A New Interpretable Neural Network-Based Rule Model for Healthcare Decision Making

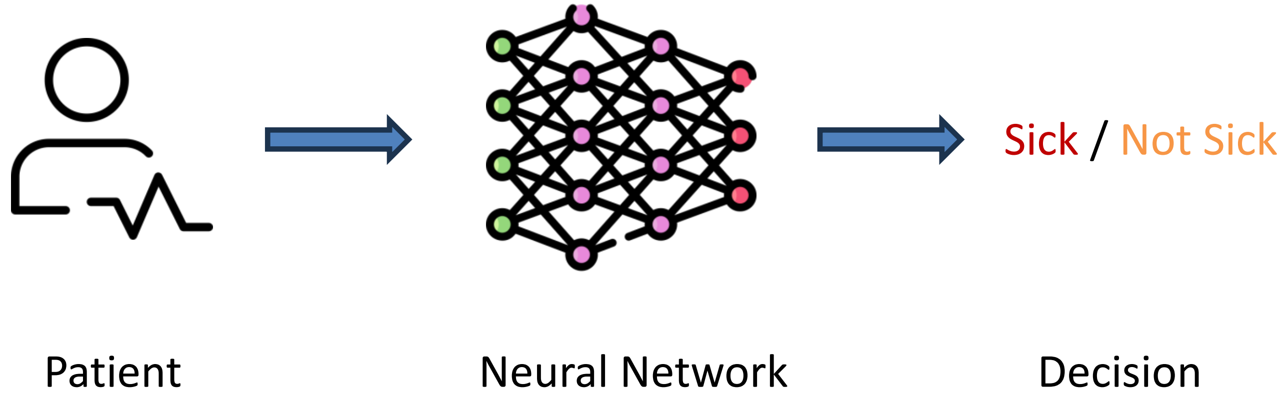
Adrien Benamira*, Tristan Guérand*, Thomas Peyrin
* main contribution

Nanyang Technological University

04/10/2023

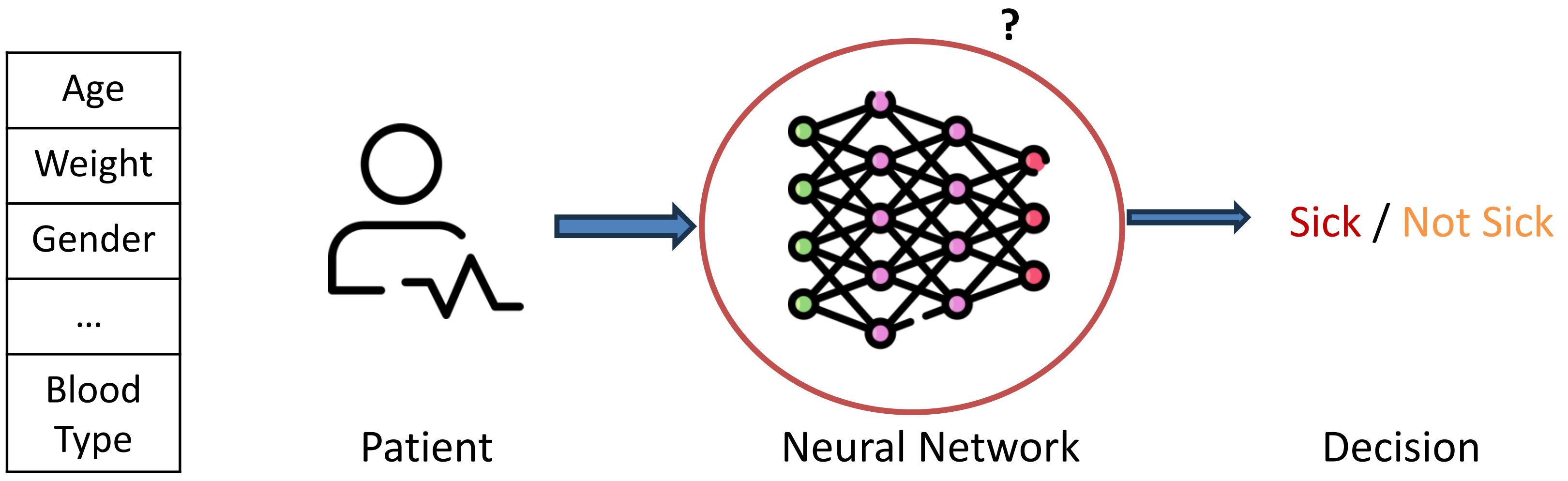
Introduction

Why Explainability ?



Introduction

Why Explainability ?



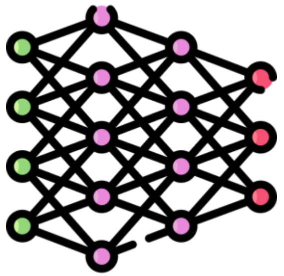
→ Why did the NN decides that the patient is sick ?
→ Which features matter ?

Explainable AI (XAI)

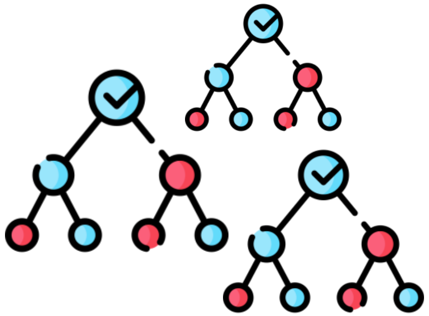
Introduction

What is XAI

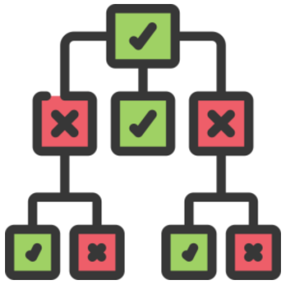
Black Box | Interpretable / Explainable methods



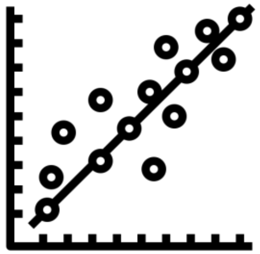
Neural Network



Random Forest



Decision tree



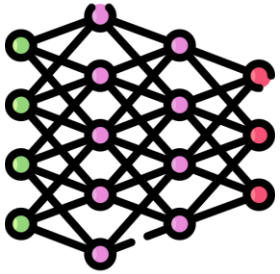
Linear Regression

	Neural Network	Random Forest	Decision tree	Linear Regression
Scalable	✓✓	✓	✗✗	✗
Performances	✓✓	✓	✗	✗✗
Interpretable	✗✗	✗	✓✓	✓✓

Less Interpretable → More Interpretable

Introduction

What do we want



Neural Network

Decision tree

Scalable
Performances
Interpretable

TTnet

From black box to truth tables

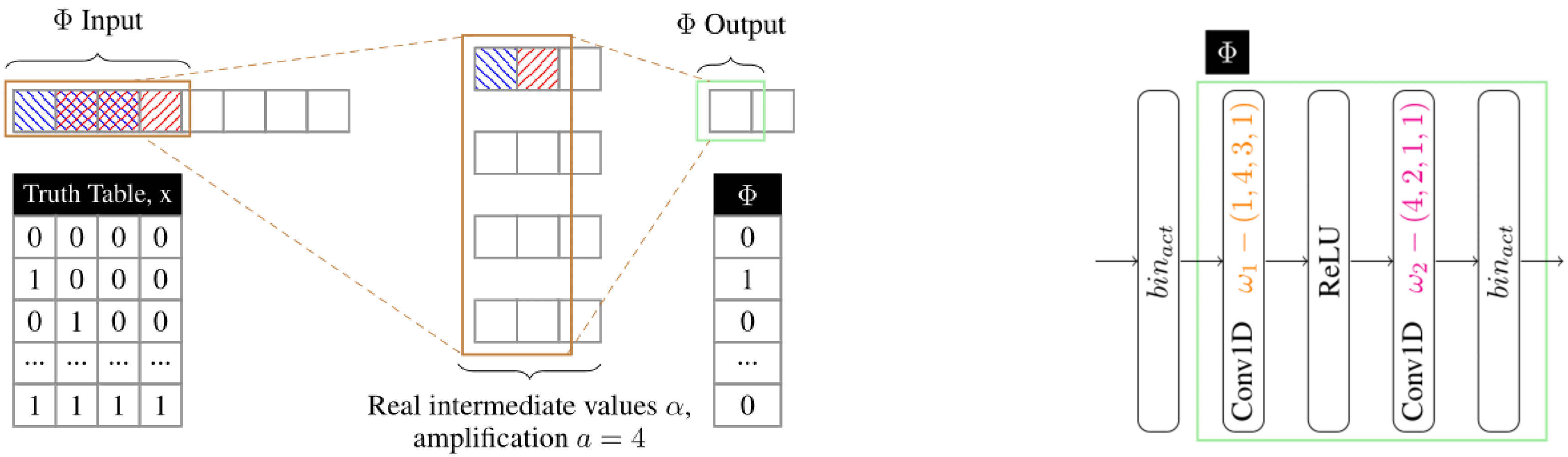


Figure from [4]

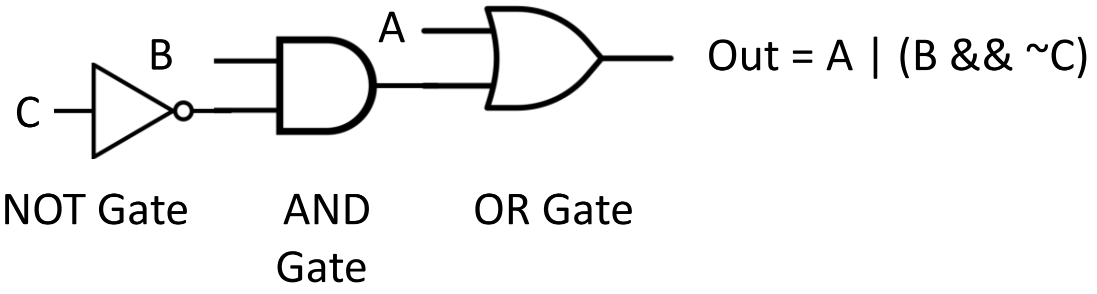
Convolution block ⇔ Truth Table

TTnet

From black box to truth tables

What is the most complete, objective, simple form of information?

→ Truth Tables (for discrete at least)



Out Function Truth Table

A	B	C	Out
0	0	0	0
0	0	1	0
0	1	0	1
0	1	1	0
1	0	0	1
1	0	1	1
1	1	0	1
1	1	1	1

TTnet

From black box to truth tables

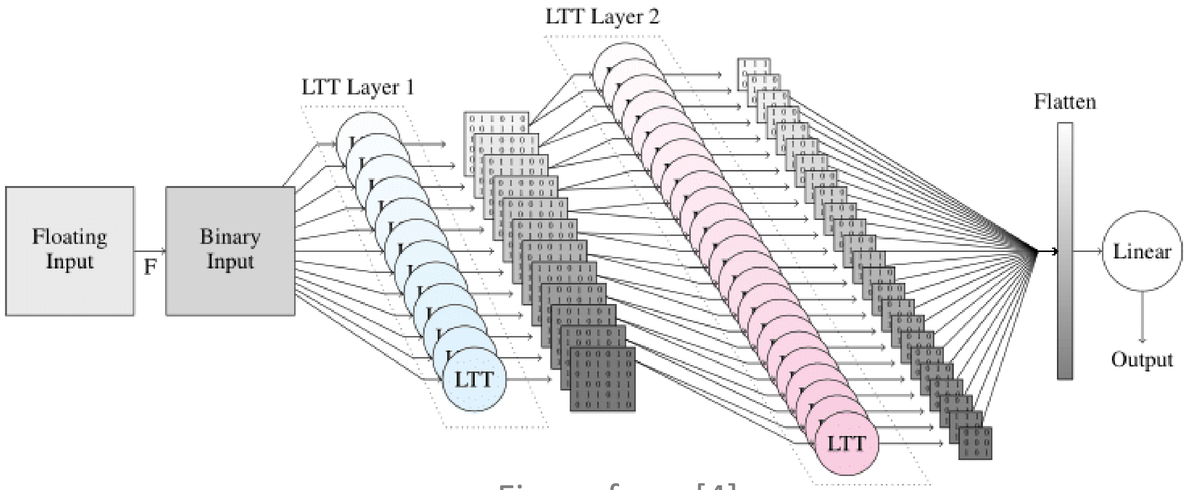


Figure from [4]

- The Neural Network is seen as an aggregate of Truth Tables
- Neural Network ⇔ Truth Tables ⇔ Boolean Expressions
- Scales to ImageNet

TT-rules

NN-based Rule Model

- Set of Truth Tables → Set of Rules
- Accurate
- Possibility to add Human Knowledge

Procedure:

1) We train our neural network TTnet on the dataset.
2) We convert TTnet in form of rules-based model.
3) We only use the rules-based model to infer.

→ All automated

TT-rules

Use Case: Breast Cancer Wisconsin dataset [5]

Goal: Is a cell malign or benign ?

Dataset Dimensions: 100 binary variables, 700 samples → very small dataset

Dataset features:

- Clump Thickness
- Uniformity of Cell Size
- Uniformity of Cell Shape
- Single Epithelial Cell Size
- Bare Nuclei
- Bland Chromatin
- Mitoses
- Marginal Adhesion
- Normal Nucleoli

TT-rules

Use Case: Breast Cancer Wisconsin dataset [5]

Metrics	Linear/ log	Decision Tree	GL	TT-rules (Ours)	Random Forest	DNNs
Accuracy	0.951	0.926	0.951	0.973	0.950	0.951
Number of Rules	-	49.8	15.8	21.6	19882.4	-

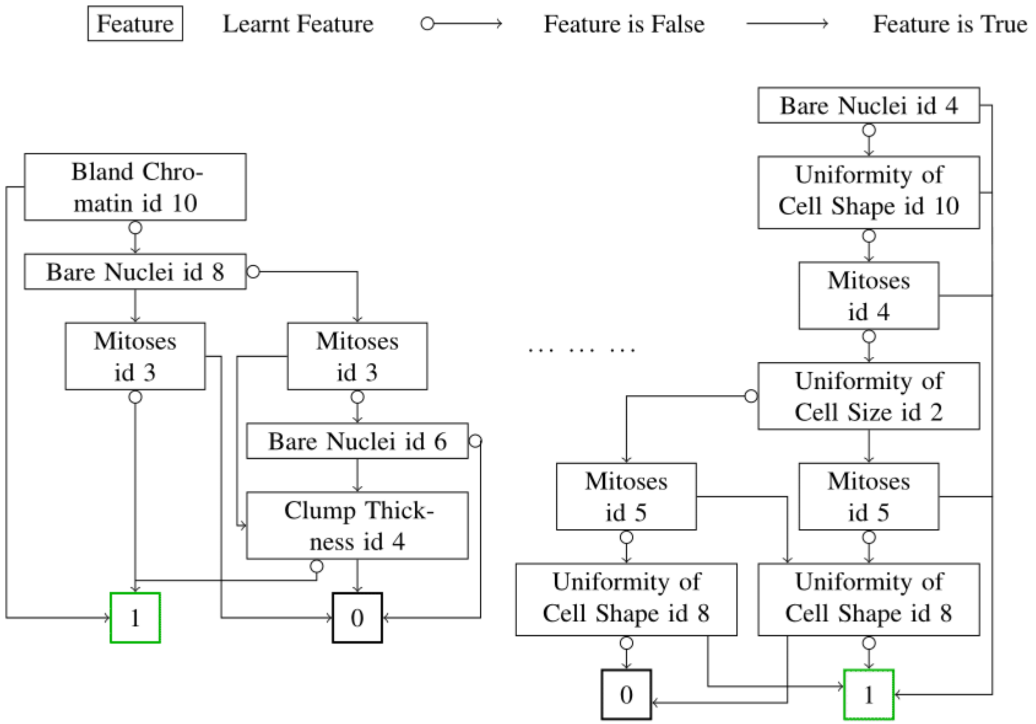
Comparison of our method to Linear/Logistic Regression[1], Decision Trees (DT)[1], GL[2] and DNNs. Our TT-rules models were trained with a final linear regression with weights as floating points. The higher the AUC the better. Means are reported from 5-fold cross validation.

TT-rules

Use Case: Breast Cancer Wisconsin dataset [5]

This model has **24 rules only**

On the same test set, Random forest had 200 trees with lower accuracy: 0.993 vs 0.957



$$\text{Classifier}(\mathrm{I}, \mathcal{R}) = \begin{cases} 1 & \text{if } \sum_{r \in \mathcal{R}} r(\mathrm{I}) > 0 \\ 0 & \text{otherwise.} \end{cases}$$

An output example of TT-rules

<u>Rule 1 is:</u>

(Bland Chromatin id = 10) | (Bare Nuclei id = 8 && Mitoses id ≠ 3) | (Bare Nuclei id ≠ 8 && Clump Thickness id ≠ 4)

Results

Example settings

Four different example settings:

- **TCCA Lung Cancer** [6]: https://bit.ly/tcga_lung_rna
- **Melanoma**: single-cell RNA-seq analysis datasets for head and neck cancer[7], melanoma cancer [8] https://bit.ly/neck_head_rna and https://bit.ly/melanoma_rna

- **Breast Cancer Wisconsin** [5] : https://archive.ics.uci.edu/dataset/15/breast+cancer+wisconsin+original
- **Diabetes 130 US Hospitals** [5] : https://bit.ly/diabetes_130_uci

Two DNA datasets with a lot of features (>20k), and two Machine Learning datasets with few features (<300)

In the melanoma cancer setup, we trained on the head and neck dataset [7] and tested on the melanoma dataset [8] following established literature [9, 10, 11, 12].

Results

Results table

	Regression		Binary Classification				Multi-classification	
	TCCA Cancer		Melanoma		Breast Cancer		Diabetes	
continous/binary #	0/20530		0/23689 features		0/81 features		43/296 features	
Metrics	RMSE	#Rules	Acc	#Rules	Acc	#Rules	Acc	#Rules
Linear/ log	0.092	-	0.833	-	0.951	-	0.581	-
Decision Tree	-	-	-	-	0.926	49.8	0.572	530.6
GL	-	-	-	-	0.951	15.8	-	-
TT-rules (Ours)	0.029	1064	0.835	9472	0.973	21.6	0.584	480
Random Forest	0.42	16377.8	0.729	13514	0.950	19882.4	0.587	4767415
DNNs	0.028	-	0.725	-	0.951	-	0.603	-

Comparison of our method to Linear/Logistic Regression[1], Decision Trees (DT)[1], Random Forest[1], GL[2] and DNNs. The lower RMSE the better, the higher AUC/Accuracy the better.Means and standard deviations are reported from 5-fold cross validation.

Results

Results table

	Regression		Binary Classification				Multi-classification	
	TCCA Cancer		Melanoma		Breast Cancer		Diabetes	
continous/binary #	0/20530		0/23689 features		0/81 features		43/296 features	
Metrics	RMSE	#Rules	Acc	#Rules	Acc	#Rules	Acc	#Rules
Linear/ log	0.092	-	0.833	-	0.951	-	0.581	-
Decision Tree	-	-	-	-	0.926	49.8	0.572	530.6
GL	-	-	-	-	0.951	15.8	-	-
TT-rules (Ours)	0.029	1064	0.835	9472	0.973	21.6	0.584	480
Random Forest	0.42	16377.8	0.729	13514	0.950	19882.4	0.587	4767415
DNNs	0.028	-	0.725	-	0.951	-	0.603	-

Comparison of our method to Linear/Logistic Regression[1], Random Forest[1] and DNNs. The lower RMSE the better, the higher AUC/Accuracy the better. Means and standard deviations are reported from 5-fold cross validation.

Our approach **scales**

It reduces the input feature set → **feature selection method**:
- Regression Problem : we generated a set of 1064 rules out of 20530 features
- Binary Classification: we generated 9472 rules, more than **halved the input size** from 23689 to 9472.

→ **drastic reduction in complexity**

Conclusion

- We obtain **all the rules** describing our model
- A Rule Model that **scales to 10k+ features** and 100GB of data
- TT-rules is a new tool for **Explainability and Decision-making** in healthcare

Perspectives

- Able to dig into the learnings of a Neural Network
- Compute the **Sufficient Reasons** and **Necessary Reasons** for a decision [13]
- Give the **most important rules**
- Presenting global and/or local explanations for diagnosis and improvement **with human feedbacks**

Contact us for collaborations on new medical datasets :
adrien002@e.ntu.edu.sg, guer0001@e.ntu.edu.sg, thomas.peyrin@ntu.sg

References

[1] Fabian Pedregosa, Gaël Varoquaux, Alexandre Gramfort, Vincent Michel, Bertrand Thirion, Olivier Grisel, Mathieu Blondel, Peter Prettenhofer, Ron Weiss, Vincent Dubourg, et al., 'Scikit-learn: Machine learning in python', *the Journal of machine Learning research*, 12, 2825–2830, (2011).
[2] Dennis Wei, Sanjeeb Dash, Tian Gao, and Oktay Gunluk, 'Generalized linear rule models', in *International Conference on Machine Learning*, pp. 6687–6696. PMLR, (2019).
[3] Tianqi Chen and Carlos Guestrin, 'XGBoost', in *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. ACM,* (aug 2016).
[4] Adrien Benamira, Thomas Peyrin, and Bryan Hooi Kuen-Yew, 'Truth-table net: A new convolutional architecture encodable by design into sat formulas', in *Computer Vision – ECCV 2022 Workshops*, eds., Leonid Karlinsky, Tomer Michaeli, and Ko Nishino, pp. 483–500, Cham, (2023). Springer Nature Switzerland
[5] Dheeru Dua and Casey Graff. UCI machine learning repository, 2017
[6] Jianfang Liu, Tara Lichtenberg, Katherine A Hoadley, Laila M Poisson, Alexander J Lazar, Andrew D Cherniack, Albert J Kovatich, Christopher C Benz, Douglas A Levine, Adrian V Lee, et al., 'An integrated tcga pan-cancer clinical data resource to drive high-quality survival outcome analytics', *Cell*, 173(2), 400–416, (2018)
[7] Siddhartha V Puram, Itay Tirosh, Akash S Parikh, Anoop P Patel, and et al., 'Single-cell transcriptomic analysis of primary and metastatic tumor ecosystems in head and neck cancer', *Cell*, 171(7), 1611–1624.e24, (Dec 2017).
[8] Itay Tirosh, Benjamin Izar, Sanjay M Prakadan, Mark H 2nd Wadsworth, and et al., 'Dissecting the multicellular ecosystem of metastatic melanoma by single-cell rna-seq', *Science*, 352(6282), 189–196, (Apr 2016).
[9] Jasleen K Grewal, Basile Tessier-Cloutier, Martin Jones, Sitanshu Gakkhar, Yussanne Ma, Richard Moore, Andrew J Mungall, Yongjun Zhao, Michael D Taylor, Karen Gelmon, et al., 'Application of a neural network whole transcriptome–based pan-cancer method for diagnosis of primary and metastatic cancers', *JAMA network open*, 2(4), e192597–e192597, (2019).
[10] Yuanyuan Li, Kai Kang, Juno M Krahn, Nicole Croutwater, Kevin Lee, David M Umbach, and Leping Li, 'A comprehensive genomic pancancer classification using the cancer genome atlas gene expression data', *BMC genomics*, 18, 1–13, (2017)
[11] Egor Revkov, Tanmay Kulshrestha, Ken Wing-Kin Sung, and Anders Jacobsen Skanderup, 'Puree: accurate pan-cancer tumor purity estimation from gene expression data', *Communications Biology*, 6(1), 394, (2023).
[12] Duc Tran, Hung Nguyen, Bang Tran, Carlo La Vecchia, Hung N Luu, and Tin Nguyen, 'Fast and precise single-cell data analysis using a hierarchical autoencoder', *Nature communications*, 12(1), 1029, (2021).
[13] Darwiche, A., & Hirth, A. (2020). On the reasons behind decisions. *arXiv preprint arXiv:2002.09284*