**Goal:**

1)Input size-genome level

2)make a visualizer

3)change input type from di(crystallized image) to tri/tetra(md simulated image)

**Understanding of the paper:**

**Test\_Seq.txt has 100 sequences of length 1001**

* Why promoter prediction is needed? Pg1

For genome annotation-to understand transcription regulatory networks

* Why has it been tough?? Pg1

Genomic sequences are highly adaptive and diversified to enable survival in diverse conditions.

Variability in the length of 5’ un-translated regions and presence of multiple TSS

High gene density, very less inter-genic spaces or having overlapping coding regions

* What is meant by parameter signals?

Structural/energetic properties like bendability, curvature etc.  
 STRUCTURAL: DNA shape is cumulation of certain parameters. So instead of taking their

Cumulative effect, we consider different parameters. (28)

ENERGETIC: hydrogen bond, stacking energy and solvation energy.

LEARNING PCA: PRINCIPAL COMPONENT ANALYSIS