Fuzzy logic SVM based classification for large astronomical data sets

Artem Poliszczuk, Aleksandra Solarz, Agnieszka Pollo

- **Support Vector Machine (SVM) Algorithm**
  
  Main idea: mapping data from its normal parameter space (input space) to the high dimensional feature space, where it becomes possible to construct a separating hyperplane. Mapping done by the kernel functions.

![Diagram](image)

Training - searching for the combination of the kernel parameters that maximize the efficiency (grid search)

No measurement uncertainties in classical SVM!

- **Fuzzy SVM (FSVM)**
  
  - Objects have fuzzy memberships.
  - Two types of FSVM: error-based, distance-based.
Fuzzy logic SVM based classification for large astronomical data sets

Artem Poliszczuk, Aleksandra Solarz, Agnieszka Pollo

- **Results**
  
  - **Accuracy** of the classification of all SVM versions stands at around 97%-99%.
  
  - **Grid search process**: error-based FSVM shows a better performance in the grid search process where it minimizes more effectively the number of support vectors (high generalization ability and smaller probability of the overfitting).
  
  - **Distance histograms**: both error-based and distance-based FSVM show better results (wide range of the distances). Clustering of the objects even at the big distances can lead to overfitting.