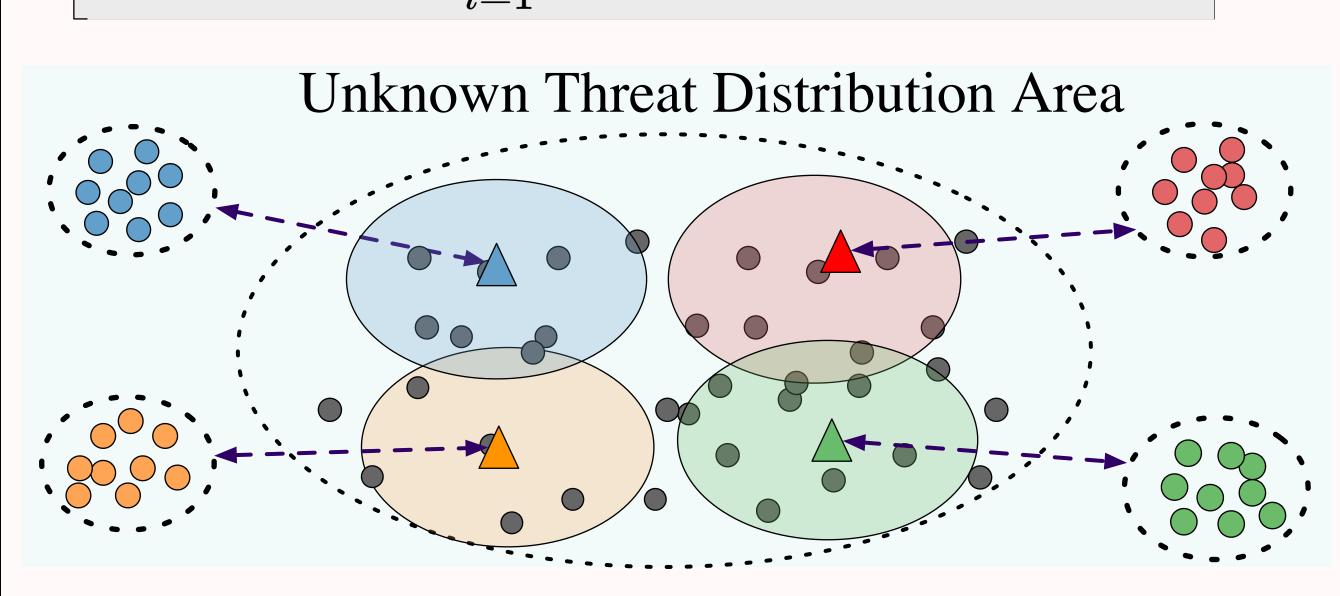
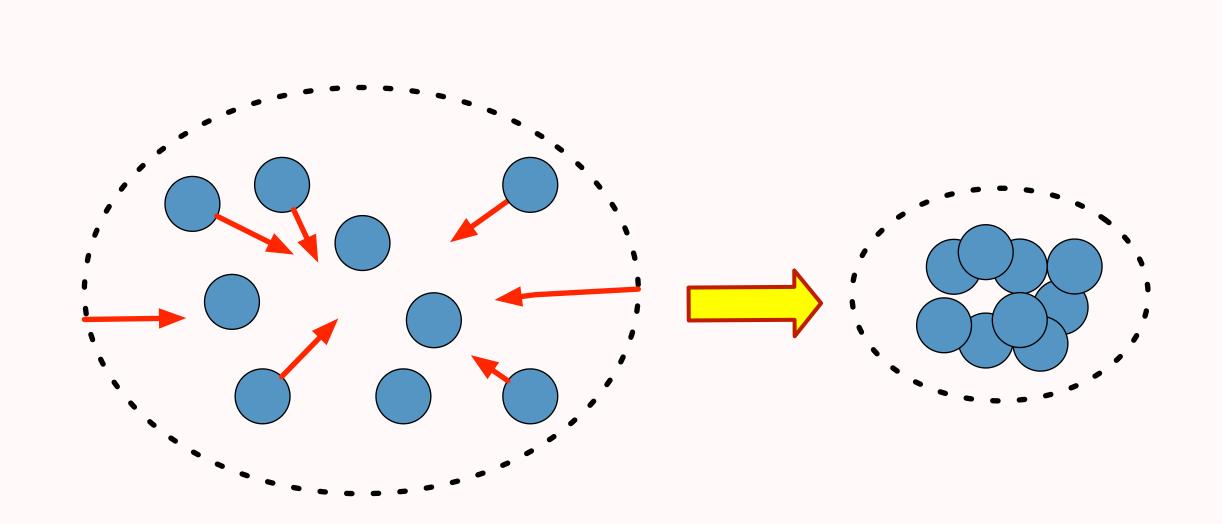




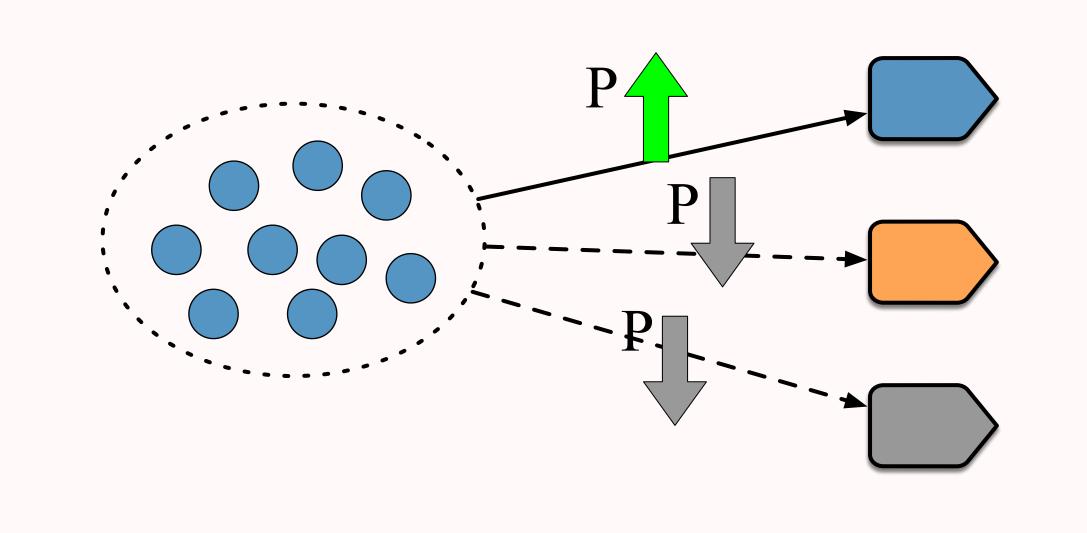
$$L_{margin} = rac{1}{N} \sum_{i=1}^{N} \max(d_e(z_i, P^{y_i}) - R^{y_i}, 0)$$



## $L_{ ext{Fisher}} = rac{1}{1 + ext{tr}\left(S_b ight)/ ext{tr}\left(S_w ight)}$



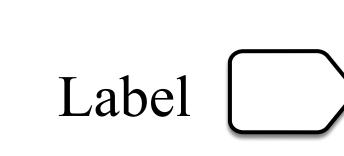
## $Classification \ Loss$ $Classification \ Classification \$

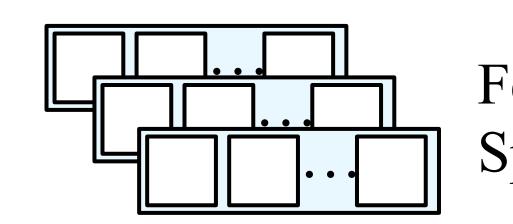




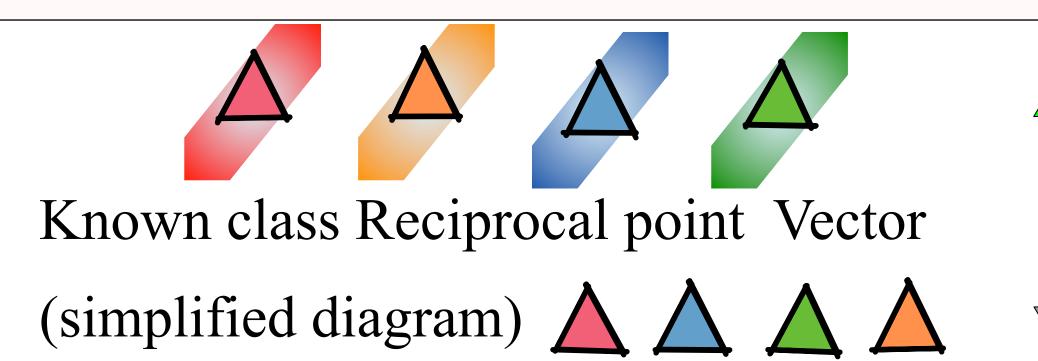
Benign

Attack2





Feature Embedding
Space



Increased probability

Probability reduced