Twitter Thread by Jesús López



Jesús López
@jsulopz

I understand ReLu in this way:

1. a neuron is a number in the model

2. not every neuronal is important for each label you want to predict

3. if you apply a “linear” activation, redundant neurons will influence the prediction a bit

@svpino 4. but the ReLu activation won't consider a neuron relevant until it overpass a threshold

5. therefore removing noise for the predictions

would you agree ?