Redundancy and Robustness in DNN Article– Main Ideas:

- Robustness -> The amount of units that can be removed without affecting accuracy (How fragile the network is to removal of units)
- Redundancy -> Having units with similar activity
    - Similarity
    - Compressibility
- DNN adjust their effective capacity by developing either robustness or redundancy
- How do networks perform well without overfitting when the amount of free parameters is greater than the amount of training examples?
- Generalization Ability does not decrease as the network degree of over parameterization increases (more neurons/layer still generalize well)
- Metrics:
    - Robustness -> the ability of the network to label correctly without being affected by ablations (removal of units)
        - How do we measure it ? we calculate the area under the curve
    - Redundancy
        - Similarity -> a measure of highly correlated units
        - Compressibility -> amount of principal components explaining 95% of the variance
- Redundancy and robustness do not imply each other !
- Results:
    - Robustness and redundancy develop on networks trained on randomly labeled data (so it does not predict generalization capability)
    - Robustness and redundancy are sensible to initialization variance, while its accuracy is not
    - Optimizers can affect how much redundancy or robustness develop, but they do not change overall trends
    - Varying learning rate and batch size have little to no effect
    - Redundancy and robustness are central to how networks auto regularize
    - Redundancy and robustness poorly predict generalization
    -

Things that weren't checked (that we can check):

- Effect of depth of networks