

- Success Criteria
  - we have successfully reproduced the results showed by Tishby and Saxe. However there are reasons to not trust either of them as they have flaws with them.
    - Tishby – used only a toy dataset
    - Saxe – Changed allot of parameters at once made the claim that no compression phase happens
  - We need better tools for MIE
    - cannot judge subtleties if something has a compression phase our MIE are not trusted
    - we have seen KDE and Discrete show inconsistent results, when n'th layers has less information about the input than the n+1'th layer.
  - $i++j$
  - performance
- 

### 0.0.1 Deterministic networks

There's a very real argument to be made against compression in neural networks. Consider a generic neural network we can think of it as a function that is a series of matrix transformation, where a matrix corresponds to weights of a specific layer. However these matrices are all random (at least at the start of training) and hence probability of them being invertible is 100%.

Knowing that every single matrix is invertible allows us to conclude that that neural network as a whole is an invertible function, which means no information is lost and compression is impossible.

### 0.0.2 Why Randomness is hard to capture