The Human in The Loop Clustering With Knowledge Augmentations (HuLCKA)

- HuLCKA project is funded by ID.UJ (U1U/P06/NO/02.16)
- **Project Leader:** dr inż. Szymon Bobek
- **Start time:** 15.11.2021
- **Duration:** 12 months

Objectives

The Human in The Loop Clustering With Knowledge Augmentations (\HULCKA) project aims at investigating methods for combining background knowledge with clustering algorithms and eXplainable Artificial Intelligence (XAI) methods in order to provide comprehensive framework for human-in-the loop data analysis. The preliminary results shown the feasibility of the approach for low dimensional, tabular data. The main objective of HuLCKA is to extend it with diverse knowledge augmentation methods, explanation mechanisms and data sources and enclose whole solution into a generic framework. The project is relevant to the DigiWorld as it concerns research in the areas of Artificial Intelligence (AI), explainable AI, data mining and machine learning and their applications and customisation to exact and natural sciences.

Research Hypotheses, and Innovativeness of the Project

Clustering aims at unfolding hidden patterns in data to discover similar instances and group them under common cluster labels. This task is often performed to either discover unknown groups, to automate the process of discovering possibly known groups or for segmentation of data points into arbitrary number of segments. Either of the above can be done in unsupervised, semi-supervised or supervised manner which depends on the availability of the prior knowledge and the ability of deriving new knowledge based on partial data and human interaction. In many practical applications, prior knowledge is available for machine learning algorithms to be utilised. However, incorporating it into the statistical learning pipeline is non trivial task and has been a matter of study for decades.

Objectives

Work within a project will be divided into four work packages with time dependencies depicted in Figure below:

- WP1 – Knowledge encoding methods and augmentation with clustering algorithms.
- WP2 – Custom explanation mechanisms exploiting nature of the data and type of background knowledge.

- WP3 – Optimisation of the guided process of merges and splits.
- WP4 – Application and evaluation of results to different types of data, such as tabular data, time series and possibly images.

From:
https://www.geist.re/ - **GEIST Research Group**

Permanent link:
https://www.geist.re/pub:projects:hulcka:start?rev=1647506318

Last update: **2022/03/17 08:38**