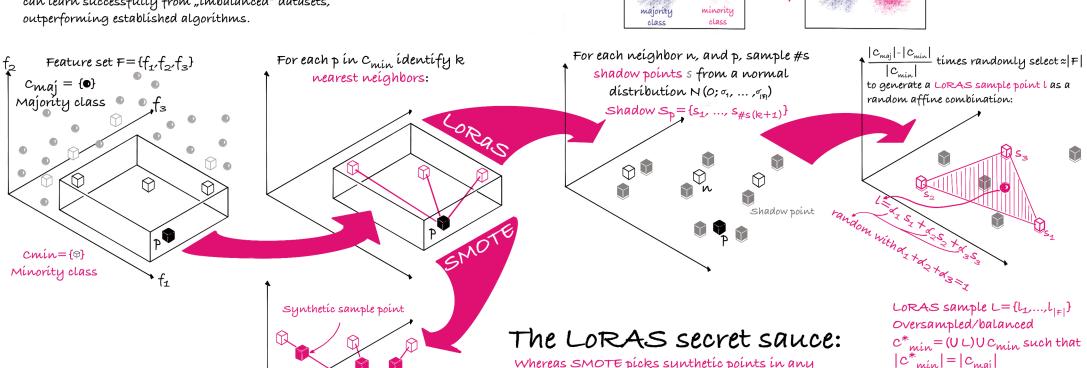
## Looking for the extraordinary, the rare, something unusual?

We are often most interested in the rare event, the things that deviate from the normal. Important examples are accidents, fraud, and the onset of disease. Algorithms can be trained to detect, classify and decide on such matters. The big challenge for these algorithms is however that we usually have lots of data for the normal, the ordinary and only small datasets for the rare event, the special, the thing we are after. We have developed an algorithm that can learn successfully from "imbalanced" datasets, outperforming established algorithms.

The problem with learning from unbalanced datasets is that precision is not a good measure since this will only reflect the underlying class distribution. A common solution is , oversampling' - generating synthetic datapoints for the minority class to rebalance the overall dataset:

Rebalanced dataset



these shadows up:)

Whereas SMOTE picks synthetic points in any dimension from the line between p and it's nearest neighbors, LoRAS points are sampled from a unique |F|-1 dimensional hyperplane. They say that special events throw their shadows ahead ... and LoRAS picks

## So, what can we predict for you?

We compared LORAS with SMOTE and its extensions and found that LORAS wins the competition in a wide range of practical applications, including the detection of credit card fraud, spam filtering, detection of thyroid diseases, arrhythmia and breast cancer, the predigen of soils from satellite data and the prediction of insurances from customer data. So, what can we do for you? Contact: www.sbi.uni-rostock.de

