Introduction to HTS/HCA informatics

CDDP/HtFCP infrastructure

Stand alone liquid handling

- Multidrops

- Multiflow

- Mantis

Automated liquid handling platforms

- Plate washers

- Techan Evo

- Labcyte Echo

Robotically integrated Detection systems

- imageXpress microconfocal

- InCell 6000

- ZE5

- Neo Synergy2

- Tecan M1000

Post-processing and secondary analysis

- Biovia Pipeline Pilot

- Deep Learning Studios

- Python

- Fiji/ImageJ

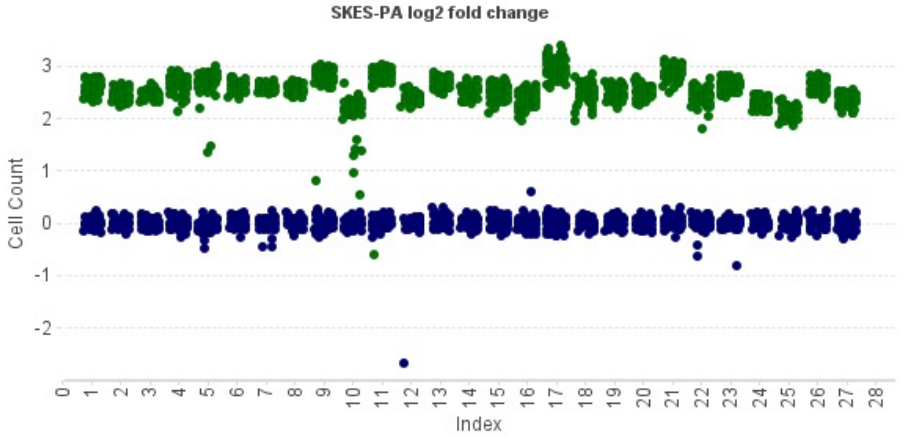
- CellProfiler

CellProfiler™
cell image analysis software

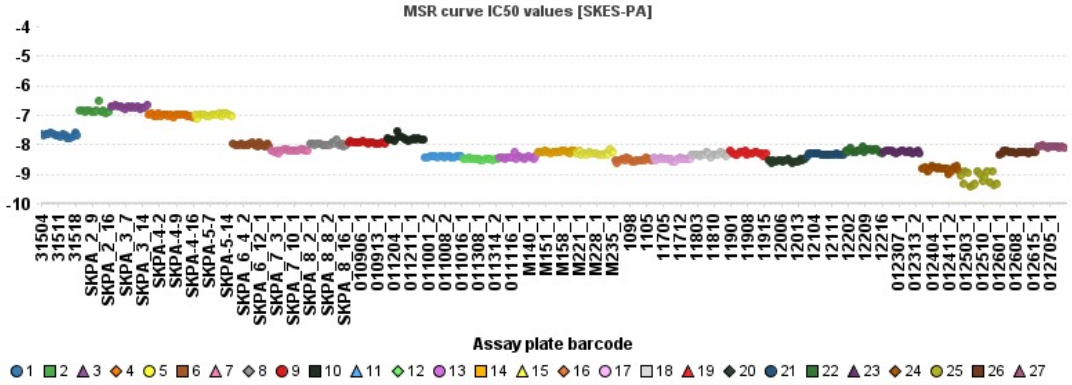
Rigor and reproducibility analysis

Knowing your data

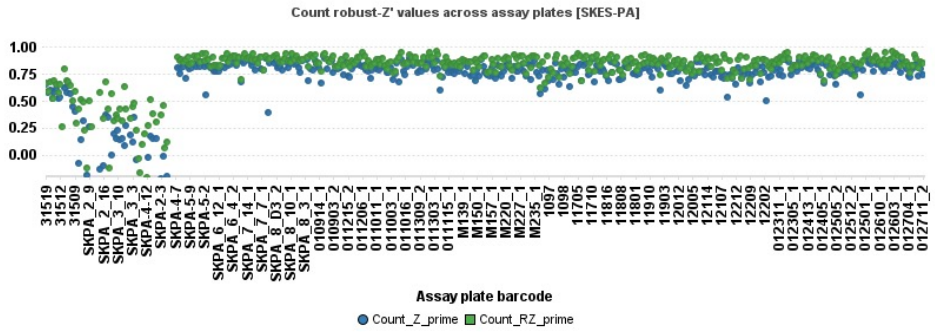
Tracking confounding variables (Rate of growth)



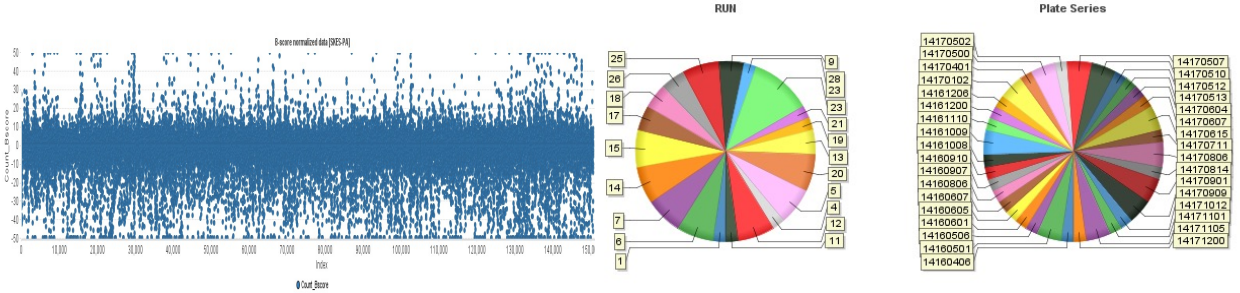
Tracking reproducibility of data (MSR)



Track robustness of the data (Z')



Systematic bias analysis



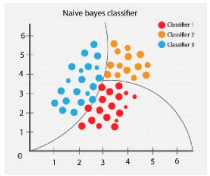
There are many uses for machine learning

Naïve Bayes model

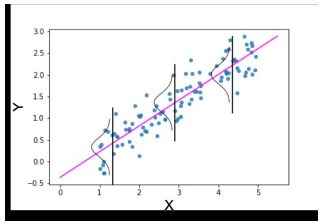
$$P(A|B) = \frac{P(B|A)\,P(A)}{P(B)}$$

using Bayesian probability terminology, the above equation can be written as

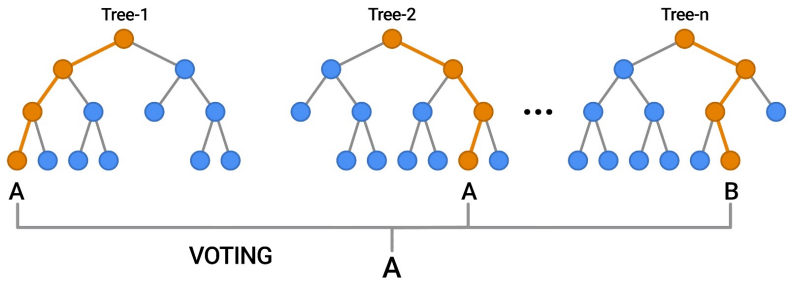
$$\text{Posterior} = \frac{\text{prior x likelihood}}{\text{evidence}}$$



General linear model



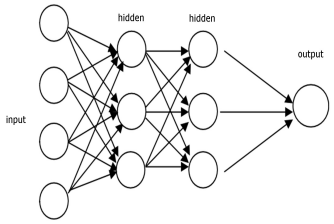
Random forest – Explainable modeling



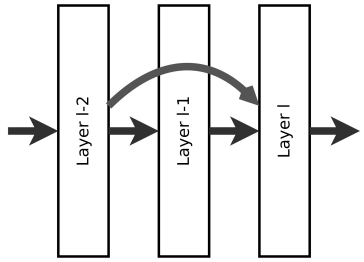
Elastic/Ridge net regression – Mutli-task learning

$$\frac{\sum_{i=1}^{n}(y_i - x_i^J \hat{\beta})^2}{2n} + \lambda \left(\frac{1-\alpha}{2} \sum_{j=1}^{m} \hat{\beta}_j^2 + \alpha \sum_{j=1}^{m} |\hat{\beta}_j| \right)$$

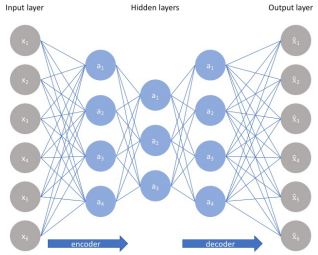
Feed forward neural network – General architecture



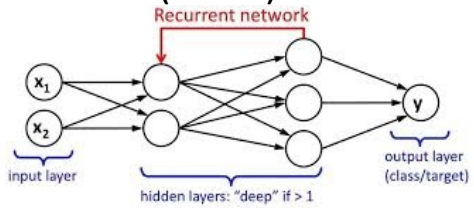
RESNET – Image classification



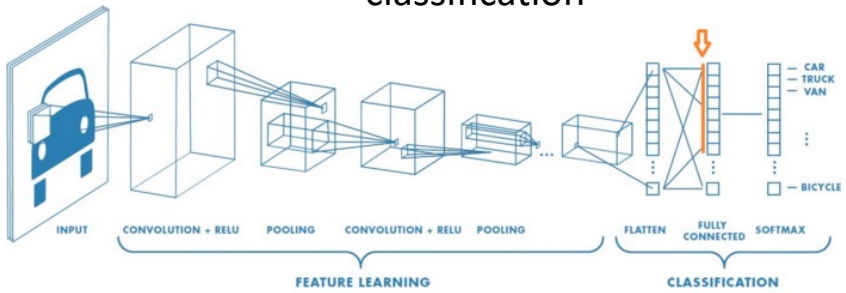
Autoencoder – compression



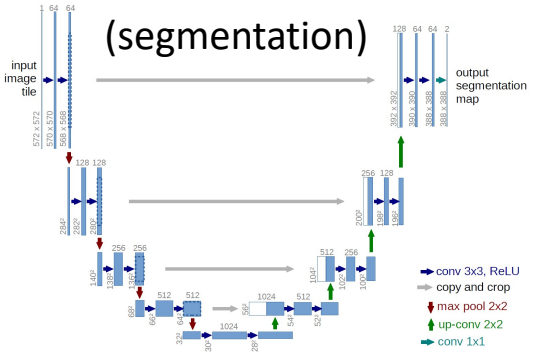
RNN (LSTM) – NLP



CNN – Image classification

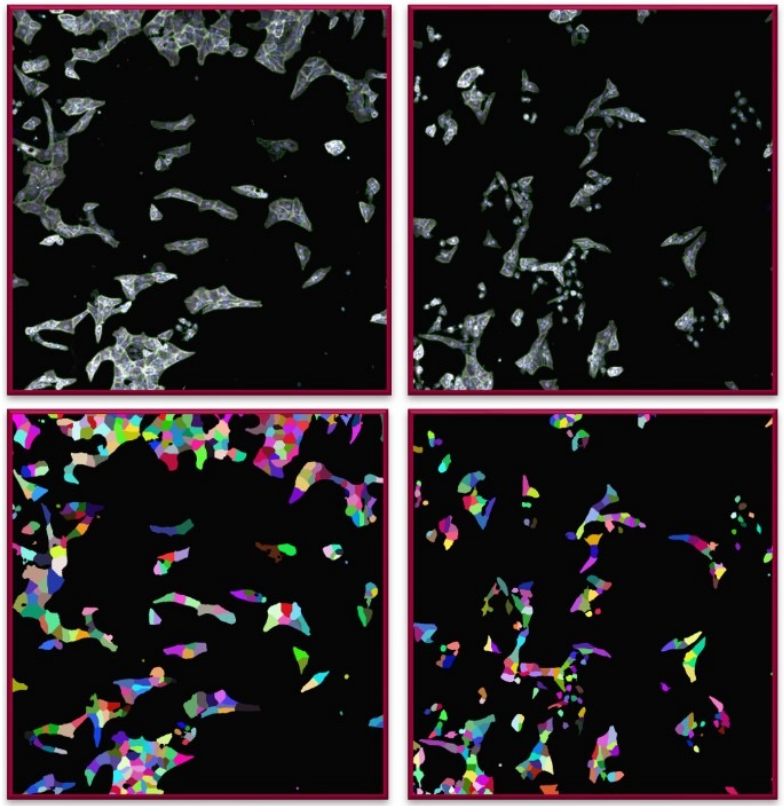


UNET – Pixel classification (segmentation)

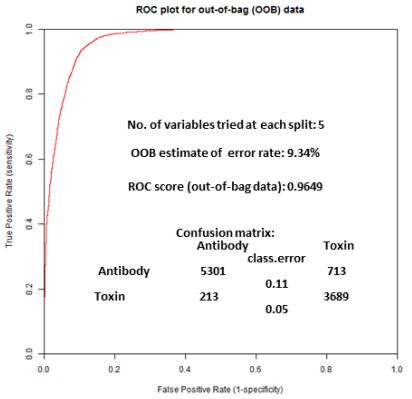


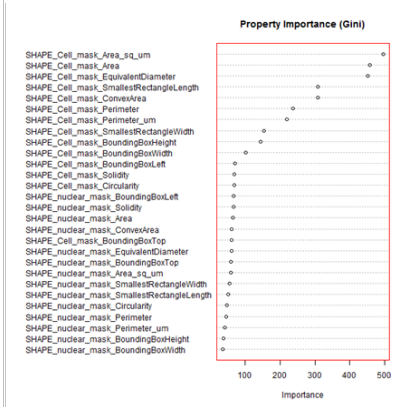
Conventional analysis using machine learned phenotypes

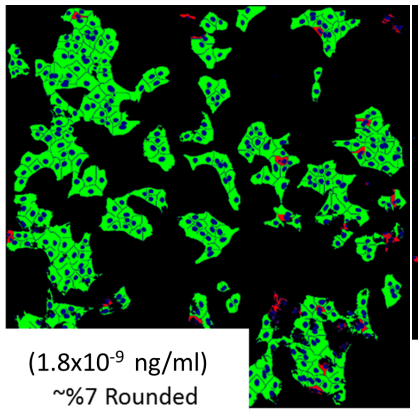
Build image library



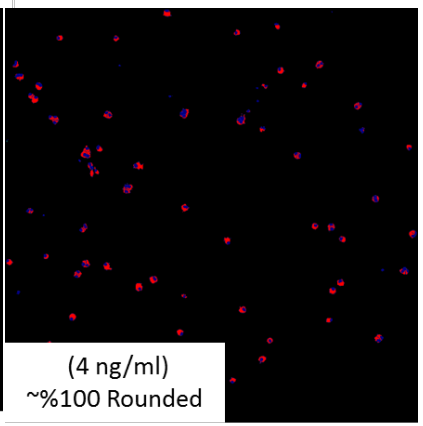
Model construction/validation





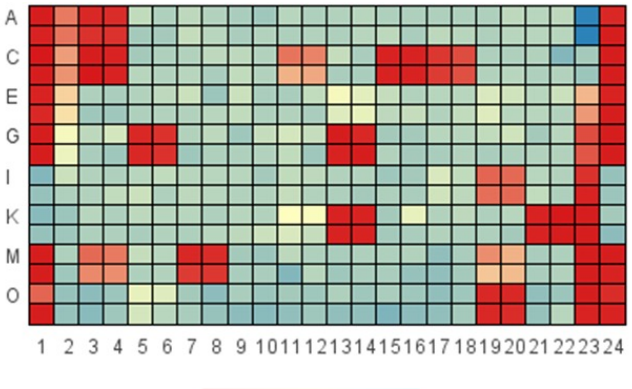


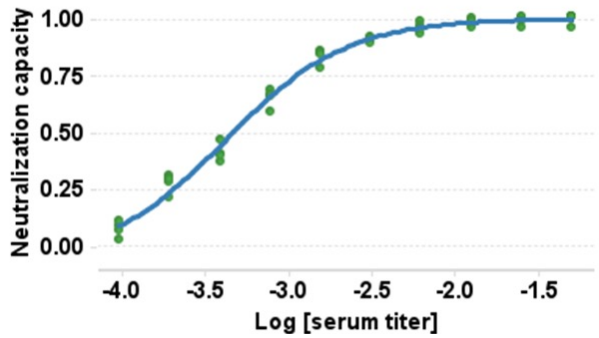
(1.8x10^{-9} ng/ml)
~%7 Rounded



(4 ng/ml)
~%100 Rounded

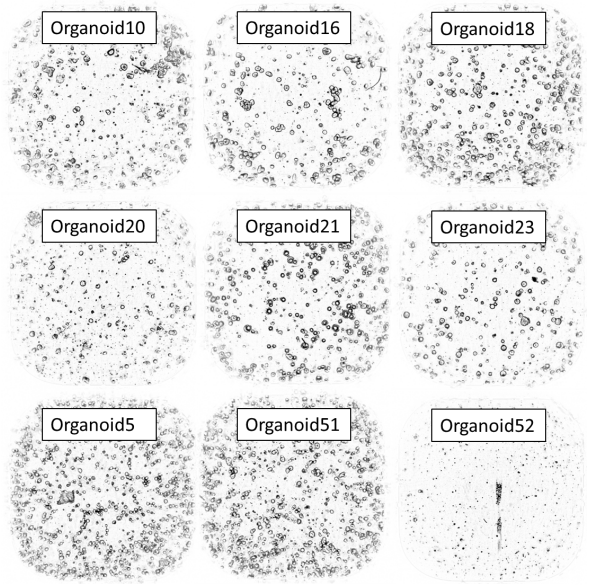
Deployment





Deep neural network driven image analysis

Validating panels of patient derived organoids



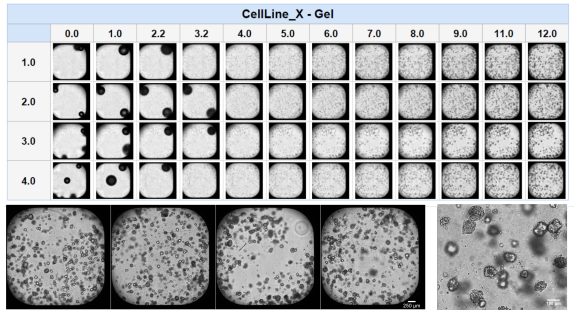
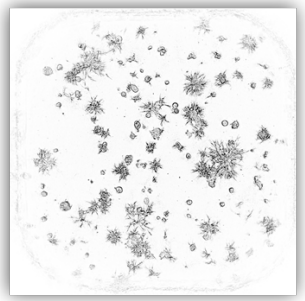
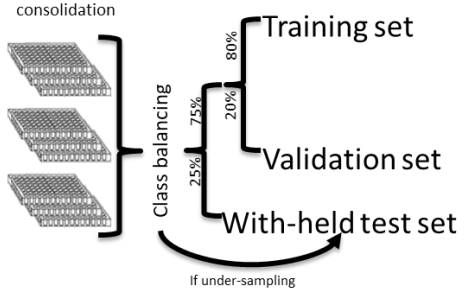
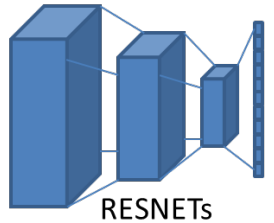


Image pre-processing → **Dataset structuring** → **Model selection** → **Output and reporting**



Projection
Intensity correction
Compression
RGB Conversation
Etc.





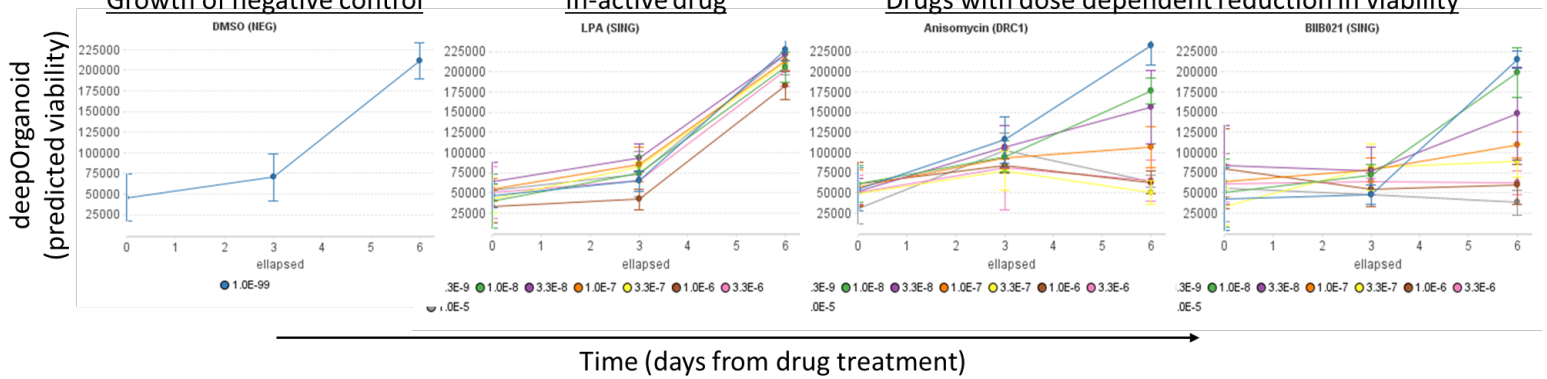
RESNETs
Inception
ModelNetv2
others

Linear embedding

Training/ Fine-tuning

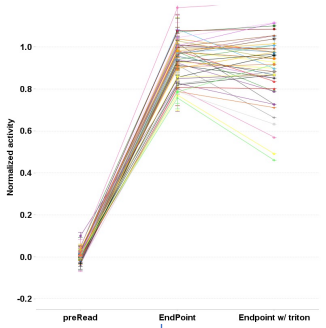
Depends on campaign stage

Growth of negative control | In-active drug | Drugs with dose dependent reduction in viability

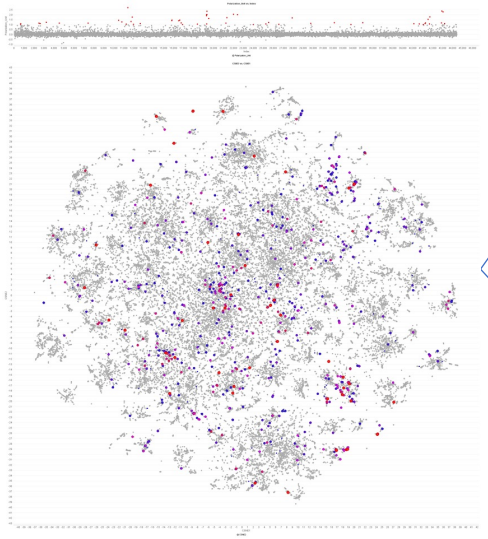


Combining chemical data with functional screening (example qSAR)

Drug activity data



Chemical fingerprints



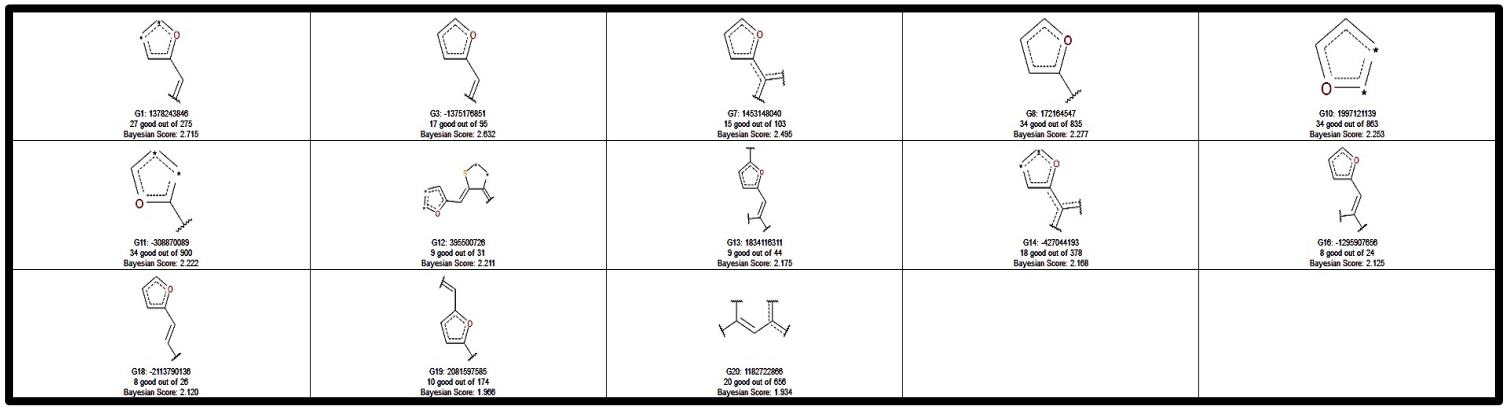
Exploratory embedding

Training (66%)

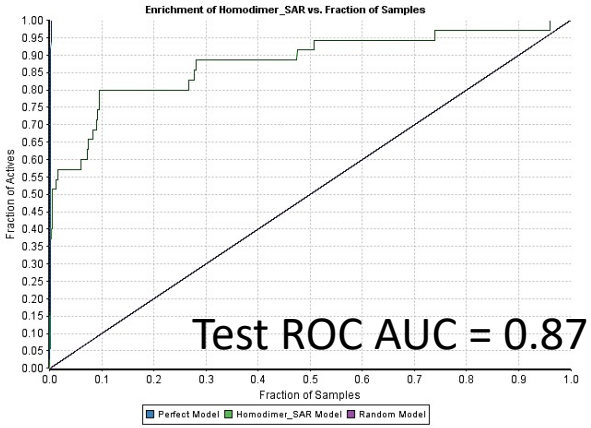
XV ROC AUC = 0.86

Testing (33%)

Explore chemical features associated with active molecules

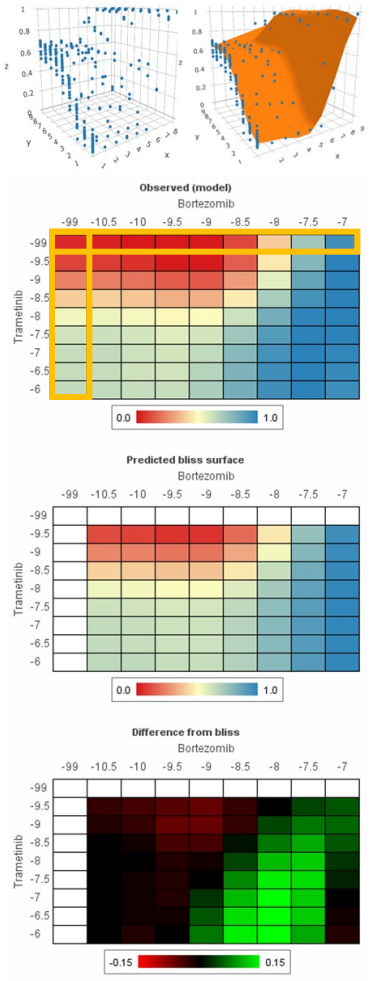


Evaluate model generalizability



Statistical methods to evaluate drug combinations

$$FA = 1 - \left(\frac{Int_{ij}}{Int_{DMSO}}\right)$$



Normalize data
0 (in-active) to 1 (full activity)

↓

Fit normalized combination surface with SVM model
(provides additional rigor and minimizes FDR from outliers)

↓

Visualize surface in 3d surface as 2d heatmap embedding
(single agents highlighted in gold)

↓

Estimate additive surface using a bliss independence model
($P_{bliss} = D_{1i} + D_{2i} - D_{1i} \times D_{2i}$)

↓

Calculate the difference (volume) between surfaces
(Red= antagonistic, Black=additive, Green=Synergistic)