Low-Dimensional Feature Learning with Kernel Construction

Artem Sokolov[1], Tanguy Urvoy[2], Hai-Son Le[1]

1 2

WWW data : millions of features / many tasks

- learning tasks unknown beforehand
- impossible to keep all data on disk
- necessity of a **learning-oriented summary**

Semi Supervised Feature Learning Challenge

- **task:** learn a dimension reduction $\mathbf{10^6 \rightarrow 100}$
- given mostly unlabeled instances
- **evaluation:** linear classifier ranking performance
- more details on poster of D. Sculley

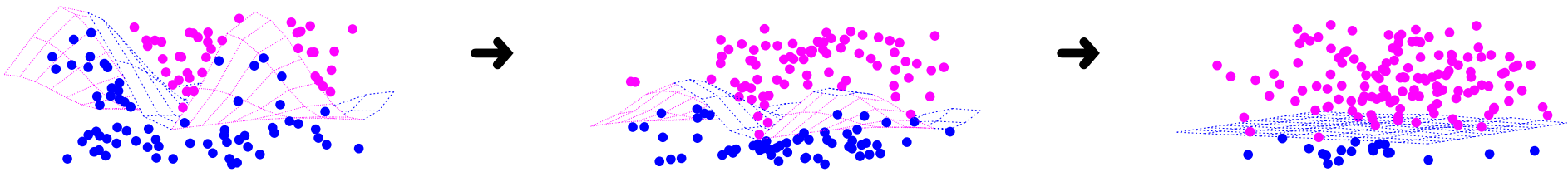
Low-Dimensional Feature Learning with Kernel Construction

Artem Sokolov[1], Tanguy Urvoy[2], Hai-Son Le[1]

1 2

1 – Use Rankboost to make data separable

- RankBoost optimizes AUC with a combination of weak rankers h_t



- Weak learners as new features: $\Phi(\boldsymbol{x}) = (\alpha_1 h_1(\boldsymbol{x}), \ldots, \alpha_T h_T(\boldsymbol{x}))$
- Provides an explicit kernel $K(\boldsymbol{x}_1, \boldsymbol{x}_2) = \langle \Phi(\boldsymbol{x}_1); \Phi(\boldsymbol{x}_2) \rangle$

2 – Use obtained kernel for low dim. embedding [Balcan et al. 2004]

- Sample pivots from dataset $\boldsymbol{p}_1, \ldots, \boldsymbol{p}_d$
- Summary of $\boldsymbol{x}$ is $[K(\boldsymbol{x}, \boldsymbol{p}_1), \ldots, K(\boldsymbol{x}, \boldsymbol{p}_d)]$
- Also tested with other similarity functions