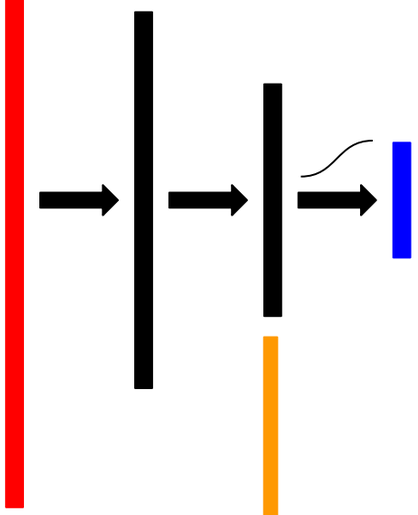
Iterative Enc-Dec Training for Binary Classification

Step 1: Train Encoder on the binary classification task on inputs **X**.

Step 6: Train Encoder on the classification task on **X**, and the classification task on **S-Inputs** labeled 0.5. ***Return to step 2.***

Encoder

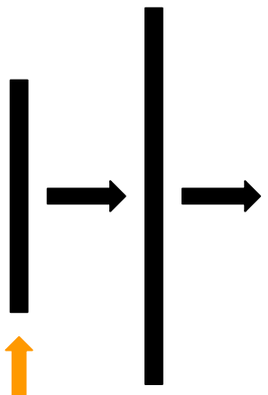
Inputs **X**: Dimension N



Loss Function: Categorization Error (categorical cross entropy, hinge loss, etc.)

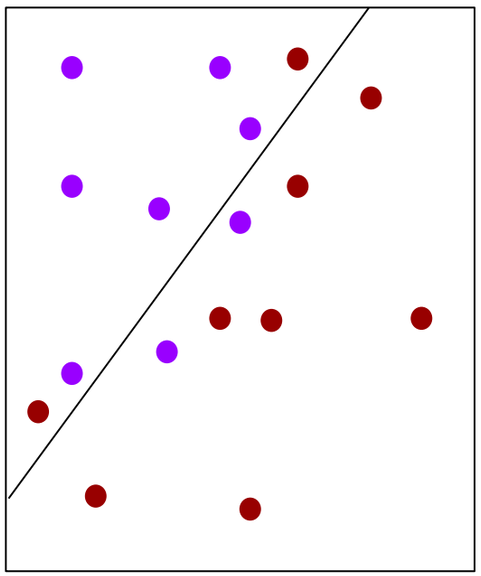
Decoder

Step 4: Train Decoder on the regression task from **Y-c** to **X**

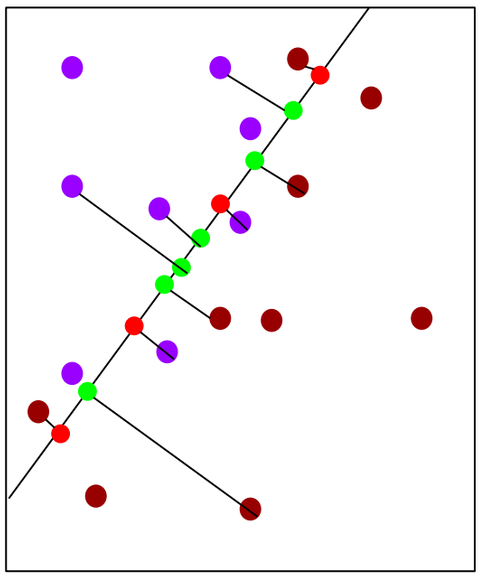


Loss Function: Regression Error (cross entropy, mean squared, etc.)

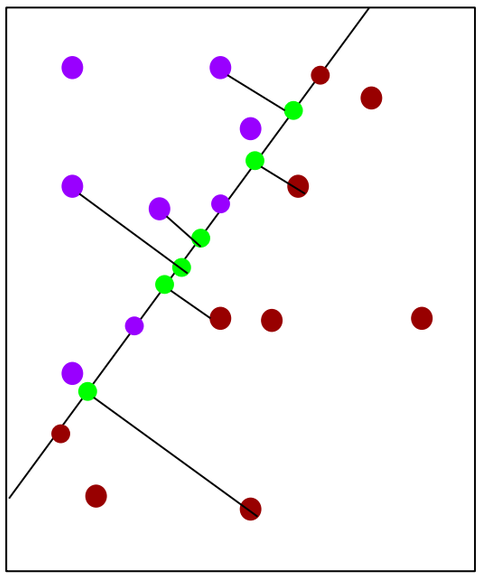
Step 5: Run **S-Points** through the decoder and denote outputs as **S-Inputs**



The final layer is just hyperplane separators with logistic squashing function. We will denote the output of the final layer **Y**.



Step 2: Find the closest points on the separating hyperplane. Correctly classified points are marked green, incorrectly classified points marked red. **Denote the green points S-Points**.



Step 3: Translate the incorrectly classified points to their corresponding locations on the separating hyperplane + epsilon distance on the other side. Denote our new collection of brown and purple points **Y-c**.