﺳﺖ ﮐﻪ DPC ﯾﮏ ﺍﺳﺘﺮﺍﺗﮋﯼ ﺧﻮﺷﻪ ﺑﻨﺪﯼ ﻣﺆﺛﺮ ﺍﺳﺖ ﻭ ﺑﺮ ﺭﻭﯼ ﮐﺎﺭﻫﺎﯼ ﻣﺨﺘﻠﻒﺩﺭ ﺩﻧﯿﺎﯼ ﻭﺍﻗﻌﯽ ﺁﺯﻣﺎﯾﺶ ﺷﺪﻩ ﺍﺳﺖ، ﺍﻣﺎ ﮐﺎﺭﺍﯾﯽ ﻫﻤﭽﻨﯿﻦ ﻣﻨﺠﺮ ﺑﻪ ﻧﻘﺺ ﻫﺎﯼ

ﺁﺷﮑﺎﺭﻣﯽ ﺷﻮﺩ. ﺍﻧﺘﺨﺎﺏ ﺩﻟﺨﻮﺍﻩ ﻣﻌﯿﺎﺭﻫﺎﯼ ﺗﺨﻤﯿﻦ ﭼﮕﺎﻟﯽ، ﻣﺮﺍﮐﺰ ﺧﻮﺷﻪ ﺍﯼ ﻭ ﺧﻄﺮ ﺍﻧﺘﺸﺎﺭ

ﯾﮏ،ﻣﺎﺗﺮﯾﺲ ﻓﺎﺻﻠﻪ = ] [ × ﺍﺑﺘﺪﺍ ﻣﺤﺎﺳﺒﻪ ﻣﯽ ﺷﻮﺩ، ﺟﺎﯾﯽ ﮐﻪ ﻓﺎﺻﻠﻪ ﺍﻗﻠﯿﺪﺳﯽ ﺍﺳﺖ

ﺧﻄﺎﻣﺤﺪﻭﺩﯾﺖ ﻫﺎﯼ ﺍﺻﻠﯽ ﺭﻭﺵ ﺍﺻﻠﯽ ﻫﺴﺘﻨﺪ. ﺑﺴﯿﺎﺭﯼ ﺍﺯ ﺍﻧﻮﺍﻉ ﺑﻬﯿﻨﻪ ﺳﺎﺯﯼ ﺷﺪﻩ DPC

ﺑﯿﻦﺩﺍﺩﻩ ﻫﺎﻣﻦﻭ.*j* ﺳﭙﺲ ﺗﺮﺍﮐﻢ ﻣﺤﻠﯽ ﺭﺍ ﻣﯽ ﺗﻮﺍﻥ ﺑﻪ ﺻﻮﺭﺕ ﺑﯿﺎﻥ ﮐﺮﺩ

)ﯾﻮ،

 ﺩﺍﺭﻧﺪ ﺑﺎﺗﻮﺟﻪ ﺑﻪ ﺍﺛﺮﺍﺕ ﻣﺨﺘﻠﻒ ﺫﮐﺮ ﺷﺪﻩ ﺩﺭ ﺍﯾﻦ ﺳﺎﻟﻬﺎ ﭘﯿﺸﻨﻬﺎﺩ ﺷﺪﻩ ﺍﺳﺖ

ﻟﯿﻮ،ﮔﻮﺉﻮ، ﻟﯿﻮ ﻭ ﯾﺎﺉﻮ،.(2019

# (1)

0 *>* 1 }

# 0≤ 0 =

( - ) ∑ [=](#_bookmark65)

ﺍﻭﻟﯿﻦﻧﻮﻉ ﺑﻬﺒﻮﺩ ﺑﺎ ﻫﺪﻑ ﭘﯿﺸﻨﻬﺎﺩ ﯾﮏ ﻣﺘﺮﯾﮏ ﻣﻮﺛﺮ ﺑﺮﺍﯼ ﺗﺨﻤﯿﻦ ﭼﮕﺎﻟﯽ ﺍﺳﺖ.ﻣﺤﻤﻮﺩ، ﮊﺍﻧﮓ،ﺑﯽ، ﺩﺍﻭﻭﺩ ﻭ ﺍﺣﻤﺪ .2016)ﺑﻪ ﻓﺎﺻﻠﻪ ﺑﺮﺵ ﺍﺳﺘﻔﺎﺩﻩ ﻣﯽ ﮐﻨﺪ DPC ﺑﺮﺍﯼ ﺣﺬﻑ ﺍﺗﮑﺎﯼ KDE( ،ﺗﺨﻤﯿﻦ ﭼﮕﺎﻟﯽ ﻫﺴﺘﻪ) ﺭﺍ ﭘﯿﺸﻨﻬﺎﺩ ﮐﺮﺩ ﮐﻪ ﺍﺯ ﯾﮏ ﺗﺨﻤﯿﻨﮕﺮ ﭼﮕﺎﻟﯽ ﻧﺎﭘﺎﺭﺍﻣﺘﺮﯾﮏ HD

-

CFSFDP ( .ﮊﻭ، ﺳﯽ، ﮊﺍﻧﮓ ﻭ ﮊﻧﮓ(2018) ﻫﻤﭽﻨﯿﻦ ﺍﺯ KDE ﺑﻪ ﻋﻨﻮﺍﻥ ﺟﺎﯾﮕﺰﯾﻨﯽ ﺑﺮﺍﯼ

ﺟﺎﯾﯽﮐﻪ ﺑﻪ ﻋﻨﻮﺍﻥ ﻓﺎﺻﻠﻪ ﻗﻄﻊ ﺗﻌﺮﯾﻒ ﻣﯽ ﺷﻮﺩ. ﺩﯾﺪﻩ ﻣﯽ ﺷﻮﺩ ﮐﻪ ﺑﺎ ﺗﻮﺟﻪ ﺑﻪ ﺗﻌﺪﺍﺩ ﻧﻘﺎﻁ [ﺩﺍﺧﻞﻫﻤﺴﺎﯾﮕﯽ ﮐﺮﻭﯼ i ﺗﺨﻤﯿﻦ ﺯﺩﻩ ﻣﯽ ﺷﻮﺩ. ﺑﺮﺍﯼ ﻣﺠﻤﻮﻋﻪ](#_bookmark52) ﺩﺍﺩﻩ ﻫﺎﯼ ﮐﻮﭼﮏ، ﮐﻪ ﺑﻪ ﺭﺍﺣﺘﯽﻣﯽ ﺗﻮﺍﻧﺪ ﺗﺤﺖ ﺗﺄﺛﯿﺮ ﺧﻄﺎﻫﺎﯼ ﺁﻣﺎﺭﯼ ﻗﺮﺍﺭ ﮔﯿﺮﺩ، ﻋﺒﺎﺭﺕ ﺩﯾﮕﺮﯼ ﺍﺯ ﻧﯿﺰ ﻣﻌﺮﻓﯽ [ﺷﺪﻩ](#_bookmark52) ﺍﺳﺖ. ﺣﺪﺍﻗﻞ ﻓﺎﺻﻠﻪ ﺑﯿﻦ i ﻭ ﻧﻘﻄﻪ ﺑﺎ ﭼﮕﺎﻟﯽ ﺑﯿﺸﺘﺮ ﺍﺳﺖ ﮐﻪ ﺑﻪ ﺻﻮﺭﺕ ﺑﯿﺎﻥ ﻣﯽ ﺷﻮﺩ

ﻣﺘﺮﯾﮏﭼﮕﺎﻟﯽ ﺩﺭ DPC ﺍﺳﺘﻔﺎﺩﻩ ﮐﺮﺩ ﻭ ﺑﺎ ﺷﻨﺎﺳﺎﯾﯽ ﺭﮔﻪ ﻫﺎﯼ ﺧﻮﺷﻪ ﻫﺎ (IVDPC)

⎧ ﺧﻮﺷﻪ ﺑﻨﺪﯼﭘﯿﮏ ﭼﮕﺎﻟﯽ ﺭﺍ ﺍﻧﺠﺎﻡ ﺩﺍﺩ. ﺑﺎ ﺍﯾﻦ ﺣﺎﻝ، ﻫﺰﯾﻨﻪ ﻣﺤﺎﺳﺒﺎﺗﯽ ﺑﺎﻻ ﺑﺮ ﮐﺎﺭﺑﺮﺩ KDE ﺩﺭ

ﺗﺨﻤﯿﻦﭼﮕﺎﻟﯽ ﻣﺠﻤﻮﻋﻪ ﺩﺍﺩﻩ ﻫﺎﯼ ﺍﺑﻌﺎﺩﯼ ﺑﺎﻻ ﺗﺄﺛﯿﺮ ﻣﯽ ﮔﺬﺍﺭﺩ.ﺩﻭ، ﺩﯾﻨﮓ ﻭ ﺟﯿﺎ.2016)ﺭﻧﺞ ﻣﯽ ﺑﺮﺩ،ﺯﯾﺮﺍ ﻫﻤﺎﻥ ﺭﻭﯾﻪ ﺭﺍ ﺩﺭ ﺗﻌﯿﯿﻦ ﻣﺮﺍﮐﺰ ﺧﻮﺷﻪ ﺍﯼ ﻭ ﺗﺨﺼﯿﺺ ﺑﻘﯿﻪ ﻧﻘﺎﻁ ﺍﻋﻤﺎﻝ ﻣﯽ ﮐﻨﺪ DPC ﺑﺮﺍﯼ ﮐﺎﻫﺶ ﺍﺑﻌﺎﺩ ﺩﺍﺩﻩ ﻫﺎ ﺍﺳﺘﻔﺎﺩﻩ ﻣﯽ ﮐﻨﻨﺪ. ﺑﺎ ﺍﯾﻦ ﺣﺎﻝ، ﺭﻭﺵ ﻫﻤﭽﻨﺎﻥ ﺍﺯ

# (2)

. . *<*∃

*ℎ*

( ) ⎪

( ) ⎪ ⎨=

# ⎩

ﻣﺤﺪﻭﺩﯾﺖ ﻫﺎﯼPCA ﺭﺍ ﺑﺮﺍﯼ ﻣﺤﺎﺳﺒﻪ ﭼﮕﺎﻟﯽ ﻣﺤﻠﯽ ﻣﻌﺮﻓﯽ ﻣﯽ ﮐﻨﺪ. ﺁﻧﻬﺎ ﻫﻤﭽﻨﯿﻦ ﺍﺯ KNN ﭘﯿﺸﻨﻬﺎﺩﺷﺪﻩ ﺍﺳﺖ ﮐﻪ ﺍﯾﺪﻩ DPC-KNN (ﺯﯼ، ﮔﺎﺉﻮ، ﺯﯼ، ﻟﯿﻮ ﻭ ﮔﺮﺍﻧﺖ .2016)ﺍﺳﺖ ﺍﻣﺎ

ﻫﻨﻮﺯﺩﺭ ﺷﻨﺎﺳﺎﯾﯽ ﻣﺮﺍﮐﺰ ﺧﻮﺷﻪ ﺧﻮﺩﮐﺎﺭ ﻧﯿﺴﺖ DPC ﺍﺻﻠﯽ ﺍﺳﺘﻔﺎﺩﻩ ﻣﯽ ﮐﻨﺪ. ﺭﻭﺵ ﻗﻮﯼ ﺗﺮ ﺍﺯDPC ﺑﺮﺍﯼ ﯾﮑﺴﺎﻥ ﺳﺎﺯﯼ ﺍﻧﺪﺍﺯﻩ ﮔﯿﺮﯼ ﻫﺎﯼ ﭼﮕﺎﻟﯽ ﺩﺭ KNN ﺭﺍ ﭘﯿﺸﻨﻬﺎﺩ ﮐﺮﺩ ﮐﻪ ﻫﻤﭽﻨﯿﻦ

ﺑﻪﻃﻮﺭ ﮐﻠﯽ ﻣﺮﮐﺰ ﺧﻮﺷﻪ ﺩﺭ ﻧﺎﺣﯿﻪ ﻣﺘﺮﺍﮐﻢ ﯾﮏ ﺧﻮﺷﻪ ﻭ ﺩﻭﺭ ﺍﺯ ﺳﺎﯾﺮ ﺧﻮﺷﻪ ﻫﺎ ﻗﺮﺍﺭ ﺩﺍﺭﺩ. ﺍﯾﻦ ﻫﻤﭽﻨﯿﻦﺩﺭ DPC ﻧﺸﺎﻥ ﺩﺍﺩﻩ ﻣﯽ ﺷﻮﺩ ﮐﻪ ﺩﺭ ﺁﻥ ﻧﻘﺎﻁ ﺑﺎ ﻣﻘﺪﺍﺭ ﻧﺴﺒﺘﺎ ًﺑﺰﺭﮔﯽ ﺍﺯ ﻭ ﺑﯿﺸﺘﺮ ﺑﻪ ﻋﻨﻮﺍﻥﻣﺮﺍﮐﺰ ﺍﻧﺘﺨﺎﺏ ﻣﯽ ﺷﻮﻧﺪ. ﭘﺲ ﺍﺯ ﻣﺤﺎﺳﺒﻪ، ) ، ( ﺩﺭ ﻧﻤﻮﺩﺍﺭ ﺗﺼﻤﯿﻢ ﮔﯿﺮﯼ ﺑﺮﺍﯼ [ﺷﻨﺎﺳﺎﯾﯽﻣﺮﺍﮐﺰ ﺧﻮﺷﻪ ﺭﺳﻢ ﻣﯽ](#_bookmark56) ﺷﻮﻧﺪ. ﻧﻘﺎﻃﯽ ﮐﻪ ﺩﺭ ﮔﻮﺷﻪ ﺳﻤﺖ ﺭﺍﺳﺖ ﺑﺎﻻﯼ ﻧﻤﻮﺩﺍﺭ

ﺍﺯFKNN-DPC (ﭼﻦ ﻭ ﻫﻤﮑﺎﺭﺍﻥ .2016)ﺑﺮﺍﯼ ﻧﻤﺎﯾﺶ ﭘﯿﮏ ﻫﺎﯼ ﭼﮕﺎﻟﯽ ﻣﯽ ﺳﺎﺯﺩ KNN

ﻗﺮﺍﺭﺩﺍﺭﻧﺪ، ﺟﺎﯾﯽ ﮐﻪ ﻣﻘﺪﺍﺭ ﻧﺴﺒﺘﺎ ًﺯﯾﺎﺩﯼ ﺭﺍ ﻧﺸﺎﻥ ﻣﯽ ﺩﻫﺪ ﻭ ، ﺑﻪ ﻋﻨﻮﺍﻥ ﻣﺮﺍﮐﺰ ﺍﻧﺘﺨﺎﺏ [ﻣﯽ](#_bookmark56)

ﭘﯿﺸﻨﻬﺎﺩﯼﮐﻪ ﺳﺘﻮﻥ ﻓﻘﺮﺍﺕ ﭼﮕﺎﻟﯽ ﺭﺍ ﺑﺎ ﺍﻧﺘﺨﺎﺏ ﻭ ﮔﺮﻭﻩ ﺑﻨﺪﯼ ﻧﻘﺎﻁ ﺑﺎ ﭼﮕﺎﻟﯽ ﻧﯿﻤﻪ ﺑﺎﻻﺗﺮ ﺩﺭ ﻫﺮﺧﻮﺷﻪ ﺍﻭﻟﯿﻪ ﺑﺮ ﺍﺳﺎﺱ CLUBﮔﻨﮓ ﻭ ﻫﻤﮑﺎﺭﺍﻥ)2018) ADPC-KNN .ﭘﺎﺭﺍﻣﺘﺮ ﺩﯾﮕﺮﯼ ﺭﺍ ﺩﺭﻓﺮﺁﯾﻨﺪ ﺧﻮﺷﻪ ﺑﻨﺪﯼ ﺷﺎﻣﻞ ﻣﯽ ﺷﻮﻧﺪ ﮐﻪ ﺍﺳﺘﺤﮑﺎﻡ ﺭﻭﺵ ﻫﺎ ﺭﺍ ﮐﺎﻫﺶ ﻣﯽ ﺩﻫﺪ

RECOME ﻭ CLUB ﻧﺴﺒﯽ ﺗﺮﺍﮐﻢ ﻫﺴﺘﻪ ﻧﺰﺩﯾﮑﺘﺮﯾﻦ ﻫﻤﺴﺎﯾﻪ ﭘﯿﺸﻨﻬﺎﺩ ﮐﺮﺩ. ﺑﺎ ﺍﯾﻦ ﺣﺎﻝ، ﻫﺮ

ﺩﻭK ﺭﺍ ﺑﺎ ﺩﺭ ﻧﻈﺮ ﮔﺮﻓﺘﻦ ﯾﮏ ﺍﻧﺪﺍﺯﻩ ﮔﯿﺮﯼ ﭼﮕﺎﻟﯽ ﺟﺪﯾﺪ ﺑﺮ ﺍﺳﺎﺱ RECOME (ﯾﺎﺉﻮﻫﻮﯾﯽ، ﮊﻧﮓﻣﯿﻨﮓ ﻭ ﻧﯿﺶ، (2017 ﻭ DPC-DLP )ﺳﯿﺪﯼ، ﻟﻄﻔﯽ، ﻣﺮﺍﺩﯼ ﻭ ﻗﺎﺩﺭ،(2019 ﻫﺮ ﺩﻭ ﺍﺯ KNN ﺩﺭ ﺗﻌﺮﯾﻒ ﻓﺎﺻﻠﻪ ﻗﻄﻊ ﺍﺳﺘﻔﺎﺩﻩ ﻣﯽ ﮐﻨﻨﺪ .ﻣﺸﺎﺑﻪ DPC، ﺍﯾﻦ ﺭﻭﺵ ﻫﺎ ﻧﻤﯽ ﺗﻮﺍﻧﻨﺪ ﺑﻪ ﻃﻮﺭﮐﺎﻣﻞ ﺷﮑﻞ ﻧﺎﻣﻨﻈﻢ ﺧﻮﺷﻪ ﻫﺎ ﺭﺍ ﺁﺷﮑﺎﺭ ﮐﻨﻨﺪ ﯾﺎ ﺧﻮﺷﻪ ﻫﺎﯼ ﻫﻤﭙﻮﺷﺎﻧﯽ ﺭﺍ ﺑﻪ ﺩﻟﯿﻞ ﺗﮑﯿﻪﺑﺮ ﺭﻭﺵ ﻫﺎﯼ ﺛﺎﺑﺖ ﺗﺸﺨﯿﺺ ﺩﻫﻨﺪ. .ﻟﯿﻮ، ﻭﺍﻧﮓ ﻭ ﯾﻮ.2018)ﺭﺍ ﭘﯿﺸﻨﻬﺎﺩ ﮐﺮﺩ ﮐﻪ ﭼﮕﺎﻟﯽ

ﻣﺤﻠﯽﺭﺍ ﺑﺮ ﺍﺳﺎﺱ ﺍﯾﺪﻩ ﺷﺒﺎﻫﺖ ﻧﺰﺩﯾﮑﺘﺮﯾﻦ ﻫﻤﺴﺎﯾﻪ ﻣﺸﺘﺮﮎ ﺗﺨﻤﯿﻦ ﻣﯽ ﺯﻧﺪ. ﺑﺎ ﺍﯾﻦ ﺣﺎﻝ،

ﻫﻨﻮﺯﺍﺯ ﻣﺤﺪﻭﺩﯾﺖ ﺍﻧﺘﺨﺎﺏ ﺩﺳﺘﯽ ﺗﻌﺪﺍﺩ ﻣﺮﺍﮐﺰ ﺧﻮﺷﻪ ﺭﻧﺞ ﻣﯽ ﺑﺮﺩ SNN-DPC (

ﻧﻮﻉﺩﻭﻡ ﺑﻬﺒﻮﺩ ﺑﺎ ﻫﺪﻑ ﺷﻨﺎﺳﺎﯾﯽ ﺧﻮﺩﮐﺎﺭ ﺗﻌﺪﺍﺩ ﻣﺮﺍﮐﺰ ﺧﻮﺷﻪ ﺍﯼ ﺍﺳﺖ.ﻟﯿﺎﻧﮓ ﻭ ﭼﻦ) 2016ﭘﯿﺸﻨﻬﺎﺩ ﺷﺪﻩ ﺍﺳﺖ ﮐﻪ ﯾﮏ ﺍﺳﺘﺮﺍﺗﮋﯼ ﺗﻘﺴﯿﻢ ﻭ ﺣﮑﻮﻣﺖ ﮐﻦ ﺭﺍ ﺑﺮﺍﯼ ﺗﻌﯿﯿﻦ ﺗﻌﺪﺍﺩ ﺍﯾﺪﻩﺁﻝ ﺧﻮﺷﻪ ﻫﺎ ﻣﻌﺮﻓﯽ ﻣﯽ ﮐﻨﺪ. ﺑﺎ ﺍﯾﻦ ﺣﺎﻝ، ﺳﺎﺧﺘﺎﺭ ﻣﺤﻠﯽ ﺭﺍ ﻧﺎﺩﯾﺪﻩ ﻣﯽ ﮔﯿﺮﺩ 3DC (

ﺷﻮﻧﺪ.ﺩﺭ ﻧﻬﺎﯾﺖ، ﺑﻘﯿﻪ ﻧﻘﺎﻁ ﺑﻪ ﻫﻤﺎﻥ ﺧﻮﺷﻪ ﺍﯼ ﺍﺧﺘﺼﺎﺹ ﻣﯽ ﯾﺎﺑﺪ ﮐﻪ ﻧﺰﺩﯾﮏ ﺗﺮﯾﻦ

[ﻫﻤﺴﺎﯾﻪﺁﻥ ﺑﺎ](#_bookmark34) ﭼﮕﺎﻟﯽ ﺑﺎﻻﺗﺮ ﺑﻪ ﺁﻥ ﺗﻌﻠﻖ ﺩﺍﺭﺩ.

ﻧﺸﺎﻥ[ﺩﺍﺩﻩ](#_bookmark44) ﺷﺪﻩ [ﺍﺳﺖ ﮐﻪ](#_bookmark44) DPC ﯾﮏ ﺭﻭﺵ ﻣﻮﺛﺮ ﺩﺭ ﺧﻮﺷﻪ ﺑﻨﺪﯼ ﻣﺠﻤﻮﻋﻪ ﻫﺎﯼ ﺩﺍﺩﻩ ﺑﺎ

ﺍﺷﮑﺎﻝﻏﯿﺮ ﮐﺮﻭﯼ ﺍﺳﺖ. ﺑﺎ ﺍﯾﻦ ﺣﺎﻝ، ﺍﺛﺮﺑﺨﺸﯽ DPC ﺑﻪ ﺗﺨﻤﯿﻦ ﺗﻮﺯﯾﻊ ﭼﮕﺎﻟﯽ ﻣﺘﮑﯽ ﺍﺳﺖ.

ﺍﮔﺮﭼﻪﻧﻮﯾﺴﻨﺪﮔﺎﻥ ﺑﺮ ﺍﺳﺎﺱ ﺭﻭﯾﮑﺮﺩﻫﺎﯼ ﺍﮐﺘﺸﺎﻓﯽ ﺩﻭ ﻣﻌﯿﺎﺭ ﺗﺮﺍﮐﻢ ﺑﺮﺍﯼ ﻣﺠﻤﻮﻋﻪ ﺩﺍﺩﻩ ﻫﺎﯼ

ﺑﺰﺭﮒﻭ ﮐﻮﭼﮏ ﺍﺭﺍﺉﻪ ﮐﺮﺩﻧﺪ، ﻫﯿﭻ ﻣﻌﯿﺎﺭﯼ ﺩﺭ ﺗﻌﺮﯾﻒ ﻣﻘﯿﺎﺱ ﻣﺠﻤﻮﻋﻪ ﺩﺍﺩﻩ ﻫﺎ ﺍﺭﺍﺉﻪ ﻧﺸﺪ. [ﯾﮏﺍﺳﺘﺮﺍﺗﮋﯼ ﻣﻨﻄﻘﯽ ﺗﺮ ﺩﺭ ﺗﺨﻤﯿﻦ](#_bookmark64) ﭼﮕﺎﻟﯽ، ﺍﺩﻏﺎﻡ ﺍﻃﻼﻋﺎﺕ ﺗﻮﺯﯾﻊ ﻣﺤﻠﯽ ﺩﺭ ﻣﻮﺭﺩ ﻣﺎﻫﯿﺖ ﻣﺠﻤﻮﻋﻪﺩﺍﺩﻩ ﻫﺎ [ﺍﺳﺖ. ﻋﻼﻭﻩ ﺑﺮ ﺍﯾﻦ، ﺷﻨﺎﺳﺎﯾﯽ ﻣﺮﺍﮐﺰ ﺧﻮﺷﻪ ﺍﯼ](#_bookmark54) ﻣﺴﺘﻠﺰﻡ ﺗﻌﯿﯿﻦ ﺁﺳﺘﺎﻧﻪ ﺑﺮﺍﯼﺩﻭ ﮐﻤﯿﺖ ) ﻭ ( ﺑﺮﺍﯼ ﺗﻌﺮﯾﻒ ﻣﻔﻬﻮﻡ "ﻧﺴﺒﺘﺎ ﺑﺰﺭﮒ." ﺍﺯ ﺁﻧﺠﺎﯾﯽ ﮐﻪ ﻧﻮﯾﺴﻨﺪﮔﺎﻥ ﺭﻭﺵ ﮐﻤﯽﺑﺮﺍﯼ ﺗﻨﻈﯿﻢ ﺁﺳﺘﺎﻧﻪ ﻫﺎ ﺍﺭﺍﺉﻪ ﻧﻤﯽ ﺩﻫﻨﺪ، ﺗﻌﯿﯿﻦ ﺗﻌﺪﺍﺩ ﺍﯾﺪﻩ ﺁﻝ ﺧﻮﺷﻪ ﻫﺎ ﺑﺎ ﺍﺳﺘﻔﺎﺩﻩ ﺍﺯ

[ﻧﻤﻮﺩﺍﺭ](#_bookmark47)ﺗﺼﻤﯿﻢ [ﺑﺮﺍﯼ ﮐﺎﺭﺑﺮﺍﻥ ﻣﺒﻬﻢ](#_bookmark47) ﺍﺳﺖ. ﻋﻼﻭﻩ ﺑﺮ ﺍﯾﻦ، ﺭﻭﺵ ﮔﺎﻡ ﺑﻪ ﮔﺎﻡ DPC ﺩﺭ ﺗﺨﺼﯿﺺ ﻧﻘﺎﻁﺑﺎﻗﯽ ﻣﺎﻧﺪﻩ ﭘﺲ ﺍﺯ ﯾﺎﻓﺘﻦ ﻣﺮﺍﮐﺰ ﺧﻮﺷﻪ ﻣﻤﮑﻦ ﺍﺳﺖ ﺑﺎﻋﺚ ﺍﻧﺘﺸﺎﺭ ﺧﻄﺎﻫﺎ ﺷﻮﺩ. ﯾﻌﻨﯽ ﻫﻨﮕﺎﻣﯽﮐﻪ ﯾﮏ ﻧﻘﻄﻪ ﺑﻪ ﯾﮏ ﺧﻮﺷﻪ ﺍﺷﺘﺒﺎﻩ ﺗﻘﺴﯿﻢ ﻣﯽ ﺷﻮﺩ، ﻫﻤﺴﺎﯾﮕﺎﻥ ﺁﻥ ﺑﺎ ﭼﮕﺎﻟﯽ ﮐﻤﺘﺮ

ﺑﻪﻧﺎﭼﺎﺭ ﺩﺭ ﮔﺮﻭﻩ ﺍﺷﺘﺒﺎﻩ ﻃﺒﻘﻪ ﺑﻨﺪﯼ ﻣﯽ ﺷﻮﻧﺪ.

ﻋﮑﺲ.1.ﻓﺮﺁﯾﻨﺪ ﺍﺻﻠﯽ .ICKDC )ﺍﻟﻒ( ﺗﻮﺯﯾﻊ ﺍﺻﻠﯽﺗﺠﻤﻊ. )ﺏ( ﻧﻮﺍﺣﯽ ﺍﺻﻠﯽ ﺩﺭ ﻫﺮ ﺧﻮﺷﻪ. ﺝ( ﻧﺘﺎﯾﺞ ﻧﻬﺎﯾﯽ

[.3.1ﺗﺨﻤﯿﻦ ﭼﮕﺎﻟﯽ](#_bookmark57) ﻣﺤﻠﯽ

ﺛﺎﺑﺖﺷﺪﻩ ﺍﺳﺖ ﮐﻪ KNN ﯾﮏ ﺗﮑﻨﯿﮏ ﻣﻮﺛﺮ ﺑﺮﺍﯼ ﺗﺨﻤﯿﻦ ﭼﮕﺎﻟﯽ ﺍﺳﺖ. ﻫﺪﻑ ﺑﺪﺳﺖ

ﺁﻭﺭﺩﻥK ﺷﺒﯿﻪ ﺗﺮﯾﻦ ﻫﻤﺴﺎﯾﮕﺎﻥ ﻧﻘﻄﻪ ﺍﺳﺖﻣﻦ، ﻣﻌﻤﻮﻻ ًﺑﺎ ﺍﻧﺘﺨﺎﺏ ﻧﻤﻮﻧﻪ ﻫﺎﯼ ﻣﺠﺎﻭﺭ ﺁﻥ ﺑﺮ

[ﺍﺳﺎﺱﻓﺎﺻﻠﻪ](#_bookmark29) ﺍﻗﻠﯿﺪﺳﯽ ﺁﻧﻬﺎ ﻭ ﺗﺨﻤﯿﻦ ﭼﮕﺎﻟﯽ ﻣﺤﻠﯽ ﺑﺎ ﺍﻃﻼﻋﺎﺕ ﺗﻮﺯﯾﻊ ﺩﺍﺩﻩ ﺷﺪﻩ ﺗﻮﺳﻂ

ﻫﻤﺴﺎﯾﮕﺎﻥﺁﻥ. ﺩﺭ ICKDC، ﺷﮑﻞ ﺟﺪﯾﺪﯼ ﺍﺯ ﺗﺨﻤﯿﻦ [ﭼﮕﺎﻟﯽ](#_bookmark29) ﺑﺮ [ﺍﺳﺎﺱ ﺍﯾﺪﻩ KNN ﭘﯿﺸﻨﻬﺎﺩ](#_bookmark29) ﺷﺪﻩﺍﺳﺖ. ﺑﻪ ﻃﻮﺭ ﮐﻠﯽ، ﺍﻃﻼﻋﺎﺕ ﺗﻮﺯﯾﻊ ﻣﺤﻠﯽ ﺭﺍ ﺑﺎ ﺍﻧﺒﺎﺷﺖ ﻓﺎﺻﻠﻪ ﺑﯿﻦ ﺁﻧﻬﺎ ﯾﮑﭙﺎﺭﭼﻪ ﻣﯽ

ﮐﻨﺪﻣﻦﻭ ﻫﻤﺴﺎﯾﮕﺎﻥ ﺁﻥ ﺑﻪ ﺻﻮﺭﺕ ﺗﺎﺑﻊ ﻫﺴﺘﻪ ﮔﺎﻭﺳﯽ.

ﻣﺠﻤﻮﻋﻪﺩﺍﺩﻩ ﻫﺎﯾﯽ ﮐﻪ ﻣﻤﮑﻦ ﺍﺳﺖ ﺑﺎﻋﺚ ﺍﺯ ﺩﺳﺖ ﺭﻓﺘﻦ ﺧﻮﺷﻪ ﻫﺎ ﺷﻮﻧﺪ.ﺷﻮ، ﻭﺍﻧﮓ ﻭ ﺩﻧﮓ.2016)ﺑﻪ ﺩﻟﯿﻞ ﺍﺳﺘﺮﺍﺗﮋﯼ ﺧﻮﺷﻪ ﺑﻨﺪﯼ ﺳﻠﺴﻠﻪ ﻣﺮﺍﺗﺒﯽ، ﺍﻧﺘﺸﺎﺭ ﺧﻄﺎﻫﺎ ﺭﺍ ﺗﺸﺪﯾﺪ ﻣﯽ ﮐﻨﻨﺪDenPEHC ﻭ DCﺭﺍ ﭘﯿﺸﻨﻬﺎﺩ ﮐﺮﺩ ﮐﻪ ﻣﯽ ﺗﻮﺍﻧﺪ ﺑﻪ ﻃﻮﺭ ﺧﻮﺩﮐﺎﺭ ﺗﻤﺎﻡ ﻣﺮﺍﮐﺰ ﻣﻤﮑﻦ ﺭﺍﺷﻨﺎﺳﺎﯾﯽ ﮐﻨﺪ ﻭ ﯾﮏ ﻧﻤﺎﯾﺶ ﺳﻠﺴﻠﻪ ﻣﺮﺍﺗﺒﯽ ﺑﺮﺍﯼ ﻣﺠﻤﻮﻋﻪ ﺩﺍﺩﻩ ﺑﺴﺎﺯﺩ. ﺑﺎ ﺍﯾﻦ ﻭﺟﻮﺩ، ﻫﺮ ﺩﻭ3 DenPEHC (ﺑﯽ، ﻣﺤﻤﻮﺩ، ﺭﻭﺍﻥ، ﺳﺎﻥ ﻭ ﺩﺍﻭﻭﺩ(2016) ﻓﺮﺁﯾﻨﺪ ﺍﻧﺘﺨﺎﺏ ﻣﺮﮐﺰ ﺭﺍ ﺑﺎ

ﺍﺳﺘﻔﺎﺩﻩﺍﺯ ﯾﮏ ﺭﻭﺵ ﺍﮐﺘﺸﺎﻓﯽ ﮐﻪ ﺗﻌﺪﺍﺩ ﺯﯾﺎﺩﯼ ﺍﺯ ﻣﺮﺍﮐﺰ ﺧﻮﺷﻪ ﺑﺎﻟﻘﻮﻩ ﺭﺍ ﺍﺩﻏﺎﻡ ﻣﯽ ﮐﻨﺪ، ﺧﻮﺩﮐﺎﺭﮐﺮﺩ. ﺑﺎ ﺍﯾﻦ ﺣﺎﻝ، ﺭﻭﺵ ﻣﻤﮑﻦ ﺍﺳﺖ ﺑﻪ ﺩﻟﯿﻞ ﺍﺗﮑﺎ ﺑﻪ ﺛﺎﺑﺖ، ﺧﻮﺷﻪ ﻫﺎﯾﯽ ﺑﺎ ﺍﺷﮑﺎﻝ ﭼﻨﺪﮔﺎﻧﻪﺭﺍ ﺗﺸﺨﯿﺺ ﻧﺪﻫﺪ. .ﮐﻪ ﺩﺭﻟﻄﻔﯽ، ﻣﺮﺍﺩﯼ ﻭ ﺑﯿﮕﯽ(2020) ﻭﺳﯿﺪﯼ ﻭ ﻫﻤﮑﺎﺭﺍﻥ2019) ﻧﻮﯾﺴﻨﺪﮔﺎﻥ،ﻫﺮ ﺩﻭ ﯾﮏ ﺗﺎﺑﻊ ﺍﻣﺘﯿﺎﺯ ﺑﺮﺍﯼ ﮐﻤﮏ ﺑﻪ ﺗﻌﯿﯿﻦ ﻣﺮﺍﮐﺰ ﺧﻮﺷﻪ ﺍﯼ ﺑﻪ ﻋﻨﻮﺍﻥ ﺟﺎﯾﮕﺰﯾﻨﯽ ﺑﺮﺍﯼﻧﻤﻮﺩﺍﺭ ﺗﺼﻤﯿﻢ ﺍﺭﺍﺉﻪ ﮐﺮﺩﻧﺪ. ﺑﺎ ﺍﯾﻦ ﺣﺎﻝ، ﺍﯾﻦ ﺭﻭﺵ ﻫﺎ ﺑﻪ ﺟﺎﯼ ﻏﻠﺒﻪ ﺑﺮ ﻣﺸﮑﻞ، ﺑﻪ

ﺍﺟﺎﺯﻩﺩﻫﯿﺪ = 1، 2،…، ﯾﮏ ﻣﺠﻤﻮﻋﻪ ﺩﺍﺩﻩ ﺣﺎﻭﯼ*n*ﻧﻤﻮﻧﻪ ﻫﺎ. ﺑﺮﺍﯼ ﻫﺮ ﻧﻘﻄﻪ ، ﻓﺎﺻﻠﻪ

ﭘﺎﺭﺍﻣﺘﺮﻗﺎﺑﻞ ﺗﻨﻈﯿﻢ ﺩﯾﮕﺮﯼ ﻧﯿﺎﺯ ﺩﺍﺭﻧﺪ.

ﺑﯿﻦ ﻭ ﺑﻘﯿﻪ ﻧﮑﺎﺕ ﺩﺭ*D*ﺑﺮﺍﯼ ﯾﺎﻓﺘﻦ ﺑﻪ ﺗﺮﺗﯿﺐ ﺻﻌﻮﺩﯼ ﻣﺮﺗﺐ ﻣﯽ ﺷﻮﻧﺪﮎﻧﺰﺩﯾﮑﺘﺮﯾﻦ ﻧﻤﻮﻧﻪ

ﻫﺎ،ﻧﺸﺎﻥ ﺩﺍﺩﻩ ﺷﺪﻩ ﺑﺎ ) (ﮐﻪ ﻣﯽ ﺗﻮﺍﻧﺪ ﺑﯿﺸﺘﺮ ﺑﻪ ﺻﻮﺭﺕ ﺑﯿﺎﻥ ﺷﻮﺩ

ﻫﺪﻑﺳﻮﻣﯿﻦ ﻧﻮﻉ ﺑﻬﺒﻮﺩ، ﺑﻬﺒﻮﺩ ﺍﺳﺘﺮﺍﺗﮋﯼ ﺗﺨﺼﯿﺺ ﺑﺮﺍﯼ ﮐﺎﻫﺶ ﺍﻧﺘﺸﺎﺭ ﺧﻄﺎ ﺍﺳﺖ.

ﺩﺭﺑﺴﯿﺎﺭﯼ ﺍﺯ ﺍﻧﻮﺍﻉ DPC، ﺍﯾﺪﻩ k ﻧﺰﺩﯾﮑﺘﺮﯾﻦ ﻫﻤﺴﺎﯾﻪ ﺩﺭ ﺍﺳﺘﺮﺍﺗﮋﯼ ﻫﺎﯼ ﺗﺠﻤﻊ ﺗﺮﮐﯿﺐ ﻣﯽ

# (3)

{(( ) ، ) ≥( ، ) ∣ ∋ } = ( )

ﺷﻮﺩ.ﺑﻪ ﻋﻨﻮﺍﻥ ﻣﺜﺎﻝ، ﺩﺭﮊﻭ ﻭ ﻫﻤﮑﺎﺭﺍﻥ(2018) ﻧﻮﯾﺴﻨﺪﮔﺎﻥ ﺭﮔﻪ ﻫﺎﯼ ﺧﻮﺷﻪ ﻫﺎ ﺭﺍ ﺑﺎ ﺍﺗﺼﺎﻝ

ﺟﺎﯾﯽﮐﻪ ) (ﻫﺴﺖﮎﻧﺰﺩﯾﮑﺘﺮﯾﻦ ﻫﻤﺴﺎﯾﻪ . ﻣﺘﺮﯾﮏ ﭼﮕﺎﻟﯽ ﺟﺪﯾﺪ

ﺳﭙﺲﺑﻪ ﻋﻨﻮﺍﻥ ﺗﻌﺮﯾﻒ ﻣﯽ ﺷﻮﺩ

ﺟﻔﺖ ﻫﺎﯾﯽﺑﺎ ﺑﯿﺸﺘﺮﯾﻦ ﺷﺒﺎﻫﺖ ﺍﺯ ﻣﻨﺎﻃﻖ ﺑﺎ ﭼﮕﺎﻟﯽ ﺑﺎﻻ ﺑﻪ ﻣﺮﺯﻫﺎﯼ ﺧﻮﺷﻪ ﺳﺎﺧﺘﻨﺪ. ﺳﭙﺲ

ﺑﻘﯿﻪﻧﻘﺎﻁ ﺑﻪ ﻧﺰﺩﯾﮑﺘﺮﯾﻦ ﺭﮔﻬﺎ ﺍﺧﺘﺼﺎﺹ ﺩﺍﺩﻩ ﻣﯽ ﺷﻮﻧﺪ. ﮐﻪ ﺩﺭﺯﯼ ﻭ ﻫﻤﮑﺎﺭﺍﻥ(2016)

ﻧﻮﯾﺴﻨﺪﮔﺎﻥﯾﮏ ﺭﻭﺵ ﺩﻭ ﻣﺮﺣﻠﻪ ﺍﯼ ﺭﺍ ﺑﺮﺍﯼ ﺍﻧﺘﺸﺎﺭ ﺑﺮﭼﺴﺐ ﭘﯿﺸﻨﻬﺎﺩ ﮐﺮﺩﻧﺪ. ﺍﺳﺘﺮﺍﺗﮋﯼ ﺍﻭﻝ،

# (4)

[2](#_bookmark56) ( [، ) -](#_bookmark56)

∑ =

( ) ∋

ﻏﯿﺮﭘﺮﺕ ﺭﺍ ﺑﺮ ﺍﺳﺎﺱ ﺟﺴﺘﺠﻮﯼ k ﻧﺰﺩﯾﮑﺘﺮﯾﻦ ﻫﻤﺴﺎﯾﻪ ﻫﺎ ﮐﻪ ﺍﺯ ﭘﯿﮏ ﻫﺎﯼ ﭼﮕﺎﻟﯽ ﺷﺮﻭﻉ

ﻣﯽ ﺷﻮﻧﺪ،ﺍﺧﺘﺼﺎﺹ ﻣﯽ ﺩﻫﺪ. ﺍﺳﺘﺮﺍﺗﮋﯼ ﺩﻭﻡ ﺳﺎﯾﺮ ﻧﻘﺎﻁ ﺭﺍ ﺑﺎ ﺍﺳﺘﻔﺎﺩﻩ ﺍﺯ ﺗﮑﻨﯿﮏ KNN

ﻋﺒﺎﺭﺕﺟﺪﯾﺪ ﺍﻃﻼﻋﺎﺕ ﺗﻮﺯﯾﻊ ﭼﮕﺎﻟﯽ ﻣﺤﻠﯽ ﺭﺍ ﺍﺯ ﻧﻤﻮﻧﻪ ﻫﺎﯼ ﻣﺠﺎﻭﺭ ﺗﺮﮐﯿﺐ ﻣﯽ ﮐﻨﺪ ﻭ ﺗﻔﺎﻭﺕ[ﺑﯿﻦ ﺍﯾﻦ ﺗﻮﺯﯾﻊ ﻫﺎ ﺭﺍ ﺑﺎ](#_bookmark44) ﺷﮑﻞ ﻣﺠﺬﻭﺭ ﻓﺎﺻﻠﻪ ﺍﻓﺰﺍﯾﺶ ﻣﯽ ﺩﻫﺪ. ﺩﺭ ﻧﺘﯿﺠﻪ، ﺗﻮﺯﯾﻊ ﭼﮕﺎﻟﯽﻣﺤﻠﯽ ﺭﺍ ﻣﯽ ﺗﻮﺍﻥ ﺑﻪ ﺧﻮﺑﯽ ﺑﺎ ﺍﺳﺘﻔﺎﺩﻩ ﺍﺯ ﻣﻌﺎﺩﻟﻪ ﺍﺳﺘﺨﺮﺍﺝ ﮐﺮﺩ(5).ﮐﻪ [ﺑﺮ ﺭﻭﺍﺑﻂ](#_bookmark54)

[ﻣﺘﻘﺎﺑﻞﺑﯿﻦ](#_bookmark47) ﻫﻤﺴﺎﯾﮕﺎﻥ ﺗﺎﮐﯿﺪ ﺩﺍﺭﺩ. ﺑﺮﺍﯼ ﻣﺠﻤﻮﻋﻪ ﺩﺍﺩﻩ ﻫﺎﯾﯽ ﺑﺎ ﺗﻮﺯﯾﻊ ﭼﮕﺎﻟﯽ [ﻧﺎﻫﻤﮕﻦ ﮐﻪ ﺩﺭ](#_bookmark54)

[ﺁﻥ](#_bookmark47)DPC ﻋﻤﻠﮑﺮﺩ ﺭﺿﺎﯾﺖ ﺑﺨﺸﯽ ﺭﺍ ﻧﺸﺎﻥ ﻣﯽ ﺩﻫﺪ ﻣﻨﺎﺳﺐ ﺍﺳﺖ.

[.3.2](#_bookmark36)ﺷﻨﺎﺳﺎﯾﯽ [ﻧﻘﺎﻁ](#_bookmark36) ﺍﺻﻠﯽ

ﭘﺲﺍﺯ ﺗﺨﻤﯿﻦ ﭼﮕﺎﻟﯽ ﺍﺯ ﻃﺮﯾﻖ ﻣﻌﺎﺩﻟﻪ(5)ﺑﺮﺍﯼ ﻫﺮ ﻧﻤﻮﻧﻪ ﺩﺭ*D*[ﺷﻨﺎﺳﺎﯾﯽ ﻧﻘﺎﻁ ﺍﺻﻠﯽ](#_bookmark56) ﻣﯽ

ﻓﺎﺯﯼﺍﺧﺘﺼﺎﺹ ﻣﯽ ﺩﻫﺪ. ﮐﻪ ﺩﺭﮔﻨﮓ ﻭ ﻫﻤﮑﺎﺭﺍﻥ(2018) ﻭ ﺳﯿﺪﯼ ﻭ ﻫﻤﮑﺎﺭﺍﻥ2019)ﻫﺮ ﺩﻭ ﻧﻮﯾﺴﻨﺪﻩﯾﮏ ﺍﺳﺘﺮﺍﺗﮋﯼ ﺍﻧﺘﺸﺎﺭ ﺑﺮﭼﺴﺐ ﻣﺒﺘﻨﯽ ﺑﺮ ﮔﺮﺍﻑ KNN ﺭﺍ ﺑﺮﺍﯼ ﺗﺨﺼﯿﺺ ﻧﻘﺎﻁ ﺑﺎﻗﯽ ﻣﺎﻧﺪﻩﭘﯿﺸﻨﻬﺎﺩ ﮐﺮﺩﻧﺪ. ﮐﻪ ﺩﺭﻟﯿﻮ ﻭ ﻫﻤﮑﺎﺭﺍﻥ 2018)ﻧﻮﯾﺴﻨﺪﮔﺎﻥ ﯾﮏ ﺭﻭﺵ ﺗﺨﺼﯿﺺ ﺩﻭ ﻣﺮﺣﻠﻪﺍﯼ ﺭﺍ ﺑﺮ ﺍﺳﺎﺱ ﺗﺨﺼﯿﺺ ﻧﺎﮔﺰﯾﺮ ﻭ ﺍﺣﺘﻤﺎﻻ ًﺗﺎﺑﻌﯽ ﺑﺮﺍﯼ ﺗﺨﺼﯿﺺ ﻧﻘﺎﻁ ﻏﯿﺮ ﻣﺮﮐﺰﯼ

ﻣﻌﺮﻓﯽﮐﺮﺩﻧﺪ.

ﺍﮔﺮﭼﻪﺑﺴﯿﺎﺭﯼ ﺍﺯ ﮔﻮﻧﻪ ﻫﺎﯼ ﻣﺒﺘﻨﯽ ﺑﺮ DPC ﭘﯿﺸﻨﻬﺎﺩ ﺷﺪﻩ ﺍﻧﺪ، ﻫﯿﭻ ﯾﮏ ﺍﺯ ﺁﻧﻬﺎ

ﻧﻤﯽ ﺗﻮﺍﻧﻨﺪﺭﻭﺵ ﺍﺻﻠﯽ ﺭﺍ ﺑﺎ ﺩﺭ ﻧﻈﺮ ﮔﺮﻓﺘﻦ ﻫﺮ ﺳﻪ ﺟﻨﺒﻪ ﺑﻪ ﻃﻮﺭ ﺧﻼﺻﻪ ﺑﻬﺒﻮﺩ ﺑﺨﺸﻨﺪ. ﺑﺎ

ﺍﻟﻬﺎﻡﺍﺯ ﮐﺎﺭ ﺍﻭﻟﯿﻪﺩﻭ ﻭ ﻫﻤﮑﺎﺭﺍﻥ(2016) ﻭﺯﯼ ﻭ ﻫﻤﮑﺎﺭﺍﻥ(2016)، ﺑﯿﺸﺘﺮ ﺍﻧﻮﺍﻉ، ﻣﻔﻬﻮﻡ KNN ﺭﺍﺩﺭ ﺗﺨﻤﯿﻦ ﭼﮕﺎﻟﯽ ﻭ ﻓﺮﺁﯾﻨﺪ ﺗﺨﺼﯿﺺ [ﺗﺮﮐﯿﺐ](#_bookmark5) ﻣﯽ ﮐﻨﻨﺪ. ﺑﻪ ﺩﻧﺒﺎﻝ ﺍﯾﻦ ﺍﯾﺪﻩ، ﺩﺭ ﺑﺨﺶ ﺑﻌﺪﯼ

ﺗﻮﺍﻧﺪﺑﻪ ﺭﺍﺣﺘﯽ ﺍﺯ ﻃﺮﯾﻖ "ﻓﺮﺍﯾﻨﺪ ﺍﻧﺘﺨﺎﺑﺎﺕ ﻋﻤﻮﻣﯽ" ﺍﺟﺮﺍ ﺷﻮﺩ. ﺍﺟﺎﺯﻩ ﺩﻫﯿﺪ )+ (ﮔﺮﻭﻩ ﺍﺯ

ﭼﺎﺭﭼﻮﺏﺟﺪﯾﺪﯼ ﺭﺍ ﺑﺮﺍﯼ ﺑﻬﺒﻮﺩ ﺭﻭﺵ ﺧﻮﺷﻪ ﺑﻨﺪﯼ DPC ﺍﺻﻠﯽ ﻣﻌﺮﻓﯽ ﺧﻮﺍﻫﯿﻢ ﮐﺮﺩ. ﺩﺭ

# (5)

ﻫﻤﺮﺍﻩﺑﺎ ﺁﻥﮎﻫﻤﺴﺎﯾﮕﺎﻥ ﻫﻤﺎﻧﻄﻮﺭ ﮐﻪ ﺩﺭ ﻣﻌﺎﺩﻟﻪ ﺗﻌﺮﯾﻒ ﺷﺪﻩ ﺍﺳﺖ(4).، ﻧﻘﻄﻪ ﺍﺻﻠﯽ

ﺍﻧﺘﺨﺎﺏﺷﺪﻩ ﺗﻮﺳﻂ ﺑﻪ ﻋﻨﻮﺍﻥ ... ﺗﻌﺮﯾﻒ ﺷﺪﻩ ﺍﺳﺖ

= ( )

( )+ ∋

ﻣﻘﺎﯾﺴﻪﺑﺎ ﺭﻭﺵ ﻫﺎﯼ ﭘﯿﺸﺮﻓﺘﻪ ﻣﺒﺘﻨﯽ ﺑﺮ DPC، ﻫﺪﻑ ﭼﺎﺭﭼﻮﺏ ﺟﺪﯾﺪ ﺍﺟﺘﻨﺎﺏ ﺍﺯ ﺳﻪ ﻧﻘﺺ

DPC ﺑﺪﻭﻥ ﭘﯿﭽﯿﺪﻩ ﺗﺮ ﮐﺮﺩﻥ ﺍﻟﮕﻮﺭﯾﺘﻢ ﺍﺳﺖ.

ﯾﻌﻨﯽﻧﻤﻮﻧﻪ ﺑﺎ ﺑﯿﺸﺘﺮﯾﻦ ﺗﺮﺍﮐﻢ ﺩﺭ )+ (ﺑﻪ ﻋﻨﻮﺍﻥ ﻧﻘﻄﻪ ﺍﺻﻠﯽ ﺩﺭ ﻧﻈﺮ ﮔﺮﻓﺘﻪ ﻣﯽ ﺷﻮﺩ . ﺭﻭﻧﺪ

.3ﺍﻟﮕﻮﺭﯾﺘﻢ ﭘﯿﺸﻨﻬﺎﺩﯼ

ﺍﻧﺘﺨﺎﺑﺎﺕﺑﺎ ﻣﺠﻤﻮﻋﻪ ﺍﯼ ﺍﺯ ﻧﮑﺎﺕ ﺍﺻﻠﯽ ﺑﻪ ﭘﺎﯾﺎﻥ ﻣﯽ ﺭﺳﺪ = } ) (1 ،) (2،… ،) .{(ﺗﻮﺟﻪ

ﺩﺍﺷﺘﻪﺑﺎﺷﯿﺪ ﮐﻪ ﯾﮏ ﻧﻘﻄﻪ ﺑﺎ ﺗﺮﺍﮐﻢ ﻣﺤﻠﯽ ﺑﻪ ﻃﻮﺭ ﮐﻠﯽ ﺑﺎﻻ ﻣﻤﮑﻦ ﺍﺳﺖ ﺗﻮﺳﻂ ﭼﻨﺪﯾﻦ ﺍﻧﺘﺨﺎﺏﮐﻨﻨﺪﻩ ﺑﻪ ﻋﻨﻮﺍﻥ ﻧﻘﻄﻪ ﺍﺻﻠﯽ ﺩﺭ ﻧﻈﺮ ﮔﺮﻓﺘﻪ ﺷﻮﺩ*CP*.ﺳﭙﺲ ﺑﺎ ﺣﺬﻑ ﻋﻨﺎﺻﺮ ﺗﮑﺮﺍﺭﯼ

ﺩﺭﺍﯾﻦ ﺑﺨﺶ، ﯾﮏ ﺭﻭﺵ ﺧﻮﺷﻪ ﺑﻨﺪﯼ ﻗﻮﯼ ﺑﻪ ﻧﺎﻡ ICKDC ﺑﺮﺍﯼ ﺭﻓﻊ ﺍﺷﮑﺎﻻﺕ DPC

ﭘﯿﺸﻨﻬﺎﺩﺷﺪﻩ ﺍﺳﺖ. ﺑﻪ ﻃﻮﺭ ﮐﻠﯽ، ICKDC ﺭﺍ ﻣﯽ ﺗﻮﺍﻥ ﺑﺎ ﭼﻬﺎﺭ ﻣﺮﺣﻠﻪ ﭘﯿﺎﺩﻩ ﺳﺎﺯﯼ ﮐﺮﺩ:ﯾﻌﻨﯽ

ﮐﻪﺑﻪ ﺻﻮﺭﺕ ﺑﯿﺎﻥ ﻣﯽ ﺷﻮﻧﺪ، ﺳﺎﺩﻩ ﻣﯽ ﺷﻮﺩ =- } 1 ،2،… ،

ﺗﺨﻤﯿﻦﭼﮕﺎﻟﯽ ﻣﺤﻠﯽ، ﺷﻨﺎﺳﺎﯾﯽ ﻧﻘﺎﻁ ﻣﺮﮐﺰﯼ، ﺍﺩﻏﺎﻡ ﻧﻘﺎﻁ ﻣﺮﮐﺰﯼ ﻭ ﻃﺒﻘﻪ ﺑﻨﺪﯼ ﺑﻘﯿﻪ ﻧﻘﺎﻁ.

{ﺟﺎﯾﯽ ﮐﻪ ﻋﺪﺩ ﺍﺳﺖ ﺗﻌﺪﺍﺩﯼﺍﺯ ﻧﮑﺎﺕ ﺍﺻﻠﯽ

ﺑﺮﺗﺮﯼﮐﻠﯿﺪﯼ ICKDC ﻧﺴﺒﺖ ﺑﻪ DPC ﺍﯾﻦ ﺍﺳﺖ ﮐﻪ (i) ﭼﮕﺎﻟﯽ ﻧﻘﻄﻪ i ﺩﻗﯿﻘﺎ ًﺑﺮ ﺍﺳﺎﺱ

ﻧﻘﺎﻁﻣﺮﮐﺰﯼ ﺩﺭ ﻧﻮﺍﺣﯽ ﻣﺘﺮﺍﮐﻢ ﺩﺭ ﻫﺮ ﺧﻮﺷﻪ ﻗﺮﺍﺭ ﺩﺍﺭﻧﺪ ﮐﻪ ﺗﻮﺯﯾﻊ ﭼﮕﺎﻟﯽ ﺧﻮﺷﻪ ﻫﺎ ﺭﺍ ﺩﺭ

ﺑﺮﻣﯽ ﮔﯿﺮﺩ. ﺑﻘﯿﻪ ﻧﻘﺎﻁ، ﺑﻪ ﻭﯾﮋﻩ ﻧﻘﺎﻁ ﻧﺰﺩﯾﮏ ﺑﻪ ﻟﺒﻪ ﺧﻮﺷﻪ ﻫﺎ ﺭﺍ ﻣﯽ ﺗﻮﺍﻥ ﺑﻪ ﺭﺍﺣﺘﯽ ﺩﺭ ﺁﺧﺮﯾﻦﻣﺮﺣﻠﻪ ﺑﺎ ﺍﻃﻼﻋﺎﺕ ﺗﻮﺯﯾﻊ ﻃﺒﻘﻪ ﺑﻨﺪﯼ ﮐﺮﺩ. ﻋﻼﻭﻩ ﺑﺮ ﺍﯾﻦ، ﻫﺮ ﻧﻘﻄﻪ ﺣﻖ ﺩﺍﺭﺩ ﻧﻘﻄﻪ ﻣﺮﮐﺰﯼﺧﻮﺩ ﺭﺍ ﺗﺤﺖ ﭼﻨﯿﻦ ﻣﮑﺎﻧﯿﺰﻣﯽ ﺍﻧﺘﺨﺎﺏ ﮐﻨﺪ ﮐﻪ ﺍﻣﮑﺎﻥ ﻭﺟﻮﺩ ﻧﻘﺎﻁ ﻫﺴﺘﻪ ﺍﯼ ﺭﺍ ﺩﺭ ﻓﻀﺎﯼﻧﻤﻮﻧﻪ ﻓﺮﺍﻫﻢ ﻣﯽ ﮐﻨﺪ.*D*. ﺍﯾﻦ ﺑﺮﺍﯼ ﻃﺒﻘﻪ ﺑﻨﺪﯼ ﻣﺠﻤﻮﻋﻪ ﻫﺎﯼ ﺩﺍﺩﻩ ﺑﺎ ﺗﻮﺯﯾﻊ ﭼﮕﺎﻟﯽ

ﻧﺎﻫﻤﮕﻦﺣﯿﺎﺗﯽ ﺍﺳﺖ ﺯﯾﺮﺍ ﺳﺎﺧﺘﺎﺭ ﺧﻮﺷﻪ ﻫﺎ ﺑﻪ ﺍﻧﺪﺍﺯﻩ ﮐﺎﻓﯽ ﺁﺷﮑﺎﺭ ﺷﺪﻩ ﺍﻧﺪ.

ﺍﻃﻼﻋﺎﺕﺩﺍﺩﻩ ﺷﺪﻩ ﺗﻮﺳﻂ ﻧﻤﻮﻧﻪ ﻫﺎﯼ ﻣﺠﺎﻭﺭ ﺁﻥ ﺗﺨﻤﯿﻦ ﺯﺩﻩ ﻣﯽ ﺷﻮﺩ، ﺑﻨﺎﺑﺮﺍﯾﻦ ﺗﺄﺛﯿﺮ ﺍﻧﺘﺨﺎﺏﻧﺎﻣﻨﺎﺳﺐ ﻓﺎﺻﻠﻪ ﺑﺮﺵ ﺩﺭ DPC ﺣﺬﻑ ﺷﺪﻩ ﺍﺳﺖ. (2) ﺗﻌﺪﺍﺩ ﻣﺮﺍﮐﺰ ﺧﻮﺷﻪ ﺍﯼ ﺑﻪ ﻃﻮﺭﺧﻮﺩﮐﺎﺭ ﺍﺯ ﻃﺮﯾﻖ ﯾﮏ ﻓﺮﺁﯾﻨﺪ ﯾﮑﭙﺎﺭﭼﻪ ﺳﺎﺯﯼ ﺟﺪﯾﺪ ﺗﺼﻤﯿﻢ ﮔﯿﺮﯼ ﻣﯽ ﺷﻮﺩ. (iii) ﻣﺮﺍﮐﺰ ﺧﻮﺷﻪ ﺍﯼﺑﻪ ﺟﺎﯼ ﯾﮏ ﻧﻘﻄﻪ ﻭﺍﺣﺪ ﺑﺮﺍﯼ ﺷﻨﺎﺳﺎﯾﯽ ﺳﺎﺧﺘﺎﺭ ﺧﻮﺷﻪ ﻫﺎ، ﻭ ﺩﺭ ﻧﺘﯿﺠﻪ ﮐﺎﻫﺶ

»ﻭﺍﮐﻨﺶﺯﻧﺠﯿﺮﻩ ﺍﯼ« ﻫﻨﮕﺎﻡ ﺗﺨﺼﯿﺺ ﺑﻘﯿﻪ ﻧﻘﺎﻁ، ﺑﺎ ﻧﻮﺍﺣﯽ ﺍﺻﻠﯽ ﻧﺸﺎﻥ ﺩﺍﺩﻩ ﻣﯽ ﺷﻮﻧﺪ.

ﻣﯿﺰ1

ﺟﺰﺉﯿﺎﺕﻣﺠﻤﻮﻋﻪ ﺩﺍﺩﻩ ﻫﺎ

.3.3ﺍﺩﻏﺎﻡ ﻧﻘﺎﻁ ﺍﺻﻠﯽ

ﻭﯾﮋﮔﯽﻫﺎﯼ

2

2

2

2

2

2

2

2

4

ﺩﺳﺘﻪﺑﻨﺪﯼ ﻫﺎ

7

15

2

2

2

6

31

15

3

ﺍﺷﯿﺎء

788

600

1480

944

240

399

3100

5000

150

ﻣﺠﻤﻮﻋﻪﺩﺍﺩﻩ ﻫﺎ

ﻣﺼﻨﻮﻋﯽ ﺗﺠﻤﻊ R15

ﺍﺳﺒﺎﺏﺑﺎﺯﯼ

ﻣﺎﺭﭘﯿﭻ ﺷﻌﻠﻪ ﺗﺮﮐﯿﺐ D31 S2

ﺩﻧﯿﺎﯼﻭﺍﻗﻌﯽ

ﻋﻨﺒﯿﻪ

ﺩﺭﺍﯾﻦ ﻣﺮﺣﻠﻪ، ﺳﺎﺧﺘﺎﺭ ﺍﺻﻠﯽ ﻣﺠﻤﻮﻋﻪ ﺩﺍﺩﻩ ﻫﺎ ﺑﺎ ﺍﺩﻏﺎﻡ ﻧﻘﺎﻁ ﺍﺻﻠﯽ ﻣﺘﻌﻠﻖ ﺑﻪ ﻫﻤﺎﻥ

ﺧﻮﺷﻪﺷﻨﺎﺳﺎﯾﯽ ﻣﯽ ﺷﻮﺩ. ﺑﻪ ﻃﻮﺭ ﮐﻠﯽ، ﻓﺮﺁﯾﻨﺪ ﻃﺒﻘﻪ ﺑﻨﺪﯼ KNN ﻓﺸﺮﺩﻩ ﺩﺭﮔﯿﺮ ﺍﺳﺖ، ﮐﻪ ﺍﮔﺮﺣﺪﺍﻗﻞ ﯾﮏ ﻧﻘﻄﻪ ﺩﺭ ﻫﻤﺴﺎﯾﮕﯽ ﺧﻮﺩ ﻣﺸﺘﺮﮎ ﺑﺎﺷﻨﺪ، ﺣﺮﯾﺼﺎﻧﻪ ﺩﻭ ﻧﻘﻄﻪ ﺍﺻﻠﯽ ﺭﺍ ﺩﺭ ﯾﮏ ﺧﻮﺷﻪﻃﺒﻘﻪ ﺑﻨﺪﯼ ﻣﯽ ﮐﻨﺪ. ﻣﻨﻄﻖ ﭘﺸﺖ ﭼﻨﯿﻦ ﺍﺳﺘﺮﺍﺗﮋﯼ ﺍﯾﻦ ﻓﺮﺽ ﺍﺳﺖ ﮐﻪ ﻓﻮﺍﺻﻞ ﺑﯿﻦ ﻧﻘﺎﻁﻫﺴﺘﻪ ﺍﺯ ﺧﻮﺷﻪ ﻫﺎﯼ ﻣﺨﺘﻠﻒ ﻧﺴﺒﺘﺎ ًﺩﻭﺭ ﺍﺯ ﻧﻘﺎﻁ ﺩﺭﻭﻥ ﯾﮏ ﺧﻮﺷﻪ ﺍﺳﺖ. ﻋﻼﻭﻩ ﺑﺮ ﺍﯾﻦ،ﺍﮔﺮ ﻣﺤﺪﻭﺩﻩ ﺁﻧﻬﺎ ﺑﻪ ﺧﻮﺑﯽ ﺗﻌﺮﯾﻒ ﺷﺪﻩ ﺑﺎﺷﺪ، ﻫﻤﺴﺎﯾﮕﯽ ﻧﻘﺎﻁ ﻫﺴﺘﻪ ﺍﺯ ﺧﻮﺷﻪ ﻫﺎﯼ

ﻣﺨﺘﻠﻒﺑﺎ ﻫﻢ ﻫﻤﭙﻮﺷﺎﻧﯽ ﻧﺪﺍﺭﻧﺪ.

7 3 210

ﺩﺍﻧﻪ

، ﺑﺮﺍﯼ ﻃﺒﻘﻪ ﺑﻨﺪﯼ ﻧﻘﺎﻁ ﺍﺻﻠﯽ ﻃﺮﺍﺣﯽ ﺷﺪﻩ

ﯾﮏﻓﺮﺁﯾﻨﺪ ﻓﺸﺮﺩﻩ KNN، ﮐﻪ ﺑﻪ ﻋﻨﻮﺍﻥ

13 3

7 8

30 2

34 6

34 2

23 2

178

336

569

366

351

195

ﺷﺮﺍﺏ ﺍﻟﯿﮑﻮ WDBC

ﭘﻮﺳﺖ

ﯾﻮﻧﻮﺷﯿﺮ

ﭘﺎﺭﮐﯿﻨﺴﻮﻥﻫﺎ

# (6)

ﺍﺳﺖ.ﺍﻭﻝ ﺍﺯ ﻫﻤﻪ، ﯾﮏ ﻋﺎﻣﻞ ﺗﻌﺪﯾﻞ ﺑﺮﺍﯼ ﮐﻨﺘﺮﻝ ﻣﺤﺪﻭﺩﻩ ﻫﻤﺴﺎﯾﮕﯽ ﺩﺭﮔﯿﺮ ﺍﺳﺖ ﮐﻪ ﺑﻪ

ﺻﻮﺭﺕﺗﻌﺮﯾﻒ ﺷﺪﻩ ﺍﺳﺖ

 ~~(1~~.0+ ) ⋅2=

19 7

210

ﺗﻘﺴﯿﻢﺑﻨﺪﯼ

ﺳﭙﺲ، ﻫﻤﺴﺎﯾﮕﺎﻥ ﻫﺮ ﻧﻘﻄﻪ ﺍﺻﻠﯽ، ﺟﺎﯾﯽ ﮐﻪ = ⋅ ، ﺍﺯ ﻃﺮﯾﻖ ﯾﮏ ﻓﺮﺁﯾﻨﺪ ﺟﺴﺘﺠﻮﯼ

ﺍﻓﺰﺍﯾﺸﯽﻋﻼﻣﺖ ﮔﺬﺍﺭﯼ ﻣﯽ ﺷﻮﻧﺪ. ﺗﻌﺪﺍﺩ ﻫﻤﺴﺎﯾﮕﺎﻥ ﺭﺍ ﻣﯽ ﺗﻮﺍﻥ ﺑﻪ ﻃﻮﺭ ﺗﻄﺒﯿﻘﯽ ﺑﺎ

.3.5ﺗﺤﻠﯿﻞ ﭘﯿﭽﯿﺪﮔﯽ ﺯﻣﺎﻧﯽ

ﺗﻮﺟﻪﺑﻪ ﻣﻘﺪﺍﺭ ﺗﻨﻈﯿﻢ ﮐﺮﺩﮎ. ﭼﻪ ﺯﻣﺎﻧﯽﮎﺑﺰﺭﮒ ﺍﺳﺖ، ﺑﻪ ﺍﻧﺪﺍﺯﻩ*CP*ﮐﻮﭼﮏ ﺧﻮﺍﻫﺪ ﺑﻮﺩ ﺯﯾﺮﺍ ﯾﮏﻧﻘﻄﻪ ﺑﺎ ﭼﮕﺎﻟﯽ ﻣﺤﻠﯽ ﺯﯾﺎﺩ ﺍﯾﻦ ﺍﻣﮑﺎﻥ ﺭﺍ ﺩﺍﺭﺩ ﮐﻪ ﻧﻘﺎﻁ ﺍﺻﻠﯽ ﻧﻤﻮﻧﻪ ﻫﺎﯼ ﺑﯿﺸﺘﺮﯼ ﺑﺎﺷﺪ

ﮐﻪﻣﻨﺠﺮ ﺑﻪ ﮐﺎﻫﺶ ﺗﻌﺪﺍﺩ ﻧﻤﻮﻧﻪ ﻫﺎ ﻣﯽ ﺷﻮﺩ. ﻫﻤﭽﻨﯿﻦ . ﺩﺭ ﻧﺘﯿﺠﻪ، ﺍﺯ ﻃﺮﯾﻖ ﺿﺮﺏ

ﭘﯿﭽﯿﺪﮔﯽﺯﻣﺎﻧﯽ ICKDC ﺍﺯ ﭼﻬﺎﺭ ﺑﺨﺶ ﺯﯾﺮ ﺗﺸﮑﯿﻞ ﺷﺪﻩ ﺍﺳﺖ.

ﺗﺜﺒﯿﺖﻣﯽ ﺷﻮﺩﮎﺗﻮﺳﻂ . ﻫﻤﺎﻥ ﻧﺘﯿﺠﻪ ﺯﻣﺎﻧﯽ ﮐﻪﮎﻧﺴﺒﺘﺎ ﮐﻮﭼﮏ ﺍﺳﺖ. ﺑﻪ ﻃﻮﺭ ﮐﻠﯽ،

(1) ﺩﺭ ﺗﺨﻤﯿﻦ ﭼﮕﺎﻟﯽ ﻣﺤﻠﯽ، ﭘﯿﭽﯿﺪﮔﯽ ﺯﻣﺎﻧﯽ ﺑﺮﺍﯼ ﺳﺎﺧﺖ ﻣﺎﺗﺮﯾﺲ ﻓﺎﺻﻠﻪ ﻭ ﺍﻧﺘﺨﺎﺏ

ﻣﻌﺮﻓﯽ ﺑﻪ ﺻﻮﺭﺕ ﺗﻄﺒﯿﻘﯽ ﻣﺤﺪﻭﺩﻩ ﻣﺤﻠﻪ ﻫﺎ ﺭﺍ ﺑﻬﯿﻨﻪ ﮐﺮﺩ، ﮐﻪ ﺩﻗﺖ ﻃﺒﻘﻪ ﺑﻨﺪﯼ ﻧﻘﺎﻁ

ﻧﺰﺩﯾﮑﺘﺮﯾﻦﻫﻤﺴﺎﯾﻪ ﻫﺎ ﻫﺮ ﺩﻭ ﻫﺴﺘﻨﺪ ) .(2 ﺗﺨﻤﯿﻦ ﭼﮕﺎﻟﯽ ﻃﻮﻝ ﻣﯽ ﮐﺸﺪ ) .( (2) ﭘﯿﭽﯿﺪﮔﯽ

ﺯﻣﺎﻧﯽﺑﺮﺍﯼ ﺍﻧﺘﺨﺎﺏ ﻧﻘﺎﻁ ﺍﺻﻠﯽ ﺍﺳﺖ ) .( (3) ﺍﺩﻏﺎﻡ ﻧﻘﺎﻁ ﺍﺻﻠﯽ ﻃﻮﻝ ﻣﯽ ﮐﺸﺪ ) ( ∗ ] ) (

ﺍﺻﻠﯽﺭﺍ ﺑﻬﺒﻮﺩ ﻣﯽ ﺑﺨﺸﺪ. ﺩﺭ ﻧﻬﺎﯾﺖ، ﻧﻘﺎﻁ ﺍﺻﻠﯽ ﺑﻪ ﺍﺷﺘﺮﺍﮎ ﮔﺬﺍﺭﯼ ﯾﮏ ﯾﺎ ﭼﻨﺪ ﻧﻘﻄﻪ ﺩﺭ ﺧﻮﺩ ﻣﺤﻠﻪ ﻫﺎ ﺑﻪ ﯾﮏ ﺧﻮﺷﻪ ﻃﺒﻘﻪ ﺑﻨﺪﯼ ﻣﯽ ﺷﻮﻧﺪ. ﮔﺮﻭﻩ ﻫﺎﯼ ﺍﻭﻟﯿﻪ ﺗﺸﮑﯿﻞ ﺷﺪﻩ ﺩﺭ ﺍﯾﻦ

+ ) ∗ [( = ) ( ∗ ) ( + ) ( ∗ ) (، ﺟﺎﯾﯽ ﮐﻪ ﺗﻌﺪﺍﺩ ﺧﻮﺷﻪ ﻫﺎ ﻭ ﻣﯿﺎﻧﮕﯿﻦ ﺗﻌﺪﺍﺩ ﻧﻘﺎﻁ ﺩﺭ

ﻣﺮﺣﻠﻪ،ﻧﻮﺍﺣﯽ ﻣﺮﮐﺰﯼ ﻫﺮ ﺧﻮﺷﻪ ﻧﺎﻣﯿﺪﻩ ﻣﯽ ﺷﻮﻧﺪ.

ﻫﺮﺧﻮﺷﻪ ﺍﺳﺖ. ﺑﺎ ﺗﻮﺟﻪ ﺑﻪ ﺍﯾﻨﮑﻪ ﻭ ﻫﺮ ﺩﻭ ﻣﺤﺪﻭﺩ ﻫﺴﺘﻨﺪ ، ﭘﯿﭽﯿﺪﮔﯽ ﺯﻣﺎﻧﯽ ﺩﺭ ﺍﯾﻦ

*>*( ) .( ﺑﻪ ﻃﻮﺭ

ﻣﺮﺣﻠﻪﺍﺳﺖ ) .(2 (4) ﭘﯿﭽﯿﺪﮔﯽ ﺯﻣﺎﻧﯽ ﺑﺮﺍﯼ ﺗﺨﺼﯿﺺ ﺑﻘﯿﻪ ﻧﻘﺎﻁ ﺍﺳﺖ ) -

ﺧﻼﺻﻪ،ﭘﯿﭽﯿﺪﮔﯽ ﮐﻠﯽ ﺯﻣﺎﻧﯽ ICKDC ﺍﺳﺖ ) (2 ﮐﻪ ﻫﻤﺎﻥ DPC ﺍﺳﺖ.

.4ﺁﺯﻣﺎﯾﺶ ﻫﺎ ﻭ ﺗﺠﺰﯾﻪ ﻭ ﺗﺤﻠﯿﻞ

.3.4ﺍﻟﮕﻮﺭﯾﺘﻢ *ICKDC*

ﻓﺮﺁﯾﻨﺪﺷﻨﺎﺳﺎﯾﯽ ﻧﻮﺍﺣﯽ ﺍﺻﻠﯽ ﺗﻮﺯﯾﻊ ﻣﺠﻤﻮﻋﻪ ﺩﺍﺩﻩ ﻫﺎ ﺭﺍ ﺑﻪ ﺷﺪﺕ ﺁﺷﮑﺎﺭ ﻣﯽ ﮐﻨﺪ. ﺑﻘﯿﻪ ﻧﻘﺎﻁﺭﺍ ﻣﯽ ﺗﻮﺍﻥ ﺩﻗﯿﻘﺎ ًﺑﻪ ﺧﻮﺷﻪ ﻫﺎﯾﯽ ﺍﺧﺘﺼﺎﺹ ﺩﺍﺩ ﮐﻪ ﻧﻘﺎﻁ ﺍﺻﻠﯽ ﺁﻧﻬﺎ ﺑﻪ ﺁﻧﻬﺎ ﺗﻌﻠﻖ

ﺩﺭﺍﯾﻦ ﺑﺨﺶ، ﻋﻤﻠﮑﺮﺩ ICKDC ﺑﺎ ﺍﺳﺘﻔﺎﺩﻩ ﺍﺯ ﻣﻮﺍﺩ ﻣﺼﻨﻮﻋﯽ ﮔﺴﺘﺮﺩﻩ ﺁﺯﻣﺎﯾﺶ ﻭ ﺍﺭﺯﯾﺎﺑﯽ ﻣﯽ ﺷﻮﺩ.ﻓﺮﺍﻧﺘﯽ ﻭ ﺳﯿﺮﺍﻧﻮﺟﺎ،(2018 ﻭ ﻣﺠﻤﻮﻋﻪ ﺩﺍﺩﻩ ﻫﺎﯼ ﺩﻧﯿﺎﯼ ﻭﺍﻗﻌﯽ )ﺩﻋﺎ ﻭ ﮔﺮﺍﻑ،2017

( ﺑﺎ ﻣﺸﺨﺼﺎﺕ ﺗﻮﺯﯾﻊ ﻣﺨﺘﻠﻒ. ﺟﺰﺉﯿﺎﺕ ﺍﯾﻦ ﻣﺠﻤﻮﻋﻪ ﺩﺍﺩﻩ ﻫﺎ ﺩﺭ ﺍﯾﻨﺠﺎ ﺁﻭﺭﺩﻩ ﺷﺪﻩ ﺍﺳﺖ ﻣﯿﺰ.1 ﻧﺘﺎﯾﺞ ﺧﻮﺷﻪ ﺑﻨﺪﯼ ICKDC ﺑﺎ ﻧﺘﺎﯾﺞ ﺍﺭﺍﺉﻪ ﺷﺪﻩ ﺗﻮﺳﻂ ﭼﻨﺪﯾﻦ ﺭﻭﺵ ﺧﻮﺷﻪ ﺑﻨﺪﯼ ﻣﻌﺮﻭﻑ،ﺍﺯ ﺟﻤﻠﻪ DPC )ﺭﻭﺩﺭﯾﮕﺰ ﻭ ﻻﯾﻮ،)2014 DBSCAN ،ﺍﺳﺘﺮ ﻭ ﻫﻤﮑﺎﺭﺍﻥ،)1996 ﺑﻪ ﻣﻌﻨﯽK- ،(ﻣﮏ ﮐﻮﺉﯿﻦ ﻭ ﻫﻤﮑﺎﺭﺍﻥ،)1967 Meanshift (,ﻓﻮﮐﻮﻧﺎﮔﺎ ﻭ ﻫﺎﺳﺘﻠﺮ،(1975 ﻭ

AGNES )ﻓﺮﺍﻟﯽ ﻭ ﺭﺍﻓﺘﺮﯼ، .(1998 ﺳﭙﺲ، ﻣﻘﺎﯾﺴﻪ ﺑﯿﻦ ICKDC ﻭ ﺳﻪ ﻧﻮﻉ DPC ﭘﯿﺸﺮﻓﺘﻪ

ﺭﺍﻧﺸﺎﻥ ﻣﯽ ﺩﻫﯿﻢ،ﯾﻌﻨﯽ) DPC-KNNﺩﻭ ﻭ ﻫﻤﮑﺎﺭﺍﻥ،)2016 FKNN-DPC ،(ﺯﯼ ﻭ ﻫﻤﮑﺎﺭﺍﻥ، (2016 ﻭ IVDPC )ﮊﻭ ﻭ ﻫﻤﮑﺎﺭﺍﻥ،.(2018 ﺩﺭ ﭘﺎﯾﺎﻥ، ﻣﺎ ﺗﺤﻘﯿﻘﺎﺕ ﺑﯿﺸﺘﺮﯼ ﺭﺍ ﺩﺭ ﻣﻘﺎﯾﺴﻪ ICKDC ﻭ ﮐﺎﺭ ﺍﺧﯿﺮ ﺍﻧﺠﺎﻡ ﻣﯽ ﺩﻫﯿﻢ )ﺑﻮﺭﺗﻮﻟﻮﺗﯽ، ﺩ ﺍﻭﻟﯿﻮﯾﺮﺍ ﻭ ﮐﯿﺎﺭﻟﯽ،(2021 ﺩﺭ ﺧﻮﺷﻪ ﺑﻨﺪﯼ

ﯾﺎﺩﮔﯿﺮﯼﻧﻈﺎﺭﺕ ﺷﺪﻩ ﮐﻪ ﺑﺎ SKDEKMeans ﻭ ISSKDEKMeans ﻣﺸﺨﺺ ﻣﯽ ﺷﻮﺩ.

ﺭﻭﺵﻫﺎﯼ ﺩﻭ ﻣﻮﺭﺩ ﺍﻭﻝ ﺑﺮ ﺍﺳﺎﺱ ﮐﺪ ﻣﻨﺒﻊ ﺍﺭﺍﺉﻪ ﺷﺪﻩ ﺗﻮﺳﻂ ﻧﻮﯾﺴﻨﺪﮔﺎﻥ ﺩﺭ ﻣﺘﻠﺐ ﭘﯿﺎﺩﻩ ﺳﺎﺯﯼﺷﺪﻩ ﺍﺳﺖ. ﻣﺎ ﺑﻪ ﺳﺎﺩﮔﯽ ﻧﺘﺎﯾﺞ ﺍﺭﺍﺉﻪ ﺷﺪﻩ ﺗﻮﺳﻂ ﻧﻮﯾﺴﻨﺪﮔﺎﻥ ﺭﺍ ﺍﻋﻤﺎﻝ ﻣﯽ ﮐﻨﯿﻢ

ﺑﻮﺭﺗﻮﻟﻮﺗﯽﻭ ﻫﻤﮑﺎﺭﺍﻥ(2021) ﺩﺭ ﻣﻮﺭﺩ ﺳﻮﻡ.

ﺩﺍﺭﺩ.ﺷﺒﻪ ﮐﺪ ICKDC ﺩﺭ ﺗﻌﻤﯿﻢ ﯾﺎﻓﺘﻪ ﺍﺳﺖﺍﻟﮕﻮﺭﯾﺘﻢ .1

[ﺍﻟﮕﻮﺭﯾﺘﻢ](#_bookmark53)ICKDC [. : ﻣﺠﻤﻮﻋﻪ](#_bookmark53) [ﺩﺍﺩﻩ](#_bookmark51) ﺍﯼ ﺣﺎﻭﯼ*n*ﻧﻤﻮﻧﻪ [ﻫﺎ](#_bookmark39)

[ﻭﺭﻭﺩﯼ](#_bookmark53): = } 1[،](#_bookmark53) 2[،](#_bookmark53)… {

[ﮎ:](#_bookmark41) [ﺗﻌﺪﺍﺩ](#_bookmark51) [ﻧﻘﺎﻁ](#_bookmark41) [ﻣﻮﺟﻮﺩ ﺩﺭ ﯾﮏ](#_bookmark51) ﻣﺤﻠﻪ [ﺧﺮﻭﺟﯽ](#_bookmark43): [ﻧﺘﯿﺠﻪ ﺧﻮﺷﻪ](#_bookmark43) [ﺑﻨﺪﯼ](#_bookmark39) [ﺍﺯ](#_bookmark43)*D*

%ﺗﺨﻤﯿﻦ [ﭼﮕﺎﻟﯽ ﻣﺤﻠﯽ](#_bookmark36)

ﻣﺎﺗﺮﯾﺲﻓﺎﺻﻠﻪ [ﺭﺍ](#_bookmark66) ﻣﺤﺎﺳﺒﻪ [ﮐﻨﯿﺪ](#_bookmark66) = ] [ ×

ﺗﺨﻤﯿﻦﭼﮕﺎﻟﯽ ﺗﻮﺳﻂ ﻣﻌﺎﺩﻟﻪ 4 ﺩﺭﺻﺪ

ﺷﻨﺎﺳﺎﯾﯽﻧﻘﺎﻁ ﺍﺻﻠﯽ

ﻧﻘﺎﻁﺍﺻﻠﯽ ﺭﺍ ﺑﺮﺍﯼ ﺗﻮﺳﻂ ﻣﻌﺎﺩﻟﻪ 5

ﺩﺭﺻﺪﺍﺩﻏﺎﻡ ﻧﻘﺎﻁ ﺍﺻﻠﯽ

ﺩﺭﺣﺎﻟﯽ ﮐﻪﺗﻤﺎﻡ ﻧﻘﺎﻁ ﺍﺻﻠﯽ ﻃﺒﻘﻪ ﺑﻨﺪﯼ ﻧﻤﯽ ﺷﻮﻧﺪﺍﻧﺠﺎﻡ ﺩﺍﺩﻥ

.4.1ﻣﻘﺪﻣﺎﺗﯽ

ﺑﺪﻭﻥ ﺑﺮﭼﺴﺐ ﺍﺯ -ﻭ ﺁﻥ ﺭﺍ ﺑﻪ

ﺗﻌﺪﺍﺩﺧﻮﺷﻪ ﻫﺎ (noc) ﺭﺍ 1 ﺍﺿﺎﻓﻪ ﮐﻨﯿﺪ ﺑﻪﻃﻮﺭ ﺗﺼﺎﺩﻓﯽ ﯾﮏ ﻧﻘﻄﻪ ﺍﺻﻠﯽ ﺭﺍ ﺍﻧﺘﺨﺎﺏ ﮐﻨﯿﺪ

ﺧﻮﺷﻪﺟﺪﯾﺪ ﺍﺧﺘﺼﺎﺹ ﺩﻫﯿﺪ

ﺳﻪﻣﻌﯿﺎﺭ ﺍﺭﺯﯾﺎﺑﯽ ﺩﺭ ﻣﻘﺎﯾﺴﻪ ﻣﻌﺮﻓﯽ ﺷﺪﻩ ﺍﻧﺪ، ﯾﻌﻨﯽ ﺷﺎﺧﺺ ﺭﻧﺪ ﺗﻌﺪﯾﻞ ﺷﺪﻩ (ARI)، ﺍﻃﻼﻋﺎﺕﻣﺘﻘﺎﺑﻞ ﺗﻨﻈﯿﻢ ﺷﺪﻩ (NMI)، ﻭ ﺷﺎﺧﺺ ﻓﺎﻭﻟﮑﺲ-ﻣﺎﻟﻮﺯ .(FMI) ﮐﺮﺍﻥ ﺑﺎﻻﯼ ﻫﻤﻪ

ﻣﻌﯿﺎﺭﻫﺎ1 ﺍﺳﺖ، ﺑﻪ ﻋﺒﺎﺭﺕ ﺩﯾﮕﺮ، ﻣﻘﺎﺩﯾﺮ ﺑﺰﺭﮔﺘﺮ ﺷﺎﺧﺺ ﻫﺎ ﻧﺸﺎﻥ ﺩﻫﻨﺪﻩ ﻧﺘﺎﯾﺞ ﺑﻬﺘﺮ ﺧﻮﺷﻪ

⋂ ∅≠ﺍﻧﺠﺎﻡ ﺩﺍﺩﻥ

ﺑﺮﺍﯼﻫﺮ ﯾﮏ ∋ -ﺍﻧﺠﺎﻡ ﺩﺍﺩﻥ

ﺍﮔﺮ ﻃﺒﻘﻪ ﺑﻨﺪﯼ ﻧﺸﺪﻩ ﺍﺳﺖ ﻭ

ﺍﺧﺘﺼﺎﺹﺩﻫﺪ ﺑﻪ

ﺑﻨﺪﯼﺍﺳﺖ. ﻗﺒﻞ ﺍﺯ ﺁﺯﻣﺎﯾﺶ، ﺗﻤﺎﻡ ﻣﺠﻤﻮﻋﻪ ﺩﺍﺩﻩ ﻫﺎ ﺑﺎ ﻧﺮﻣﺎﻝ ﺳﺎﺯﯼ ﺣﺪﺍﻗﻞ ﺗﺎ ﺣﺪﺍﮐﺜﺮ

ﺑﺮﺍﯼﻫﺮ ﯾﮏ ∋ ﺍﻧﺠﺎﻡ ﺩﺍﺩﻥ

ﺗﻨﻈﯿﻢﻣﯽ ﺷﻮﻧﺪ ﺗﺎ ﺗﺄﺛﯿﺮ ﺩﺍﻣﻨﻪ ﻫﺎﯼ ﻋﺪﺩﯼ ﻣﺨﺘﻠﻒ ﺭﺍ ﺍﺯ ﺑﯿﻦ ﺑﺒﺮﻧﺪ.

⋂ ∅≠ﺍﻧﺠﺎﻡ ﺩﺍﺩﻥ

ﺑﺮﺍﯼﻫﺮ ﯾﮏ ∋ -ﺍﻧﺠﺎﻡ ﺩﺍﺩﻥ

ﺍﮔﺮ ﻃﺒﻘﻪ ﺑﻨﺪﯼ ﻧﺸﺪﻩ ﺍﺳﺖ ﻭ

ﺍﺧﺘﺼﺎﺹﺩﻫﺪ ﺑﻪ

%ﺗﺨﺼﯿﺺ ﺑﻘﯿﻪ ﻧﻘﺎﻁ

# (7)

 ( ⋅ ) -

= "

ﻣﺘﻌﻠﻖ ﺑﻪ

ﺑﺮﺍﯼﻫﺮ ﺑﻘﯿﻪ ﻧﻘﻄﻪ ﺍﻧﺠﺎﻡ ﺩﺍﺩﻥ

( ⋅ ) - ( ⋅ )

ﺍﺧﺘﺼﺎﺹﺩﻫﺪ ﺑﻪ ﺧﻮﺷﻪ ﻫﺎﯾﯽ ﮐﻪ ﺁﻧﻬﺎ ) (

ﺗﻨﻈﯿﻤﺎﺕﭘﺎﺭﺍﻣﺘﺮ ﺩﺭ ﻣﺠﻤﻮﻋﻪ ﺩﺍﺩﻩ ﻫﺎﯼ ﻣﺼﻨﻮﻋﯽ

ﻣﺠﻤﻮﻋﻪﺩﺍﺩﻩ ﻫﺎ )ﻣﺼﻨﻮﻋﯽ(

ﭘﺎﺭ

ﺭﻭﺵ

1.5

30/0.04

15

0.11

15

0.6

38/0.04

31

0.071

31

1

1/0.05

6

0.22

6

1

9/0.10

2

0.25

2

2

2/0.10

2

0.25

2

2

20/0.07

2

0.135

2

2

14/0.04

15

0.09

15

1

11/0.06

7

0.155

7

/*MinPts*

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| S2 | D31 | ﺗﺮﮐﯿﺐ | ﺷﻌﻠﻪ | ﻣﺎﺭﭘﯿﭻ | ﺍﺳﺒﺎﺏﺑﺎﺯﯼ | R15 | ﺗﺠﻤﻊ |  |
| 1.1 | 1.2 | 1 | 11 | 1 | 3.5 | 2.1 | 4 | ICKDC |

*ℎ*

DPC

DBSCAN

ﮎ-ﯾﻌﻨﯽ ﺗﻐﯿﯿﺮﻣﻌﻨﯽ AGNES

ﺟﺪﻭﻝ3

ﺗﻨﻈﯿﻤﺎﺕﭘﺎﺭﺍﻣﺘﺮ ﺩﺭ ﻣﺠﻤﻮﻋﻪ ﺩﺍﺩﻩ ﻫﺎﯼ ﺩﻧﯿﺎﯼ ﻭﺍﻗﻌﯽ.

ﻣﺠﻤﻮﻋﻪﺩﺍﺩﻩ ﻫﺎ )ﺩﻧﯿﺎﯼ ﻭﺍﻗﻌﯽ(

ﭘﺎﺭ

ﺭﻭﺵ

7

0.638

7

1.5

7.4

27

2

–

2

0.62

2

1.2

2.7

7

2

–

2

1.31

2

0.5

4.8

26

1

–

6

1.487

6

1

3.5

35

0.5

–

2

1.2

2

1

1

9

0.2

–

8

0.335

8

2

2

9

2

–

3

0.665

3

2

11

9

6

–

3

0.46

3

1

2.5

9

2

–

3

0.27

3

1

15

22

1

–

ﮎ-ﯾﻌﻨﯽ

ﺗﻐﯿﯿﺮﻣﻌﻨﯽ *ℎ*

AGNES

ﻭﺍﻧﻮﺍﻉ ﺁﻥ *DPC* DPC

ICKDC FKNN-DPC DPC-KNN

– IVDPC

.ﺷﺎﻣﻞ15 ﺧﻮﺷﻪ ﻣﺤﺪﺏ ﺷﮑﻞ ﺑﺎ ﺗﻮﺯﯾﻊ ﮔﺎﻭﺳﯽ ﺍﺳﺖ R15 ﺷﮑﻞ 4ﻧﺸﺎﻥ ﻣﯽ ﺩﻫﺪ ﮐﻪ

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| ﺗﻘﺴﯿﻢﺑﻨﺪﯼ | ﭘﺎﺭﮐﯿﻨﺴﻮﻥﻫﺎ | ﯾﻮﻧﻮﺷﯿﺮ | ﭘﻮﺳﺖ | WDBC | ﺍﻟﯿﮑﻮ | ﺷﺮﺍﺏ | ﺩﺍﻧﻪ | ﻋﻨﺒﯿﻪ |  |
| 2/0.15 | 17/0.5 | 9/0.78 | 4/0.95 | 22/0.41 | 4/0.11 | 18/0.48 | 16/0.24 | 5/0.12 | ﺭﻭﺵﻫﺎﯼ ﺳﻨﺘﯽ/ DBSCAN |

ﺗﻤﺎﻡﺭﻭﺵ ﻫﺎ ﺑﻪ ﺟﺰ K-means ﻭ AGNES ﺑﺎ ﻣﻮﻓﻘﯿﺖ ﺳﺎﺧﺘﺎﺭ ﺧﻮﺷﻪ ﻫﺎ ﺭﺍ ﺗﺸﺨﯿﺺ ﻣﯽ

) ⋅ ( ﻣﻘﺎﺩﯾﺮ ﺷﺪﯾﺪ ﻫﻤﻪ ﻧﻤﻮﻧﻪ ﻫﺎ ﺩﺭ

ﻭ( ⋅ )

ﺟﺎﯾﯽﮐﻪ ﻣﻘﺪﺍﺭ ﻧﻤﻮﻧﻪ ﺍﺳﺖﻣﻦﮐﻪ ﺩﺭ ﺑﻌﺪ ﺍﻡ،

ﻫﺴﺘﻨﺪ ﺑﻌﺪ ﺍﻡ

ﺩﻫﻨﺪICKDC. ﻭ DPC ﺩﺭ R15 ﻋﻤﻠﮑﺮﺩ ﯾﮑﺴﺎﻧﯽ ﺩﺍﺭﻧﺪ، ﺍﻣﺎ 10 ﻧﻘﻄﻪ ﺍﺻﻠﯽ ﺑﻪ ﻋﻨﻮﺍﻥ ﻧﻮﯾﺰ

ﺗﻮﺳﻂDPC ﻣﺸﺨﺺ ﻣﯽ ﺷﻮﻧﺪ.

ﺍﺳﺒﺎﺏﺑﺎﺯﯼ ﺗﻮﺳﻂ ﺟﻌﺒﻪ ﺍﺑﺰﺍﺭ scikit-learn ﺑﺎ ﺗﻮﺯﯾﻊ ﻣﻨﯿﻔﻮﻟﺪ ﺗﻮﻟﯿﺪ ﻣﯽ ﺷﻮﺩ. ﺍﺯ ﺟﺎﻧﺐ

ﺷﮑﻞ5ﻣﯽ ﺑﯿﻨﯿﻢ ﮐﻪ ﻣﺠﻤﻮﻋﻪ ﺩﺍﺩﻩ ﺷﺎﻣﻞ ﺗﻮﺯﯾﻊ ﭼﮕﺎﻟﯽ ﻧﺎﻫﻤﮕﻦ ﺩﺭ ﯾﮏ ﺧﻮﺷﻪ ﺍﺳﺖ.

DBSCAN ،DPC ،ICKDC ﻭ Meanshift ﻣﻌﻤﻮﻻ ًﻣﯽ ﺗﻮﺍﻧﻨﺪ ﺧﻮﺷﻪ ﺑﻨﺪﯼ ﺭﺍ ﺍﻧﺠﺎﻡ ﺩﻫﻨﺪ ﺩﺭﺣﺎﻟﯽ ﮐﻪ K-means ﻧﻤﯽ ﺗﻮﺍﻧﺪ. ﺑﺎ ﺍﯾﻦ ﺣﺎﻝ، ﻧﻘﺎﻁ ﺩﺍﺩﻩ ﻧﺰﺩﯾﮏ ﺑﻪ ﺩﻡ ﺧﻮﺷﻪ ﻫﺎ ﺗﻮﺳﻂ DBSCAN ،DPC ﻭ Meanshift ﺑﻪ ﺍﺷﺘﺒﺎﻩ ﻃﺒﻘﻪ ﺑﻨﺪﯼ ﻣﯽ ﺷﻮﻧﺪ. ﺍﺳﺎﺳﺎ، ﺍﺳﺒﺎﺏ ﺑﺎﺯﯼ ﺷﺎﻣﻞﭼﻨﺪﯾﻦ ﻧﺎﺣﯿﻪ ﻣﺘﺮﺍﮐﻢ ﺩﺭ ﯾﮏ ﺧﻮﺷﻪ ﺍﺳﺖ ﮐﻪ ﺗﻮﺳﻂ ﻣﻨﺎﻃﻖ ﭘﺮﺍﮐﻨﺪﻩ ﺍﺯ ﻫﻢ ﺟﺪﺍ ﺷﺪﻩ ﺍﻧﺪ.ﻓﺮﺁﯾﻨﺪ ﻣﻨﺤﺼﺮ ﺑﻪ ﻓﺮﺩ ﺷﻨﺎﺳﺎﯾﯽ ﻭ ﺍﺩﻏﺎﻡ ﻧﻘﺎﻁ ﺍﺻﻠﯽ ﺑﻪ ICKDC ﺍﺟﺎﺯﻩ ﻣﯽ ﺩﻫﺪ ﺗﺎ ﺍﯾﻦ

ﻣﻨﺎﻃﻖﻣﺘﺮﺍﮐﻢ ﺭﺍ ﺍﺩﻏﺎﻡ ﮐﻨﺪ ﻭ ﺩﺭ ﻧﺘﯿﺠﻪ ﻧﺘﺎﯾﺞ ﺧﻮﺷﻪ ﺑﻨﺪﯼ ﺛﺎﺑﺘﯽ ﺭﺍ ﺍﺭﺍﺉﻪ ﺩﻫﺪ. ﺩﺭ ﻧﺘﯿﺠﻪ،

ICKDC ﺑﻬﺘﺮﯾﻦ AMI ،ARI ﻭ FMI ﺭﺍ ﺩﺭ ﻣﻘﺎﯾﺴﻪ ﺑﺎ ﺗﻤﺎﻡ ﺭﻗﺒﺎ ﺍﺭﺍﺉﻪ ﻣﯽ ﺩﻫﺪ.

ﻣﺎﺭﭘﯿﭻﺷﺎﻣﻞ 2 ﺧﻮﺷﻪ ﻣﺎﺭﭘﯿﭽﯽ ﺷﮑﻞ ﺍﺳﺖ.ﺷﮑﻞ 6ﻧﺸﺎﻥ ﻣﯽ ﺩﻫﺪ ﮐﻪ ﻫﻤﻪ ﺭﻭﺵ ﻫﺎ

ﺑﻪﺟﺰ K-means ﻭ Meanshift [ﻣﯽ](#_bookmark11) ﺗﻮﺍﻧﻨﺪ ﻧﺘﺎﯾﺞ ﺻﺤﯿﺢ ﺭﺍ ﺑﻪ ﺩﺳﺖ ﺁﻭﺭﻧﺪ. ﺑﺎ ﺍﯾﻦ ﺣﺎﻝ، ﻓﺮﻣﻮﻝﺗﺨﻤﯿﻦ ﭼﮕﺎﻟﯽ ﺩﺭ DPC ﺑﺎﯾﺪ ﺑﺎ ﻓﺮﻡ [ﮔﺎﻭﺳﯽ ﻗﺒﻞ](#_bookmark10) ﺍﺯ ﺍﻋﻤﺎﻝ ﺟﺎﯾﮕﺰﯾﻦ ﺷﻮﺩ. ﻫﻤﺎﻧﻄﻮﺭ ﮐﻪﺩﺭ ﺑﺨﺶ ﺫﮐﺮ ﺷﺪ1، ﻣﻌﯿﺎﺭﻫﺎﯼ ﭼﮕﺎﻟﯽ ﻣﺨﺘﻠﻒ ﺑﻪ ﺻﻮﺭﺕ ﺩﺳﺘﯽ ﺑﺎ ﺗﻮﺟﻪ ﺑﻪ ﺍﻧﺪﺍﺯﻩ

ﻣﺠﻤﻮﻋﻪﺩﺍﺩﻩ ﻫﺎ ﺩﺭ DPC ﺍﻧﺘﺨﺎﺏ ﻣﯽ ﺷﻮﻧﺪ ﮐﻪ ﺍﺳﺘﺤﮑﺎﻡ ﺭﻭﺵ ﺭﺍ ﮐﺎﻫﺶ ﻣﯽ ﺩﻫﺪ.

ﺷﻌﻠﻪﺑﺎ ﭼﮕﺎﻟﯽ ﯾﮑﻨﻮﺍﺧﺖ ﺗﻮﻟﯿﺪ ﻣﯽ ﺷﻮﺩ.ﺷﮑﻞ 7ﻧﺸﺎﻥ ﻣﯽ ﺩﻫﺪ ﮐﻪ ﻓﻘﻂ DBSCAN ،

ICKDC ﻭ Meanshift ﻣﯽ ﺗﻮﺍﻧﻨﺪ ﺗﻮﺯﯾﻊ ﺭﺍ ﺷﻨﺎﺳﺎﯾﯽ ﮐﻨﻨﺪ ﺩﺭ ﺣﺎﻟﯽ ﮐﻪ DPC ﺣﺘﯽ ﭘﺲ ﺍﺯ

ﺁﻥﻧﯿﺰ ﻧﻤﯽ ﺗﻮﺍﻧﺪ ﺗﻨﻈﯿﻢ ﺷﺪﻩ ﺍﺳﺖ. ﺗﺮﮐﯿﺐﺷﺎﻣﻞ ﺧﻮﺷﻪ ﻫﺎﯾﯽ ﺑﺎ ﺗﻮﺯﯾﻊ ﭼﮕﺎﻟﯽ ﻣﺨﺘﻠﻒ ﺍﺳﺖ. ﻧﻤﻮﻧﻪ ﻫﺎﯼ [ﺳﻤﺖ](#_bookmark13) ﭼﭗ ﺑﺎﻻ

ﻧﺸﺎﻥﺩﻫﻨﺪﻩ ﺧﻮﺷﻪ ﻫﺎﯾﯽ ﺑﺎ ﭼﮕﺎﻟﯽ ﻏﯿﺮ ﯾﮑﻨﻮﺍﺧﺖ ﻫﺴﺘﻨﺪ. ﺧﻮﺷﻪ ﺣﻠﻘﻪ ﺍﯼ ﺷﮑﻞ ﺩﺭ ﭘﺎﯾﯿﻦ ﺳﻤﺖﭼﭗ ﺑﯿﺶ ﺍﺯ ﯾﮏ ﻣﺮﮐﺰ ﺩﺍﺭﺩ. ﻧﻤﻮﻧﻪ ﻫﺎﯼ ﺳﻤﺖ ﺭﺍﺳﺖ ﺍﺯ ﺩﻭ ﺧﻮﺷﻪ ﺑﺎ ﺍﻧﺪﺍﺯﻩ ﻭ ﺗﻮﺯﯾﻊ [ﻧﺎﻣﺘﻌﺎﺩﻝ](#_bookmark12)ﺗﺸﮑﯿﻞ ﺷﺪﻩ ﺍﺳﺖ. ﻫﻤﺎﻧﻄﻮﺭ ﮐﻪ ﻣﯽ ﺑﯿﻨﯿﻢ ﺍﺯﺷﮑﻞ .8ﻣﯽ ﺗﻮﺍﻧﺪ ﺗﻮﺯﯾﻊ ﺑﻨﯿﺎﺩﯼ ﻫﻤﻪ

ﺧﻮﺷﻪ ﻫﺎﺭﺍ ﺷﻨﺎﺳﺎﯾﯽ ﮐﻨﺪ، ﺑﺎ ﺍﯾﻦ ﺗﻔﺎﻭﺕ ﮐﻪ ﺑﺮﺧﯽ ﺍﺯ ﻧﻤﻮﻧﻪ ﻫﺎ ﺩﺭ ﻧﺎﺣﯿﻪ ﭘﺮﺍﮐﻨﺪﻩ ﺳﻤﺖ

ﺭﺍﺳﺖﺑﻪ ﺍﺷﺘﺒﺎﻩ ﻃﺒﻘﻪ ﺑﻨﺪﯼ ﺷﺪﻩ ﺍﻧﺪ ICKDC ،ﻧﻤﯽ ﺗﻮﺍﻧﻨﺪ ﺗﻮﺯﯾﻊ ﭼﮕﺎﻟﯽ ﭘﯿﭽﯿﺪﻩ ﺭﺍ ﺁﺷﮑﺎﺭ ﮐﻨﻨﺪ.ﺍﺯ ﺳﻮﯼ ﺩﯾﮕﺮ AGNES ﻭ Meanshift ،K-means [.ﺑﺰﺭﮔﯽ ﺭﺍ](#_bookmark14) ﺑﺪﺳﺖ ﻣﯽ ﺁﻭﺭﺩ FMI ARI ﻧﻤﯽ ﺗﻮﺍﻧﺪ ﺗﻌﺪﺍﺩ ﻭﺍﻗﻌﯽ ﺧﻮﺷﻪ ﻫﺎ ﺭﺍ ﺗﺸﺨﯿﺺ ﺩﻫﺪ، ﺍﮔﺮﭼﻪ DBSCAN .ﺑﻪ ﺳﻪ ﻗﺴﻤﺖﺗﻘﺴﯿﻢ ﻣﯽ ﺷﻮﺩ DPC ﺗﻘﺮﯾﺒﺎ ًﻣﯽ ﺗﻮﺍﻧﻨﺪ ﺗﻮﺯﯾﻊ ﺍﯾﻦ ﻣﺠﻤﻮﻋﻪ ﺩﺍﺩﻩ ﺭﺍ ﺷﻨﺎﺳﺎﯾﯽ ﮐﻨﻨﺪ.ﺑﺎ ﺍﯾﻦ ﺣﺎﻝ، ﺧﻮﺷﻪ ﭘﺎﯾﯿﻦ ﺳﻤﺖ ﭼﭗ ﺑﻪ ﺩﻟﯿﻞ ﻭﺟﻮﺩ ﭼﻨﺪ ﻣﺮﮐﺰ ﺗﻮﺳﻂ DBSCAN

ﻭ

ﻭ

، DPC

ﺍﺳﺖﮐﻪ ﺷﺎﻣﻞ ﺧﻮﺷﻪ ﻫﺎﯼ ﺷﮑﻞ ﻣﺤﺪﺏ ﺍﺳﺖ. ﻫﻤﺎﻧﻄﻮﺭ ﮐﻪ ﻣﯽ ﺑﯿﻨﯿﻢ ﺍﺯ R15 ﺷﺒﯿﻪ D31ﺷﮑﻞ 9ﮐﻪ ﻫﻤﻪ ﺭﻭﺵ ﻫﺎ ﺑﻪ ﺟﺰ DBSCAN ﺷﮑﻞ ﺧﻮﺷﻪ ﻫﺎﯾﯽ ﺭﺍ ﮐﻪ ﺑﻪ ﺷﺪﺕ

[ﻫﻤﭙﻮﺷﺎﻧﯽﺩﺍﺭﻧﺪ](#_bookmark12) ﺗﺸﺨﯿﺺ ﻣﯽ ﺩﻫﻨﺪ. ﻧﺘﺎﯾﺞ ﺍﺭﺯﯾﺎﺑﯽ ﺩﺭ ﺟﺪﻭﻝ 4ﻫﻤﭽﻨﯿﻦ ﻧﺸﺎﻥ ﻣﯽ ﺩﻫﺪ ﮐﻪ

ICKDC ﺍﺯ ﻧﻈﺮ ﻫﻤﻪ ﻣﻌﯿﺎﺭﻫﺎ ﺑﻬﺘﺮ ﻋﻤﻞ ﻣﯽ ﮐﻨﺪ ﻭ ﺑﻪ ﺩﻧﺒﺎﻝ ﺁﻥ Meanshift ﻭ DPC ﻗﺮﺍﺭ

ﺩﺍﺭﻧﺪ.

ﻋﻤﻠﮑﺮﺩﺭﻭﺵ ﻫﺎﯼ ﻣﺨﺘﻠﻒ ﺑﺎ ﭘﺎﺭﺍﻣﺘﺮﻫﺎﯼ ﻭﺭﻭﺩﯼ ﻣﺮﺑﻮﻃﻪ ﺁﻧﻬﺎ ﺗﻌﯿﯿﻦ ﻣﯽ ﺷﻮﺩ. ﺑﺮﺍﯼ ICKDC، ﺗﻌﺪﺍﺩ ﻧﺰﺩﯾﮑﺘﺮﯾﻦ ﻫﻤﺴﺎﯾﻪ ﻫﺎ ﺑﺎ ﺗﻌﯿﯿﻦ ﻣﯽ ﺷﻮﺩ = ⋅ ﺟﺎﯾﯽ ﮐﻪ*n*ﺗﻌﺪﺍﺩ ﮐﻞ ﺍﻣﺘﯿﺎﺯﺍﺕ

ﻭ ﯾﮏ ﺩﺭﺻﺪ ﻗﺎﺑﻞ ﺗﻨﻈﯿﻢ ﺩﺭ ﻣﻮﺭﺩ ﻣﺠﻤﻮﻋﻪ ﺩﺍﺩﻩ ﻫﺎﯼ ﻣﺨﺘﻠﻒ ﺍﺳﺖ. ﺩﺭ ﺗﺴﺖ ﻫﺎ ﺍﺯ ﺟﺴﺘﺠﻮﯼﺗﺼﺎﺩﻓﯽ ﺑﺮﺍﯼ ﺍﻧﺘﺨﺎﺏ ﻣﻨﺎﺳﺐ ﺍﺳﺘﻔﺎﺩﻩ ﻣﯽ ﮐﻨﯿﻢ ﺍﺯ 1 ﺗﺎ 15 ﺑﺮﺍﯼ ﻫﺮ ﻣﺠﻤﻮﻋﻪ

ﺩﺍﺩﻩ.ﺑﺮﺍﯼ ﺟﻠﻮﮔﯿﺮﯼ ﺍﺯ ﻫﻤﻮﺍﺭ ﺷﺪﻥ ﺗﻮﺯﯾﻊ ﭼﮕﺎﻟﯽ ﺗﺨﻤﯿﻦ ﺯﺩﻩ ﺷﺪﻩ، ﮐﺮﺍﻥ [ﭘﺎﯾﯿﻦ](#_bookmark16) ﺑﺎ [1](#_bookmark16) ﻣﺤﺪﻭﺩ ﻣﯽﺷﻮﺩ. ﺍﺯ ﻃﺮﻑ ﺩﯾﮕﺮ، ﮐﺮﺍﻥ ﺑﺎﻻﯾﯽ 15 ﺗﻨﻈﯿﻢ ﺷﺪﻩ ﺍﺳﺖ ﺯﯾﺮﺍ ﺑﺰﺭﮒ ﺍﺳﺖ ﻣﻤﮑﻦ ﺍﺳﺖ

ﺑﺎﻋﺚﺍﺩﻏﺎﻡ ﺧﻮﺷﻪ ﻫﺎﯼ ﻣﺨﺘﻠﻒ ﺷﻮﺩ ﮐﻪ ﺑﺎﻋﺚ ﺑﻬﺒﻮﺩ ﻋﻤﻠﮑﺮﺩ ﺧﻮﺷﻪ ﺑﻨﺪﯼ ﻧﻤﯽ ﺷﻮﺩ.

ﺑﺮﺍﯼDPC، ﻣﺎ ﺑﻬﯿﻨﻪ ﺭﺍ ﺍﻧﺘﺨﺎﺏ ﻣﯽ ﮐﻨﯿﻢ ﺍﺯ 0 ﺗﺎ 2 ﺩﺭﺻﺪ ﻭ ﺍﺯ ﻫﺴﺘﻪ ﮔﺎﻭﺳﯽ ﺩﺭ ﻓﺮﺁﯾﻨﺪ ﺗﺨﻤﯿﻦﭼﮕﺎﻟﯽ ﺍﺳﺘﻔﺎﺩﻩ ﮐﻨﯿﺪ. ﺑﺮﺍﯼ DBSCAN ﻭ Meanshift، ﺍﺯ ﺟﺴﺘﺠﻮﯼ ﺷﺒﮑﻪ ﺍﯼ ﺑﺮﺍﯼ ﯾﺎﻓﺘﻦﺑﻬﺘﺮﯾﻦ ﭘﯿﮑﺮﺑﻨﺪﯼ ﭘﺎﺭﺍﻣﺘﺮ ﺍﺳﺘﻔﺎﺩﻩ ﻣﯽ ﮐﻨﯿﻢ. ﺑﺮﺍﯼ K-means ﻭ AGNES، ﻣﺎ ﺑﻪ ﺳﺎﺩﮔﯽﺗﻌﺪﺍﺩ ﺻﺤﯿﺢ ﺧﻮﺷﻪ ﻫﺎ ﺭﺍ ﻣﯽ ﺩﻫﯿﻢ. ﺑﺮﺍﯼ ﻫﻤﻪ ﺍﻧﻮﺍﻉ DPC، ﻣﺎ ﭘﺎﺭﺍﻣﺘﺮﻫﺎﯼ ﻭﺭﻭﺩﯼ ﺭﺍﻫﻤﺎﻧﻄﻮﺭ ﮐﻪ ﺗﻮﺳﻂ ﻧﻮﯾﺴﻨﺪﮔﺎﻥ ﭘﯿﺸﻨﻬﺎﺩ ﺷﺪﻩ ﺍﺳﺖ ﺗﻨﻈﯿﻢ ﻣﯽ ﮐﻨﯿﻢ.ﺟﺪﺍﻭﻝ 2ﻭ3

ﺗﻨﻈﯿﻤﺎﺕﭘﺎﺭﺍﻣﺘﺮ ﻫﻤﻪ ﺭﻭﺵ ﻫﺎ ﺭﺍ ﺩﺭ ﻣﻮﺍﺭﺩ ﻣﺨﺘﻠﻒ ﺧﻼﺻﻪ ﮐﻨﯿﺪ.

.4.2ﻣﻮﺭﺩ ﺍﻭﻝ: ﻣﻘﺎﯾﺴﻪ ﺑﺎ ﺭﻭﺵ ﻫﺎﯼ ﮐﻼﺳﯿﮏ ﺧﻮﺷﻪ ﺑﻨﺪﯼ

ﻧﺘﺎﯾﺞﺧﻮﺷﻪ ﺑﻨﺪﯼ IKDPC ﻭ ﻫﻤﻪ ﺭﻭﺵ ﻫﺎﯼ ﺧﻮﺷﻪ ﺑﻨﺪﯼ ﮐﻼﺳﯿﮏ ﺑﺮ ﺭﻭﯼ 8 ﻣﺠﻤﻮﻋﻪ ﺩﺍﺩﻩﺩﻭ ﺑﻌﺪﯼ ﻣﺼﻨﻮﻋﯽ ﺩﺭﺍﻧﺠﯿﺮ. 10–2، ﺟﺎﯾﯽ ﮐﻪﺷﮑﻞ 2ﺣﻘﯿﻘﺖ ﺍﺻﻠﯽ ﻫﻤﻪ ﻣﺠﻤﻮﻋﻪ ﺩﺍﺩﻩﻫﺎ ﺭﺍ ﻧﺸﺎﻥ ﻣﯽ ﺩﻫﺪ. ﺧﻮﺷﻪ ﻫﺎﯼ ﻣﺨﺘﻠﻒ ﺑﺎ ﺭﻧﮓ ﻫﺎ ﻣﺘﻤﺎﯾﺰ ﻣﯽ ﺷﻮﻧﺪ ﺩﺭ ﺣﺎﻟﯽ ﮐﻪ ﻧﻘﺎﻁﺳﯿﺎﻩ ﻧﻤﺎﯾﺎﻧﮕﺮ ﻧﻘﺎﻁ ﭘﺮﺕ ﯾﺎ ﻧﻮﯾﺰ ﻫﺴﺘﻨﺪ. ﻣﻌﯿﺎﺭﻫﺎﯼ ﺍﺭﺯﯾﺎﺑﯽ ﻋﻤﻠﮑﺮﺩ AMI ،ARI، ﻭ

FMI ﺑﻪ ﻫﻤﺮﺍﻩ ﺯﻣﺎﻥ ﺍﺟﺮﺍﯼ ﺗﻤﺎﻣﯽ ﺭﻗﺒﺎ ﺩﺭ ﺍﯾﻦ ﻣﻮﺍﺭﺩ ﺧﻼﺻﻪ ﺷﺪﻩ ﺍﺳﺖ.ﺟﺪﻭﻝ .4 )ﻧﺘﯿﺠﻪ

ﺧﻮﺷﻪﺑﻨﺪﯼ ﻭ ﺯﻣﺎﻥ ﺍﺟﺮﺍﯼ ﻣﺘﻨﺎﻇﺮ AGNES ﺭﻭﺷﻦ ﺍﺳﺖ 2ﻣﺠﻤﻮﻋﻪ ﺩﺍﺩﻩ ﻫﺎ ﺑﻪ ﺩﻟﯿﻞ ﺯﻣﺎﻥ

ﻣﺤﺎﺳﺒﺎﺕﺑﺎﻻ ﺩﺍﺩﻩ ﻧﻤﯽ ﺷﻮﻧﺪ(.

ﻫﻤﺎﻧﻄﻮﺭﮐﻪ ﺩﺭ ﻧﺸﺎﻥ ﺩﺍﺩﻩ ﺷﺪﻩ ﺍﺳﺖﺷﮑﻞ 3ﺑﺪﺗﺮﯾﻦ ﺧﻮﺷﻪ ﺑﻨﺪﯼ ﺭﺍ ﻓﺮﺍﻫﻢ ﻣﯽ ﮐﻨﺪ ﮐﻪ ﺧﻮﺷﻪﻫﺎﯼ ﻣﺨﺘﻠﻒ ﺭﺍ ﺩﺭ ﯾﮏ ﺧﻮﺷﻪ ﺍﺩﻏﺎﻡ ﻣﯽ ﮐﻨﺪ. ﻣﻨﺠﺮ ﻣﯽ ﺷﻮﺩ ﺑﻪ AGNES .ﻧﻤﯽ

ﺗﻮﺍﻧﻨﺪﺗﻮﺯﯾﻊ ﻭﺍﻗﻌﯽ ﺭﺍ ﺷﻨﺎﺳﺎﯾﯽ ﮐﺮﺩﻩ ﻭ ﺧﻮﺷﻪ ﻫﺎﯼ ﭘﺎﯾﯿﻦ ﻭ ﭘﺎﯾﯿﻦ ﺳﻤﺖ ﺭﺍﺳﺖ ﺭﺍ ﺑﻪ ﺩﻭ

ﺧﻮﺷﻪﺗﻘﺴﯿﻢ ﮐﻨﻨﺪ Meanshift ﻭ K-means .ﺭﺍ ﻫﻨﮕﺎﻡ ﺑﺮﺧﻮﺭﺩ ﺑﺎ ﻣﺮﺯ ﺍﺭﺍﺉﻪ ﻣﯽ ﺩﻫﺪ DPC ﻧﺘﯿﺠﻪﻣﺸﺎﺑﻪ DBSCAN .ﺍﺯ 1٪ ﺑﻪ 2٪ ﺗﻨﻈﯿﻢ ﺷﺪﻩ ﺍﺳﺖ dc ﻣﺸﺨﺺ ﻣﯽ ﺷﻮﻧﺪ، ﺣﺘﯽ DPC ﻣﯽ ﺗﻮﺍﻧﻨﺪ ﺳﺎﺧﺘﺎﺭ ﻣﺠﻤﻮﻋﻪ ﺩﺍﺩﻩ ﺗﺠﻤﻌﯽ ﺭﺍ ﮐﻪ ﺍﺯ ﺧﻮﺷﻪ ﻫﺎﯼ ﻏﯿﺮ ﮐﺮﻭﯼ ﺑﺎ ﺗﻮﺯﯾﻊ ﺩﻟﺨﻮﺍﻩﺗﺸﮑﯿﻞ ﺷﺪﻩ ﺍﺳﺖ، ﺗﺸﺨﯿﺺ ﺩﻫﻨﺪ. ﺑﺎ ﺍﯾﻦ ﺣﺎﻝ، ﺑﺮﺧﯽ ﺍﺯ ﻧﻘﺎﻁ ﻧﺰﺩﯾﮏ ﺑﻪ ﻣﺮﺯ ﺧﻮﺷﻪﻫﺎ ﺑﻪ ﻋﻨﻮﺍﻥ ﻧﻮﯾﺰ ﺗﻮﺳﻂ DBSCAN ﻭ DPC ICKDC, ,[ﺟﺪﻭﻝ 4](#_bookmark20)ﻫﻤﭽﻨﯿﻦ ﭘﯿﺸﻨﻬﺎﺩ ﻣﯽ ﮐﻨﺪﮐﻪ ICKDC ﻭ AMI ،ARI DPC ﻭ FMI ﺑﻬﺘﺮﯼ ﺭﺍ ﺩﺭ ﻣﺠﻤﻮﻋﻪ ﺩﺍﺩﻩ ﻫﺎﯼ

Aggregation ﺍﺭﺍﺉﻪ ﮐﻨﻨﺪ.

ﺟﺪﻭﻝ4

ﻋﻤﻠﮑﺮﺩﻫﻤﻪ ﺭﻗﺒﺎ ﺩﺭ ﻣﺠﻤﻮﻋﻪ ﺩﺍﺩﻩ ﻫﺎﯼ ﻣﻌﯿﺎﺭ.

ﺍﻟﮕﻮﺭﯾﺘﻢ ﻣﻌﯿﺎﺭﻫﺎﯼﺍﺭﺯﯾﺎﺑﯽ

ﺑﺎﺭ(

0.493

0.183

0.009

0.006

0.050

310 *<*

0.216

0.069

0.002

0.007

0.025

54.575

0.008

0.033

0.001

0.003

0.023

4.838

0.019

0.022

0.001

0.003

0.020

8.343

0.101

0.016

0.001

0.003

0.024

6.703

FMI

0.984

0.958

0.978

0.869

0.945

0.776

0.902

0.646

0.965

0.828

0.829

0.854

0.878

0.783

0.634

0.898

0.684

0.000

0.676

0.626

0.281

0.595

0.560

0.474

0.828

0.814

0.569

0.594

0.722

0.000

AMI

0.932

0.849

0.913

0.633

0.817

0.447

0.901

0.759

0.898

0.785

0.795

0.802

0.807

0.707

0.552

0.833

0.525

0.000

0.684

0.665

0.292

0.623

0.647

0.507

0.325

0.247

0.004

0.228

0.143

0.000

ARI

ﺍﺳﺒﺎﺏﺑﺎﺯﯼ

0.968

0.915

0.957

0.739

0.890

0.552

ﺗﺮﮐﯿﺐ 0.862

0.535

0.953

0.769

0.739

0.792

ﺷﺮﺍﺏ

0.816

0.672

0.447

0.847

0.418

0.000

ﺗﻘﺴﯿﻢﺑﻨﺪﯼ

0.624

0.556

0.048

0.508

0.423

0.242

ﭘﺎﺭﮐﯿﻨﺴﻮﻥﻫﺎ

0.456

0.269

0.014

0.052

0.230

0.000

ﺑﺎﺭ(

0.234

0.097

0.003

0.027

0.024

215.624

0.064

0.046

0.002

0.004

0.028

11.335

4.496

1.526

0.115

0.041

0.047

–

0.094

0.028

0.001

0.003

0.017

7.710

0.252

0.051

0.002

0.003

0.021

216.816

0.110

0.023

0.001

0.002

0.028

47.009

FMI

0.993

0.993

0.983

0.910

0.901

0.990

0.992

0.874

0.976

0.771

0.934

0.731

0.945

0.939

0.770

0.942

0.914

–

0.852

0.823

0.671

0.803

0.817

0.817

0.897

0.786

0.725

0.877

0.728

0.728

0.644

0.643

0.857

0.605

0.737

0.733

AMI

0.994

0.994

0.985

0.960

0.942

0.992

0.963

0.689

0.889

0.465

0.793

0.481

0.948

0.944

0.872

0.946

0.928

–

0.740

0.721

0.586

0.671

0.697

0.703

0.678

0.461

0.344

0.623

0.007

0.007

0.294

0.137

0.571

0.133

0.136

0.003

ARI R15

0.993

0.993

0.982

0.902

0.893

0.989

ﺷﻌﻠﻪ

0.983

0.735

0.950

0.524

0.854

0.442

S2

0.942

0.935

0.745

0.938

0.908

–

ﺩﺍﻧﻪ

0.778

0.734

0.529

0.705

0.725

0.725

WDBC

0.773

0.471

0.440

0.730

0.005

0.005

ﯾﻮﻧﻮﺷﯿﺮ 0.263

0.218

0.684

0.173

0.346

0.004

ﺑﺎﺭ(

0.207

0.092

0.003

0.044

0.029

602.510

0.353

0.096

0.004

0.005

0.031

960.144

1.437

0.672

0.046

0.028

0.041

310 *<*

0.007

0.023

0.001

0.002

0.018

2.944

0.137

0.028

0.001

0.004

0.022

34.896

0.108

0.040

0.001

0.004

0.024

46.862

FMI

0.996

0.997

0.991

0.804

0.914

0.993

1000

1000

1000

0.538

0.526

1000

0.953

0.938

0.814

0.911

0.934

0.911

0.900

0.816

0.729

0.811

0.760

0.816

0.796

0.753

0.526

0.758

0.683

0.821

0.820

0.691

0.534

0.784

0.720

0.777

AMI

0.991

0.992

0.983

0.851

0.929

0.987

1000

1000

1000

0.056

0.035

1000

0.965

0.955

0.907

0.950

0.954

0.950

0.829

0.781

0.634

0.739

0.708

0.781

0.651

0.593

0.391

0.597

0.543

0.707

0.840

0.757

0.583

0.877

0.743

0.823

ARI

ﺗﺠﻤﻊ

0.995

0.996

0.988

0.752

0.888

0.991

ﻣﺎﺭﭘﯿﭻ

1000

1000

1000

0.075

0.048

1000

D31

0.952

0.936

0.803

0.908

0.932

0.907

ﻋﻨﺒﯿﻪ

0.851

0.720

0.612

0.716

0.562

0.720

ﺍﻟﯿﮑﻮ 0.701

0.662

0.339

0.675

0.456

0.745

ﭘﻮﺳﺖ

0.765

0.575

0.371

0.729

0.577

0.675

ICKDC DPC DBSCAN

ﮎ-ﯾﻌﻨﯽ ﺗﻐﯿﯿﺮﻣﻌﻨﯽ AGNES

ICKDC DPC DBSCAN

ﮎ-ﯾﻌﻨﯽ ﺗﻐﯿﯿﺮﻣﻌﻨﯽ AGNES

ICKDC DPC DBSCAN

ﮎ-ﯾﻌﻨﯽ ﺗﻐﯿﯿﺮﻣﻌﻨﯽ AGNES

ICKDC DPC DBSCAN

ﮎ-ﯾﻌﻨﯽ ﺗﻐﯿﯿﺮﻣﻌﻨﯽ AGNES

ICKDC DPC DBSCAN

ﮎ-ﯾﻌﻨﯽ ﺗﻐﯿﯿﺮﻣﻌﻨﯽ AGNES

ICKDC DPC DBSCAN

ﮎ-ﯾﻌﻨﯽ ﺗﻐﯿﯿﺮﻣﻌﻨﯽ AGNES

ﮐﺎﻫﺶﻗﺎﺑﻞ ﺗﺤﻤﻞ ﺩﺭ ﮐﺎﺭﺍﯾﯽ ﻧﺴﺒﺖ ﺑﻪ DPC ﺩﺭ ﻣﻮﺭﺩ ﻣﺠﻤﻮﻋﻪ ﺩﺍﺩﻩ ﻧﺴﺒﺘﺎ ًﺑﺰﺭﮒ S2 ) 4.496ﺩﺭ ﻣﻘﺎﺑﻞ .(1.526 ﺑﺎ ﺗﻮﺟﻪ ﺑﻪ ﺳﺎﺧﺖ ﻧﺎﺣﯿﻪ ﻫﺴﺘﻪ، ICKDC ﻗﺎﺩﺭ [ﺍﺳﺖ ﻧﻤﻮﻧﻪ](#_bookmark21) ﺍﻭﻟﯿﻪ ﻫﺮﺧﻮﺷﻪ ﺭﺍ ﺑﺎ ﭼﻨﺪﯾﻦ ﻧﻘﻄﻪ ﻫﺴﺘﻪ ﻧﺸﺎﻥ ﺩﻫﺪ ﺗﺎ ﺗﻮﺯﯾﻊ ﭼﮕﺎﻟﯽ ﺭﺍ ﺑﻬﺘﺮ ﺁﺷﮑﺎﺭ ﮐﻨﺪ. ﺑﺮﺍﯼ

ﺧﻮﺷﻪﻫﺎﯾﯽ ﺑﺎ [ﺑﯿﺶ ﺍﺯ ﯾﮏ](#_bookmark12) ﻣﺮﮐﺰ ﺗﺮﺍﮐﻢ ﺑﺴﯿﺎﺭ ﻣﻬﻢ ﺍﺳﺖ. ﺑﺎ ﺍﯾﻦ ﺣﺎﻝ، ﻣﺼﺎﻟﺤﻪ ﺍﯼ ﺑﯿﻦ

ﮐﺎﺭﺍﯾﯽﻭ ﺩﻗﺖ ﺻﻮﺭﺕ ﮔﺮﻓﺘﻪ ﺍﺳﺖ.

ﺑﻪﻃﻮﺭ ﺧﻼﺻﻪ، DPC ﻣﯽ ﺗﻮﺍﻧﺪ ﺩﺭ [ﺑﺮﺧﯽ](#_bookmark12) ﻣﻮﺍﺭﺩ ﻧﺘﺎﯾﺞ ﻗﺎﺑﻞ ﻗﺒﻮﻟﯽ ﺍﺭﺍﺉﻪ ﺩﻫﺪ ﺍﻣﺎ ﻧﻤﯽ ﺗﻮﺍﻧﺪﺧﻮﺷﻪ ﻫﺎﯼ ﺑﯿﺶ ﺍﺯ ﯾﮏ ﻣﺮﮐﺰ ﺭﺍ ﺷﻨﺎﺳﺎﯾﯽ ﮐﻨﺪ. DBSCAN ﻧﻤﯽ ﺗﻮﺍﻧﺪ ﺑﺎ ﺧﻮﺷﻪ ﺑﻨﺪﯼ

ﻧﻤﻮﻧﻪﻫﺎﯾﯽ ﺑﺎ ﺗﻮﺯﯾﻊ ﭼﮕﺎﻟﯽ ﻧﺎﻣﺘﻌﺎﺩﻝ ﻣﺎﻧﻨﺪ Compound ﻣﻘﺎﺑﻠﻪ ﮐﻨﺪ. Meanshift ،

K-means ﻭ AGNES ﻓﻘﻂ ﻣﯽ ﺗﻮﺍﻧﻨﺪ ﺧﻮﺷﻪ ﻫﺎﯼ ﻣﺤﺪﺏ ﺭﺍ ﭘﯿﺪﺍ ﮐﻨﻨﺪ. ﺍﺯ ﺳﻮﯼ ﺩﯾﮕﺮ،

ICKDC ﻣﯽ ﺗﻮﺍﻧﺪ ﺧﻮﺷﻪ ﻫﺎﯾﯽ ﺭﺍ ﺑﺎ ﺷﮑﻞ ﻫﺎ ﻭ ﭼﮕﺎﻟﯽ ﻫﺎﯼ ﺩﻟﺨﻮﺍﻩ ﺗﺸﺨﯿﺺ ﺩﻫﺪ ﻭ

ﺑﻬﺘﺮﯾﻦﻧﺘﺎﯾﺞ ﺧﻮﺷﻪ ﺑﻨﺪﯼ ﺭﺍ ﺍﺯ ﻧﻈﺮ ﺳﻪ ﻣﻌﯿﺎﺭ ﺍﺭﺯﯾﺎﺑﯽ ﺩﺭ ﺍﮐﺜﺮ ﻣﻮﺍﺭﺩ ﺑﻪ ﺩﺳﺖ ﺁﻭﺭﺩ.

.4.3ﻣﻮﺭﺩ ﺩﻭﻡ: ﻣﻘﺎﯾﺴﻪ ﺑﺎ ﺍﻧﻮﺍﻉ *DPC* ﺳﭙﺲ،ﻣﺎ ﻧﺘﺎﯾﺞ ﺧﻮﺷﻪ ﺑﻨﺪﯼ ICKDC ﺭﺍ ﺑﺎ ﺳﻪ ﻧﻮﻉ ﺟﺪﯾﺪ DPC ﻣﻘﺎﯾﺴﻪ ﻣﯽ ﮐﻨﯿﻢ.ﯾﻌﻨﯽ

) DPC-KNNﺩﻭ ﻭ ﻫﻤﮑﺎﺭﺍﻥ،)2016 FKNN-DPC ،(ﺯﯼ ﻭ ﻫﻤﮑﺎﺭﺍﻥ،(2016 ﻭ IVDPC )ﮊﻭ ﻭ ﻫﻤﮑﺎﺭﺍﻥ،(2018 ﺩﺭ ﻣﺠﻤﻮﻋﻪ ﺩﺍﺩﻩ ﻫﺎﯼ ﺩﻧﯿﺎﯼ ﻭﺍﻗﻌﯽ. ﻫﻤﺎﻧﻄﻮﺭ ﮐﻪ ﺩﺭ ﺑﺨﺶ ﺑﺤﺚ ﮐﺮﺩﯾﻢ2 ،ﺍﮐﺜﺮ ﺍﻧﻮﺍﻉ ﺑﻬﯿﻨﻪ ﺷﺪﻩ DPC ﺳﻌﯽ ﻣﯽ ﮐﻨﻨﺪ ﺭﻭﺵ ﺍﺻﻠﯽ ﺭﺍ ﺑﺎ ﺩﺭ ﻧﻈﺮ [ﮔﺮﻓﺘﻦ](#_bookmark8) ﯾﮏ ﯾﺎ ﺩﻭ

ﻧﻘﺺﺑﻬﺒﻮﺩ ﺑﺨﺸﻨﺪ. ﻫﺮ ﺳﻪ ﻧﻮﻉ DPC ﺩﺭ ﻧﻈﺮ ﮔﺮﻓﺘﻪ ﺷﺪﻩ ﺩﺭ ﺁﺯﻣﺎﯾﺶ ﺟﺪﯾﺪ ﺭﺍ ﻣﻌﺮﻓﯽ ﻣﯽ

ﮐﻨﻨﺪ

ﻣﻘﺎﯾﺴﻪﺑﯿﻦ ﺗﻤﺎﻡ ﺭﻗﺒﺎ ﺩﺭ S2 ﺷﺒﯿﻪ D31 ﺑﻮﺩ.ﺷﮑﻞ 10ﻧﺸﺎﻥ ﻣﯽ ﺩﻫﺪ ﮐﻪ ICKDC ﺑﻬﺘﺮﯾﻦﻧﺘﯿﺠﻪ ﺧﻮﺷﻪ ﺑﻨﺪﯼ ﺭﺍ ﺑﻪ ﺩﺳﺖ ﻣﯽ ﺁﻭﺭﺩ، K-means ،DPC ﻭ Meanshift ﻋﻤﻠﮑﺮﺩ

ﻗﺎﺑﻞﻗﺒﻮﻟﯽ ﺩﺍﺭﻧﺪ، ﺩﺭ ﺣﺎﻟﯽ ﮐﻪ DBSCAN ﻧﻤﯽ ﺗﻮﺍﻧﺪ ﺳﺎﺧﺘﺎﺭ ﻭﺍﻗﻌﯽ ﺭﺍ ﺷﻨﺎﺳﺎﯾﯽ ﮐﻨﺪ.

ﻣﻘﺎﯾﺴﻪﻫﺎ ﺩﺭﺟﺪﻭﻝ 4ﻫﻤﭽﻨﯿﻦ ﭘﯿﺸﻨﻬﺎﺩ ﻣﯽ ﮐﻨﺪ ﮐﻪ ICKDC ﺍﺯ ﻧﻈﺮ ﺳﻪ ﻣﻌﯿﺎﺭ ﻧﺴﺒﺖ ﺑﻪ S2

ﺍﺯﺳﺎﯾﺮ ﺭﻗﺒﺎ ﺑﻬﺘﺮ ﻋﻤﻞ ﮐﻨﺪ.

ﻋﻼﻭﻩﺑﺮ ﺍﯾﻦ، ﻣﺎ ﻫﻤﭽﻨﯿﻦ ﻣﯽ ﺗﻮﺍﻧﯿﻢ ﺍﺯﺟﺪﻭﻝ 4ﮐﻪ ICKDC ﻫﻨﮕﺎﻡ ﺑﺮﺧﻮﺭﺩ ﺑﺎ ﻣﺠﻤﻮﻋﻪ ﺩﺍﺩﻩﻫﺎﯼ ﺩﻧﯿﺎﯼ ﻭﺍﻗﻌﯽ، ﺑﻪ ﺟﺰ Elico ،Wine ﻭ Ionoshe، ﺍﺯ ﺭﻭﺵ ﻫﺎﯼ ﮐﻼﺳﯿﮏ ﺑﻬﺘﺮ

ﻋﻤﻞﻣﯽ ﮐﻨﺪ. ﺑﺎ ﺍﯾﻦ ﺣﺎﻝ، ﺑﺮﺍﯼ ﺍﯾﻦ ﺳﻪ ﻣﺠﻤﻮﻋﻪ ﺩﺍﺩﻩ، ICKDC ﺩﺭ ﺭﺗﺒﻪ ﺩﻭﻡ Wine ﻭ Elico ﻭﺩﺭ Ionoshere ﺍﺯ ﻧﻈﺮ AMI ،ARI ﻭ FMI ﺩﺭ ﺭﺗﺒﻪ ﺳﻮﻡ ﻗﺮﺍﺭ ﺩﺍﺭﺩ. ﻧﺘﺎﯾﺞ ﻧﺸﺎﻥ ﻣﯽ ﺩﻫﺪ ﮐﻪICKDC ﻣﯽ ﺗﻮﺍﻧﺪ ﺭﻭﯼ ﻣﺠﻤﻮﻋﻪ ﺩﺍﺩﻩ ﻫﺎﯼ ﭼﻨﺪ ﺑﻌﺪﯼ ﻧﯿﺰ ﮐﺎﺭ ﮐﻨﺪ. ﺩﺭ ﻣﻘﺎﯾﺴﻪ ﺑﺎ

ﺭﻭﺵ ﻫﺎﯼﺧﻮﺷﻪ ﺑﻨﺪﯼ ﭘﺮﮐﺎﺭﺑﺮﺩ، ﻣﯽ ﺗﻮﺍﻧﺪ ﻋﻤﻠﮑﺮﺩ ﮐﻠﯽ ﺑﻬﺘﺮﯼ ﺩﺍﺷﺘﻪ ﺑﺎﺷﺪ.

ﺯﻣﺎﻥﺍﺟﺮﺍﯼ ﻫﻤﻪ ﺭﻭﺵ ﻫﺎ ﺩﺭ ﻫﺮ ﻣﻮﺭﺩ ﻧﯿﺰ ﺩﺭﺝ ﺷﺪﻩ ﺍﺳﺖ ﺟﺪﻭﻝ .4 ﻣﯽ ﺗﻮﺍﻧﯿﻢ ﺑﺒﯿﻨﯿﻢ ﮐﻪDBSCAN ﻭ K-means ﺩﺭ ﺑﯿﺸﺘﺮ ﻣﻮﺍﺭﺩ ﺑﻪ ﻗﯿﻤﺖ ﮐﯿﻔﯿﺖ ﺧﻮﺷﻪ ﺑﻨﺪﯼ ﺑﺮﻧﺪﻩ

ﻣﯽ ﺷﻮﻧﺪ.ﺍﺯ ﺳﻮﯼ ﺩﯾﮕﺮ، ICKDC ﻣﺎ ﻭ DPC ﺍﺻﻠﯽ ﻋﻤﻠﮑﺮﺩ ﭘﺎﯾﺪﺍﺭﺗﺮﯼ ﺑﺎ ﺍﻓﺰﺍﯾﺶ ﺁﺷﮑﺎﺭ ﺯﻣﺎﻥﺍﺟﺮﺍ ﻧﺸﺎﻥ ﻣﯽ ﺩﻫﻨﺪ. ﻋﻼﻭﻩ ﺑﺮ ﺍﯾﻦ، 2x DPC ﺗﺎ 3 ﺑﺮﺍﺑﺮ ﺳﺮﯾﻌﺘﺮ ﺍﺯ ICKDC ﺑﻪ ﻃﻮﺭ ﮐﻠﯽ ﺑﺪﻭﻥﺗﻮﺟﻪ ﺑﻪ ﺍﻧﺪﺍﺯﻩ ﻣﺠﻤﻮﻋﻪ [ﺩﺍﺩﻩ، ﺑﺎ](#_bookmark36) ﺗﻮﺟﻪ ﺑﻪ ﭘﯿﭽﯿﺪﮔﯽ ﺯﻣﺎﻧﯽ ﯾﮑﺴﺎﻥ ﺍﺳﺖ )ﺑﻪ ﺑﺨﺶ ﻣﺮﺍﺟﻌﻪﮐﻨﯿﺪ.(3.5 ﺍﯾﻦ ﻋﻤﺪﺗﺎ ًﻧﺎﺷﯽ [ﺍﺯ ﻓﺮﺁﯾﻨﺪ ﺗﮑﺮﺍﺭ](#_bookmark66) ﺍﺿﺎﻓﯽ ﺩﺭ ﻃﻮﻝ ﺳﺎﺧﺖ ﻧﺎﺣﯿﻪ [ﻫﺴﺘﻪ](#_bookmark56) ﺍﺳﺖ.ﺑﺎ ﺍﯾﻦ ﺣﺎﻝ، ﮐﺎﻫﺶ ﺳﺮﻋﺖ ﺑﺎ ﺗﻮﺟﻪ ﺑﻪ ﺍﻧﺪﺍﺯﻩ [ﻣﺠﻤﻮﻋﻪ](#_bookmark2) ﺩﺍﺩﻩ ﻫﺎ ﺛﺎﺑﺖ ﻣﯽ ﻣﺎﻧﺪ. ﺑﺮﺍﯼ

ﻣﺜﺎﻝ،ICKDC ﺍﻟﻒ ﺭﺍ ﻧﺸﺎﻥ ﻣﯽ ﺩﻫﺪ

ﺷﮑﻞ.2ﻃﺒﻘﻪ ﺑﻨﺪﯼ ﻣﺠﻤﻮﻋﻪ ﺩﺍﺩﻩ ﻫﺎﯼ ﻣﺼﻨﻮﻋﯽ )ﺣﻘﯿﻘﺖ ﺯﻣﯿﻨﯽ.( )ﺍﻟﻒ( ﺗﺠﻤﻊ. )ﺏ( .R15 ﺝ( ﺍﺳﺒﺎﺏ ﺑﺎﺯﯼ ﺩ( ﻣﺎﺭﭘﯿﭻ. )ﻩ( ﺷﻌﻠﻪ. )ﻭ( .A1 )ﺯ( .D31 )ﺡ( .S2

ﺷﮑﻞ.3ﻧﺘﺎﯾﺞ ﺧﻮﺷﻪ ﺑﻨﺪﯼ ﺍﺯﺗﺠﻤﻊ. )ﺍﻟﻒ( ﻧﺘﯿﺠﻪ ﺧﻮﺷﻪ ﺑﻨﺪﯼ .ICKDC )ﺏ( ﻧﺘﯿﺠﻪ ﺧﻮﺷﻪ ﺑﻨﺪﯼ .DPC )ﺝ( ﻧﺘﯿﺠﻪ ﺧﻮﺷﻪ ﺑﻨﺪﯼ .DBSCAN )ﺩ( ﻧﺘﯿﺠﻪ ﺧﻮﺷﻪ ﺑﻨﺪﯼ .K-means )ﻩ( ﻧﺘﯿﺠﻪ ﺧﻮﺷﻪ ﺑﻨﺪﯼ .Meanshift )ﻭ( ﻧﺘﯿﺠﻪ ﺧﻮﺷﻪ ﺑﻨﺪﯼ

.AGNES

ﻋﻼﻭﻩﺑﺮ ﺍﯾﻦ، ﻣﺎ ﻧﺘﺎﯾﺞ ﺧﻮﺷﻪ ﺑﻨﺪﯼ ICKDC ﺭﺍ ﺑﺎ ﺳﺎﯾﺮ ﺭﻗﺒﺎ ﺩﺭ ﻧﻤﻮﺩﺍﺭﻫﺎﯼ ﭘﺮﺍﮐﻨﺪﻩ ﻣﻘﺎﯾﺴﻪﻣﯽ ﮐﻨﯿﻢ. ﻫﻤﺎﻧﻄﻮﺭ ﮐﻪ ﺩﺭﺷﮑﻞ 11، the ﻭ ICKDC ﺍﺯ ARI ﻣﺨﻔﻒ -axis -ﻣﺤﻮﺭ ﺑﺮﺍﯼﺳﺎﯾﺮ ﺭﻗﺒﺎ. ﻧﻘﺎﻃﯽ ﮐﻪ ﺩﺭ ﻧﺎﺣﯿﻪ ﺳﺎﯾﻪ ﻧﯿﺴﺘﻨﺪ ﻧﺸﺎﻥ ﻣﯽ ﺩﻫﻨﺪ ﮐﻪ ICKDC ﺍﺯ ﻧﻈﺮ ARI ﺩﺭﺁﻥ ﻣﺠﻤﻮﻋﻪ ﺩﺍﺩﻩ ﺑﻬﺘﺮ ﻋﻤﻞ ﻣﯽ ﮐﻨﺪ، ﺩﺭ ﺣﺎﻟﯽ ﮐﻪ ﻧﻘﺎﻁ ﺭﻭﯼ ﻣﻮﺭﺏ ﻋﻤﻠﮑﺮﺩﯼ ﻣﻌﺎﺩﻝ ﺭﺍ ﻧﺸﺎﻥ[ﻣﯽ ﺩﻫﻨﺪ.](#_bookmark22) ﻣﺸﺎﻫﺪﻩ ﻣﯽ ﮐﻨﯿﻢ ﮐﻪ ﺑﯿﺸﺘﺮ ﻧﻘﺎﻁ ﺩﺭ ﻧﺎﺣﯿﻪ ﺑﺪﻭﻥ ﺳﺎﯾﻪ ﻗﺮﺍﺭ ﻣﯽ ﮔﯿﺮﻧﺪ ﻭ ﻧﺴﺒﺘﺎ ًﺍﺯ ﻣﻮﺭﺏ ﻓﺎﺻﻠﻪ ﺩﺍﺭﻧﺪ. ﺍﯾﻦ ﻧﺸﺎﻥ ﻣﯽ ﺩﻫﺪ ﮐﻪ ICKDC ﻣﯽ ﺗﻮﺍﻧﺪ ﺧﻮﺷﻪ ﺑﻨﺪﯼ ﺑﻬﺘﺮﯼ

ﺭﺍﺩﺭ ﻣﻘﺎﯾﺴﻪ ﺑﺎ ﻫﻤﻪ ﺭﻗﺒﺎ ﺍﺭﺍﺉﻪ ﺩﻫﺪ.

ﻣﻌﯿﺎﺭﻫﺎﺩﺭ ﺗﻌﺮﯾﻒ ﺗﻮﺯﯾﻊ ﭼﮕﺎﻟﯽ ﻣﺤﻠﯽ، ﺩﺭ ﺣﺎﻟﯽ ﮐﻪ FKNN-DPC ﻭ IVDPC ﺷﺎﻣﻞ ﺍﺳﺘﺮﺍﺗﮋﯼ ﻫﺎﯼ[ﺍﻧﺘﺴﺎﺏ](#_bookmark24) ﺟﺪﯾﺪ ﻫﺴﺘﻨﺪ. ﺍﺯ ﺳﻮﯼ ﺩﯾﮕﺮ، ICKDC ﻣﯽ ﺗﻮﺍﻧﺪ ﺍﺯ ﺳﻪ ﻣﺤﺪﻭﺩﯾﺖ DPC ﺑﻪ ﻃﻮﺭ ﻫﻤﺰﻣﺎﻥ ﺍﺟﺘﻨﺎﺏ ﮐﻨﺪ. ﺑﻨﺎﺑﺮﺍﯾﻦ، ﻣﺎ ﺍﺯ ﻧﻈﺮ ﻣﻌﯿﺎﺭﻫﺎﯼ ﺍﺭﺯﯾﺎﺑﯽ ﺧﻮﺷﻪ ﺑﻨﺪﯼ ﺍﻧﺘﻈﺎﺭﺑﺮﺗﺮﯼ ICKDC ﺭﺍ ﻧﺴﺒﺖ ﺑﻪ ﺳﺎﯾﺮ ﺍﻧﻮﺍﻉ DPC ﺩﺍﺭﯾﻢ. ﻣﺎ ﻣﯽ ﺗﻮﺍﻧﯿﻢ ﺍﺯﺟﺪﻭﻝ 5ﮐﻪ

ICKDC ﺑﻬﺘﺮﯾﻦ ﻧﺘﺎﯾﺞ ﺭﺍ ﺩﺭ ﺩﺍﻧﻪ ﻫﺎ، ﺷﺮﺍﺏ، Ionoshhere ،Segmentation ،WDBC ﻭ Parkinsons ﺑﻪ ﺩﺳﺖ ﻣﯽ ﺁﻭﺭﺩ ﻭ ﺩﺭ ﺳﻪ ﻣﻮﺭﺩ ﺩﯾﮕﺮ ﻣﻘﺎﻡ ﺩﻭﻡ ﺭﺍ ﮐﺴﺐ ﻣﯽ ﮐﻨﺪ. ﻧﺘﺎﯾﺞ ﻧﺸﺎﻥﻣﯽ ﺩﻫﺪ ﮐﻪ ICKDC ﻧﺴﺒﺖ ﺑﻪ ﺳﺎﯾﺮ ﮔﻮﻧﻪ ﻫﺎﯼ DPC ﺩﺭ ﺑﺮﺧﻮﺭﺩ ﺑﺎ ﻣﺠﻤﻮﻋﻪ ﺩﺍﺩﻩ ﻫﺎﯼ

ﺩﻧﯿﺎﯼﻭﺍﻗﻌﯽ ﺑﺮﺗﺮﯼ ﺩﺍﺭﺩ.

ﺷﮑﻞ.4ﻧﺘﺎﯾﺞ ﺧﻮﺷﻪ ﺑﻨﺪﯼ ﺍﺯ.*R15* )ﺍﻟﻒ( ﻧﺘﯿﺠﻪ ﺧﻮﺷﻪ ﺑﻨﺪﯼ .ICKDC )ﺏ( ﻧﺘﯿﺠﻪ ﺧﻮﺷﻪ ﺑﻨﺪﯼ .DPC )ﺝ( ﻧﺘﯿﺠﻪ ﺧﻮﺷﻪ ﺑﻨﺪﯼ .DBSCAN )ﺩ( ﻧﺘﯿﺠﻪ ﺧﻮﺷﻪ ﺑﻨﺪﯼ .K-means )ﻩ( ﻧﺘﯿﺠﻪ ﺧﻮﺷﻪ ﺑﻨﺪﯼ .Meanshift )ﻭ( ﻧﺘﯿﺠﻪ ﺧﻮﺷﻪ ﺑﻨﺪﯼ

.AGNES

ﺷﮑﻞ.5ﻧﺘﺎﯾﺞ ﺧﻮﺷﻪ ﺑﻨﺪﯼ ﺍﺯﺍﺳﺒﺎﺏ ﺑﺎﺯﯼ. )ﺍﻟﻒ( ﻧﺘﯿﺠﻪ ﺧﻮﺷﻪ ﺑﻨﺪﯼ .ICKDC )ﺏ( ﻧﺘﯿﺠﻪ ﺧﻮﺷﻪ ﺑﻨﺪﯼ .DPC )ﺝ( ﻧﺘﯿﺠﻪ ﺧﻮﺷﻪ ﺑﻨﺪﯼ .DBSCAN )ﺩ( ﻧﺘﯿﺠﻪ ﺧﻮﺷﻪ ﺑﻨﺪﯼ .K-means )ﻩ( ﻧﺘﯿﺠﻪ ﺧﻮﺷﻪ ﺑﻨﺪﯼ .Meanshift )ﻭ( ﻧﺘﯿﺠﻪ ﺧﻮﺷﻪ

ﺑﻨﺪﯼ.AGNES

ﺑﺎﯾﮑﺪﯾﮕﺮ ﻣﺘﻔﺎﻭﺕ ﺍﺳﺖ. ﺩﺭ ﺁﺯﻣﺎﯾﺶ ﻫﺎﯼ ﻣﺎ، ﺳﻌﯽ ﻣﯽ ﮐﻨﺪ ﻓﺮﺿﯿﻪ ﺻﻔﺮ (H0) ﺭﺍ ﺗﺄﯾﯿﺪ ﮐﻨﺪﮐﻪ ﻣﯿﺎﻧﻪ ARI ﺍﺭﺍﺉﻪ ﺷﺪﻩ ﺗﻮﺳﻂ ICKDC ﺑﯿﺸﺘﺮ ﺍﺯ ﺁﻥ ﭼﯿﺰﯼ ﻧﯿﺴﺖ ﮐﻪ ﺗﻮﺳﻂ ﺭﻗﺒﺎ

ﺍﺭﺍﺉﻪﺷﺪﻩ ﺍﺳﺖ. ﻭﯾﻠﮑﺎﮐﺴﻮﻥ ﺁﻣﺎﺭ ) -ﺍﺭﺯﺵ( ﻫﺴﺘﻨﺪ

ﺑﺮﺍﯼﺍﻃﻤﯿﻨﺎﻥ ﺍﺯ ﺍﻫﻤﯿﺖ ﺑﺮﺗﺮﯼ، ﺁﺯﻣﻮﻥ ﺭﺗﺒﻪ ﺑﻨﺪﯼ ﻋﻼﻣﺖ ﻭﯾﻠﮑﺎﮐﺴﻮﻥ ﺑﺮﺍﯼ ﻣﻘﺎﯾﺴﻪ ﺑﯿﺸﺘﺮﻣﻌﺮﻓﯽ ﺷﺪﻩ ﺍﺳﺖ. ﺍﺯ ﺍﯾﻦ ﺁﺯﻣﻮﻥ ﺑﺮﺍﯼ ﺗﻌﯿﯿﻦ ﺍﯾﻨﮑﻪ ﺁﯾﺎ ﻣﯿﺎﻧﮕﯿﻦ ﺭﺗﺒﻪ ﺟﺎﻣﻌﻪ ﺩﻭ ﻧﻤﻮﻧﻪ

ﻣﺮﺗﺒﻂﺍﺳﺘﻔﺎﺩﻩ ﻣﯽ ﺷﻮﺩ

ﺷﮑﻞ.6ﻧﺘﺎﯾﺞ ﺧﻮﺷﻪ ﺑﻨﺪﯼ ﺍﺯﻣﺎﺭﭘﯿﭻ. )ﺍﻟﻒ( ﻧﺘﯿﺠﻪ ﺧﻮﺷﻪ ﺑﻨﺪﯼ .ICKDC )ﺏ( ﻧﺘﯿﺠﻪ ﺧﻮﺷﻪ ﺑﻨﺪﯼ .DPC )ﺝ( ﻧﺘﯿﺠﻪ ﺧﻮﺷﻪ ﺑﻨﺪﯼ .DBSCAN )ﺩ( ﻧﺘﯿﺠﻪ ﺧﻮﺷﻪ ﺑﻨﺪﯼ .K-means )ﻩ( ﻧﺘﯿﺠﻪ ﺧﻮﺷﻪ ﺑﻨﺪﯼ .Meanshift )ﻭ( ﻧﺘﯿﺠﻪ ﺧﻮﺷﻪ ﺑﻨﺪﯼ

.AGNES

ﺷﮑﻞ.7ﻧﺘﺎﯾﺞ ﺧﻮﺷﻪ ﺑﻨﺪﯼ ﺍﺯﺷﻌﻠﻪ. )ﺍﻟﻒ( ﻧﺘﯿﺠﻪ ﺧﻮﺷﻪ ﺑﻨﺪﯼ .ICKDC )ﺏ( ﻧﺘﯿﺠﻪ ﺧﻮﺷﻪ ﺑﻨﺪﯼ .DPC )ﺝ( ﻧﺘﯿﺠﻪ ﺧﻮﺷﻪ ﺑﻨﺪﯼ .DBSCAN )ﺩ( ﻧﺘﯿﺠﻪ ﺧﻮﺷﻪ ﺑﻨﺪﯼ .K-means )ﻩ( ﻧﺘﯿﺠﻪ ﺧﻮﺷﻪ ﺑﻨﺪﯼ .Meanshift )ﻭ( ﻧﺘﯿﺠﻪ ﺧﻮﺷﻪ ﺑﻨﺪﯼ

.AGNES

ﻣﯿﺎﻧﻪARI ﺩﺍﺩﻩ ﺷﺪﻩ ﺗﻮﺳﻂ ICKDC ﺑﻪ ﻃﻮﺭ ﻗﺎﺑﻞ ﺗﻮﺟﻬﯽ [ﺑﯿﺸﺘﺮ ﺍﺯ ﺁﻥ](#_bookmark23) ﺍﺳﺖ

ﻫﺮﺭﻗﯿﺒﯽ

ﺧﻼﺻﻪﺷﺪﻩ ﺩﺭﺟﺪﻭﻝ 6، ﺟﺎﯾﯽ ﮐﻪ ﻣﻘﺪﺍﺭ ﺑﺤﺮﺍﻧﯽ ﮐﺮﺍﻥ ﺑﺎﻻﯾﯽ ﺭﺍ ﻣﯽ ﺩﻫﺪ -ﻣﻘﺪﺍﺭ ﺑﺎ ﺳﻄﺢ ﻣﻌﻨﯽﺩﺍﺭﯼ 95 ﺩﺭﺻﺪ. ﻣﯽ ﺑﯿﻨﯿﻢ ﮐﻪ H0 ﺩﺭ ﻫﻤﻪ ﻣﻮﺍﺭﺩ ﺩﺭ ﺳﻄﺢ ﻣﻌﻨﯽ ﺩﺍﺭﯼ 0.05 ﺭﺩ

ﻣﯽ ﺷﻮﺩ،ﮐﻪ ﻧﺸﺎﻥ ﻣﯽ ﺩﻫﺪ ﮐﻪ

ﺷﮑﻞ.8ﻧﺘﺎﯾﺞ ﺧﻮﺷﻪ ﺑﻨﺪﯼ ﺍﺯﺗﺮﮐﯿﺐ. )ﺍﻟﻒ( ﻧﺘﯿﺠﻪ ﺧﻮﺷﻪ ﺑﻨﺪﯼ .ICKDC )ﺏ( ﻧﺘﯿﺠﻪ ﺧﻮﺷﻪ ﺑﻨﺪﯼ .DPC )ﺝ( ﻧﺘﯿﺠﻪ ﺧﻮﺷﻪ ﺑﻨﺪﯼ .DBSCAN )ﺩ( ﻧﺘﯿﺠﻪ ﺧﻮﺷﻪ ﺑﻨﺪﯼ .K-means )ﻩ( ﻧﺘﯿﺠﻪ ﺧﻮﺷﻪ ﺑﻨﺪﯼ .Meanshift )ﻭ( ﻧﺘﯿﺠﻪ ﺧﻮﺷﻪ ﺑﻨﺪﯼ

.AGNES

ﺷﮑﻞ.9ﻧﺘﺎﯾﺞ ﺧﻮﺷﻪ ﺑﻨﺪﯼ ﺍﺯ.*D31* )ﺍﻟﻒ( ﻧﺘﯿﺠﻪ ﺧﻮﺷﻪ ﺑﻨﺪﯼ .ICKDC )ﺏ( ﻧﺘﯿﺠﻪ ﺧﻮﺷﻪ ﺑﻨﺪﯼ .DPC )ﺝ( ﻧﺘﯿﺠﻪ ﺧﻮﺷﻪ ﺑﻨﺪﯼ .DBSCAN )ﺩ( ﻧﺘﯿﺠﻪ ﺧﻮﺷﻪ ﺑﻨﺪﯼ .K-means )ﻩ( ﻧﺘﯿﺠﻪ ﺧﻮﺷﻪ ﺑﻨﺪﯼ .Meanshift )ﻭ( ﻧﺘﯿﺠﻪ ﺧﻮﺷﻪ ﺑﻨﺪﯼ

.AGNES

ﺷﮑﻞ.10ﻧﺘﺎﯾﺞ ﺧﻮﺷﻪ ﺑﻨﺪﯼ ﺍﺯ.*S2* )ﺍﻟﻒ( ﻧﺘﯿﺠﻪ ﺧﻮﺷﻪ ﺑﻨﺪﯼ .ICKDC )ﺏ( ﻧﺘﯿﺠﻪ ﺧﻮﺷﻪ ﺑﻨﺪﯼ .DPC )ﺝ( ﻧﺘﯿﺠﻪ ﺧﻮﺷﻪ ﺑﻨﺪﯼ .DBSCAN )ﺩ( ﻧﺘﯿﺠﻪ ﺧﻮﺷﻪ ﺑﻨﺪﯼ .K-means )ﻩ( ﻧﺘﯿﺠﻪ ﺧﻮﺷﻪ ﺑﻨﺪﯼ .Meanshift

ﺟﺪﻭﻝ5

ﻣﻘﺎﯾﺴﻪﺑﯿﻦ ﺍﻧﻮﺍﻉ ICKDC ﻭ DPC ﺩﺭ ﻣﺠﻤﻮﻋﻪ ﺩﺍﺩﻩ ﻫﺎﯼ ﺩﻧﯿﺎﯼ ﻭﺍﻗﻌﯽ.

|  |  |  |
| --- | --- | --- |
|  | ﻣﻌﯿﺎﺭﻫﺎﯼﺍﺭﺯﯾﺎﺑﯽ | ﺍﻟﮕﻮﺭﯾﺘﻢ |
| FMI | AMI | ARIﺷﺮﺍﺏ | FMI | AMI | ARIﺩﺍﻧﻪ | FMI | AMI | ARIﻋﻨﺒﯿﻪ |  |
| 0.878 | 0.807 | 0.816 | 0.852 | 0.740 | 0.778 | 0.900 | 0.829 | 0.851 | ICKDC |
| 0.8670.8010.731 | 0.8040.7230.656 | 0.7990.6990.589 | 0.8280.8260.852 | 0.6970.6870.752 | 0.7420.7400.778 | 0.9360.7590.900 | 0.8830.7090.835 | 0.9040.6250.851 | FKNN-DPCDPC-KNN IVDPC |
|  |  | ﺗﻘﺴﯿﻢﺑﻨﺪﯼ |  |  | WDBC |  |  | ﺍﻟﯿﮑﻮ |  |
| 0.676 | 0.684 | 0.624 | 0.897 | 0.678 | 0.773 | 0.796 | 0.651 | 0.701 | ICKDC |

0.558

0.588

0.625

0.828

0.658

0.814

0.764

0.583

0.673

0.675

0.325

0.073

0.247

0.095

0.437

0.453

0.553

ﭘﺎﺭﮐﯿﻨﺴﻮﻥﻫﺎ

0.456

0.160

0.269

0.207

0.766

0.857

0.864

0.644

0.584

0.643

0.610

0.356

0.602

0.586

0.294

0.131

0.137

0.139

0.401

0.703

0.707

ﯾﻮﻧﻮﺷﯿﺮ

0.263

0.132

0.218

0.187

0.692

0.831

0.738

0.820

0.850

0.782

0.790

0.476

0.695

0.543

0.840

0.836

0.849

0.721

0.554

0.756

0.644

ﭘﻮﺳﺖ

0.765

0.813

0.725

0.730

FKNN-DPC DPC-KNN IVDPC

ICKDC FKNN-DPC DPC-KNN IVDPC

.4.4ﻣﻮﺭﺩ ﺳﻮﻡ: ﻣﻘﺎﯾﺴﻪ ﺑﺎ ﺭﻭﺵ ﻫﺎﯼ ﺧﻮﺷﻪ ﺑﻨﺪﯼ ﻧﻈﺎﺭﺕ ﺷﺪﻩ

ﮐﻪﺩﺭﺑﻮﺭﺗﻮﻟﻮﺗﯽ ﻭ ﻫﻤﮑﺎﺭﺍﻥ2021)ﻧﻮﯾﺴﻨﺪﮔﺎﻥ ﯾﮏ ﺍﺳﺘﺮﺍﺗﮋﯼ ﺧﻮﺷﻪ ﺑﻨﺪﯼ ﺟﺪﯾﺪ ﺑﺎ

ﻧﺘﯿﺠﻪ

ﺍﺭﺯﺵﺑﺤﺮﺍﻧﯽ

ﺟﺪﻭﻝ6

ﻧﺘﺎﯾﺞﺁﺯﻣﻮﻥ ﺭﺗﺒﻪ ﺍﻣﻀﺎ ﺷﺪﻩ ﻭﯾﻠﮑﺎﮐﺴﻮﻥ.

ﻧﻈﺎﺭﺕﺩﻭ ﻣﺮﺣﻠﻪ ﺍﯼ ﺭﺍ ﭘﯿﺸﻨﻬﺎﺩ ﮐﺮﺩﻧﺪ. ﺩﺭ ﻣﺮﺣﻠﻪ ﺁﻓﻼﯾﻦ، k-means ﻧﻈﺎﺭﺕ ﺷﺪﻩ ﺍﺻﻼﺡ ﺷﺪﻩﺑﺮ ﺭﻭﯼ ﺗﻌﺪﺍﺩ ﮐﻤﯽ ﺍﺯ ﻣﺠﻤﻮﻋﻪ ﺩﺍﺩﻩ ﻫﺎﯼ ﺑﺮﭼﺴﺐ ﮔﺬﺍﺭﯼ ﺷﺪﻩ ﺑﺮﺍﯼ ﺳﺎﺧﺖ

ﺧﻮﺷﻪ ﻫﺎﯼﺍﻭﻟﯿﻪ ﺍﻧﺠﺎﻡ ﻣﯽ ﺷﻮﺩ، ﮐﻪ ﺗﺎ ﺣﺪﻭﺩﯼ ﺷﺒﯿﻪ ﺑﻪ ﺳﺎﺧﺖ ﻣﻨﺎﻃﻖ ﺍﺻﻠﯽ ﺩﺭ ICKDC ﺍﺳﺖ.ﺩﺭ ﺣﺎﻟﯽ ﮐﻪ ﺩﺭ ﻓﺎﺯ ﺁﻧﻼﯾﻦ ﺍﺳﺖ، ﺍﺯ ﯾﮏ k-means ﻧﯿﻤﻪ ﻧﻈﺎﺭﺗﯽ ﺍﻓﺰﺍﯾﺸﯽ ﺑﺮﺍﯼ

ﯾﺎﺩﮔﯿﺮﯼﺍﺯ ﻧﻤﻮﻧﻪ ﻫﺎﯼ ﺟﺪﯾﺪ ﻭ ﺑﻪ ﺭﻭﺯ ﺭﺳﺎﻧﯽ ﺳﺎﺧﺘﺎﺭ ﺍﻭﻟﯿﻪ ﺧﻮﺷﻪ ﻫﺎ ﺍﺳﺘﻔﺎﺩﻩ ﻣﯽ ﮐﻨﺪ.

ﺭﻭﺵﮐﺎﻣﻞ ISSKDEKMeans ﻧﺎﻣﯿﺪﻩ ﻣﯽ ﺷﻮﺩ ﺩﺭ ﺣﺎﻟﯽ ﮐﻪ ﺭﻭﺵ ﺑﺪﻭﻥ ﻓﺎﺯ ﺁﻧﻼﯾﻦ

ﺭﺩﺷﺪ ﺭﺩﺷﺪ ﺭﺩﺷﺪ ﺭﺩﺷﺪ ﺭﺩﺷﺪ ﺭﺩﺷﺪ ﺭﺩﺷﺪ

-ﻣﻘﺪﺍﺭ

35 19

41 3

41 5

30 3

8 0

8 7

8 4

ﺭﻭﺵ

DBSCAN

ﮎ-ﯾﻌﻨﯽ ﺗﻐﯿﯿﺮﻣﻌﻨﯽ

AGNES DPC FKNN-DPC DPC-KNN

SKDEKMeans ﻧﺎﻣﯿﺪﻩ ﻣﯽ ﺷﻮﺩ. ﺩﺭ ﺍﯾﻦ ﺑﺨﺶ، ﻋﻤﻠﮑﺮﺩ ICKDC ﻭ ﻫﺮ ﺩﻭ ﺭﻭﺵ

ﭘﯿﺸﻨﻬﺎﺩﯼﺭﺍ ﺑﺎ ﻫﻢ ﻣﻘﺎﯾﺴﻪ ﻣﯽ ﮐﻨﯿﻢﺑﻮﺭﺗﻮﻟﻮﺗﯽ ﻭ ﻫﻤﮑﺎﺭﺍﻥ(2021) ﺩﺭ ﻣﻮﺭﺩ ﺩﻗﺖ ﺧﻮﺷﻪ ﺑﻨﺪﯼ،ﮐﻪ ﺍﺳﺎﺳﺎ ًﻣﻘﺎﯾﺴﻪ ﺑﯿﻦ ICKDC ﻭ ﺭﻭﺵ ﻫﺎﯼ ﺧﻮﺷﻪ ﺑﻨﺪﯼ ﻧﻈﺎﺭﺕ ﺷﺪﻩ ﻣﻌﻤﻮﻟﯽ

ﺍﺳﺖ.

 ﺭﺩﺷﺪ 3 0 IVDPC

ﺡ) .DPC-KNN ﻭ ICKDC (ﺯ) .FKNN-DPC ﻭ DPC. )f( ICKDC ﻭ ICKDC (ﻩ) .AGNES ﻭ ICKDC (ﺩ) .Meanshift ﻭ ICKDC (ﺝ) .K-means ﻭ ICKDC (ﺏ) .DBSCAN ﻭ ICKDC (ﺍﻟﻒ) .ﺭﻗﺒﺎ ﺳﺎﯾﺮ ﻭ ICKDC ﺑﯿﻦ ﻣﻘﺎﯾﺴﻪ.11ﺷﮑﻞ

.IVDPC ﻭ ICKDC (

ﺑﺮﺭﻭﯼ ﺗﻌﯿﯿﻦ ﺧﻮﺩﮐﺎﺭ ﭘﺎﺭﺍﻣﺘﺮ ﺑﺮ ﺍﺳﺎﺱ ﺭﻭﺵ ﻫﺎﯼ ﺑﺪﻭﻥ ﻧﻈﺎﺭﺕ ﯾﺎ ﻧﯿﻤﻪ ﻧﻈﺎﺭﺕ ﺗﻤﺮﮐﺰ ﮐﻨﯿﺪ.

ICKDC ﻭ ﺭﻭﺵ ﻫﺎﯼ ﺧﻮﺷﻪ ﺑﻨﺪﯼ ﻧﻈﺎﺭﺕ ﺷﺪﻩ ﺍﺯ ﻧﻈﺮ ﺩﻗﺖ ﺧﻮﺷﻪ ﺑﻨﺪﯼ ﺩﺭ ﻣﺠﻤﻮﻋﻪ ﺩﺍﺩﻩ ﻫﺎﯼ

ﺟﺪﻭﻝ7

ﻣﻘﺎﯾﺴﻪﺑﯿﻦ

ﺩﻧﯿﺎﯼﻭﺍﻗﻌﯽ ﺍﺯ ﻣﺨﺰﻥ .UCI

ﺑﯿﺎﻧﯿﻪﻣﺸﺎﺭﮐﺖ ﻧﻮﯾﺴﻨﺪﻩ **CRediT**

ﺩﻗﺖ)٪(

ﺍﻟﮕﻮﺭﯾﺘﻢ

ﮊﻭﮊﻭ:ﻣﻔﻬﻮﻡ ﺳﺎﺯﯼ، ﺭﻭﺵ ﺷﻨﺎﺳﯽ، ﻧﮕﺎﺭﺵ – ﭘﯿﺶ ﻧﻮﯾﺲ ﺍﺻﻠﯽ، ﻧﮕﺎﺭﺵ – ﺑﺮﺭﺳﯽ ﻭ

ﻭﯾﺮﺍﯾﺶ**Si:**. **Gangquan**ﻧﻈﺎﺭﺕ، ﻣﺪﯾﺮﯾﺖ ﭘﺮﻭﮊﻩ، ﮐﺴﺐ ﻣﻨﺎﺑﻊ ﻣﺎﻟﯽ.ﻫﺎﺉﻮﺩﻭﻧﮓ ﺳﺎﻥ:

ﻣﺪﯾﺮﯾﺖﺩﺍﺩﻩ، ﻧﺮﻡ ﺍﻓﺰﺍﺭ.ﮐﺎﯼ ﮐﻮ:ﺍﻋﺘﺒﺎﺭ ﺳﻨﺠﯽ.ﻭﺍﯾﭽﻨﮓ ﻫﻮ:ﺗﺤﻘﯿﻖ، ﺗﺤﻠﯿﻞ ﺭﺳﻤﯽ.

ﺑﻪﻣﻌﻨﯽ ISSKDEK

93.51

74.62

79.71

94.33

93.85

ﺑﻪﻣﻌﻨﯽ SKDEK

94.08

73.68

81.21

94.33

92.76

ICKDC

94.02

79.40

82.62

92.00

88.28

ﺳﺮﻃﺎﻥ ﻗﻠﺐ ﯾﻮﻧﻮﺷﯿﺮ

ﻋﻨﺒﯿﻪ

ﺭﺍﯼ

ﺍﻋﻼﻣﯿﻪﻣﻨﺎﻓﻊ ﺭﻗﺎﺑﺘﯽ

ﻧﻮﯾﺴﻨﺪﮔﺎﻥﺍﻋﻼﻡ ﻣﯽ ﮐﻨﻨﺪ ﮐﻪ ﻫﯿﭻ ﻣﻨﺎﻓﻊ ﻣﺎﻟﯽ ﺭﻗﯿﺐ ﯾﺎ ﺭﻭﺍﺑﻂ ﺷﺨﺼﯽ ﺷﻨﺎﺧﺘﻪ ﺷﺪﻩ

ﺍﯼﻧﺪﺍﺭﻧﺪ ﮐﻪ ﺑﻪ ﻧﻈﺮ ﻣﯽ ﺭﺳﺪ ﺑﺮ ﮐﺎﺭ ﮔﺰﺍﺭﺵ ﺷﺪﻩ ﺩﺭ ﺍﯾﻦ ﻣﻘﺎﻟﻪ ﺗﺄﺛﯿﺮ [ﺑﮕﺬﺍﺭﺩ.](#_bookmark25)

ﻗﺪﺭﺩﺍﻧﯽ

ﺍﯾﻦﮐﺎﺭ ﺗﺎ ﺣﺪﯼ ﺗﻮﺳﻂ ﺑﺮﻧﺎﻣﻪ ﺗﺤﻘﯿﻖ ﻭ ﺗﻮﺳﻌﻪ ﮐﻠﯿﺪ ﻣﻠﯽ، ﭼﯿﻦ (2018AAA0101501)

ﻭﺑﻨﯿﺎﺩ ﻣﻠﯽ ﻋﻠﻮﻡ ﻃﺒﯿﻌﯽ ﭼﯿﻦ )ﺷﻤﺎﺭﻩ ﮔﺮﻧﺖ: (61304118 ﭘﺸﺘﯿﺒﺎﻧﯽ ﻣﯽ ﺷﻮﺩ.

ﻣﻨﺎﺑﻊ

ﺗﺸﺨﯿﺺﺟﺎﻣﻌﻪ ﻭ ﻣﺪﻝ ﻫﺎﯼ ﺑﻠﻮﮎ ﺗﺼﺎﺩﻓﯽ: ﺍﺧﯿﺮ )2017(. E. Abbe,

ﺗﺤﻮﻻﺕﻣﺠﻠﻪ ﺗﺤﻘﯿﻘﺎﺕ ﯾﺎﺩﮔﯿﺮﯼ ﻣﺎﺷﯿﻦ،18ﺍﭘﺘﯿﮏ: ﻧﻘﺎﻁ ﺳﻔﺎﺭﺵ )1999(. J. Sander,

)1(، 6446-6531. Ankerst, M., Breunig, MM, Kriegel, H.-P., &

ﺑﺮﺍﯼﺷﻨﺎﺳﺎﯾﯽ ﺳﺎﺧﺘﺎﺭ ﺧﻮﺷﻪ ﺑﻨﺪﯼ*Record Sigmod ACM*،(2)28، .49-60 ﺑﯽ،ﺁﺭ.، ﻣﺤﻤﻮﺩ، ﺭ.، ﺭﻭﺍﻥ، ﺱ.، ﺳﺎﻥ، ﯼ.، ﻭ ﺩﺍﻭﻭﺩ، ﺍﭺ .(2016) ﺧﻮﺷﻪ ﺑﻨﺪﯼ ﻓﺎﺯﯼ ﺗﻄﺒﯿﻘﯽ

ﺑﺎﺟﺴﺘﺠﻮﯼ ﺳﺮﯾﻊ ﻭ ﯾﺎﻓﺘﻦ ﻗﻠﻪ ﻫﺎﯼ ﭼﮕﺎﻟﯽ.ﻣﺤﺎﺳﺒﺎﺕ ﺷﺨﺼﯽ ﻭ ﻫﻤﻪ ﺟﺎ ﺣﺎﺿﺮ،(5)20، .785-793

ﺍﻟﮕﻮﺭﯾﺘﻤﯽﺑﺮﺍﯼ ﺧﻮﺷﻪ ﺑﻨﺪﯼ ﻣﮑﺎﻧﯽ-ﺯﻣﺎﻧﯽ ST-DBSCAN: )2007(. A. Kut, & D., Birant,

ﺩﺍﺩﻩﻫﺎ.ﻣﻬﻨﺪﺳﯽ ﺩﺍﺩﻩ ﻭ ﺩﺍﻧﺶ،(1)60، .208-221

Bondi, L., Lameri, S., Guera, D., Bestagini, P., Delp, EJ, & Tubaro, S. )2017(.

ﺗﺸﺨﯿﺺﺩﺳﺘﮑﺎﺭﯼ ﻭ ﻣﺤﻠﯽ ﺳﺎﺯﯼ ﺍﺯ ﻃﺮﯾﻖ ﺧﻮﺷﻪ ﺑﻨﺪﯼ ﻭﯾﮋﮔﯽ ﻫﺎﯼ CNN ﻣﺒﺘﻨﯽ ﺑﺮ ﺩﻭﺭﺑﯿﻦ. ﮐﻪ ﺩﺭ

ﮐﻨﻔﺮﺍﻧﺲ2017 *ieee* ﺩﺭ ﮐﺎﺭﮔﺎﻩ ﻫﺎﯼ ﺁﻣﻮﺯﺷﯽ ﺑﯿﻨﺎﯾﯽ ﮐﺎﻣﭙﯿﻮﺗﺮ ﻭ ﺗﺸﺨﯿﺺ ﺍﻟﮕﻮIEEE.*(cvprw)* .

ﺗﺮﺍﮐﻢﻫﺴﺘﻪ ﻧﻈﺎﺭﺕ ﺷﺪﻩ )2021(. PM ،Ciarelli & ،E. ،Oliveira de ،FD ،Bortoloti

ﺗﺨﻤﯿﻦ.K-meansﺳﯿﺴﺘﻢ ﻫﺎﯼ ﺧﺒﺮﻩ ﺑﺎ ﺑﺮﻧﺎﻣﻪ ﻫﺎﯼ ﮐﺎﺭﺑﺮﺩﯼ،168، ﻣﺎﺩﻩ .114350 )2009( PM Pardalos,

& WA, Chaovalitwongse, S., .Butenko,ﭼﺎﻟﺶ ﻫﺎﯼ ﺧﻮﺷﻪ ﺑﻨﺪﯼ ﺩﺭ

ﺷﺒﮑﻪﻫﺎﯼ ﺑﯿﻮﻟﻮﮊﯾﮑﯽ. ﻋﻠﻤﯽ ﺟﻬﺎﻧﯽ

 97.78 96.11 93.82 ﺷﺮﺍﺏ

ﻣﻘﺎﯾﺴﻪﻣﺠﻤﻮﻋﻪ ﺩﺍﺩﻩ ﻫﺎﯼ ﺩﻧﯿﺎﯼ ﻭﺍﻗﻌﯽ ﺍﺯ UCI )ﺑﺮﺧﯽ ﻣﺘﻔﺎﻭﺕ ﺍﺯ ﺁﻧﭽﻪ ﺩﺭ ﺩﻭ ﻣﻮﺭﺩ ﺍﻭﻝﺑﻪ ﻣﻨﻈﻮﺭ ﻣﻘﺎﯾﺴﻪ ﺍﺳﺘﻔﺎﺩﻩ ﮐﺮﺩﯾﻢ( ﺩﺭﺟﺪﻭﻝ .7 ﻫﻤﺎﻧﻄﻮﺭ ﮐﻪ ﻣﯽ ﺑﯿﻨﯿﻢ، ﺳﺎﺧﺖ ﺧﻮﺷﻪ

ﻫﺎﯼﺍﻭﻟﯿﻪ ﺍﺯ ﻃﺮﯾﻖ ﯾﺎﺩﮔﯿﺮﯼ ﻧﻈﺎﺭﺕ ﺷﺪﻩ ﻣﻨﺠﺮ ﺑﻪ ﻧﺘﺎﯾﺞ ﺧﻮﺷﻪ ﺑﻨﺪﯼ ﺩﻗﯿﻖ ﺗﺮﯼ ﺑﻪ ﻃﻮﺭ

ﮐﻠﯽﻣﯽ ﺷﻮﺩ. ISSKDEKMeans ﺑﻪ ﺩﻟﯿﻞ ﺑﻪ ﺭﻭﺯ ﺭﺳﺎﻧﯽ ﺗﺪﺭﯾﺠﯽ ﺳﺎﺧﺘﺎﺭ ﺧﻮﺷﻪ ﻫﺎ ﺑﻬﺘﺮﯾﻦ ﻋﻤﻠﮑﺮﺩﺭﺍ ﻧﺸﺎﻥ ﻣﯽ ﺩﻫﺪ. ﻋﻼﻭﻩ ﺑﺮ ﺍﯾﻦ، ICKDC ﺩﺭ ﺍﮐﺜﺮ ﻣﻮﺍﺭﺩ ﻧﺘﺎﯾﺞ ﺭﻗﺎﺑﺘﯽ ﺑﻪ ﺩﺳﺖ ﻣﯽ ﺁﻭﺭﺩﮐﻪ ﺍﺛﺮﺑﺨﺸﯽ ﻣﺮﺣﻠﻪ ﺷﻨﺎﺳﺎﯾﯽ ﻧﻘﺎﻁ ﺍﺻﻠﯽ ﺭﺍ ﺩﺭ ﺷﻨﺎﺳﺎﯾﯽ ﺗﻮﺯﯾﻊ ﺩﺍﺩﻩ ﻫﺎ ﻧﺸﺎﻥ

ﻣﯽ ﺩﻫﺪ.

.5ﻧﺘﯿﺠﻪ ﮔﯿﺮﯼ ﻫﺎ

ﺩﺭﺍﯾﻦ ﻣﻘﺎﻟﻪ، ﯾﮏ ﺭﻭﺵ ﺧﻮﺷﻪ ﺑﻨﺪﯼ ﻣﺒﺘﻨﯽ ﺑﺮ ﭼﮕﺎﻟﯽ ﻫﺴﺘﻪ KNN ﺟﺪﯾﺪ ICKDC [ﭘﯿﺸﻨﻬﺎﺩﺷﺪﻩ ﺍﺳﺖ. ﭼﺎﺭﭼﻮﺏ ﺟﺪﯾﺪ ﺳﻌﯽ ﻣﯽ ﮐﻨﺪ ﯾﮏ ﺧﻮﺷﻪ ﺭﺍ ﺑﺎ ﻣﺠﻤﻮﻋﻪ ﺍﯼ ﺍﺯ ﻧﻘﺎﻁ](http://refhub.elsevier.com/S0957-4174%2822%2900070-7/sb1) [ﻫﺴﺘﻪﺑﻪ ﺟﺎﯼ ﯾﮏ ﺷﯽ ﻭﺍﺣﺪ ﺑﺮﺍﯼ ﻣﺮﮐﺰ ﻧﺸﺎﻥ ﺩﻫﺪ. ﺍﯾﻦ ﺑﺮﺍﯼ ﻃﺒﻘﻪ ﺑﻨﺪﯼ ﻣﺠﻤﻮﻋﻪ ﻫﺎﯼ](http://refhub.elsevier.com/S0957-4174%2822%2900070-7/sb2)

[ﺩﺍﺩﻩﺑﺎ ﺗﻮﺯﯾﻊ ﭼﮕﺎﻟﯽ ﻧﺎﻫﻤﮕﻦ ﺣﯿﺎﺗﯽ ﺍﺳﺖ ﺯﯾﺮﺍ ﺳﺎﺧﺘﺎﺭ ﺧﻮﺷﻪ ﻫﺎ ﺑﻪ ﺍﻧﺪﺍﺯﻩ ﮐﺎﻓﯽ ﺁﺷﮑﺎﺭ](http://refhub.elsevier.com/S0957-4174%2822%2900070-7/sb2)

[ﺷﺪﻩ ﺍﻧﺪ.ﻋﻼﻭﻩ ﺑﺮ ﺍﯾﻦ، ﺍﯾﺪﻩ KNN ﻧﯿﺰ ﺩﺭ ﻓﺮﺁﯾﻨﺪ ﺗﺨﻤﯿﻦ ﭼﮕﺎﻟﯽ ﻣﻌﺮﻓﯽ ﺷﺪﻩ ﺍﺳﺖ ﮐﻪ ﺷﮑﻞ](http://refhub.elsevier.com/S0957-4174%2822%2900070-7/sb3) [ﻣﺘﺮﯾﮏﭼﮕﺎﻟﯽ ﺭﺍ ﺩﺭ DPC ﯾﮑﺴﺎﻥ ﻣﯽ ﮐﻨﺪ. ﺑﯿﺎﻥ ﺟﺪﯾﺪ ﻫﺴﺘﻪ KNN ﺍﻃﻼﻋﺎﺕ ﺗﻮﺯﯾﻊ](http://refhub.elsevier.com/S0957-4174%2822%2900070-7/sb3) ﭼﮕﺎﻟﯽ

[ﻣﺤﻠﯽﺍﺯ ﻧﻤﻮﻧﻪ ﻫﺎﯼ ﻣﺠﺎﻭﺭ ﺭﺍ ﺗﺮﮐﯿﺐ ﻣﯽ ﮐﻨﺪ ﻭ ﺗﻔﺎﻭﺕ ﺑﯿﻦ ﺍﯾﻦ ﺗﻮﺯﯾﻊ ﻫﺎ ﺭﺍ ﺑﺰﺭﮒ ﻣﯽ ﮐﻨﺪ.](http://refhub.elsevier.com/S0957-4174%2822%2900070-7/sb4)

[ﺁﺯﻣﺎﯾﺶ ﻫﺎﺑﺮ ﺭﻭﯼ ﻣﺠﻤﻮﻋﻪ ﺩﺍﺩﻩ ﻫﺎﯼ ﻣﺼﻨﻮﻋﯽ ﻭ ﺩﻧﯿﺎﯼ ﻭﺍﻗﻌﯽ ﺍﺯ](http://refhub.elsevier.com/S0957-4174%2822%2900070-7/sb4) UCI [ﻧﺸﺎﻥ ﻣﯽ ﺩﻫﺪ ﮐﻪ](http://refhub.elsevier.com/S0957-4174%2822%2900070-7/sb4)

[ICKDC ﻧﺴﺒﺖ ﺑﻪ ﺭﻭﺵ ﻫﺎﯼ ﺧﻮﺷﻪ ﺑﻨﺪﯼ ﭘﺮﮐﺎﺭﺑﺮﺩ ﻭ ﺍﻧﻮﺍﻉ ﺟﺪﯾﺪ DPC ﺑﺮﺗﺮﯼ ﺩﺍﺭﺩ.](http://refhub.elsevier.com/S0957-4174%2822%2900070-7/sb5)

[ﺩﺭﺁﺯﻣﻮﻥ ﻣﺎ، ﺗﻌﺪﺍﺩ ﮐﻪ ﻣﺤﺪﻭﺩﻩ ﻫﻤﺴﺎﯾﮕﯽ ﺭﺍ ﺩﺭ ICKDC ﺗﻌﯿﯿﻦ ﻣﯽ ﮐﻨﺪ، ﺑﺎ ﺩﺳﺖ ﺍﺯ](http://refhub.elsevier.com/S0957-4174%2822%2900070-7/sb7)

ﻗﺒﻞﻣﺸﺨﺺ ﺷﺪﻩ ﺍﺳﺖ. ﺁﺛﺎﺭ ﺑﻌﺪﯼ ﺧﻮﺍﻫﺪ ﺷﺪ

[ﻟﯿﺎﻧﮓ،ﺯﯼ، ﻭ ﭼﻦ، ﭘﯽ .(2016) ﺧﻮﺷﻪ ﺑﻨﺪﯼ ﻣﺒﺘﻨﯽ ﺑﺮ ﭼﮕﺎﻟﯽ ﺩﻟﺘﺎ ﺑﺎ ﺗﻘﺴﯿﻢ ﻭ ﻏﻠﺒﻪ](http://refhub.elsevier.com/S0957-4174%2822%2900070-7/sb8)

[ﺍﺳﺘﺮﺍﺗﮋﯼ:ﺧﻮﺷﻪ ﺑﻨﺪﯼ .3dcﻧﺎﻣﻪ ﻫﺎﯼ ﺗﺸﺨﯿﺺ ﺍﻟﮕﻮ،73، .52-59](http://refhub.elsevier.com/S0957-4174%2822%2900070-7/sb8) [ﻟﯿﻮ،ﺁﺭ.، ﻭﺍﻧﮓ، ﺍﭺ، ﻭ ﯾﻮ، ﺍﯾﮑﺲ .(2018) ﺧﻮﺷﻪ ﺑﻨﺪﯼ ﻣﺒﺘﻨﯽ ﺑﺮ ﻧﺰﺩﯾﮑﺘﺮﯾﻦ ﻫﻤﺴﺎﯾﻪ ﺗﻮﺳﻂ ﺳﺮﯾﻊ](http://refhub.elsevier.com/S0957-4174%2822%2900070-7/sb9)

[ﺟﺴﺘﺠﻮﻭ ﯾﺎﻓﺘﻦ ﻗﻠﻪ ﻫﺎﯼ ﭼﮕﺎﻟﯽﻋﻠﻮﻡ ﺍﻃﻼﻋﺎﺕ،450، .200-226](http://refhub.elsevier.com/S0957-4174%2822%2900070-7/sb9)

ﻟﻄﻔﯽ،ﻉ.، ﻣﺮﺍﺩﯼ، ﺹ، ﻭ ﺑﯿﮕﯽ، ﺡ. .(1399) ﭼﮕﺎﻟﯽ ﺑﺮ ﺍﺳﺎﺱ ﭼﮕﺎﻟﯽ ﺧﻮﺷﻪ ﺑﻨﺪﯼ ﻣﯽ ﺷﻮﺩ

[ﺳﺘﻮﻥﻓﻘﺮﺍﺕ ﻭ ﻣﺤﻠﻪ ﻓﺎﺯﯼ.ﺗﺸﺨﯿﺺ ﺍﻟﮕﻮ، ﻣﺎﺩﻩ .107449](http://refhub.elsevier.com/S0957-4174%2822%2900070-7/sb10)

[Ma, T., Wang, Y., Tang, M., Cao, J., Tian, Y., Al-Dhelaan, A., et al. )2016(. LED: A](http://refhub.elsevier.com/S0957-4174%2822%2900070-7/sb10)

ﺍﻟﮕﻮﺭﯾﺘﻢﺗﺸﺨﯿﺺ ﺟﻮﺍﻣﻊ [ﻫﻤﭙﻮﺷﺎﻧﯽ ﺳﺮﯾﻊ ﺑﺮ ﺍﺳﺎﺱ ﺧﻮﺷﻪ ﺑﻨﺪﯼ ﺳﺎﺧﺘﺎﺭﯼ. ﮐﺎﻣﭙﯿﻮﺗﺮﻫﺎﯼ ﻋﺼﺒﯽ،207،](http://refhub.elsevier.com/S0957-4174%2822%2900070-7/sb11)

[.488-500](http://refhub.elsevier.com/S0957-4174%2822%2900070-7/sb12)

ﻣﮏ[ﮐﻮﺉﯿﻦ، ﺟﯽ، ﻭ ﻫﻤﮑﺎﺭﺍﻥ. .(1967) ﭼﻨﺪ ﺭﻭﺵ ﺑﺮﺍﯼ ﻃﺒﻘﻪ ﺑﻨﺪﯼ ﻭ ﺗﺠﺰﯾﻪ ﻭ ﺗﺤﻠﯿﻞ ﭼﻨﺪ ﻣﺘﻐﯿﺮﻩ](http://refhub.elsevier.com/S0957-4174%2822%2900070-7/sb12) [ﻣﺸﺎﻫﺪﺍﺕﮐﻪ ﺩﺭﻣﺠﻤﻮﻋﻪ ﻣﻘﺎﻻﺕ ﭘﻨﺠﻤﯿﻦ ﺳﻤﭙﻮﺯﯾﻮﻡ ﺑﺮﮐﻠﯽ ﺩﺭ ﻣﻮﺭﺩ ﺁﻣﺎﺭ ﻭ ﺍﺣﺘﻤﺎﻝ ﺭﯾﺎﺿﯽ، ﺟﻠﺪ. )1ﺻﺺ](http://refhub.elsevier.com/S0957-4174%2822%2900070-7/sb13)

[.(297-281 ﺍﻭﮐﻠﻨﺪ، ﮐﺎﻟﯿﻔﺮﻧﯿﺎ، ﺍﯾﺎﻻﺕ ﻣﺘﺤﺪﻩ ﺁﻣﺮﯾﮑﺎ](http://refhub.elsevier.com/S0957-4174%2822%2900070-7/sb13) ﻣﺤﻤﻮﺩ،ﺭ.، ﮊﺍﻧﮓ، ﺟﯽ، ﺑﯽ، ﺁﺭ، ﺩﺍﻭﻭﺩ، ﺍﭺ، ﻭ ﺍﺣﻤﺪ، ﺍﭺ .(2016) ﺧﻮﺷﻪ ﺑﻨﺪﯼ ﺑﺎ ﺳﺮﻋﺖ

[ﺟﺴﺘﺠﻮﻭ ﯾﺎﻓﺘﻦ ﭘﯿﮏ ﻫﺎﯼ ﭼﮕﺎﻟﯽ ﺍﺯ ﻃﺮﯾﻖ ﺍﻧﺘﺸﺎﺭ ﮔﺮﻣﺎ.ﮐﺎﻣﭙﯿﻮﺗﺮﻫﺎﯼ ﻋﺼﺒﯽ،208، .210-217 ﺭﻭﺩﺭﯾﮕﺰ، ﺍ.، ﻭ](http://refhub.elsevier.com/S0957-4174%2822%2900070-7/sb14)

ﻻﯾﻮ،ﺍ. .(2014) ﺧﻮﺷﻪ ﺑﻨﺪﯼ ﺑﺎ ﺟﺴﺘﺠﻮﯼ ﺳﺮﯾﻊ ﻭ [ﯾﺎﻓﺘﻦ ﻗﻠﻪ ﻫﺎﯼ ﭼﮕﺎﻟﯽ.](http://refhub.elsevier.com/S0957-4174%2822%2900070-7/sb14)

[ﻋﻠﻮﻡﭘﺎﯾﻪ،(6191)344، .1492-1496](http://refhub.elsevier.com/S0957-4174%2822%2900070-7/sb15)

[ﺳﯿﺪﯼ،ﺱ.ﺍ.، ﻟﻄﻔﯽ، ﻉ.، ﻣﺮﺍﺩﯼ، ﺹ.، ﻭ ﻗﺎﺩﺭ، )1398( .NN ﻣﺒﺘﻨﯽ ﺑﺮ ﻧﻤﻮﺩﺍﺭ ﭘﻮﯾﺎ](http://refhub.elsevier.com/S0957-4174%2822%2900070-7/sb15)

[ﺍﻧﺘﺸﺎﺭﺑﺮﭼﺴﺐ ﺑﺮﺍﯼ ﺧﻮﺷﻪ ﺑﻨﺪﯼ ﻗﻠﻪ ﻫﺎﯼ ﭼﮕﺎﻟﯽ.ﺳﯿﺴﺘﻢ ﻫﺎﯼ ﺧﺒﺮﻩ ﺑﺎ ﺑﺮﻧﺎﻣﻪ ﻫﺎﯼ ﮐﺎﺭﺑﺮﺩﯼ،115،](http://refhub.elsevier.com/S0957-4174%2822%2900070-7/sb16)

.314-328

[ﺷﺎﻩ،ﺟﯽ ﺍﭺ، ﺑﻬﻨﺴﺪﺍﺩﯾﺎ، ﺳﯽ، ﻭ ﮔﺎﻧﺎﺗﺮﺍ، )2012( .AP ﯾﮏ ﺍﺭﺯﯾﺎﺑﯽ ﺗﺠﺮﺑﯽ ﺍﺯ](http://refhub.elsevier.com/S0957-4174%2822%2900070-7/sb17)

[ﺗﮑﻨﯿﮏﻫﺎﯼ ﺧﻮﺷﻪ ﺑﻨﺪﯼ ﻣﺒﺘﻨﯽ ﺑﺮ ﭼﮕﺎﻟﯽﻣﺠﻠﻪ ﺑﯿﻦ ﺍﻟﻤﻠﻠﯽ ﻣﺤﺎﺳﺒﺎﺕ ﻭ ﻣﻬﻨﺪﺳﯽ ﻧﺮﻡ *(IJSCE) ISSN*،](http://refhub.elsevier.com/S0957-4174%2822%2900070-7/sb17)

.216-223 ،22312307

[Xie, J., Gao, H., Xie, W., Liu, X., & Grant, PW )2016(. ﺗﺸﺨﯿﺺ ﺑﺎ ﻗﻮﯼ ﺧﻮﺷﻪﺑﻨﺪﯼ](http://refhub.elsevier.com/S0957-4174%2822%2900070-7/sb18)

[ﭘﯿﮏﻫﺎﯼ ﭼﮕﺎﻟﯽ ﻭ ﺗﺨﺼﯿﺺ ﻧﻘﺎﻁ ﺑﺮ ﺍﺳﺎﺱ ﻭﺯﻥ ﻓﺎﺯﯼ -kﻧﺰﺩﯾﮏ ﺗﺮﯾﻦ ﻫﻤﺴﺎﯾﻪ ﻫﺎ. ﻋﻠﻮﻡ ﺍﻃﻼﻋﺎﺕ،354](http://refhub.elsevier.com/S0957-4174%2822%2900070-7/sb18)،

.19-40

ﺳﻠﺴﻠﻪ[ﻣﺮﺍﺗﺒﯽ ﮐﺎﺭﺁﻣﺪ ﻣﺒﺘﻨﯽ ﺑﺮ ﺍﻭﺝ ﭼﮕﺎﻟﯽ Denpehc: )2016(. W. Deng, & G., Wang, J., Xu,](http://refhub.elsevier.com/S0957-4174%2822%2900070-7/sb19)

[ﺧﻮﺷﻪﺑﻨﺪﯼ ﻣﻮﺛﺮ )2016(. X. Chen, & L., Pan, J., Cheng, B., Wang, L., Li, M., Chen,](http://refhub.elsevier.com/S0957-4174%2822%2900070-7/sb26) ﺑﺎﯾﺎﻓﺘﻦ ﺳﺘﻮﻥ ﻓﻘﺮﺍﺕ [ﭼﮕﺎﻟﯽ ﺑﺮ ﺍﺳﺎﺱ .knnﺗﺸﺨﯿﺺ ﺍﻟﮕﻮ،60ﺑﺮﺭﺳﯽ ﺭﻭﺵ ﻫﺎﯼ ﺗﻘﺴﯿﻢ ﺑﻨﺪﯼ ﺗﺼﻮﯾﺮ](http://refhub.elsevier.com/S0957-4174%2822%2900070-7/sb26)

[، 486-498. Dhanachandra، N.، & Chanu، YJ )2017(.](http://refhub.elsevier.com/S0957-4174%2822%2900070-7/sb27)

ﺑﺎﺍﺳﺘﻔﺎﺩﻩ ﺍﺯ [ﺗﮑﻨﯿﮏ ﻫﺎﯼ ﺧﻮﺷﻪ ﺑﻨﺪﯼﻣﺠﻠﻪ ﺍﺭﻭﭘﺎﯾﯽ ﺗﺤﻘﯿﻘﺎﺕ ﻭ ﻋﻠﻮﻡ ﻣﻬﻨﺪﺳﯽ، (1)2، .15](http://refhub.elsevier.com/S0957-4174%2822%2900070-7/sb27)

ﺩﻭ،ﺍﻡ.، ﺩﯾﻨﮓ، [ﺍﺱ.، ﻭ ﺟﯿﺎ، ﺍﭺ .(2016) ﻣﻄﺎﻟﻌﻪ ﺧﻮﺷﻪ ﺑﻨﺪﯼ ﭘﯿﮏ ﻫﺎﯼ ﭼﮕﺎﻟﯽ ﺑﺮ ﺍﺳﺎﺱ -kﻧﺰﺩﯾﮏ ﺗﺮﯾﻦ](http://refhub.elsevier.com/S0957-4174%2822%2900070-7/sb28) [ﺗﺠﺰﯾﻪﻭ ﺗﺤﻠﯿﻞ ﻫﻤﺴﺎﯾﮕﺎﻥ ﻭ ﻣﺆﻟﻔﻪ ﻫﺎﯼ ﺍﺻﻠﯽﺳﯿﺴﺘﻢ ﻫﺎﯼ ﺩﺍﻧﺶ ﺑﻨﯿﺎﻥ،UCI.99 ﻣﺨﺰﻥ ﯾﺎﺩﮔﯿﺮﯼ ﻣﺎﺷﯿﻦ](http://refhub.elsevier.com/S0957-4174%2822%2900070-7/sb29)

[، 135-145. Dua, D., & Graff, C. )2017(.](http://refhub.elsevier.com/S0957-4174%2822%2900070-7/sb29)

ﺍﻟﺒﺎﺗﺎ،ﺍﻡ ﺗﯽ ﻭ ﺁﺷﻮﺭ، ﺩﺑﻠﯿﻮ ﺍﻡ .(2013) ﺭﻭﺷﯽ ﭘﻮﯾﺎ ﺑﺮﺍﯼ ﮐﺸﻒ ﭼﮕﺎﻟﯽ

[ﺧﻮﺷﻪﻫﺎﯼ ﻣﺘﻨﻮﻉﺭﻭﺷﯽ ﭘﻮﯾﺎ ﺑﺮﺍﯼ ﮐﺸﻒ ﺧﻮﺷﻪ ﻫﺎﯼ ﻣﺘﻨﻮﻉ ﭼﮕﺎﻟﯽ،.(1)6 ﺍﺳﺘﺮ، ﺍﻡ.، ﮐﺮﯾﮕﻞ، X ،Xu ،J. ،](http://refhub.elsevier.com/S0957-4174%2822%2900070-7/sb30)

[Sander ،.H.-P.، ﻭ ﻫﻤﮑﺎﺭﺍﻥ. .(1996) ﯾﮏ ﺍﻟﮕﻮﺭﯾﺘﻢ ﻣﺒﺘﻨﯽ ﺑﺮ ﭼﮕﺎﻟﯽ](http://refhub.elsevier.com/S0957-4174%2822%2900070-7/sb30)

ﺑﺮﺍﯼﮐﺸﻒ ﺧﻮﺷﻪ ﻫﺎ ﺩﺭ ﭘﺎﯾﮕﺎﻩ [ﺩﺍﺩﻩ ﻫﺎﯼ ﻓﻀﺎﯾﯽ ﺑﺰﺭﮒ ﺑﺎ ﻧﻮﯾﺰ. ﮐﻪ ﺩﺭﺟﻠﺪ. 96 ،)*Kdd*ﺻﺺ .(231-226](http://refhub.elsevier.com/S0957-4174%2822%2900070-7/sb30)

[ﺧﻮﺷﻪﺑﻨﺪﯼ ﺳﺮﯼ ﻫﺎﯼ ﺯﻣﺎﻧﯽ ﺍﺯ ﻃﺮﯾﻖ ﺗﺸﺨﯿﺺ ﺟﺎﻣﻌﻪ ﺩﺭ )2016(. L. ،Zhao ﻭ ،LN ،Ferreira](http://refhub.elsevier.com/S0957-4174%2822%2900070-7/sb31)

[ﺷﺒﮑﻪﻫﺎﯼ.ﻋﻠﻮﻡ ﺍﻃﻼﻋﺎﺕ،326، .227-242](http://refhub.elsevier.com/S0957-4174%2822%2900070-7/sb32)

ﭼﻨﺪﺧﻮﺷﻪ؟ ﮐﺪﺍﻡ ﺭﻭﺵ ﺧﻮﺷﻪ ﺑﻨﺪﯼ؟ )1998(. AE ،Raftery ﻭ ،C. [،Fraley](http://refhub.elsevier.com/S0957-4174%2822%2900070-7/sb32) [ﭘﺎﺳﺦ ﻫﺎﺍﺯ ﻃﺮﯾﻖ ﺗﺤﻠﯿﻞ ﺧﻮﺷﻪ ﺍﯼ ﻣﺒﺘﻨﯽ ﺑﺮ ﻣﺪﻝ.ﻣﺠﻠﻪ ﮐﺎﻣﭙﯿﻮﺗﺮ،(8)41، .578-588 ﻓﺮﺍﻧﺘﯽ، ﭘﯽ، ﻭ](http://refhub.elsevier.com/S0957-4174%2822%2900070-7/sb33)

[ﺳﯿﺮﺍﻧﻮﺟﺎ،ﺍﺱ. .(2018) -Kﺑﻪ ﻣﻌﻨﯽ ﻭﯾﮋﮔﯽ ﻫﺎ ﺩﺭ ﺷﺶ ﻣﻌﯿﺎﺭ ﺧﻮﺷﻪ ﺑﻨﺪﯼ ﺍﺳﺖ](http://refhub.elsevier.com/S0957-4174%2822%2900070-7/sb33)

ﻣﺠﻤﻮﻋﻪﺩﺍﺩﻩ ﻫﺎ

[ﺗﺨﻤﯿﻦﮔﺮﺍﺩﯾﺎﻥ ﯾﮏ ﭼﮕﺎﻟﯽ )1975(. L. Hostetler, & K., Fukunaga,](http://refhub.elsevier.com/S0957-4174%2822%2900070-7/sb34)

[ﺗﺎﺑﻊ،ﺑﺎ ﺑﺮﻧﺎﻣﻪ ﻫﺎﯼ ﮐﺎﺭﺑﺮﺩﯼ ﺩﺭ ﺗﺸﺨﯿﺺ ﺍﻟﮕﻮ.ﻣﻌﺎﻣﻼﺕ *IEEE* ﺩﺭ ﺗﺉﻮﺭﯼ ﺍﻃﻼﻋﺎﺕ،(1)21، .32-40](http://refhub.elsevier.com/S0957-4174%2822%2900070-7/sb34)

[Geng, Y.-a., Li, Q., Zheng, R., Zhuang, F., He, R., & Xiong, N. )2018(. ﺟﺪﯾﺪ:ﭘﯿﺸﻨﻬﺎﺩ](http://refhub.elsevier.com/S0957-4174%2822%2900070-7/sb35)

[ﺍﻟﮕﻮﺭﯾﺘﻢﺧﻮﺷﻪ ﺑﻨﺪﯼ ﻣﺒﺘﻨﯽ ﺑﺮ ﭼﮕﺎﻟﯽ ﺑﺎ ﺍﺳﺘﻔﺎﺩﻩ ﺍﺯ ﭼﮕﺎﻟﯽ ﻧﺴﺒﯽ ﻫﺴﺘﻪ .knnﻋﻠﻮﻡ ﺍﻃﻼﻋﺎﺕ،436،](http://refhub.elsevier.com/S0957-4174%2822%2900070-7/sb35) .13-30

[)2011(. M. Kamber, & J., Pei, J., Han,ﺩﺍﺩﻩ ﮐﺎﻭﯼ: ﻣﻔﺎﻫﯿﻢ ﻭ ﺗﮑﻨﯿﮏ ﻫﺎ. ﺍﻟﺰﻭﯾﺮ. ﻫﯿﻨﻬﺒﻮﺭﮒ، ﺍ.، ﻭ ﮔﺎﺑﺮﯾﻞ، ﺍﭺ.](http://refhub.elsevier.com/S0957-4174%2822%2900070-7/sb36)

[ﺭﻭﺷﻦ ﺍﺳﺖ](http://refhub.elsevier.com/S0957-4174%2822%2900070-7/sb20)

[ﺧﻮﺷﻪﺑﻨﺪﯼﻋﻠﻮﻡ ﺍﻃﻼﻋﺎﺕ،373، .200-218](http://refhub.elsevier.com/S0957-4174%2822%2900070-7/sb20)

[.ﺑﺮﺭﺳﯽﺍﻟﮕﻮﺭﯾﺘﻢ ﻫﺎﯼ ﺧﻮﺷﻪ ﺑﻨﺪﯼ )2005(. D. Wunsch, & R., Xu,ﻣﻌﺎﻣﻼﺕ *IEEE*](http://refhub.elsevier.com/S0957-4174%2822%2900070-7/sb20)

ﺷﺒﮑﻪﻫﺎﯼ ﻋﺼﺒﯽ،(3)16، .645-678

.(2007) 2.0 :Denclue ﺧﻮﺷﻪ ﺑﻨﺪﯼ ﺳﺮﯾﻊ ﺑﺮ [ﺍﺳﺎﺱ ﻫﺴﺘﻪ](http://refhub.elsevier.com/S0957-4174%2822%2900070-7/sb36)

[ﺗﺨﻤﯿﻦﭼﮕﺎﻟﯽ ﮐﻪ ﺩﺭﺳﻤﭙﻮﺯﯾﻮﻡ ﺑﯿﻦ ﺍﻟﻤﻠﻠﯽ ﺗﺤﻠﯿﻞ ﺩﺍﺩﻩ ﻫﺎﯼ ﻫﻮﺷﻤﻨﺪ)ﺻﺺ .(80-70 ﺍﺳﭙﺮﯾﻨﮕﺮ.](http://refhub.elsevier.com/S0957-4174%2822%2900070-7/sb37)

[ﺑﺮﺍﺳﺎﺱ ﺧﻮﺷﻪ ﺑﻨﺪﯼ ﭘﯿﮏ ﭼﮕﺎﻟﯽ ﺗﻄﺒﯿﻘﯽ )2017(. Y. Fang, & M., Zhengming, L., Yaohui,](http://refhub.elsevier.com/S0957-4174%2822%2900070-7/sb21)

ﺩﺭ-Kﻧﺰﺩﯾﮑﺘﺮﯾﻦ [ﻫﻤﺴﺎﯾﻪ ﺑﺎ ﺍﺳﺘﺮﺍﺗﮋﯼ ﺗﺠﻤﯿﻊ.ﺳﯿﺴﺘﻢ ﻫﺎﯼ ﺩﺍﻧﺶ ﺑﻨﯿﺎﻥ،133، .208-220](http://refhub.elsevier.com/S0957-4174%2822%2900070-7/sb21)

ﯾﻮ،ﺩﯼ، ﻟﯿﻮ، ﺟﯽ، ﮔﻮﺉﻮ، ﺍﻡ.، ﻟﯿﻮ، ﺍﯾﮑﺲ، ﻭ ﯾﺎﺉﻮ، ﺍﺱ. [.](http://refhub.elsevier.com/S0957-4174%2822%2900070-7/sb22)(2019) [ﺍﻭﺝ ﭼﮕﺎﻟﯽ ﺧﻮﺷﻪ ﺑﻨﺪﯼ ﺑﺮ ﺍﺳﺎﺱ](http://refhub.elsevier.com/S0957-4174%2822%2900070-7/sb22)

[ﺗﻮﺍﻟﯽﭼﮕﺎﻟﯽ ﻣﺤﻠﯽ ﻭﺯﻥ ﺩﺍﺭ ﻭ ﻧﺰﺩﯾﮑﺘﺮﯾﻦ ﻫﻤﺴﺎﯾﻪ.ﺩﺳﺘﺮﺳﯽ *IEEE*،7، .34317–34301](http://refhub.elsevier.com/S0957-4174%2822%2900070-7/sb23)

ﺧﻮﺷﻪﺑﻨﺪﯼ [ﻗﻮﯼ ﺑﺎ ﺷﻨﺎﺳﺎﯾﯽ )2018(. K. ،Zheng ﻭ ،Y. ،Zhang ،G. ،Si ،Z. ،Zhou](http://refhub.elsevier.com/S0957-4174%2822%2900070-7/sb23)

[ﺭﮔﻪﻫﺎﯼ ﺧﻮﺷﻪ ﺑﺮ ﺍﺳﺎﺱ ﺗﺨﻤﯿﻦ ﺗﺮﺍﮐﻢ ﻫﺴﺘﻪﺳﯿﺴﺘﻢ ﻫﺎﯼ ﺩﺍﻧﺶ ﺑﻨﯿﺎﻥ،159، .320–309](http://refhub.elsevier.com/S0957-4174%2822%2900070-7/sb24)

[Zou, Q., Lin, G., Jiang, X., Liu, X., & Zeng, X. )2018(. ﺩﺭ ﺗﻮﺍﻟﯽ ﺧﻮﺷﻪﺑﻨﺪﯼ](http://refhub.elsevier.com/S0957-4174%2822%2900070-7/sb25)

[ﺑﯿﻮﺍﻧﻔﻮﺭﻣﺎﺗﯿﮏ:ﯾﮏ ﻣﻄﺎﻟﻌﻪ ﺗﺠﺮﺑﯽﺟﻠﺴﺎﺕ ﺗﻮﺟﯿﻬﯽ ﺩﺭ ﺑﯿﻮﺍﻧﻔﻮﺭﻣﺎﺗﯿﮏ.](http://refhub.elsevier.com/S0957-4174%2822%2900070-7/sb25)

[ﯾﮏﺭﻭﯾﮑﺮﺩ ﮐﻠﯽ ﺑﺮﺍﯼ ﺧﻮﺷﻪ ﺑﻨﺪﯼ ﺩﺭ ﮐﻞ )2003(. DA ،Keim & ،A. ،Hinneburg](http://refhub.elsevier.com/S0957-4174%2822%2900070-7/sb38)

[ﭘﺎﯾﮕﺎﻩﻫﺎﯼ ﺩﺍﺩﻩ ﺑﺎ ﻧﻮﯾﺰﺩﺍﻧﺶ ﻭ ﺳﯿﺴﺘﻢ ﻫﺎﯼ ﺍﻃﻼﻋﺎﺗﯽ،(4)5، .387-415 ﺟﯿﻦ، AK، ﻣﻮﺭﺗﯽ، ﺍﻡ.ﺍﻥ، ﻭ](http://refhub.elsevier.com/S0957-4174%2822%2900070-7/sb38) ﻓﻠﯿﻦ،

ﭘﯽﺟﯽ .(1999) ﺧﻮﺷﻪ ﺑﻨﺪﯼ ﺩﺍﺩﻩ ﻫﺎ: ﯾﮏ ﺑﺮﺭﺳﯽ*ACM*

[ﺑﺮﺭﺳﯽﻫﺎﯼ ﻣﺤﺎﺳﺒﺎﺗﯽ،(3)31، .264-323](http://refhub.elsevier.com/S0957-4174%2822%2900070-7/sb39)

[ﻧﺎﻫﻨﺠﺎﺭﯼﻣﺒﺘﻨﯽ ﺑﺮ ﺧﻮﺷﻪ ﺑﻨﺪﯼ ﺩﺍﺩﻩ ﻫﺎ )2014(. G. Sebestyén, & P., Haller, B., Genge, I., Kiss,](http://refhub.elsevier.com/S0957-4174%2822%2900070-7/sb39)

ﺗﺸﺨﯿﺺﺩﺭ ﺳﯿﺴﺘﻢ ﻫﺎﯼ ﮐﻨﺘﺮﻝ ﺻﻨﻌﺘﯽ ﮐﻪ ﺩﺭ*)iccp(* ﺩﻫﻤﯿﻦ ﮐﻨﻔﺮﺍﻧﺲ ﺑﯿﻦ ﺍﻟﻤﻠﻠﯽ [ﺍﺭﺗﺒﺎﻃﺎﺕ ﻭ ﭘﺮﺩﺍﺯﺵ](http://refhub.elsevier.com/S0957-4174%2822%2900070-7/sb39)

[ﮐﺎﻣﭙﯿﻮﺗﺮﯼﻫﻮﺷﻤﻨﺪ *ieee* )2014ﺹ .(281-275 )2011( A. Zimek,](http://refhub.elsevier.com/S0957-4174%2822%2900070-7/sb40)

[ﻭﺍﯾﻠﯽ](http://refhub.elsevier.com/S0957-4174%2822%2900070-7/sb40)

[& J., Sander, P., Kröger, H.-P., Kriegel, .IEEE. ﺧﻮﺷﻪ ﺑﻨﺪﯼ ﻣﺒﺘﻨﯽ ﺑﺮ ﭼﮕﺎﻟﯽ](http://refhub.elsevier.com/S0957-4174%2822%2900070-7/sb40)

ﺑﺮﺭﺳﯽﻫﺎﯼ ﻣﯿﺎﻥ ﺭﺷﺘﻪ ﺍﯼ: ﺩﺍﺩﻩ ﮐﺎﻭﯼ ﻭ ﮐﺸﻒ ﺩﺍﻧﺶ،O.1 ،Piven ﻭ ،L. ،Myrutenko ،[V. ،](http://refhub.elsevier.com/S0957-4174%2822%2900070-7/sb40)

[)3(، 231-240. Lakhno، V.، Kravchuk، P.، Malyukov، V.، Domrachev](http://refhub.elsevier.com/S0957-4174%2822%2900070-7/sb41)

[ﻣﺠﻠﻪ](http://refhub.elsevier.com/S0957-4174%2822%2900070-7/sb41)

.(2017) ﺗﻮﺳﻌﻪ ﺳﯿﺴﺘﻢ [ﺍﻣﻨﯿﺖ ﺳﺎﯾﺒﺮﯼ ﻣﺒﺘﻨﯽ ﺑﺮ ﺧﻮﺷﻪ ﺑﻨﺪﯼ ﻭ ﺗﺸﮑﯿﻞ ﻋﻼﺉﻢ ﺍﻧﺤﺮﺍﻑ ﮐﻨﺘﺮﻟﯽ.](http://refhub.elsevier.com/S0957-4174%2822%2900070-7/sb41)

ﻓﻨﺎﻭﺭﯼﺍﻃﻼﻋﺎﺕ ﻧﻈﺮﯼ ﻭ ﮐﺎﺭﺑﺮﺩﯼ، (21)95، .5778-5786