DeepImportance Proof

Theorem 1. *The importance-driven coverage (IDC) test adequacy criterion given by (1) is sound by construction.*

$$IDC(Y) = \frac{|\{INCC(j) | \exists y \in Y : \forall V_n^i \in INCC(j) \bullet \min d(\phi(y,n), V_n^i)\}|}{|INCC|} \quad (1)$$

where

- $INCC$ is the important neurons cluster combinations
- Y is the test set that is part of the input domain X, i.e., $Y \subseteq X$
- $\phi(y, n)$ is the activation value of neuron n given input $y \in Y$
- V_n^i is the neuron's n cluster centroid
- $d(\phi(y, n), V_n^i)$ is the Euclidean distance between $\phi(y, n)$ and V_n^i

Proof. We need to show that the IDC score will either remain the same or increase when adding a new test input y' to the test set Y.

According to (1), we have that Y covers a finite number of INCC with the score $IDC(Y)$ given by the set $Z \subseteq Y$ such that

$$\min d(\phi(y,n), V_n^i) \text{ holds for each } y \in Z, \forall V_n^i \in INCC(j). \quad (2)$$

Assume we have $Y' = Y \bigcup \{y'\}$, i.e., extending test suite Y with test input y'. We have two cases for y'

case I: the i-th INCC covered by y' as per (2) is already included in the INCC covered by Y. Hence, $IDC(Y) = IDC(Y')$

case II: the i-th INCC covered by y' as per (2) is not included in the INCC covered by Y. Thus, $IDC(Y) < IDC(Y')$

□