Support selection algorithms

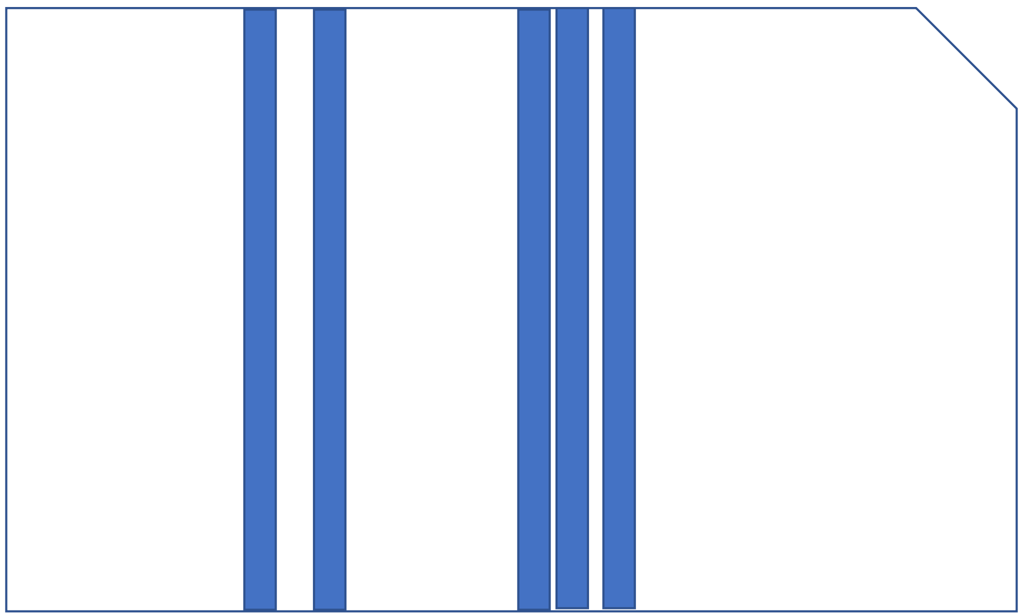
Rajiv Khanna

In collaboration with

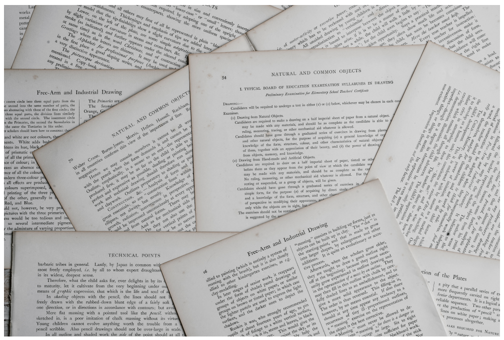
Michael Mahoney (UC Berkeley), Alex Dimakis(UT Austin), Joydeep Ghosh (UT Austin), Oluwasanmi Koyejo (UIUC), Sahand Negaban (Yale), Michal Derezinski (UMich), Been Kim (Google Brain), Russell Poldrack (Stanford)

Sparsification by support selection

- Interpretability by modeling choice: Given a matrix, choose a few columns

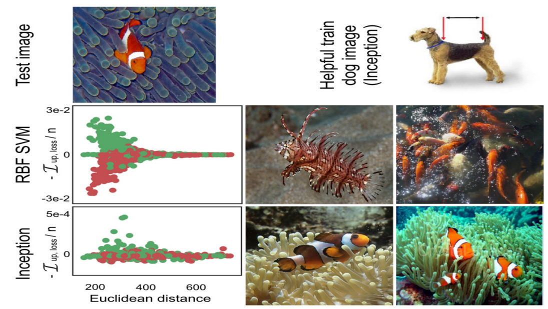


Greedy selection in practice



Document summarization.
[Vanderwende et. al. 2007]

Gene Analysis
[Paschou et. al. 2007]



Interpretability
[Koh et. al. 2017]

Greedy support selection

- Goal:

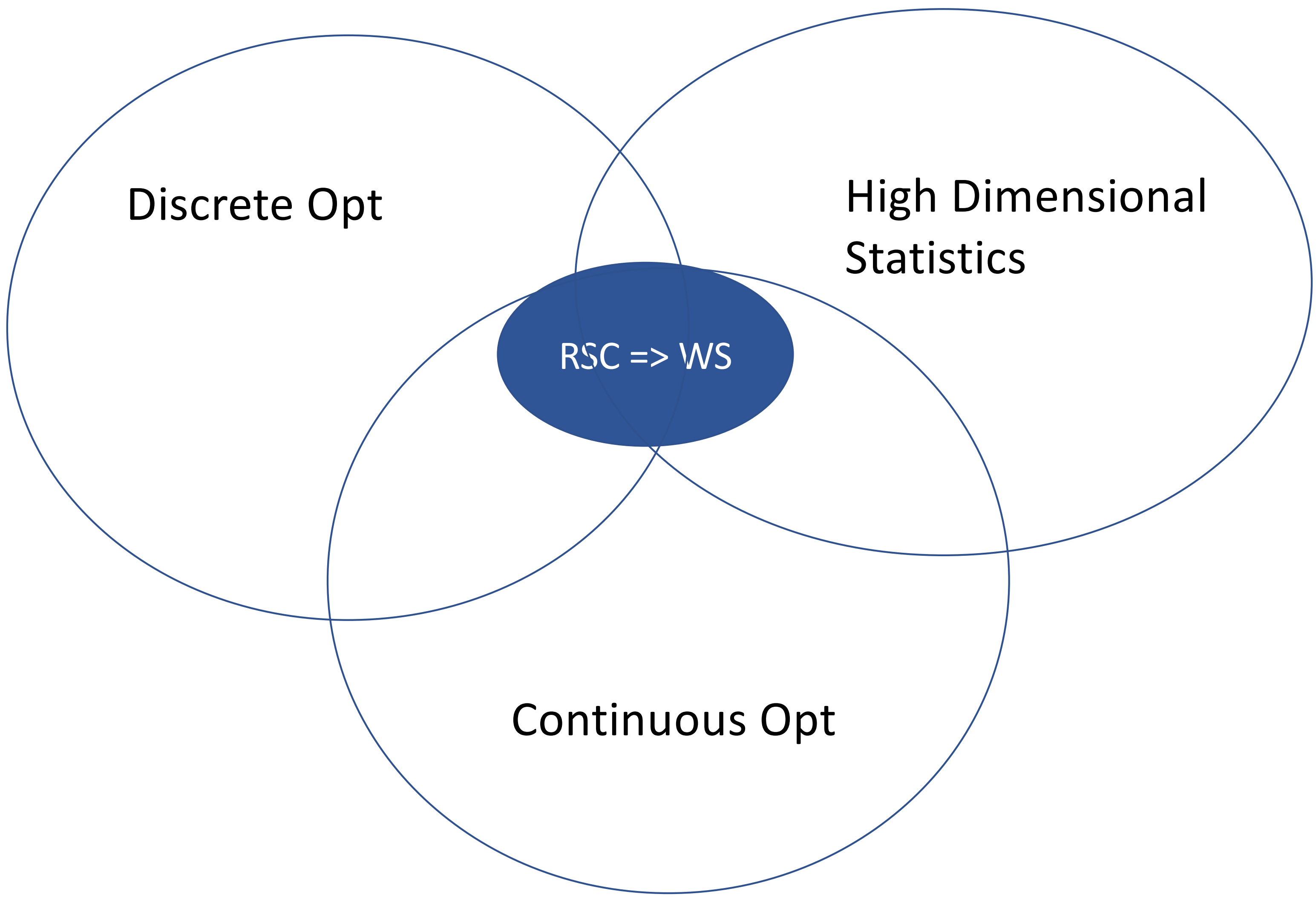
$$\max_{\mathsf{S}:|\mathsf{S}|\le k} f(\mathsf{S}) \Leftrightarrow \max_{\substack{\boldsymbol{\beta}:\boldsymbol{\beta}_{\mathsf{S}^c=0} \\ |\mathsf{S}|\le k}} l(\boldsymbol{\beta}) - l(\mathbf{0})$$

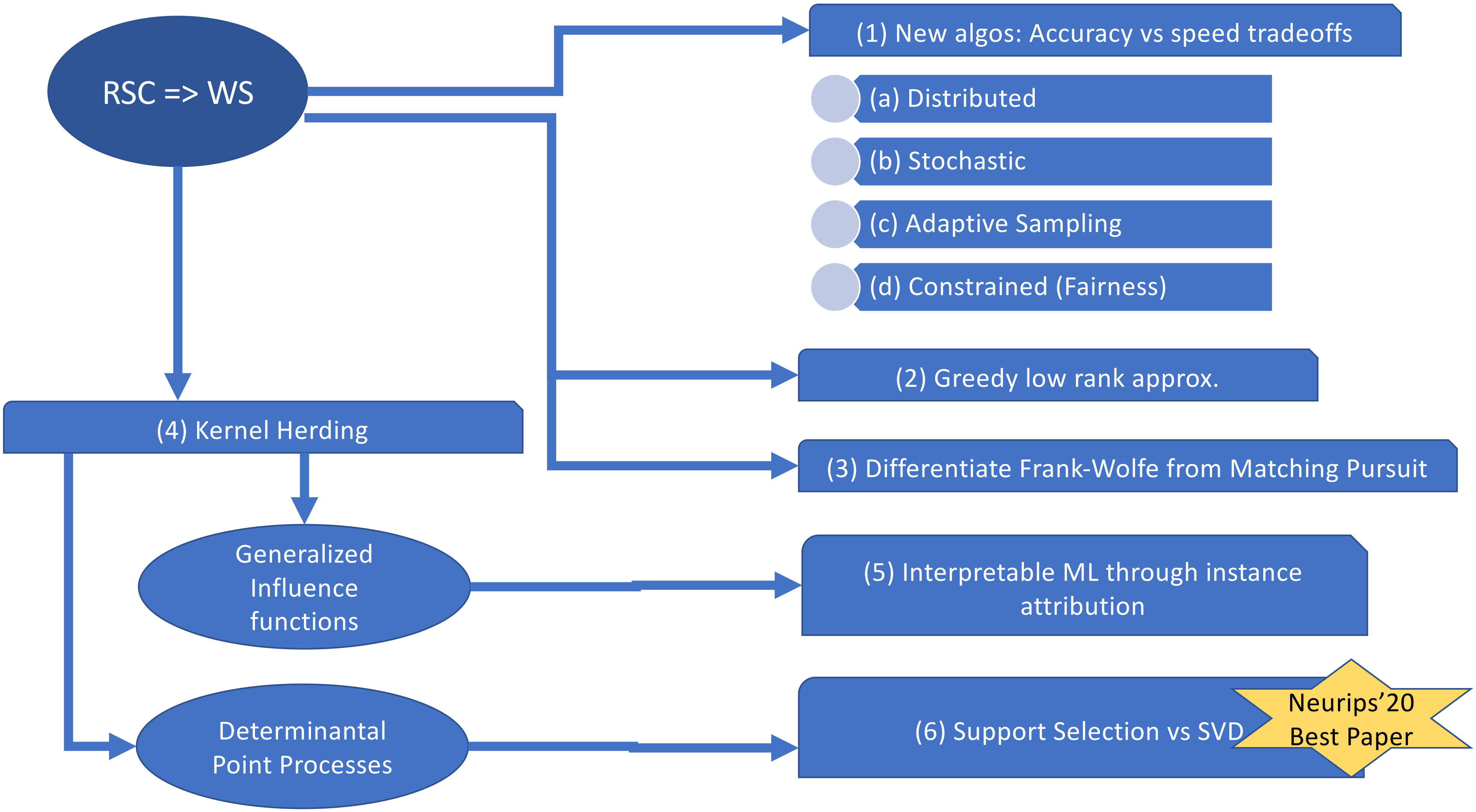
- If $l(\,)$ is:
  - m-Restricted strong concave on a certain subdomain, and
  - M-smooth on another subdomain,

  then,

RSC => WS

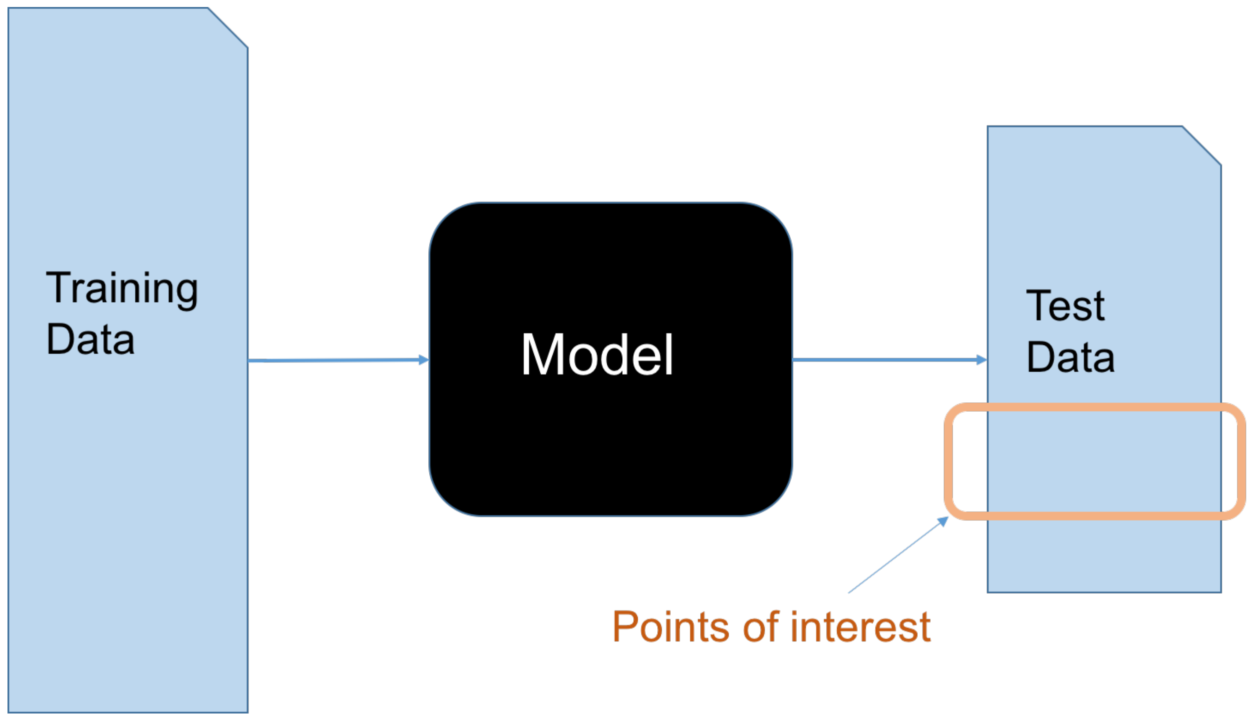
$$\underbrace{f(G_k)}_{\text{Greedy utility}} \ge \left(1 - \frac{1}{e\hat{}\{m/M\}}\right) \underbrace{f(S_k^*)}_{\text{Optimal utility}}$$


Discrete Opt
High Dimensional Statistics
RSC => WS
Continuous Opt



RSC => WS
(1) New algos: Accuracy vs speed tradeoffs
(a) Distributed
(b) Stochastic
(c) Adaptive Sampling
(d) Constrained (Fairness)
(2) Greedy low rank approx.
(3) Differentiate Frank-Wolfe from Matching Pursuit
(4) Kernel Herding
Generalized Influence functions
(5) Interpretable ML through instance attribution
Determinantal Point Processes
(6) Support Selection vs SVD
Neurips'20 Best Paper

Black-box interpretability



Which training data points are “most” responsible for predicting on points of interest ?

Research direction in near future: <u>Interpretability guided data-centric optimization</u>

- Sparse-training (coreset selection)
- Effect of increased generalization on privacy