Project in Practical Machine Learning

Radar image clustering
Marko Hassinen

Machine Learning

- Clustering
- Predictive clustering so far
- Implemented algorithms: k-means, k-means++ and k-medoids

Data provider

- Data from Finnish Meteorological Institute (FMI)

- Offers open data in many formats

- FMI uses Web Map Service (WMS) to provide Weather Radar images

- WMS has different options to shape the data: for example style, resolution and format

FMI's Radars

- Ten weather radars around Finland

- Cover almost all of Finland

- Four different measures: radar reflectivity, radial velocity, rain classification and cloud top height

- Anjalankoski, Ikaalinen,

  Kesälahti, Korpo, Kuopio,

  Luosto, Petäjävesi,

  Utajärvi, Vantaa ja Vimpeli
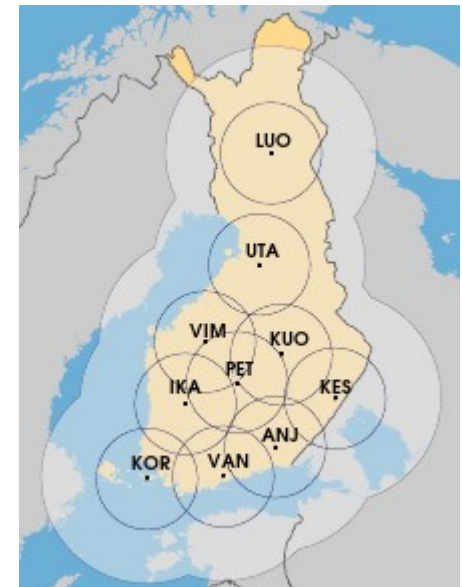
Data

- Image resolution: 300x250 and 720x480

- Different resolution is used to compare machine learning algorithms performance

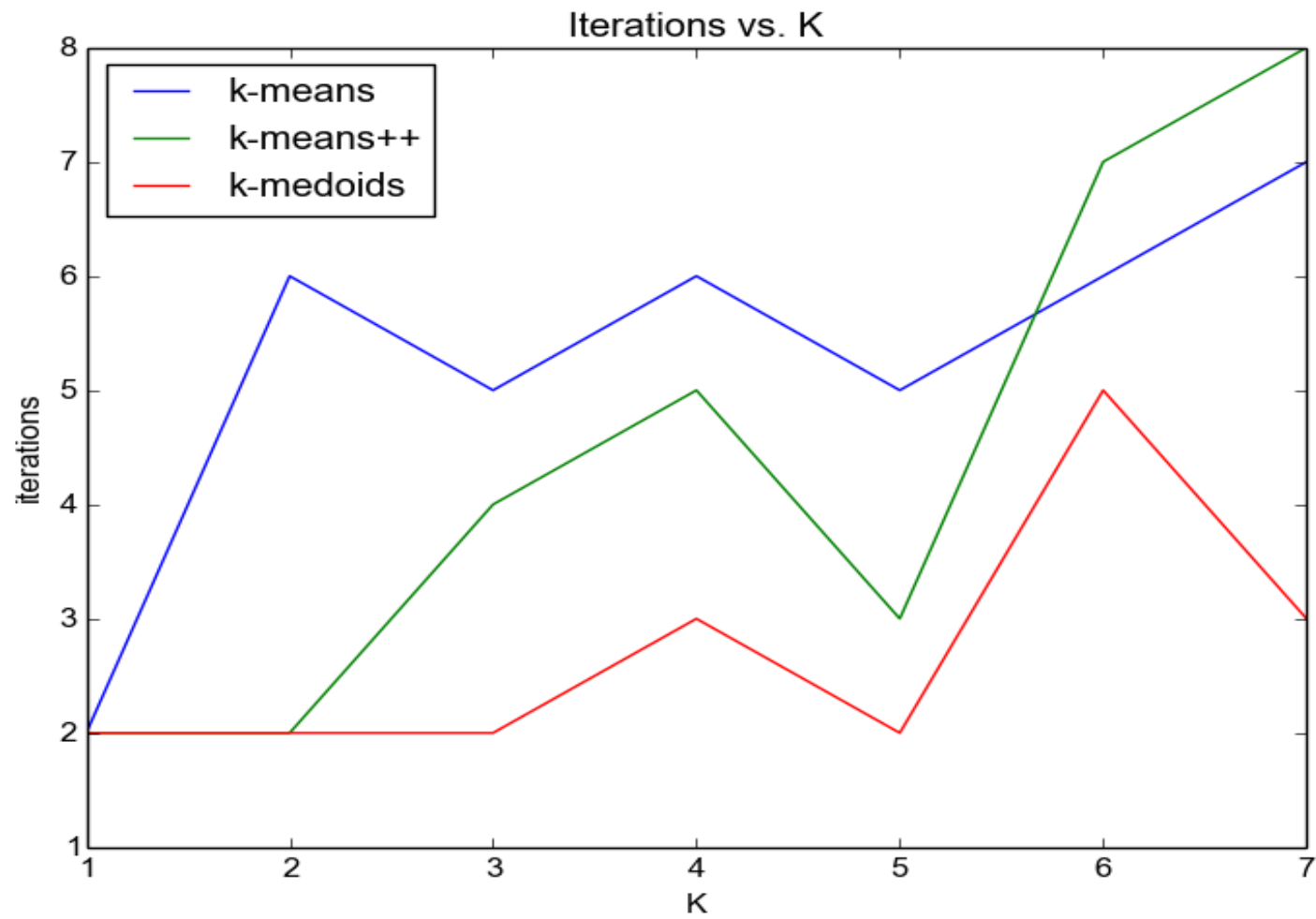- Style options: raster and Radar dbz Summer
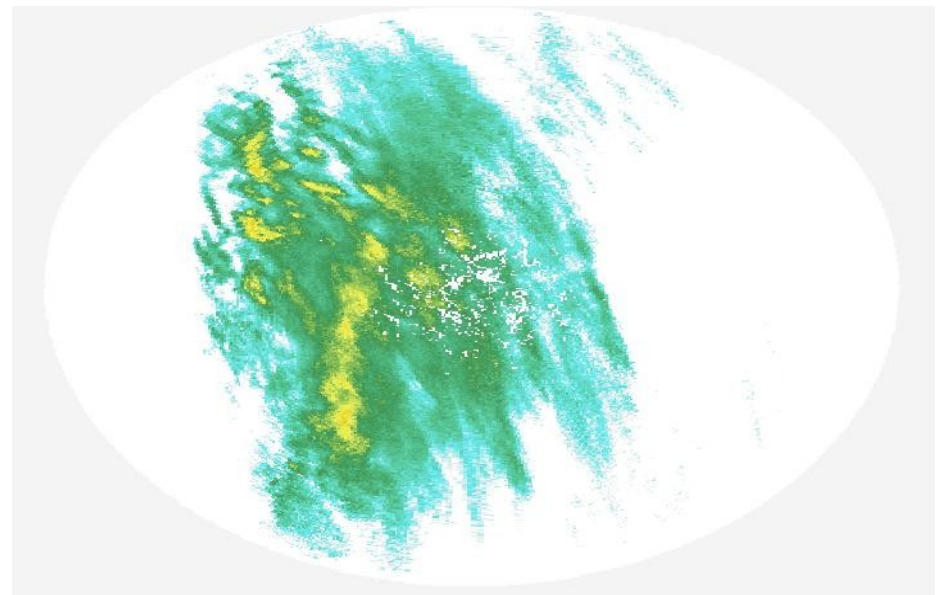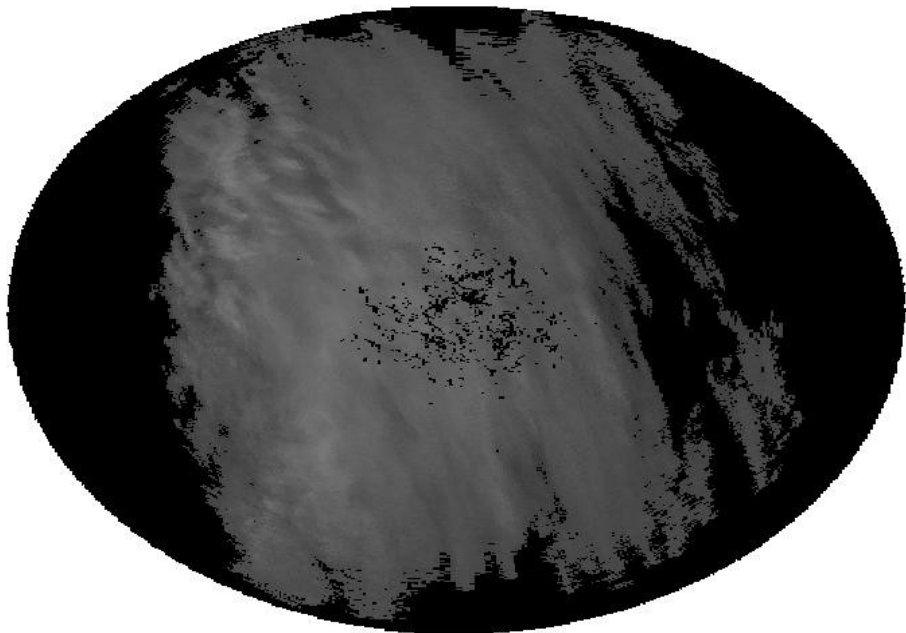
Machine Learning algorithms

- K-means with maximum iterations and initial cluster centers are selected randomly

- K-means++: only difference compared to k-means is the way that the inital cluster centers are selected

- K-medoids, the slowest (this fact was also discussed briefly in IML)

Similarity within clusters
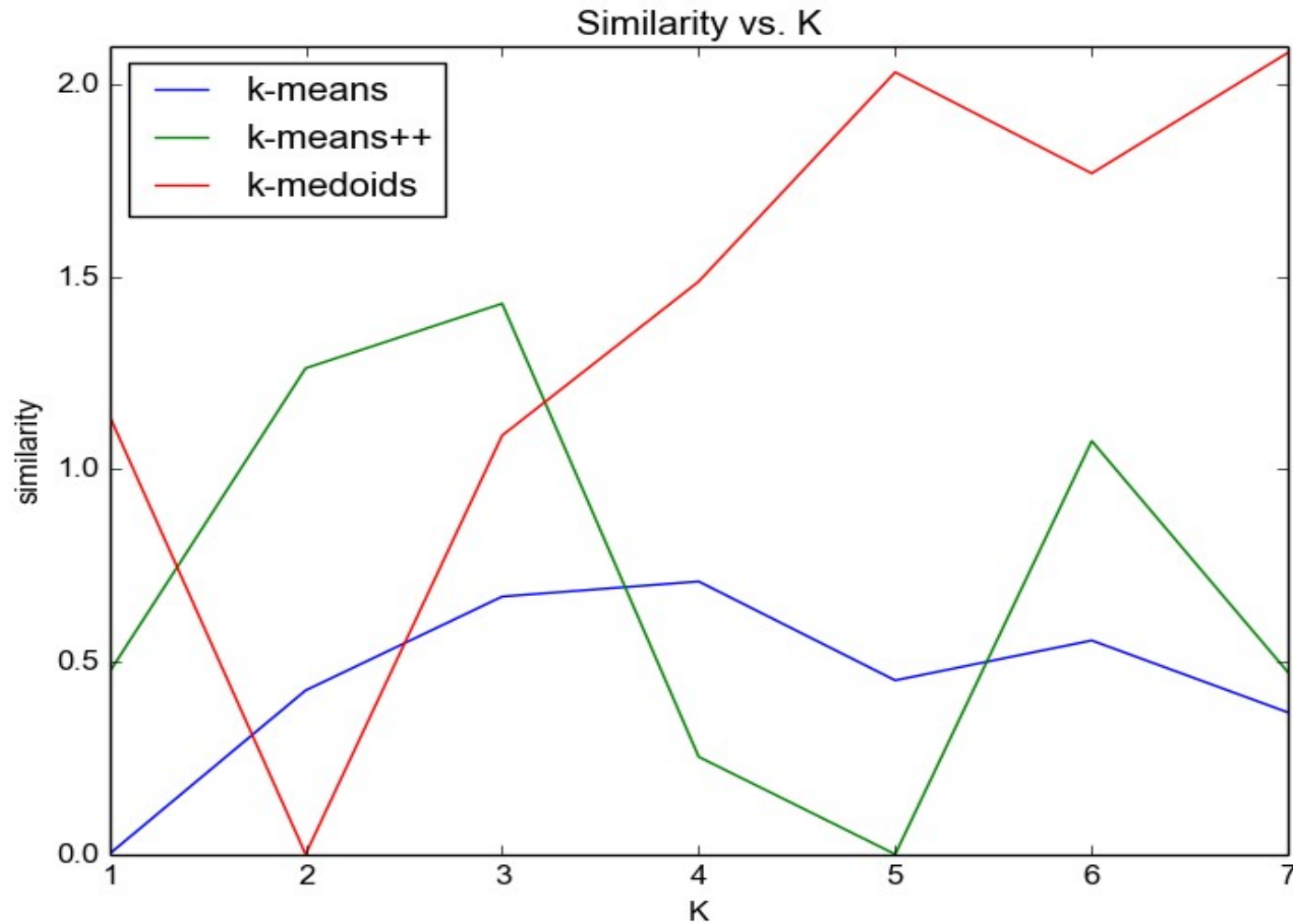
Iterations

Sources

- ilmatieteenlaitos.fi
- en.wikipedia.org
- http://theory.stanford.edu/~sergei/papers/vldb12-kmpar.pdf
- Flach, P. (2012). *Machine Learning: The Art and Scinece of Algorithms that Make Sense of Data.* Cambridge: Cambridge Univ. Press. (IML course book)