Applications of multi-class machine learning models to drug design

Marvin Waldman, Michael Lawless, Pankaj R. Daga, Robert D. Clark

Simulations Plus, Inc.

Lancaster CA, USA

SCIENCE + SOFTWARE = SUCCESS

Overview
Applications of multi-class machine learning models to ADMET

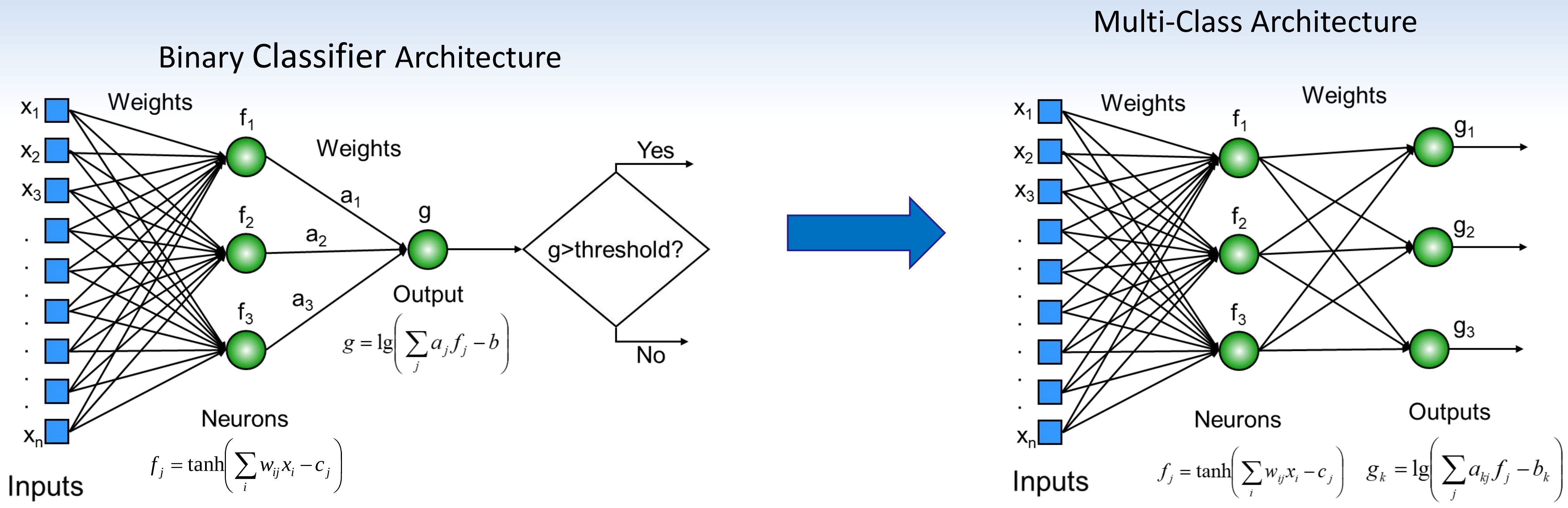
- Motivation
- Approaches/Methodology
- Application
- Comparison
- Summary/Conclusions

Motivation

- Until recently, machine learning classification models in Cheminformatics literature have generally modeled binary endpoints (active/inactive, substrate/non-substrate, toxic/non-toxic, etc.)
- Recent examples of multi-class models and/or endpoints relevant to drug discovery
  - Mode of Action of 220 phenols in T. pyriformis toxicity assay (4 class decision tree model)
    - Schüürmann et al, Chem Res Tox, **16**, 974 (2003)
  - Extended Clearance Classification System
    - Predicts 1 of 3 dominant clearance mechanisms via a 6 class decision tree scheme
    - Varma et al, Pharm Res, **32**, 3785 (2015) and subsequent publications
  - Acute Rat Toxicity based on LD50
    - GHS (Globally Harmonized System of Classification and Labelling of Chemicals)
      - 5 toxic classes
      - https://en.wikipedia.org/wiki/Globally_Harmonized_System_of_Classification_and_Labelling_of_Chemicals
    - EPA uses 4 toxicity classes derived from essentially the same data
      - https://en.wikipedia.org/wiki/Toxicity_category_rating
  - AMES Mutagenicity
    - NIHS Japan uses 3 categories, strongly positive, positive, negative

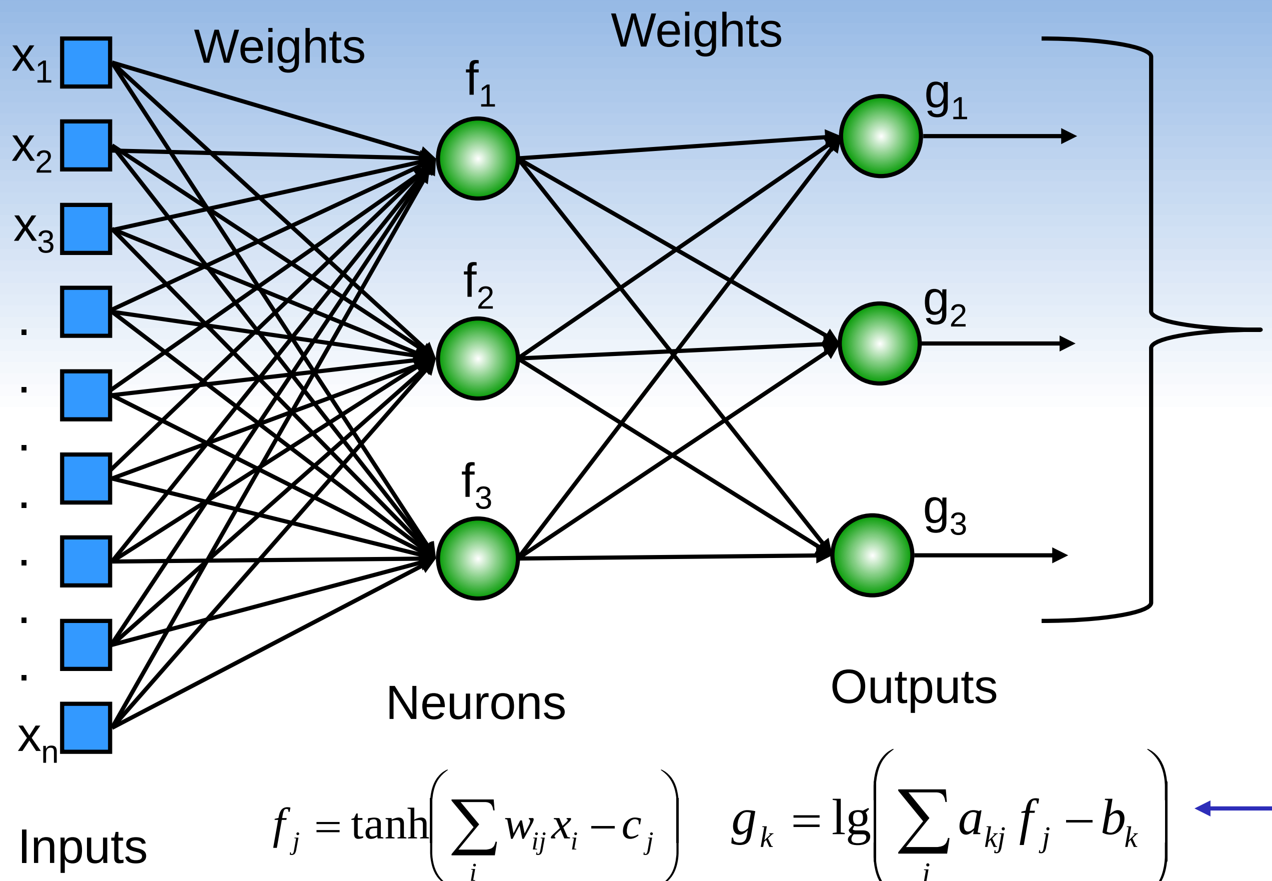
Approach

- Extend the machine learning algorithms of ADMET Modeler™ and ADMET Predictor™ to train and deploy Artificial Neural Network Ensemble (ANNE) models for predicting multi-class endpoints



SimulationsPlus
SCIENCE + SOFTWARE = SUCCESS

Multi-Classifier ANN Architecture - Deployment



Largest output is winning class. For example, if $g_1 > (g_2 \text{ or } g_3)$, we predict class 1.

$$f_j = \tanh\left(\sum_i w_{ij} x_i - c_j\right) \qquad g_k = \lg\left(\sum_j a_{kj} f_j - b_k\right)$$

Logistic Output

Weights optimized to improve training set model performance

SimulationsPlus
SCIENCE + SOFTWARE = SUCCESS

Multi-Class ANN Objective Function - Training

- Let g_{ij} be the i'th output for observation j of the training set. Let observation j belong to class k. Let the number of observations in class k be N_k. Let the number of classes be K. Then:

$$F = \sum_{k=1}^{K} \left\{ \frac{1}{N_k} \sum_{j \in k} \overbrace{\left[(1 - g_{kj})^2 + \sum_{i \neq k} {g_{ij}}^2 \right]}^{\text{Squared Error Loss}} \right\}$$

Sum over classes

Sum over Observations in class k

Output of correct class should approach 1

Outputs of incorrect classes should approach 0

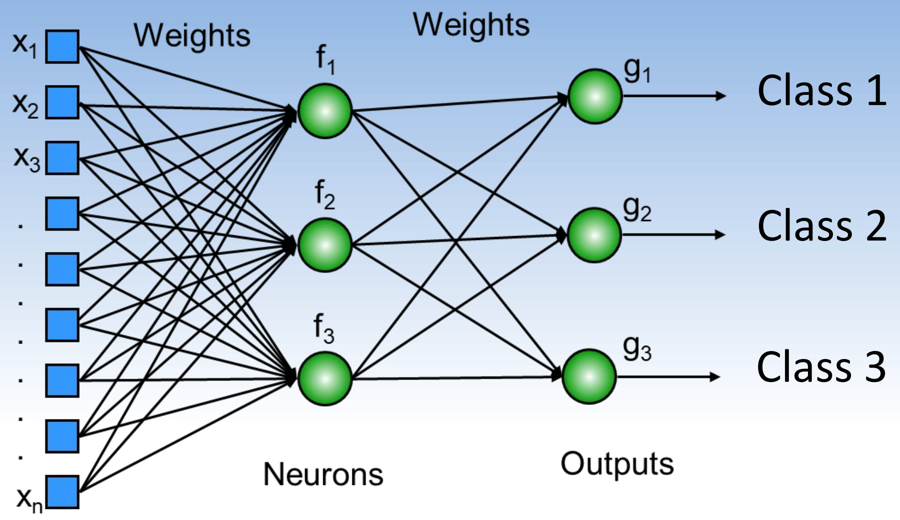
SimulationsPlus
SCIENCE + SOFTWARE = SUCCESS

Combining Individual Network Predictions

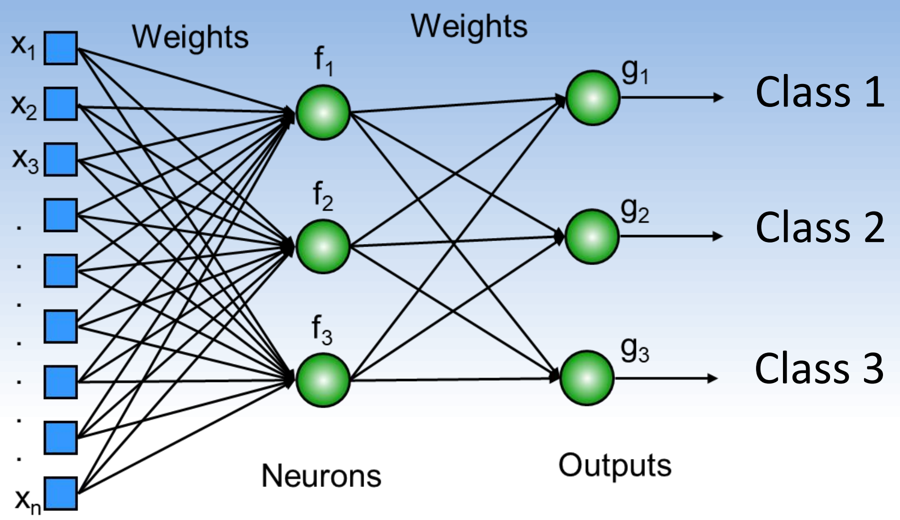
- Averaging

$$G_i = \frac{1}{N_{nets}} \sum_{j=1}^{N_{nets}} g_{ij}$$

Largest G_i is winning class



Network 1



Network 2

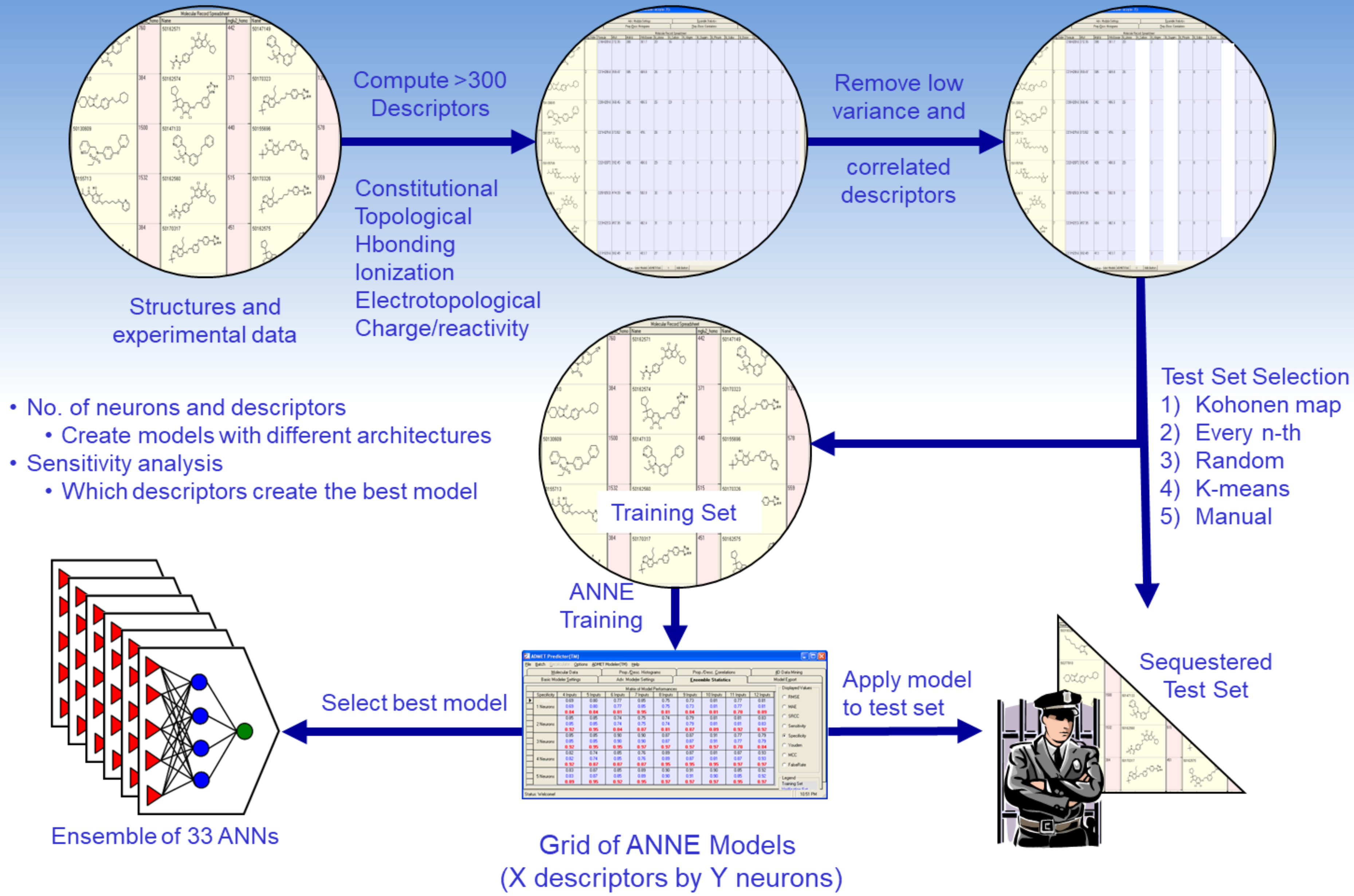
- Voting
  - Plurality with elimination (also known as instant runoff voting)
    - https://en.wikipedia.org/wiki/Instant-runoff_voting

SCIENCE + SOFTWARE = SUCCESS

Some Alternative ANN Approaches

- Softmax Output
  - https://en.wikipedia.org/wiki/Softmax_function
- Cross-Entropy Loss function
  - https://en.wikipedia.org/wiki/Cross_entropy
- Did not offer any improvement over logistic output and squared error loss on data sets we investigated

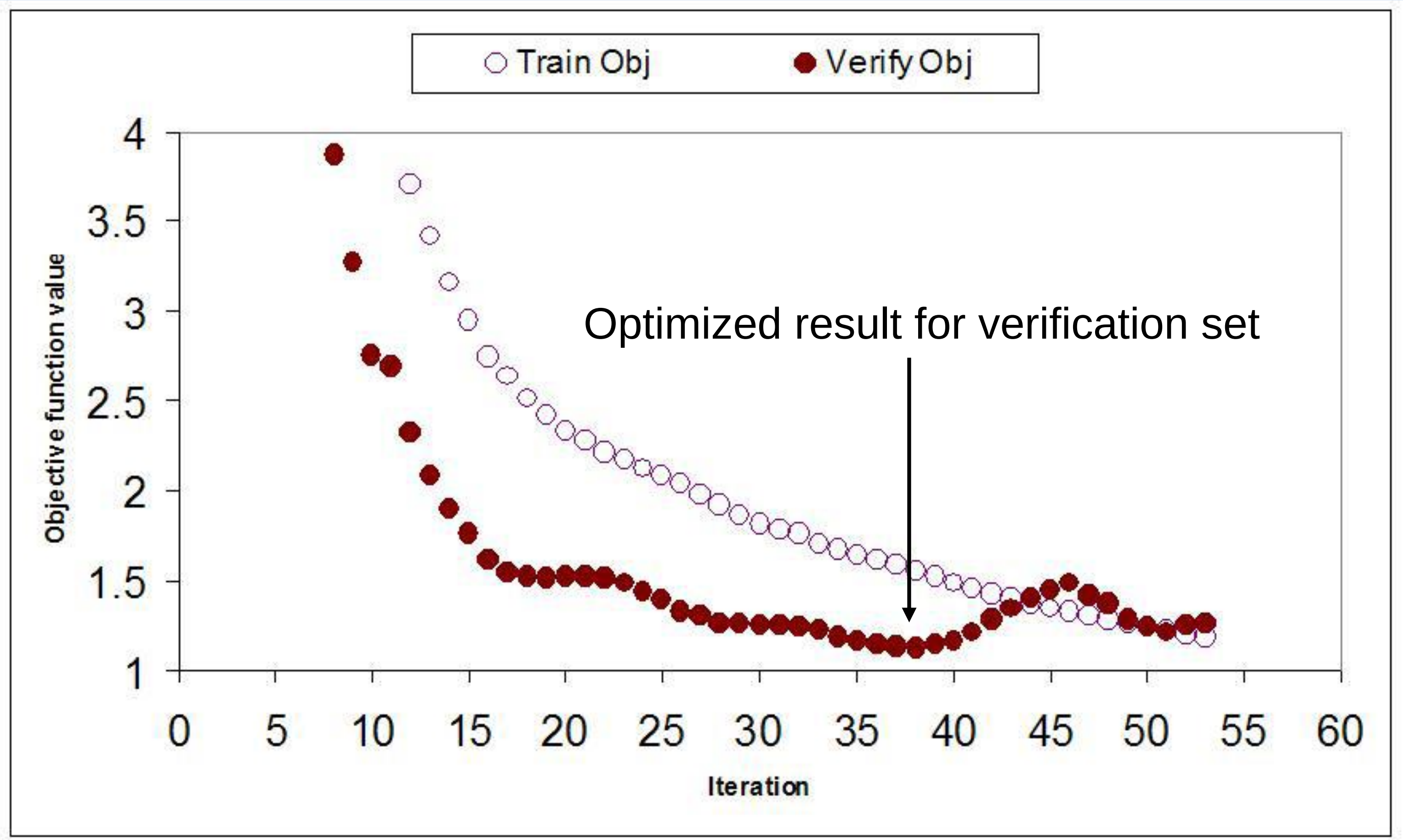
Model Building - Overview



SimulationsPlus
SCIENCE + SOFTWARE = SUCCESS

Early Stopping Avoids Overtraining

1. Split training set into training and verification sets
2. Optimize network weights to improve training set performance
3. Monitor performance of verification set – determines stopping point



SimulationsPlus
SCIENCE + SOFTWARE = SUCCESS

Some Metrics for Two-Class models

- Accuracy

$$Acc = N_{correct}/N_{total}$$

- Matthews Correlation Coefficient (perfect = 1, random = 0, worst = -1)
  - Matthews, Biochem Biophys Acta, **405**, 442, (1975)

$$MCC = \frac{TP * TN - FP * FN}{\sqrt{(TP + FP)(TP + FN)(TN + FP)(TN + FN)}}$$

- Youden's Index (perfect = 1, random = 0, worst = -1)
  - Youden, *Cancer, **3**, 32 (1950)*

$$Y = \frac{TP}{TP + FN} + \frac{TN}{TN + FP} = \text{sensitivity} + \text{specificity} - 1$$

$$\text{Note: Balanced Accuracy (BA)} = \frac{Y + 1}{2}$$

Some Metrics for Multi-Class models

- Accuracy

$$Acc = N_{correct}/N_{total}$$

- Generalized Matthews Correlation Coefficient (perfect = 1, random = 0)
  - Gorodkin, Comp. Biol. Chem., **28**, 367 (2004)

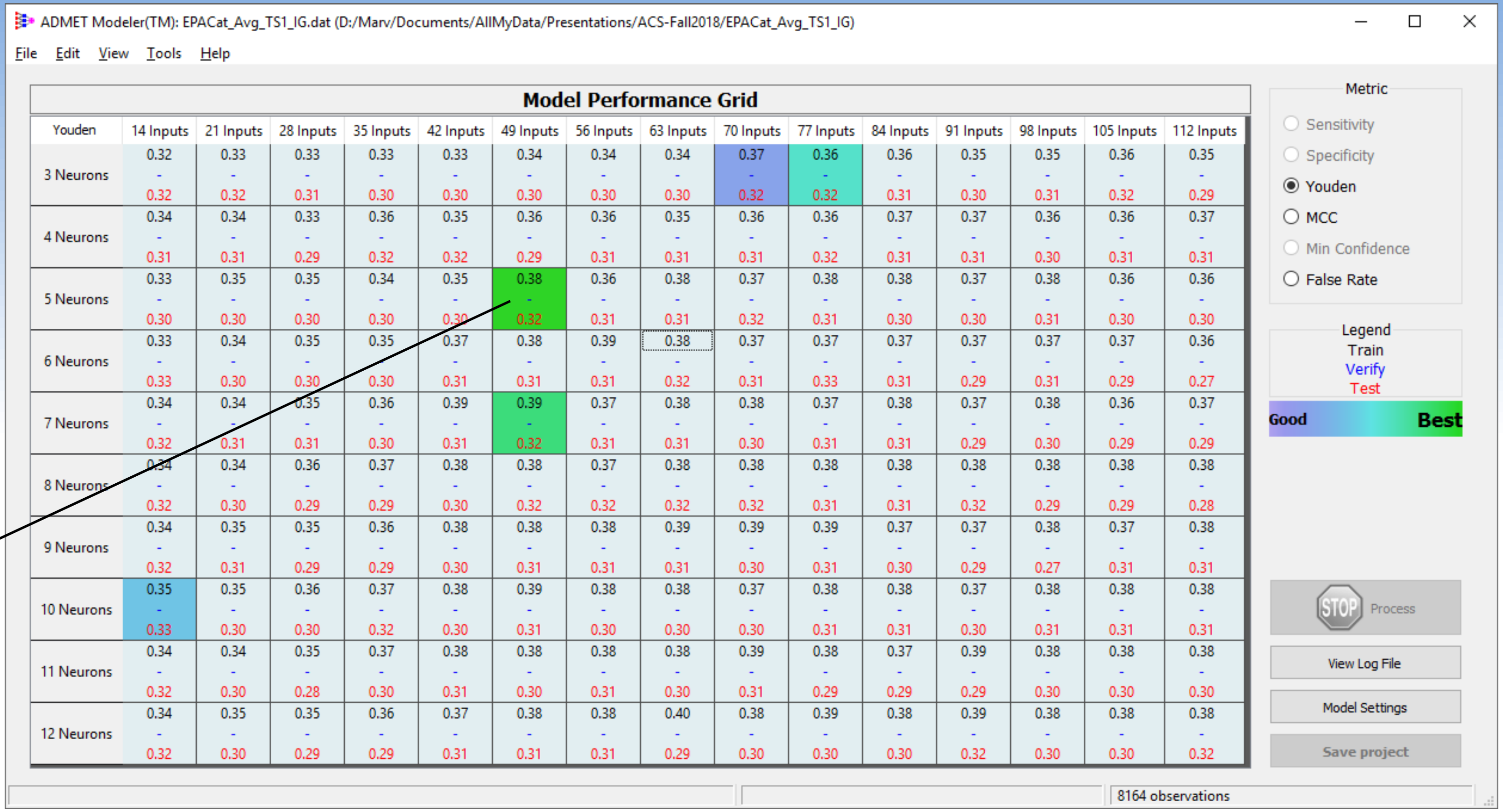
$$MCC = \frac{N_{tot}N_{corr} - \sum o_k p_k}{\sqrt{(N_{tot}^2 - \sum o_k o_k)(N_{tot}^2 - \sum p_k p_k)}}$$

o_k, p_k : Number observed, predicted in class k

- Generalized Youden's Index (perfect = 1, random = 0)

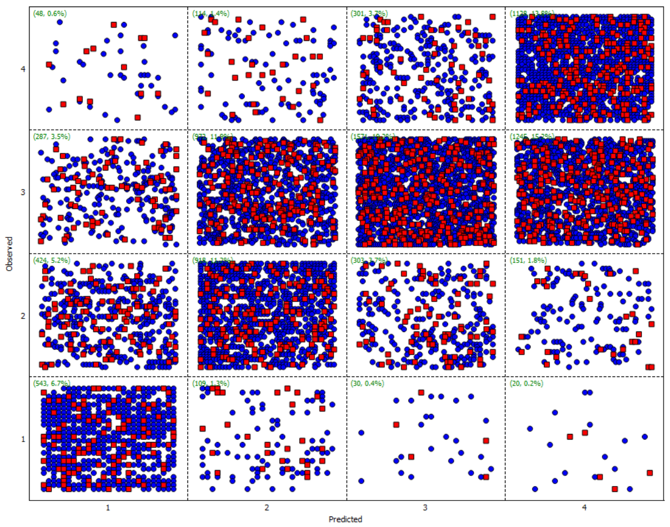
$$Y = \frac{N_{tot}N_{corr} - \sum o_k p_k}{N_{tot}^2 - \sum o_k o_k}$$

Evaluate Performance of ANNE Models



Youden	14 Inputs	21 Inputs	28 Inputs	35 Inputs	42 Inputs	49 Inputs	56 Inputs	63 Inputs	70 Inputs	77 Inputs	84 Inputs	91 Inputs	98 Inputs	105 Inputs	112 Inputs
3 Neurons	0.32 - 0.32	0.33 - 0.32	0.33 - 0.31	0.33 - 0.30	0.33 - 0.30	0.34 - 0.30	0.34 - 0.30	0.34 - 0.30	0.37 - 0.32	0.36 - 0.32	0.36 - 0.31	0.35 - 0.30	0.35 - 0.31	0.36 - 0.32	0.35 - 0.29
4 Neurons	0.34 - 0.31	0.34 - 0.31	0.33 - 0.29	0.36 - 0.32	0.35 - 0.32	0.36 - 0.29	0.36 - 0.31	0.35 - 0.31	0.36 - 0.31	0.36 - 0.32	0.37 - 0.31	0.37 - 0.31	0.36 - 0.30	0.36 - 0.31	0.37 - 0.31
5 Neurons	0.33 - 0.30	0.35 - 0.30	0.35 - 0.30	0.34 - 0.30	0.35 - 0.30	0.38 - 0.32	0.36 - 0.31	0.38 - 0.31	0.37 - 0.32	0.38 - 0.31	0.38 - 0.30	0.37 - 0.30	0.38 - 0.31	0.36 - 0.30	0.36 - 0.30
6 Neurons	0.33 - 0.33	0.34 - 0.30	0.35 - 0.30	0.35 - 0.30	0.37 - 0.31	0.38 - 0.31	0.39 - 0.31	0.38 - 0.32	0.37 - 0.31	0.37 - 0.33	0.37 - 0.31	0.37 - 0.29	0.37 - 0.31	0.37 - 0.29	0.36 - 0.27
7 Neurons	0.34 - 0.32	0.34 - 0.31	0.35 - 0.31	0.36 - 0.30	0.39 - 0.31	0.39 - 0.32	0.37 - 0.31	0.38 - 0.31	0.38 - 0.30	0.37 - 0.31	0.38 - 0.31	0.37 - 0.29	0.38 - 0.30	0.36 - 0.29	0.37 - 0.29
8 Neurons	0.34 - 0.32	0.34 - 0.30	0.36 - 0.29	0.37 - 0.29	0.38 - 0.30	0.38 - 0.32	0.37 - 0.32	0.38 - 0.32	0.38 - 0.32	0.38 - 0.31	0.38 - 0.31	0.38 - 0.32	0.38 - 0.29	0.38 - 0.29	0.38 - 0.28
9 Neurons	0.34 - 0.32	0.35 - 0.31	0.35 - 0.29	0.36 - 0.29	0.38 - 0.30	0.38 - 0.31	0.38 - 0.31	0.39 - 0.31	0.39 - 0.30	0.39 - 0.31	0.37 - 0.30	0.37 - 0.29	0.38 - 0.27	0.37 - 0.31	0.38 - 0.31
10 Neurons	0.35 - 0.33	0.35 - 0.30	0.36 - 0.30	0.37 - 0.32	0.38 - 0.30	0.39 - 0.31	0.38 - 0.30	0.38 - 0.30	0.37 - 0.30	0.38 - 0.31	0.38 - 0.31	0.37 - 0.30	0.38 - 0.31	0.38 - 0.31	0.38 - 0.31
11 Neurons	0.34 - 0.32	0.34 - 0.30	0.35 - 0.28	0.37 - 0.30	0.38 - 0.31	0.38 - 0.30	0.38 - 0.31	0.38 - 0.30	0.39 - 0.31	0.38 - 0.29	0.37 - 0.29	0.39 - 0.29	0.38 - 0.30	0.38 - 0.30	0.38 - 0.30
12 Neurons	0.34 - 0.32	0.35 - 0.30	0.35 - 0.29	0.36 - 0.29	0.37 - 0.31	0.38 - 0.31	0.38 - 0.31	0.40 - 0.29	0.38 - 0.30	0.39 - 0.30	0.38 - 0.30	0.39 - 0.32	0.38 - 0.30	0.38 - 0.30	0.38 - 0.32

Performance for 49 inputs and 5 neurons



ALL: Youden=0.370 MCC=0.350 False Rate=0.490
TRAIN: Youden=0.383 MCC=0.362 False Rate=0.481
TEST: Youden=0.319 MCC=0.300 False Rate=0.529

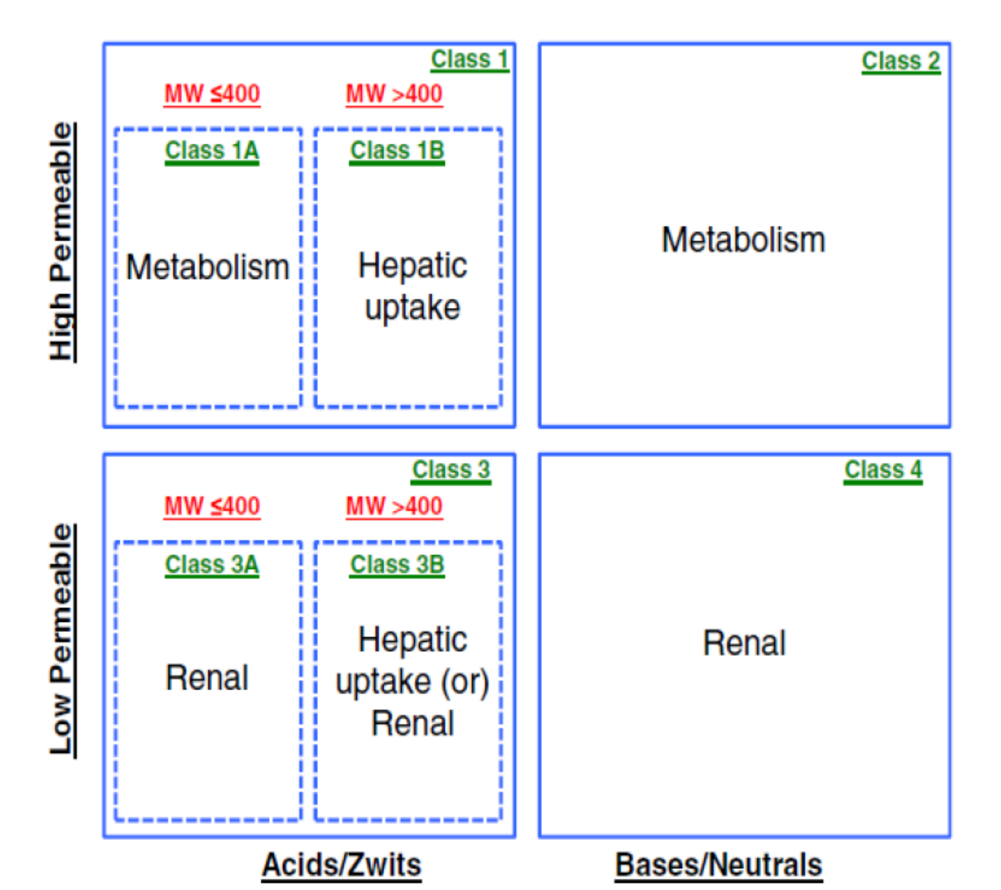
SimulationsPlus
SCIENCE + SOFTWARE = SUCCESS

Data Sets

- Dominant Clearance Mechanism
  - Compare to ECCS scheme (Extended Clearance Classification System)
- Acute Rat Toxicity Class
  - Based on LD50 cutoffs
  - 4 class scheme used by EPA
  - 5 class scheme used by GHS (Globally Harmonized System)
    - https://en.wikipedia.org/wiki/Globally_Harmonized_System_of_Classification_and_Labelling_of_Chemicals
- AMES Mutagenicity (provided by NIHS Japan)
  - 3 class model (strongly positive, positive, negative)

Extended Clearance Classification System

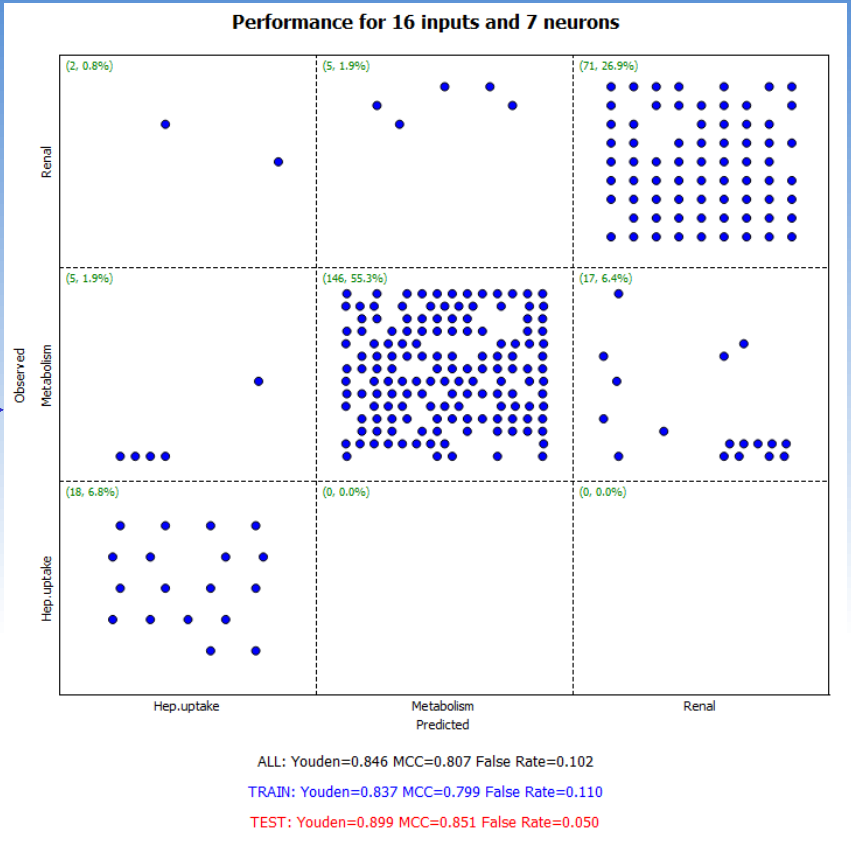
- Predicts dominant clearance mechanism of drugs
  - Varma et al, Pharm Res, 32, 3785 (2015) and subsequent publications
  - ~300 compounds



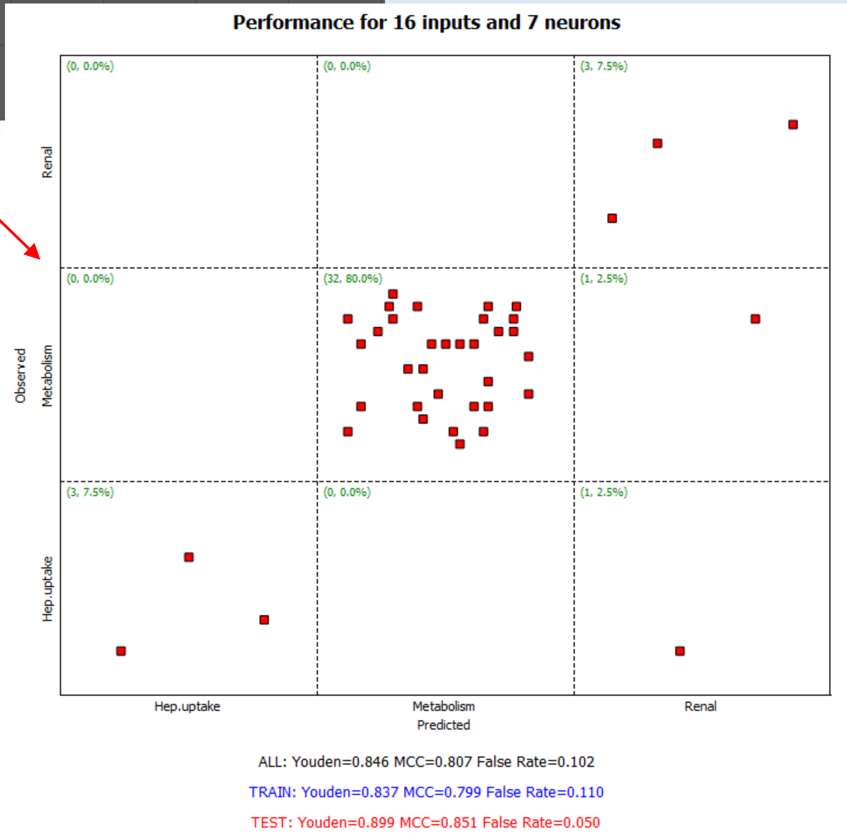
ECCS – ANNE Model

Model Performance Grid

Youden	10 Inputs	12 Inputs	14 Inputs	16 Inputs	18 Inputs	20 Inputs	22 Inputs	24 Inputs	26 Inputs	28 Inputs	30 Inputs
3 Neurons	0.76 - 0.79	0.78 - 0.83	0.82 - 0.79	0.83 - 0.81	0.83 - 0.79	0.80 - 0.79	0.80 - 0.81	0.82 - 0.81	0.81 - 0.90	0.81 - 0.81	0.81 - 0.81
5 Neurons	0.78 - 0.85	0.79 - 0.83	0.82 - 0.77	0.83 - 0.88	0.82 - 0.88	0.82 - 0.86	0.83 - 0.90	0.83 - 0.90	0.84 - 0.90	0.81 - 0.90	0.81 - 0.90
7 Neurons	0.81 - 0.90	0.80 - 0.81	0.81 - 0.90	0.84 - 0.90	0.80 - 0.81	0.80 - 0.81	0.80 - 0.79	0.80 - 0.79	0.81 - 0.79	0.80 - 0.79	
9 Neurons	0.82 - 0.88	0.82 - 0.79	0.80 - 0.88	0.81 - 0.88	0.81 - 0.79	0.81 - 0.86					
11 Neurons	0.71 - 0.51	0.71 - 0.61	0.77 - 0.77	0.78 - 0.77							



Training Set



Test Set

SimulationsPlus
SCIENCE + SOFTWARE = SUCCESS

ECCS Comparisons

Statistic	ECCS	ANNE-Train	ANNE-Test	ANNE-All	SVME-Train	SVME-Test	SVME-All
Youden	0.83	0.84	0.90	0.85	0.98	0.90	0.97
MCC	0.82	0.80	0.85	0.81	0.98	0.85	0.97
Accuracy	91%	89%	95%	90%	99%	95%	99%

ECCS : Benefit of the doubt for Class 3B (Hep. Uptake or Renal)

ANNE Model: 7 neurons, 16 descriptors
Descriptors selected by Input Gradient method
Some Key Descriptors: FAnion, S+logP, FZwitter, QAvgNeg, QAvgPos

SVME Model: 23 Descriptors selected by Genetic Algorithm
Some Key Descriptors: S+logP, No. Pi systems, T_Dipole, E-states, No. acidic atoms

SCIENCE + SOFTWARE = SUCCESS

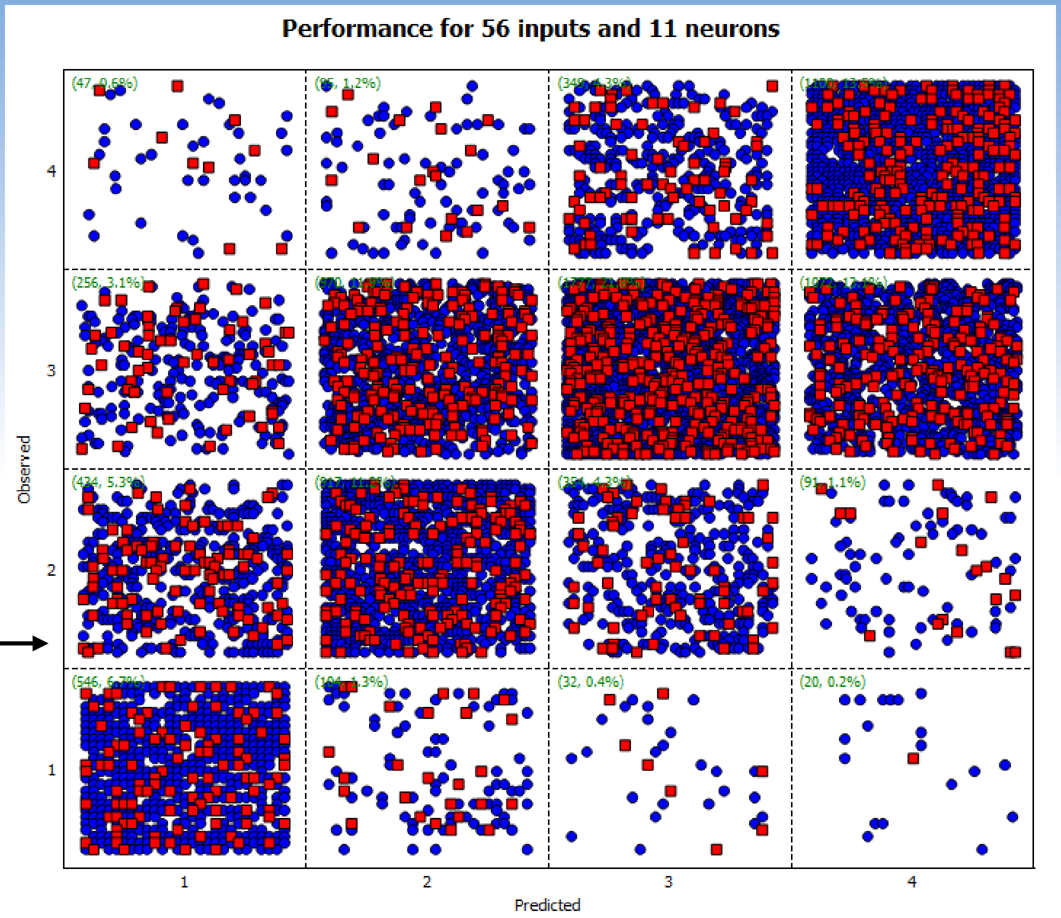
Multi-Class Toxicity Models

- Workshop on Predictive Models for Acute Oral Systemic Toxicity (April 2018)
  - Sponsor: National Toxicology Program Interagency Center for the Evaluation of Alternative Toxicological Methods (NICEATM)
    - https://ntp.niehs.nih.gov/pubhealth/evalatm/3rs-meetings/past-meetings/tox-models-2018/index.html
  - Various toxicity data sets available to develop *in silico* models and present at workshop
  - 2 Multi-class toxicity datasets were included
    - Both based on acute rat LD50 data with cutoffs
    - 4 Category model using cutoffs based on EPA guidelines
      - Category 1 : LD50 ≤ 50 mg/kg
      - Category 2 : 50 mg/kg < LD50 ≤ 500 mg/kg
      - Category 3 : 500 mg/kg < LD50 ≤ 5000 mg/kg
      - Category 4 : LD50 > 5000 mg/kg
    - 5 Category model using cutoffs based on GHS guidelines
      - Category 1 : LD50 ≤ 5 mg/kg
      - Category 2 : 5 mg/kg < LD50 ≤ 50 mg/kg
      - Category 3 : 50 mg/kg < LD50 ≤ 300 mg/kg
      - Category 4 : 300 mg/kg < LD50 ≤ 2000 mg/kg
      - Category 5 : LD50 > 2000 mg/kg
    - ~8000 compounds in each data set
    - ~4000 blind compounds for prediction

EPA 4 Class ANNE Model Performance

Model Performance Grid

Youden	14 Inputs	21 Inputs	28 Inputs	35 Inputs	42 Inputs	49 Inputs	56 Inputs	63 Inputs	70 Inputs	77 Inputs	84 Inputs	91 Inputs	98 Inputs	105 Inputs	112 Inputs
3 Neurons	0.32 - 0.33	0.33 - 0.32	0.34 - 0.33	0.34 - 0.31	0.35 - 0.32	0.34 - 0.30	0.35 - 0.32	0.35 - 0.32	0.35 - 0.31	0.36 - 0.33	0.35 - 0.32	0.35 - 0.31	0.35 - 0.32	0.36 - 0.32	0.36 - 0.33
4 Neurons	0.33 - 0.33	0.34 - 0.32	0.34 - 0.32	0.34 - 0.32	0.35 - 0.32	0.35 - 0.31	0.36 - 0.32	0.36 - 0.32	0.36 - 0.33	0.36 - 0.32	0.36 - 0.33	0.37 - 0.32	0.36 - 0.32	0.36 - 0.31	0.36 - 0.33
5 Neurons	0.33 - 0.33	0.35 - 0.33	0.35 - 0.33	0.36 - 0.32	0.37 - 0.32	0.37 - 0.33	0.36 - 0.32	0.37 - 0.33	0.38 - 0.34	0.37 - 0.34	0.37 - 0.35	0.37 - 0.34	0.38 - 0.33	0.37 - 0.33	0.37 - 0.32
6 Neurons	0.33 - 0.34	0.36 - 0.33	0.36 - 0.33	0.35 - 0.33	0.36 - 0.33	0.37 - 0.33	0.38 - 0.34	0.38 - 0.34	0.38 - 0.34	0.37 - 0.34	0.37 - 0.34	0.38 - 0.33	0.38 - 0.32	0.37 - 0.33	0.38 - 0.33
7 Neurons	0.34 - 0.34	0.36 - 0.34	0.36 - 0.33	0.36 - 0.35	0.37 - 0.33	0.37 - 0.33	0.38 - 0.34	0.38 - 0.34	0.38 - 0.34	0.38 - 0.35	0.39 - 0.34	0.38 - 0.33	0.38 - 0.33	0.38 - 0.34	0.38 - 0.33
8 Neurons	0.34 - 0.34	0.36 - 0.35	0.36 - 0.34	0.36 - 0.34	0.38 - 0.34	0.38 - 0.32	0.38 - 0.34	0.38 - 0.35	0.39 - 0.34	0.39 - 0.34	0.39 - 0.34	0.38 - 0.34	0.39 - 0.32	0.38 - 0.32	0.38 - 0.33
9 Neurons	0.35 - 0.34	0.37 - 0.35	0.37 - 0.35	0.36 - 0.33	0.38 - 0.34	0.38 - 0.34	0.39 - 0.34	0.39 - 0.35	0.40 - 0.35	0.39 - 0.34	0.39 - 0.34	0.40 - 0.34	0.39 - 0.34	0.39 - 0.34	0.40 - 0.34
10 Neurons	0.35 - 0.34	0.36 - 0.34	0.36 - 0.34	0.37 - 0.35	0.38 - 0.34	0.38 - 0.33	0.39 - 0.34	0.39 - 0.35	0.39 - 0.34	0.40 - 0.33	0.40 - 0.35	0.40 - 0.34	0.40 - 0.32	0.39 - 0.34	0.40 - 0.33
11 Neurons	0.35 - 0.35	0.37 - 0.34	0.36 - 0.34	0.37 - 0.34	0.39 - 0.35	0.38 - 0.33	0.39 - 0.36	0.39 - 0.34	0.40 - 0.34	0.40 - 0.34	0.39 - 0.34	0.40 - 0.35	0.39 - 0.34	0.40 - 0.34	0.39 - 0.33
12 Neurons	0.35 - 0.35	0.37 - 0.34	0.36 - 0.35	0.37 - 0.35	0.39 - 0.35	0.40 - 0.35	0.39 - 0.35	0.40 - 0.35	0.40 - 0.35	0.40 - 0.35	0.40 - 0.35	0.40 - 0.34	0.40 - 0.34	0.40 - 0.35	0.40 - 0.34

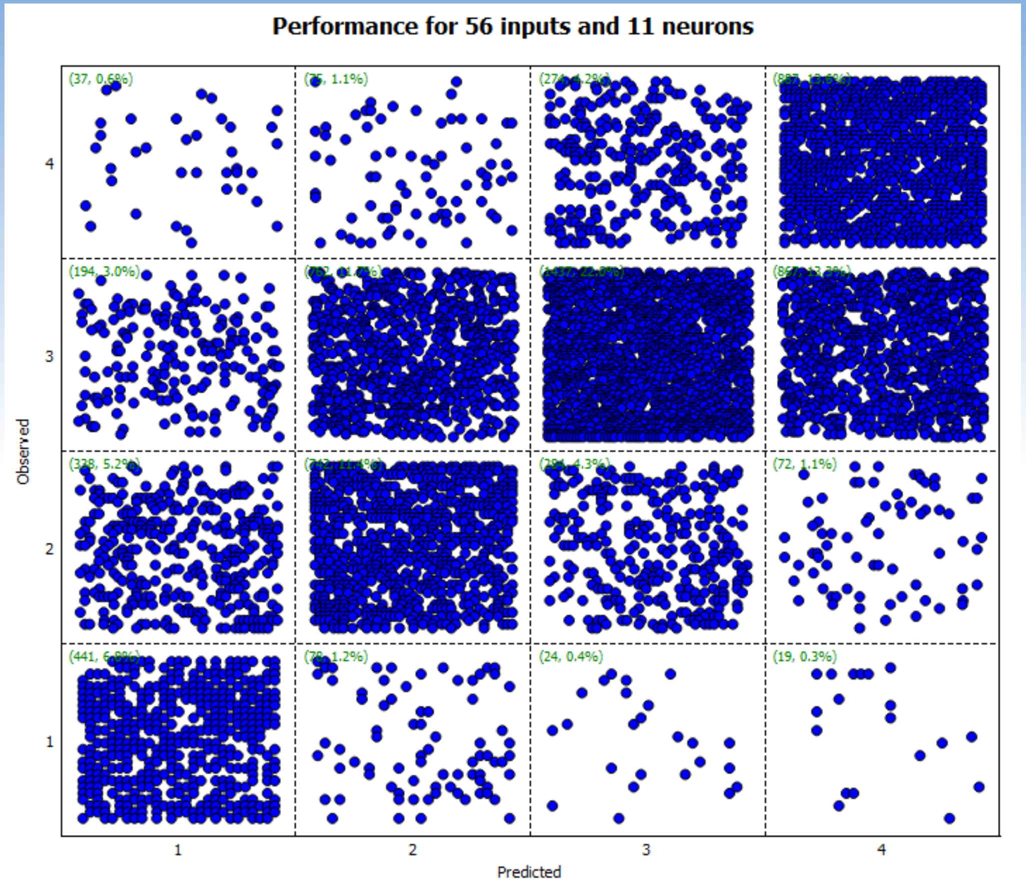


Training time for entire grid : ~3.5 hours

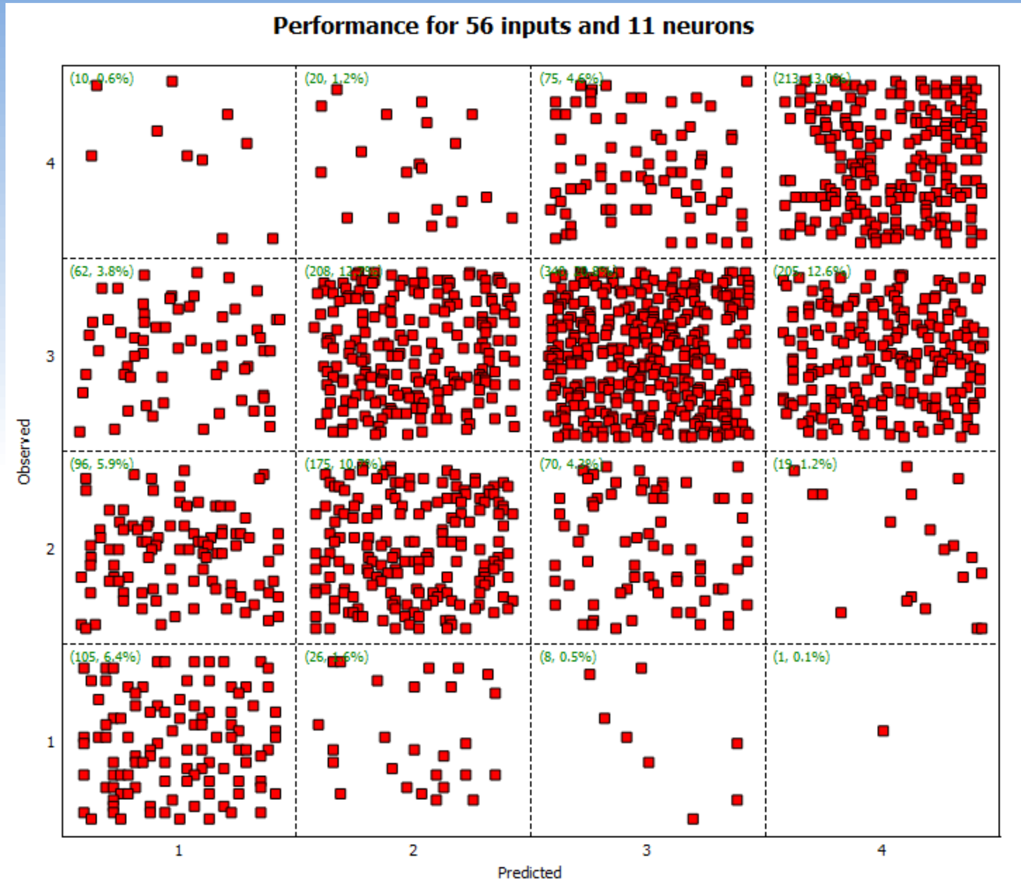
Descriptors: Standard + ANNE regression model of LD50 using NICEATM data

SimulationsPlus
SCIENCE + SOFTWARE = SUCCESS

EPA 4 Class ANNE Model Performance



Training Set



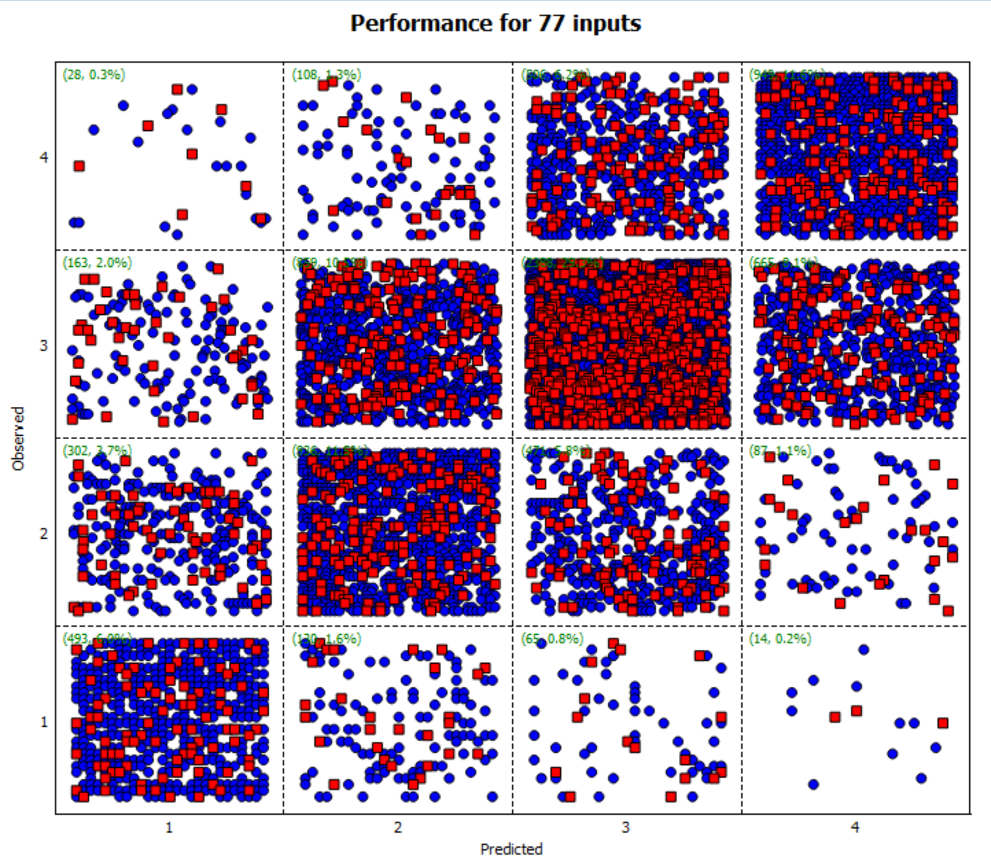
Test Set

SimulationsPlus
SCIENCE + SOFTWARE = SUCCESS

SVM Model Performance

Model Performance Grid

Youden	14 Inputs	21 Inputs	28 Inputs	35 Inputs	42 Inputs	49 Inputs	56 Inputs	63 Inputs	70 Inputs	77 Inputs	84 Inputs	91 Inputs	98 Inputs	105 Inputs	112 Inputs
CSVM	0.34 - 0.30	0.36 - 0.29	0.39 - 0.28	0.40 - 0.31	0.41 - 0.32	0.40 - 0.30	0.42 - 0.32	0.41 - 0.33	0.40 - 0.33	0.42 - 0.34	0.43 - 0.32	0.42 - 0.33	0.45 - 0.32	0.45 - 0.31	0.47 - 0.33



Training time: ~11 hours

SimulationsPlus
SCIENCE + SOFTWARE = SUCCESS

EPA 4 Class Model Performance

ANNE Training Set

Obs/Pred	1	2	3	4
1	441	78	24	19
2	338	742	284	72
3	194	762	1427	867
4	37	75	274	887

ANNE Test Set

Obs/Pred	1	2	3	4
1	105	26	8	1
2	96	175	70	19
3	62	280	340	205
4	10	20	75	213

SVM Training Set

Obs/Pred	1	2	3	4
1	400	102	49	11
2	239	768	366	63
3	122	681	1927	530
4	20	87	383	783

SVM Test Set

Obs/Pred	1	2	3	4
1	93	28	16	3
2	63	168	105	24
3	41	178	461	135
4	8	21	123	166

SimulationsPlus

SCIENCE + SOFTWARE = SUCCESS

EPA 4 Class Model Performance Metrics

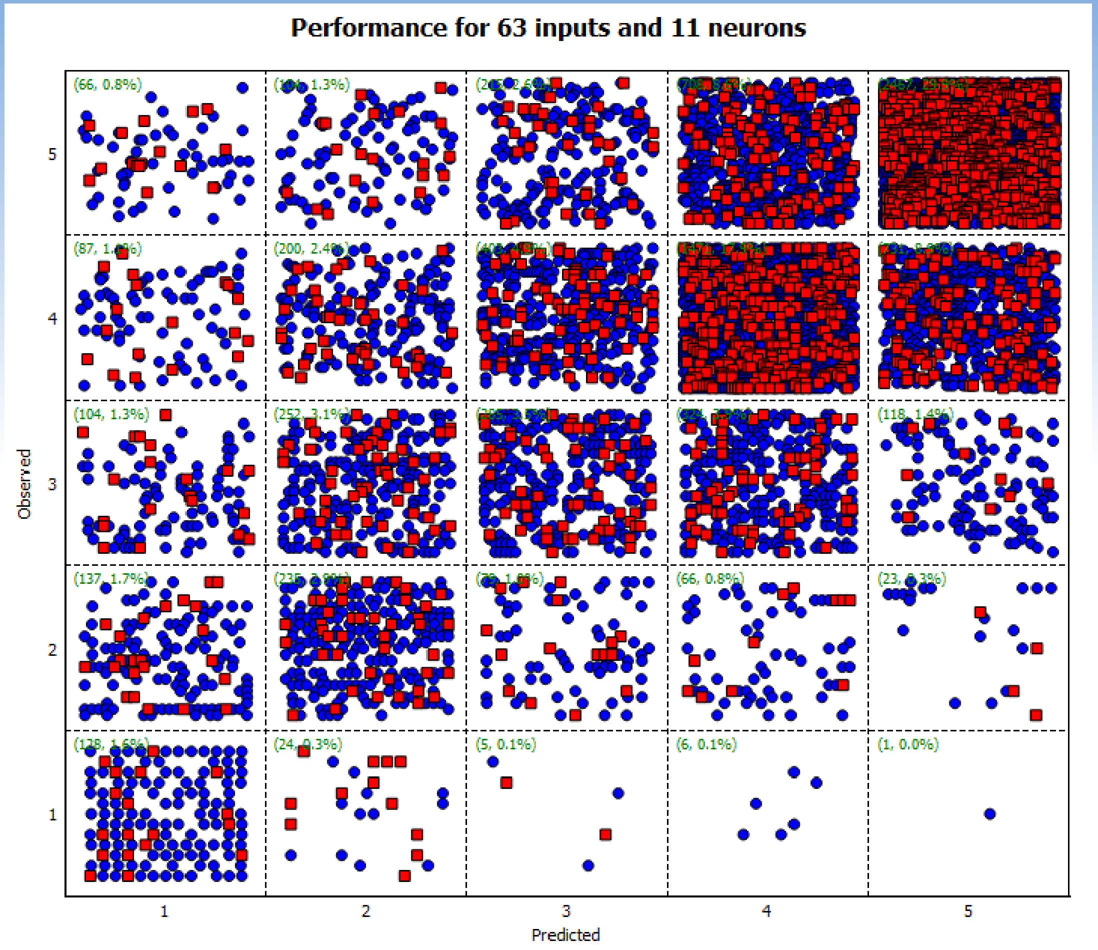
	ANNE Train	ANNE Test	ANNE All	SVM Train	SVM Test	SVM All
Youden	0.39	0.39	0.37	0.42	0.34	0.41
MCC	0.36	0.38	0.36	0.41	0.33	0.39
Acc	54%	52%	53%	59%	54%	58%

Overtrained

GHS 5 Class Model Performance

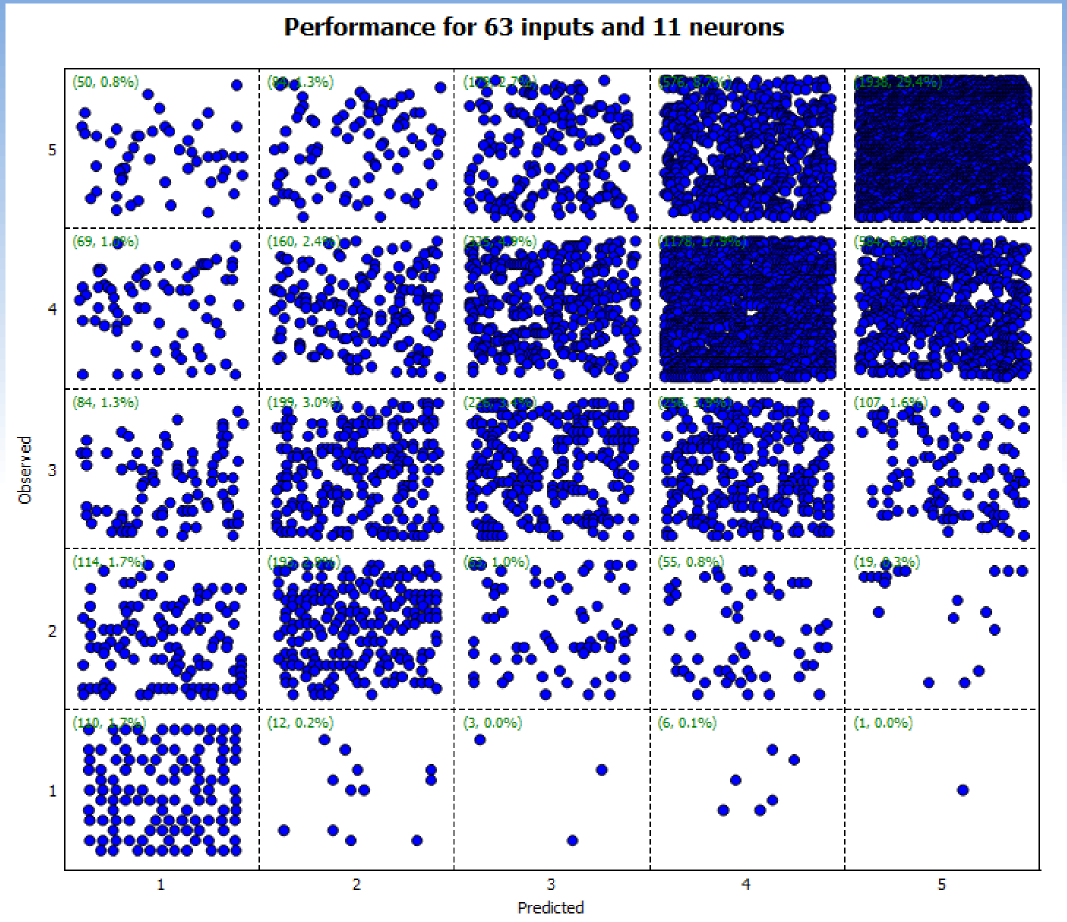
Model Performance Grid

Youden	14 Inputs	21 Inputs	28 Inputs	35 Inputs	42 Inputs	49 Inputs	56 Inputs	63 Inputs	70 Inputs	77 Inputs	84 Inputs	91 Inputs	98 Inputs
3 Neurons	0.31 - 0.33	0.32 - 0.34	0.32 - 0.36	0.33 - 0.36	0.34 - 0.36	0.34 - 0.34	0.34 - 0.34	0.34 - 0.35	0.34 - 0.37	0.35 - 0.35	0.35 - 0.36	0.35 - 0.35	0.36 - 0.37
5 Neurons	0.32 - 0.32	0.33 - 0.35	0.33 - 0.37	0.34 - 0.36	0.34 - 0.36	0.34 - 0.36	0.34 - 0.34	0.34 - 0.36	0.35 - 0.37	0.36 - 0.37	0.36 - 0.37	0.36 - 0.37	0.35 - 0.36
7 Neurons	0.32 - 0.33	0.33 - 0.34	0.34 - 0.36	0.34 - 0.37	0.34 - 0.37	0.35 - 0.35	0.35 - 0.35	0.35 - 0.38	0.35 - 0.37	0.36 - 0.36	0.36 - 0.36	0.36 - 0.37	0.37 - 0.37
9 Neurons	0.33 - 0.34	0.34 - 0.35	0.34 - 0.35	0.34 - 0.37	0.35 - 0.35	0.35 - 0.35	0.35 - 0.36	0.36 - 0.37	0.36 - 0.37	0.37 - 0.36	0.37 - 0.36	0.38 - 0.36	0.37 - 0.36
11 Neurons	0.33 - 0.34	0.33 - 0.36	0.34 - 0.36	0.34 - 0.35	0.34 - 0.36	0.35 - 0.36	0.36 - 0.36	0.37 - 0.39	0.36 - 0.37	0.37 - 0.37	0.37 - 0.36	0.37 - 0.37	0.37 - 0.36
13 Neurons	0.33 - 0.35	0.34 - 0.35	0.34 - 0.37	0.34 - 0.35	0.35 - 0.36	0.35 - 0.35	0.36 - 0.36	0.36 - 0.36	0.36 - 0.36	0.37 - 0.36	0.36 - 0.36	0.38 - 0.37	0.36 - 0.37
15 Neurons	0.33 - 0.34	0.34 - 0.35	0.34 - 0.37	0.35 - 0.36	0.35 - 0.36	0.35 - 0.37	0.35 - 0.35	0.36 - 0.37	0.37 - 0.38	0.37 - 0.37	0.38 - 0.36	0.36 - 0.36	0.36 - 0.36
17 Neurons	0.32 - 0.34	0.33 - 0.36	0.35 - 0.36	0.34 - 0.35	0.35 - 0.35	0.35 - 0.35	0.35 - 0.35	0.36 - 0.38	0.36 - 0.38				

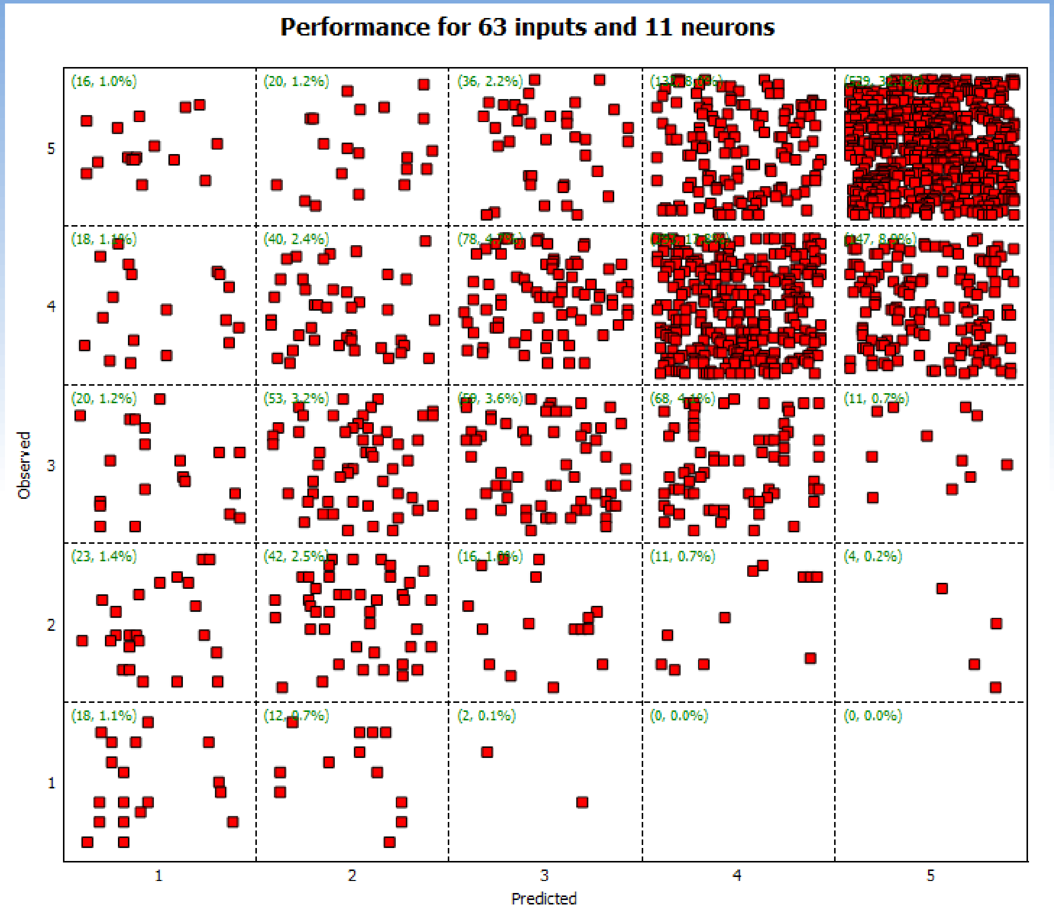


SimulationsPlus
SCIENCE + SOFTWARE = SUCCESS

GHS 5 Class Model Performance



Training Set



Test Set

SCIENCE + SOFTWARE = SUCCESS

GHS 5 Class Model Performance

Obs/Pred	1	2	3	4	5
1	110	12	3	6	1
2	114	193	63	55	19
3	84	199	226	256	107
4	69	160	325	1178	584
5	50	84	179	576	1938

Training Set

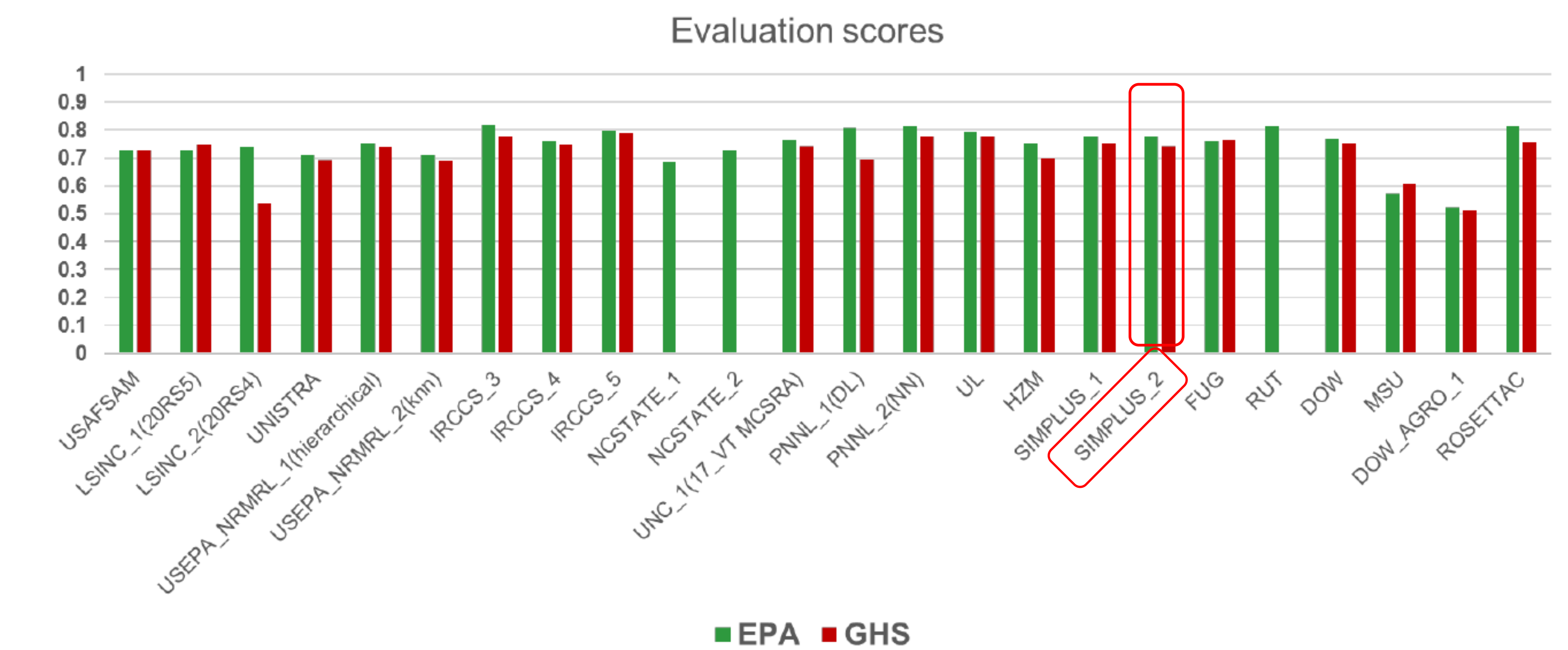
Obs/Pred	1	2	3	4	5
1	18	12	2	0	0
2	23	42	16	11	4
3	20	53	59	68	11
4	18	40	78	293	147
5	16	20	36	132	529

Test Set

	Train	Test	All
Youden	0.37	0.39	0.37
MCC	0.36	0.38	0.36
Acc	55%	57%	56%

SimulationsPlus
SCIENCE + SOFTWARE = SUCCESS

Comparison with other Participants



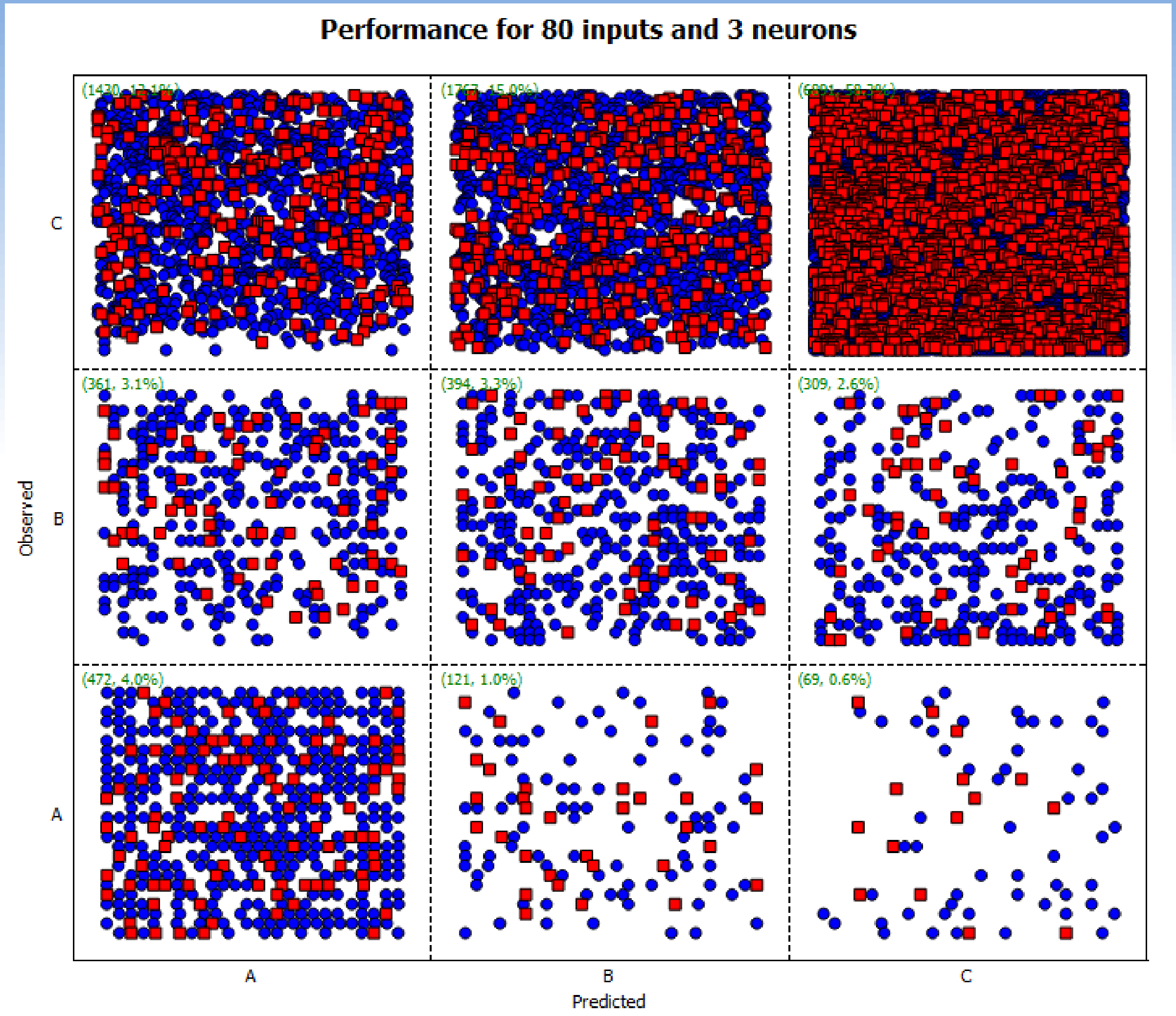
https://ntp.niehs.nih.gov/iccvam/meetings/at-models-2018/ppt/5-mansouri.pdf

SimulationsPlus
SCIENCE + SOFTWARE = SUCCESS

AMES Mutagenicity 3 class model

- Data from NIHS Japan
  - http://www.nihs.go.jp/dgm/index-e.html
  - http://www.nihs.go.jp/dgm/amesqsar.html
  - ~12000 data points
  - 3 classes
    - Class A : Strongly positive
      - Induces >1000 mutated colonies/mg in at least one AMES strain (+/- rat S9)
    - Class B : Positive
      - Induces >2-fold increase in mutated colonies relative to negative control in at least one AMES strain
    - Class C : Negative
      - Not A or B

Mut 3 Class Model Performance



Obs/Pred	A	B	C
A	373	89	54
B	292	319	251
C	1161	1424	5488

Training Set

Obs/Pred	A	B	C
A	99	32	15
B	69	75	58
C	269	343	1403

Test Set

Obs/Pred	A	B	C
A	0	0	662
B	0	0	1064
C	0	0	10088

Brain Dead
All Predictions =
Class C, Negative

SimulationsPlus
SCIENCE + SOFTWARE = SUCCESS

Mut 3 Class Model Performance Comparison

	Train	Test	All	Brain Dead
Youden	0.39	0.41	0.40	0
MCC	0.27	0.28	0.27	0/0
Acc	65%	67%	66%	85%

Summary

- ADMET Multi-Class models are seeing increased usage in industry and government
- ANNE Multi-Class methodology provides good model performance and training time performance compared to SVM and other approaches
- Use of proper metrics is critical in assessing quality of multi-class models, especially for heavily imbalanced data sets

Acknowledgements

- Co-authors
  - Pankaj Daga
  - Michael Lawless
  - Robert D. Clark
- David Miller
- Michael Bolger

SCIENCE + SOFTWARE = SUCCESS