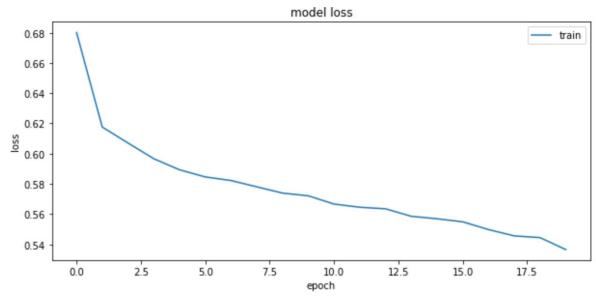
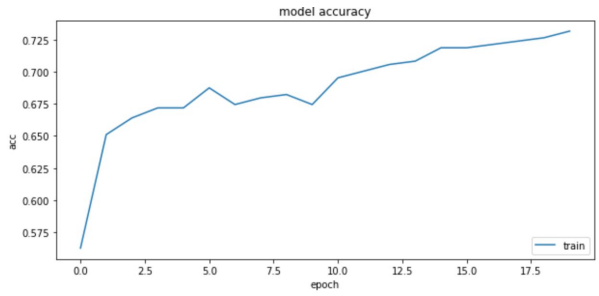
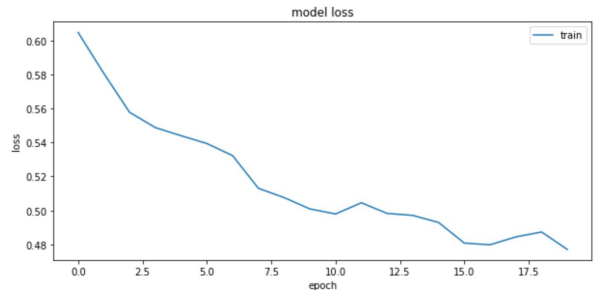
以下Epoch皆為20

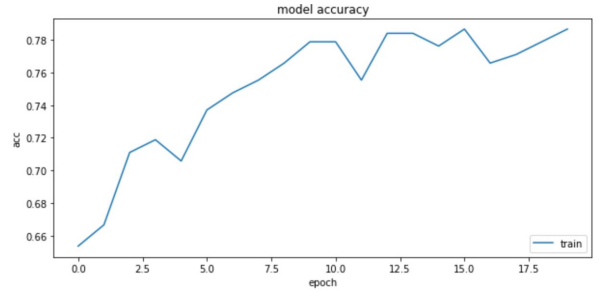
神經元層數：1(10個)
Loss func: kullback_leibler_divergence
time:17.66, best acc:0.73
Dense
Activation(relu)



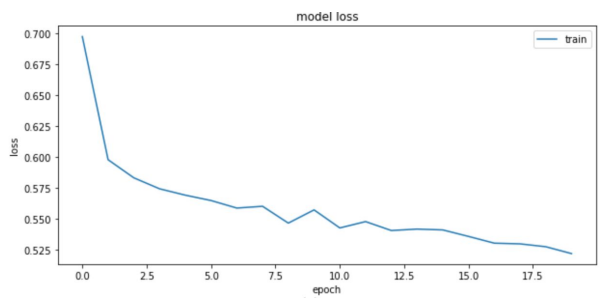


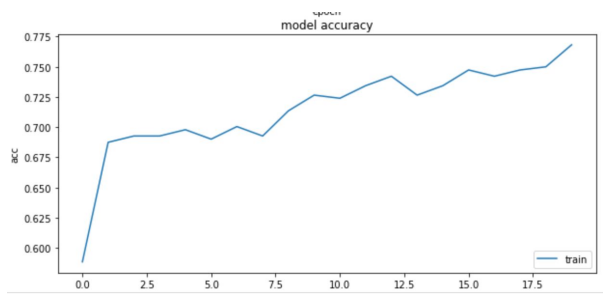
神經元層數：1(100個)
Loss func: kullback_leibler_divergence
time: 17.609, best acc:0.7865
Dense
Activation(relu)



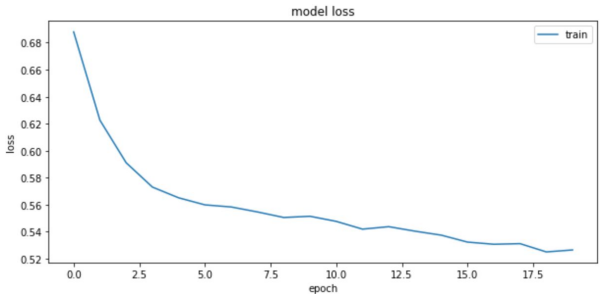


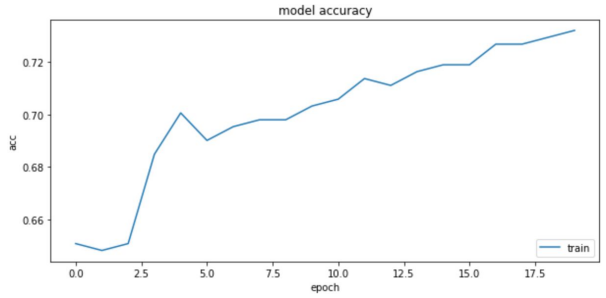
神經元層數：3(10個)
Loss func: kullback_leibler_divergence
time:21.73, best acc:0.7682
Dense
Activation(relu)
Dense
Activation(relu)
Dense
Activation(relu)



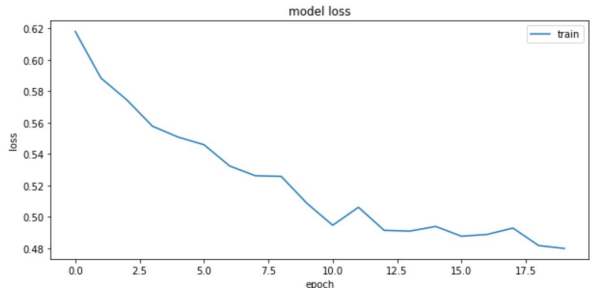


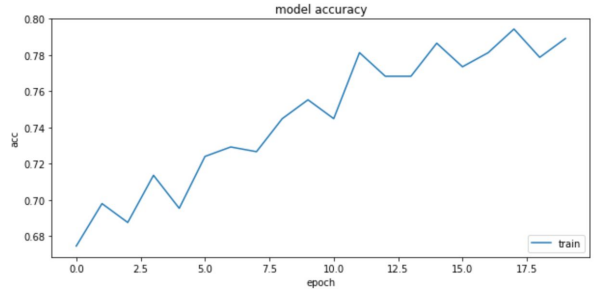
神經元層數：1(10個)
Loss func: binary_crossentropy
time:20.257, best acc:0.7318
Dense
Activation(relu)



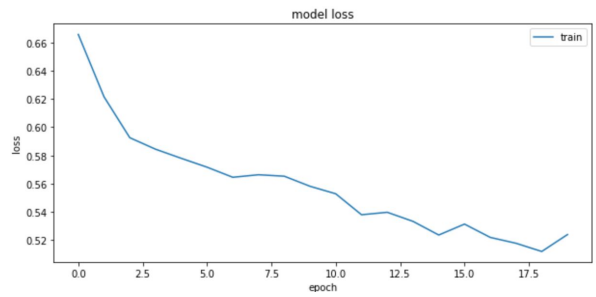
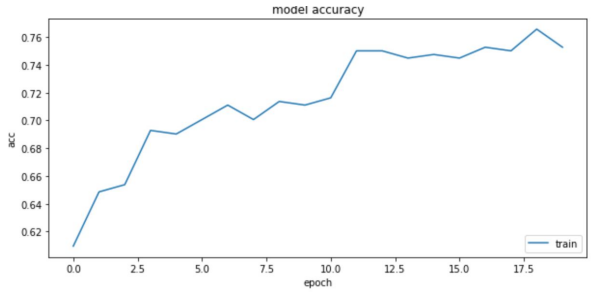


神經元層數：1(100個)
Loss func: binary_crossentropy
time:27.548, best acc:0.78
Dense
Activation(relu)





神經元層數：3(10個)
Loss func: binary_crossentropy
time:31.85, best acc:0.7656
Dense
Activation(relu)
Dense

Activation(relu) Dense Activation(relu)
 

1. Loss function使用kullback_leibler_divergence後，加多單層內的神經元數效果最顯著，達0.78，加深也會有好的訓練效果。
2. Loss function使用binary_crossentropye後，加多單層內的神經元數效果最顯著，達0.78比較kullback_leibler_divergence跟binary_crossentropye後。兩者不管在加多單層神經元及加深模型深度訓練時間皆較短。因KLD的loss的資訊量會盡量接近真實的資訊量。