#### Main Research method: *DBSCAN Algorithm*

*Your full name, ID:* **Nguyễn Thành Thiện – ITITIU14089**

**1.** Summarize reviewed relevant papers in the below table. The reviewed papers should have the same purpose and context/application. Write references in the Sources box. Methods, datasets, evaluation and improvement might be explained/presented corresponding to the reviewed papers.

|  |  |
| --- | --- |
| **Purpose** |  |
| **Context & Applications** |  |
| **Reasons/methods** | \* Uses 2 parameters: neighborhood (ε) and minimum points (minPts).  It starts with an arbitrary starting point that has not been visited.  • A point’s ε is scanned for other points.  + If ε contains minPts, the center point and all points within ε become a cluster.  + Else, the scanned point becomes noise.Note that this point might later be found in a sufficiently sized ε-environment of a different point and hence be made part of a cluster.  + Continues until the density-connected cluster is completely found.  • This model is called Density Reachability.  • Works well for unknown data.  \* Summarize implement:  for object i  if the i-th object is not a member of a given cluster yet  create a new cluster C  while the neighboring objects satisfy the cluster condition  expand cluster C  end  end  end |
| **Evidence/datasets** |  |
| **Evaluation (Pros & Cons)** | \* Pros: + Given number of clusters not needed.  + Can form unusual shapes.  + Mostly insensitive to order.  \* Cons: + It relies on Eps and minPts.  + It has issues with varying densities.  + Sometimes has issues separating nearby clusters properly. |
| **Improvement** |  |
| **Sources** | <https://towardsdatascience.com/how-dbscan-works-and-why-should-i-use-it-443b4a191c80>  <https://www.researchgate.net/publication/287527085_Fuzzy_Core_DBScan_Clustering_Algorithm>  <https://github.com/kelvins/DBSCAN>  <https://www.researchgate.net/publication/257035378_Revised_DBSCAN_algorithm_to_cluster_data_with_dense_adjacent_clusters> |

**2.** Apply the review methods to a recommender system using the Book-crossing dataset. Present the proposed framework/ flowchart and explanation in 1 page.

[10pt font, single line]