

Task-2: Build an ontology for a DM sub-process

Task: Build an ontology for a DM sub-process based on the concept “semantic meta mining”.

1. Use Protégé: <https://protege.stanford.edu/products.php>
2. The simplified schema is shown in Figure 1.

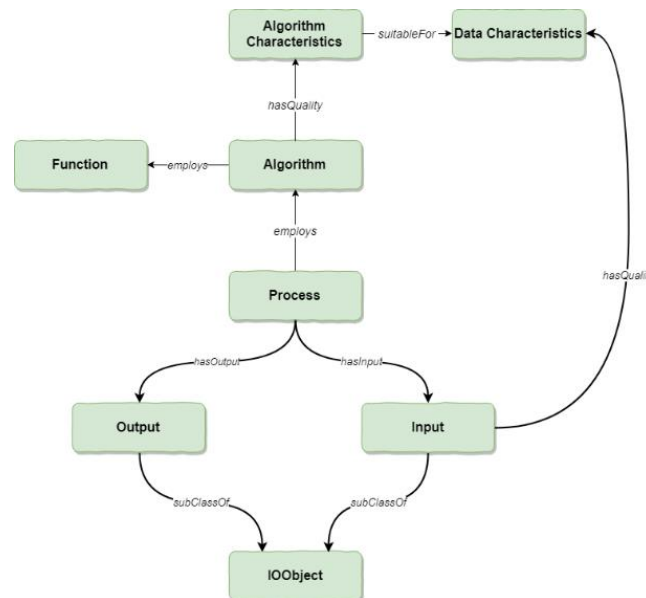


Figure 1. The schema of the ontology for a DM sub-process

3. Choose one DM sub-process as your variant. A possible way is to follow the CRISP-DM model. Or choose one of following sub-processes:
 - a) Classification
 - b) Clustering
 - c) Processing missing values
 - d) Discretization
 - e) Normalization
 - f) Any other sub-process in DM
4. Collect the related materials to find the algorithm characteristics and the corresponding data characteristics.
 - a) Present the output and input of the sub-process.
 - b) List the available algorithms.
 - c) Describe how to choose the algorithm according to the data characteristics.
 - d) List the key components of the algorithms.
5. Write a report which includes:
 - a) The description of the chosen sub-process.
 - b) The references where you collect the knowledge about this sub-process.
 - c) The design of your ontology.
 - d) The screenshot of the DL query results in your ontology.

Requirements:

1. In English, too.
2. Before 07.05
3. A Zip file including the report and the ontology.
4. Send to mantx@itmo.ru with subject form: ("Task-2"+your group+your name.)