# Incorporating Prior Domain Knowledge into Deep Neural Networks

http://people.cs.vt.edu/ramakris/papers/PID5657885.pdf

## Abstract


- domain adapted neural network (DANN)
  - knowledge is incorporated in loss terms
    - monotonicity constraints
    - approximation constraints
  - tested on synthetic and real world data
    - noisy data
    - limited data


## Problem Formulation and Solution Approach

### Problem Statement

> leverage domain knowledge to train a robust accurate learning model that yields good model performance even with sparse, noisy training data

### Approximation Constraints

$$g(\hat{Y}) = \begin{cases} 0 & \text{if } \hat{Y} \in [y_l, y_u] \\ |y_l - \hat{Y}| & \text{if } \hat{Y} < y_l \\ |y_u - \hat{Y}| & \text{if } \hat{Y} > y_u \end{cases}$$
$$\text{Loss}_D(\hat{Y}) = \sum_{i=1}^{m} ReLU(y_l - y^i) + ReLU(y^i - y_u)$$
$$ReLU(z) = z^+ = max(0, z)$$

- constraints that specify a quantitative range of operation of the target variable
  - ReLU term ensures that the output is non-zero if the input is positive

## Monotonicity Constraints

$$\text{Loss}_D(\hat{Y}_1, \hat{Y}_2) = \sum_{i=1}^{m} \mathbb{I}\Big((x_1^i < x_2^i) \wedge (\hat{y}_1^i > \hat{y}_2^i)\Big) \cdot ReLU(\hat{y}_1^i - \hat{y}_2^i)$$

- h(x) = y such that x1 > x2 → h(x1) > h(x2) then x1, x2 and h(x1), h(x2) share a monotonic relationship
- I represents the identity function → true if the boolean statement is true, else false
  - serves as a boolean mask which captures errors only of the instances where the domain constraint is violated

## Synthetic Datasets

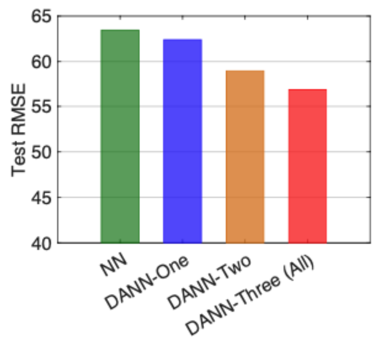
- use Bohachevsky function as the basis for generating synthetic datasets
  - **approximation constraints**: dataset X = [x_i,1, x_i,2], Y
    - randomly select a subset of rows in X and interchange the values of x1 and x2 in those rows → value of f(x1, x2) in Y for those rows is outside of the approximate range
  - **monotonicity constraints**: X', X'', X''' such that x_i,1 < x'_i,1 < x''_i,1

- randomly sample a subset of rows and switch the values of y and y' to create data where the monotonicity constraints are violated

### Real Datasets

- prediction of oxygen solubility in water
  - influenced by water temperature, salinity, pressure

## Experimental Findings




- DANN and vanilla NN are compared with respect to noise and amount of training data
- see paper for details