Agastya Gupta

Ahsan Azim

Convolutional Neural Networks for DNA Sequence Classification

**Aim / Hypothesis**

The aim of our project remains identical to that introduced previously – i.e. exploring the possibilities of Convolutional Neural Networks (CNNs) for DNA Sequence Classification. We continue to move forward with our experimentation into the efficacy of CNNs this task, and have received encouraging results on that front in the past week. Indeed, our prediction that CNNs may prove effective at such classification has already been vindicated – our architecture in some cases surpasses, and other cases matches those of both non-CNN and existing CNN-based efforts at such DNA Sequence classification problems.

Thus the first part of our initial hypothesis – i.e. proving the efficacy of our algorithm on smaller canonical datasets – has been cleared. We shall now move on to extending our research to relatively untouched datasets, as well as setting up a user-friendly web-based interface to classify relevant sequences using our architecture.

**Computational approaches developed­­­­­­**

Convolutional Neural Network – describe specific architecture

**Data used**

Splice Junction etc – describe and give an excerpt

**Current results**

Create table of results -- say we beat their time!!!! Say that we’re consulting with Deep Learning experts at Dartmouth to discuss how this happened haha!

**To do list**

Try even more varying parameters etc, different datasets

Finish off our custom implementation?

Try different ML algorithms – though idk how useful, since current results so good

**Expected results**

Random stuff about increasing/decreasing performance, etc

**References (not included in 3-page limit)**

Not many