

GRADUATE SEMINAR

Ryan Tessier

Sequence-Based Sample Compression Schemes with relations to design theory

MSc Student supervised by Sandra Zilles and Shaun Fallat

Tuesday July 16th

10:00

Math Lounge (CW307.20)

Abstract: A sample compression scheme of size k for a concept class \mathcal{C} is a pair of functions (f, g) called the compression function and the reconstruction function so that for any sample S consistent with a concept in \mathcal{C} f compresses S to some subset of S , while g returns a set of domain points, labelled consistently with the original sample S . An important open question in this area, that has implications for machine learning, is if every concept class has a sample compression scheme of size $O(d)$, where d is the VC dimension of \mathcal{C} .

This talk examines a modification of sample compression schemes, specifically, for a concept class \mathcal{C} we define a sequence-based sample compression scheme for \mathcal{C} as a pair of functions (f^*, g^*) where the items we compress to are now sequences instead of sets. We discuss a few sequence-based sample compression schemes and how they improve compression bounds over the original set based compression schemes. Finally we discuss connections between set and sequence-based sample compression schemes and design theory.