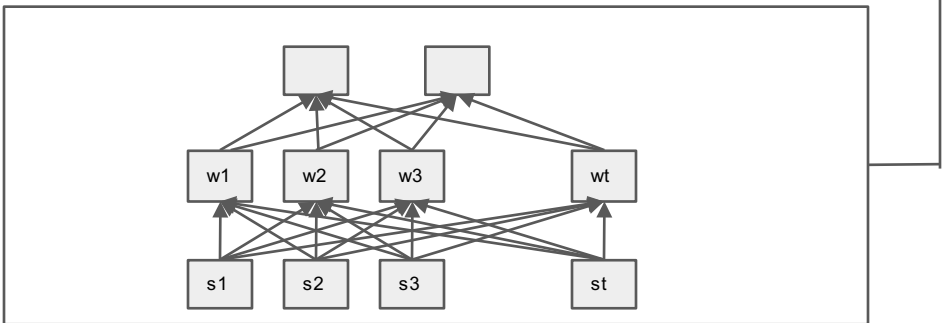


Policy Net to select words in sentence

delayed reward:
 $P(y|x)$



Classification Net to predict binary sentiment