

# Virtual Adversarial Training

Unlabeled; model  
guesses it's probably  
a bird, maybe a plane



New guess should  
match old guess  
(probably bird, maybe plane)



→  
Adversarial  
perturbation  
intended to  
change the guess