Artificial Neural Networks in Cardiac Safety Assessment: Classification of Chemotherapeutic Compound Effects on hiPSC-derived Cardiomyocyte Contractility

J. L. Hunker[1,2], M. Goßmann[1], A. H. Raman[2], P. Linder[1]

[1] innoVitro GmbH, Artilleriestraße 2, 52428 Jülich, Germany
[2] Institute of Bioengineering, FH Aachen University of Applied Sciences, Heinrich-Mußmann-Str. 1, 52428 Jülich, Germany

FH AACHEN UNIVERSITY OF APPLIED SCIENCES

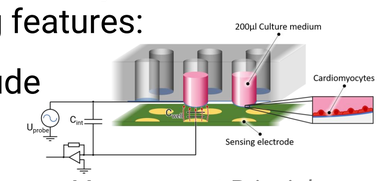
Abstract

In the last decades, technological progress in computing capacity, data acquisition and its storage has enabled new possibilities to analyze big data using machine learning, which is incomprehensible for manual analysis. Therefore, big data analysis and machine learning are well suited in the realms of drug development, to analyze the complexity of cellular processes. The objective of the present study is to develop and train an artificial neural network (NN) based on contractility parameters that differentiates compounds according to their influence on cellular pathways in human induced pluripotent stem cell-derived cardiomyocytes (hiPSC-CMs).

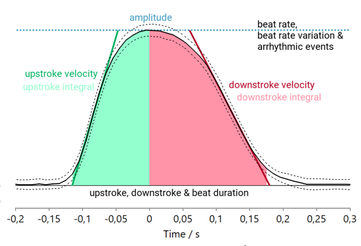
Data Acquisition

The FLEXcyte technology enables high resolution contractility measurement of hiPSC-derived cardiomyocytes cultured on hyper-elastic silicone membranes, mimicking the physiological mechanics of cardiac tissue [1, 2]. Analysis of the biohybrid's deflection provides the following features:

1. Contraction amplitude
2. Beating rate
3. Upstroke duration
4. Downstroke duration
5. Upstroke velocity
6. Downstroke velocity
7. Upstroke integral
8. Downstroke integral
9. Day of measurement



Measurement Principle



Mean Beat Analysis

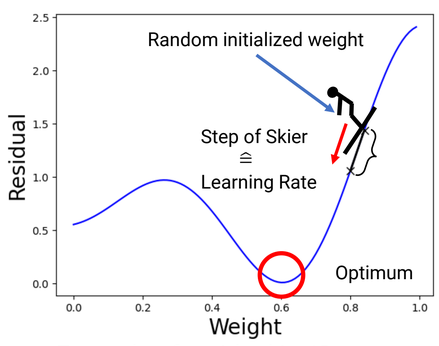
These features are used by a NN to classify 5 tyrosine kinases inhibitors (TKIs) into 11 signaling pathways (SPs) that are influenced by their effects.

No. Samples	Compound	VEGFR	EGFR	PDGRF	RET	AKT	CSF1R	FLT3	KIT	ABL	ARG	RAF
40	Sunitinib	1	0	1	0	0	1	1	1	0	0	0
19	Sorafenib	1	0	1	1	0	0	1	1	0	0	1
39	Gefitinib	0	1	0	0	0	0	0	0	0	0	0
20	Nilotinib	0	0	1	0	0	0	0	1	1	1	0
20	A-674563	0	0	0	0	1	0	0	0	0	0	0
138		59	39	79	19	20	40	59	79	20	20	19

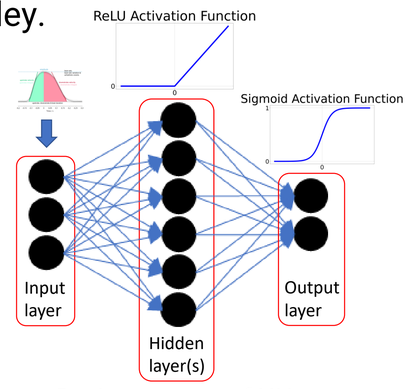
Data for NN with 5 TKIs and their effect on 11 SPs

Neural Network Design (Hyperparameters)

NNs consist of several application-dependent hyperparameters, e.g. hidden layer architecture (HLA), activation functions of neurons, learning rate (LR), and regularization. The aim is to optimize the weights of each neuron in the NN for a specific task. An analogy for this would be like skiing through the mountains (residual) in one-dimensional space (single weight) trying to reach the lowest valley.

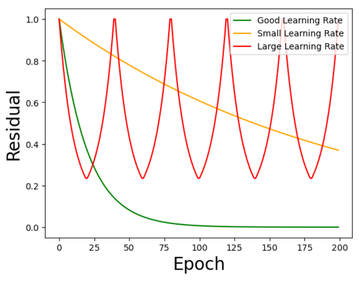


Example of residual in 1D space



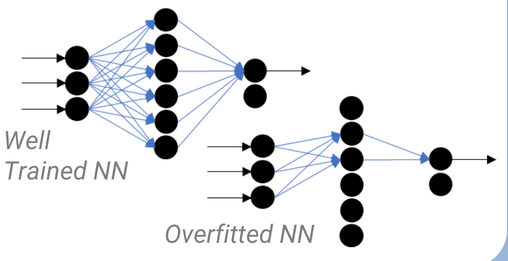
Basic structure of NN

The LR plays a critical role in reaching the optimal weights [3].



Effect of Learning Rate

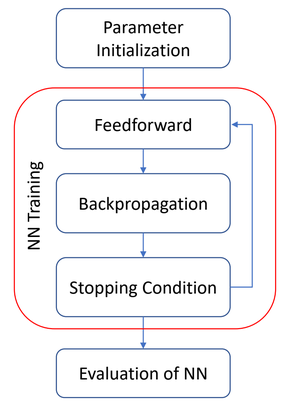
During training, dropout regularization prevents the NN from overfitting via random deactivation of neurons with a defined dropout rate (DR) [4].



Overfitted NN

Optimization (Training) of NN

Optimization starts with random initial weights [5] at each neuron to calculate the overall residual (Feedforward [3]). These random weights are optimized to attain the minimum residual (Back-propagation [3]). Adam algorithm [6] is employed in our case to optimize the weights. These new weights are then used for the next Feedforward step. This process is repeated until the stopping condition is met, e.g. the improvement of the residual drops below a certain tolerance level [7]. Afterwards the NN can be evaluated.



Performance Evaluation of the NN

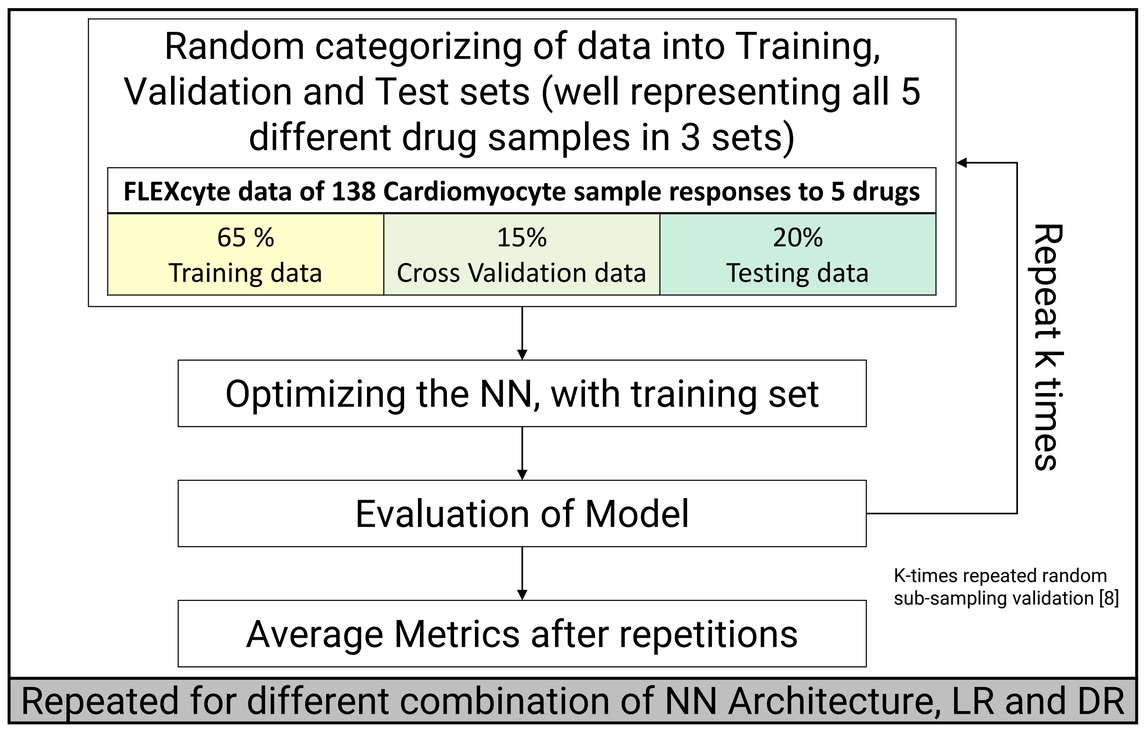
Metrics for Evaluation are based on the measures TP, FP, FN, and TN [8].

- $\text{Accuracy} = \frac{\sum(TP+TN)}{\sum(TP+FP+FN+TN)}$
  percentage of correct predictions
- $\text{ExactMatch}Ratio$
  percentage of samples with all SPs correctly predicted
- $\text{Precision} = \frac{\sum TP}{\sum(TP+FP)}$
  ratio of TPs to all predicted positives
- $\text{Recall} = \frac{\sum TP}{\sum(TP+FN)}$
  ratio of TPs to all actual positives

Prediction (H) \ Observation (Y)	P	N
P	True Positive (TP)	False Positive (FP)
N	False Negative (FN)	True Negative (TN)

Binary Confusion Matrix

Optimization of Hyperparameters



Results

Optimization of hyperparameters was performed in 2 consecutive steps:

1) Optimization of HLA
2) Optimization of LR and DR using HLA from step 1

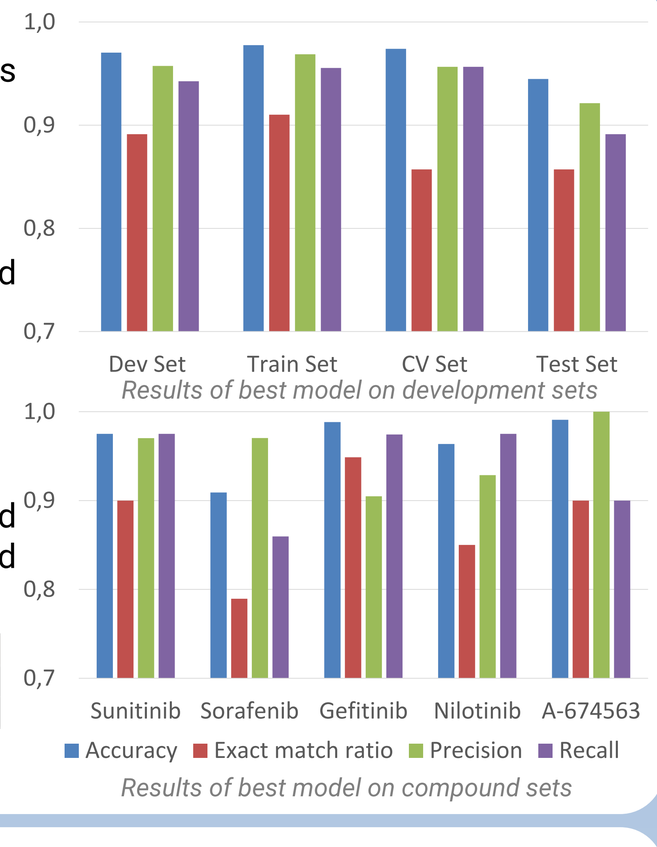
The best performance was achieved with the following hyperparameters:

- HLA with 2 hidden layers with 99 ReLU-activated neurons each
- LR of 0.025
- DR of 0.2

The best performance was achieved in run 72 of the 100-times repeated random sub-sampling validation.

	Dev Set	Train Set	CV Set	Test Set
Accuracy	0,93	0,96	0,91	0,87
Exact match ratio	0,77	0,83	0,72	0,63

Average results of 100-times random sub-sampling validation



Conclusion

On average our developed NN was able to correctly classify >90% of the corresponding target pathways using 65% of the development data for training. The remaining cross-validation (15%) and test (20%) sets ensured the functionality of the algorithm with predictive powers of 91% and 87%, respectively. In summary, the NN has shown to be a reliable analysis tool for in vitro drug screening and cardiotoxicity assays using beat shape data of hiPSC-CMs obtained with the FLEXcyte technology, and thus offers insight into the effects of compounds on the target signaling pathways. Further experiments are planned with compounds not part of the development set to show the predicting power of this NN.

References:

[1] **S. Stoelzle-feix, K. Juhasz, U. Thomas, P. Linder, M. George, M. Gossman:** Higher throughput in vitro cardiac contractility recordings under physiological mechanical conditions. 2019
[2] **M. Goßmann:** Entwicklung eines autokontraktilen Herzmuskelmodells zur funktionalen Medikamenten-und Toxinforschung. PhD Thesis University Duisburg-Essen, 2015
[3] **J. Han, M. Kamber, J. Pei, J. Han, M. Kamber, and J. Pei:** Data Mining Concepts and Techniques, Chapter 9 – Classification: Advanced Methods. Elsevier 2012
[4] **N. Srivastava, G. Hinton, A. Krizhevsky, and R. Salakhutdinov:** Dropout: A Simple Way to Prevent Neural Networks from Overfitting. 2014
[5] **K. He, X. Zhang, S. Ren, and J. Sun:** Delving Deep into Rectifiers: Surpassing Human-Level Performance on ImageNet Classification. Proceedings IEEE International Conference on Computer Vision (ICCV), 2015, pp. 1026–1034.
[6] **D. P. Kingma and J. L. Ba:** Adam: A method for stochastic optimization. 3rd Int. Conf. Learn. Represent. ICLR 2015 - Conf. Track Proc., pp. 1–15, 2015.
[7] **Prechelt L.:** Early Stopping — But When?. In: Montavon G., Orr G.B., Müller KR. (eds) Neural Networks: Tricks of the Trade. Lecture Notes in Computer Science, vol 7700. Springer, Berlin, Heidelberg. 2012
[8] **M. Sokolova and G. Lapalme:** A systematic analysis of performance measures for classification tasks," Inf. Process. Manag., vol. 45, no. 4, pp. 427–437, 2009
[9] **J. Han, M. Kamber, and J. Pei,** Data Mining Concepts and Techniques, Chapter 8 - Classification: Basic Concepts. Elsevier 2012

FLEXcyte 96 Whitepaper

Scan me

This research project was supported by:

Federal Ministry of Education and Research

JOHNS HOPKINS UNIVERSITY