Multi-Scale Graph KRS

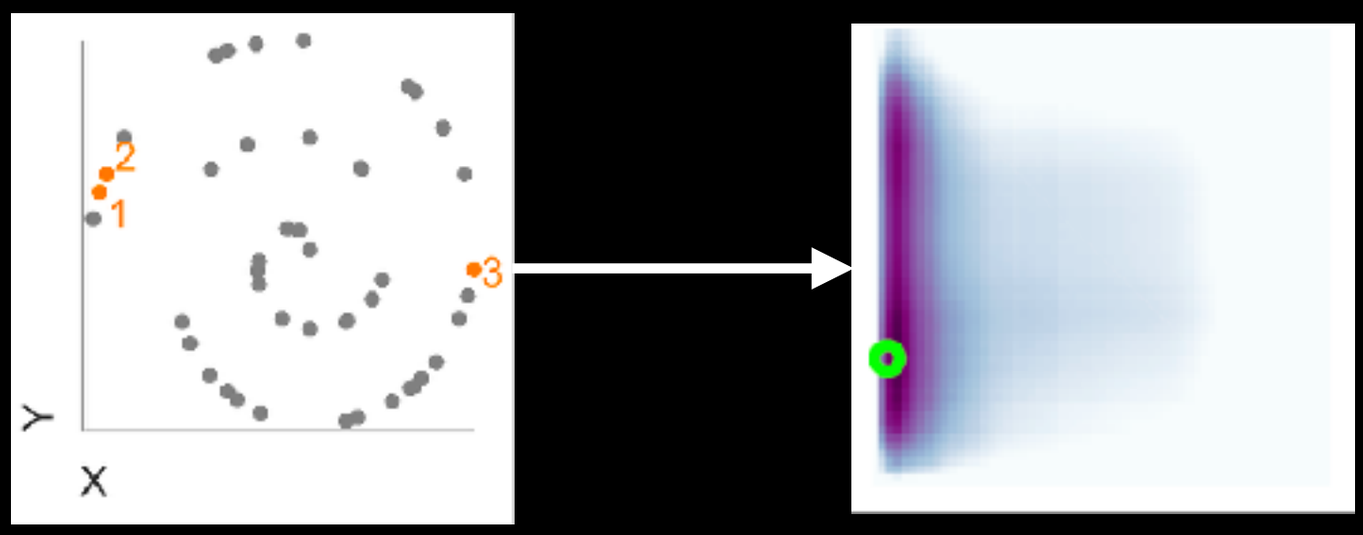
Basic Idea: Many approaches use an graph-based knowledge representation that requires a choice of scale. We *efficiently check all scales and find the best one*.

Breakthrough:

1. achieve the same performance levels as previous methods with provably *fewer samples*,
2. quantitatively characterize the *multiscale geometry* of the function (regardless of types of data or relationships)

So Far: We have *proven and demonstrated empirically superior performance* to current state of the art, including independence testing, equality testing, clustering, topology vs. attribute testing, and feature selection.

DoD Success: Mitigate *batch* effect in fusing multiple disparate heterogeneous data to improve predictive accuracy on a DoD relevant dataset.



A Few Additional Proposed Datasets

- Something from D3M: as we have no background doing time-series or NLP with MGC, any of the proposed challenge problems are a reasonably large pivot.

- Given spatial transcriptomics data from multiple different labs using disparate experimental modalities, find a “consensus clustering” fusing all of the missing & noisy data with batch effects. This is the 1st challenge associated with the International BRAIN Initiative.

- Given psychological evaluations of 1000’s of individuals including various questionnaires, physical exams, and more, acquired from different labs/hospitals, find “types” of patients that respond effectively to different treatment plans. This is the biggest grand challenge in modern clinical psychiatry.