Active Learning

Problem Setup

Pool-based: Large unlabeled dataset, non-incremental

Stream-based: Data keeps coming in, to label or to discard

Member query synthesis: synthesize new data for labeleling

Selection mode

Instance-by-instance

Batch-mode

Subset

Query strategy

Representatives

Pediction uncertainty

Pediction disagreement

Expected parameter change

Generating synthetic data