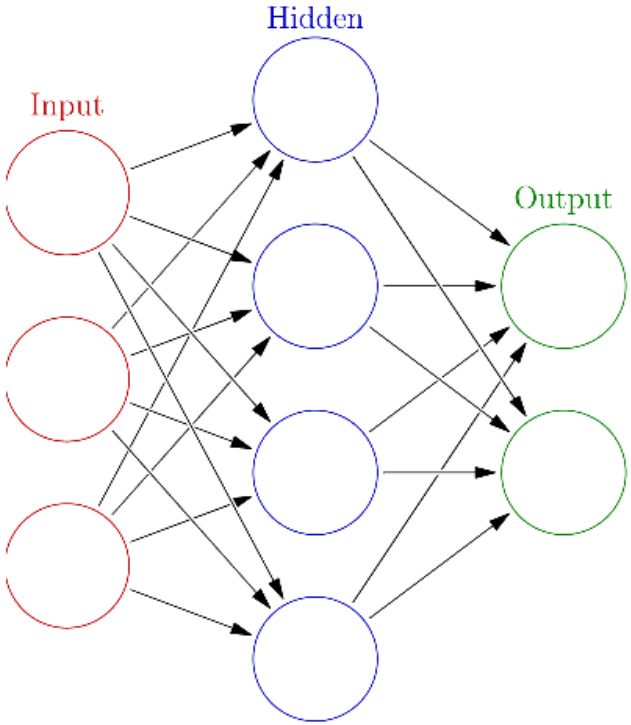
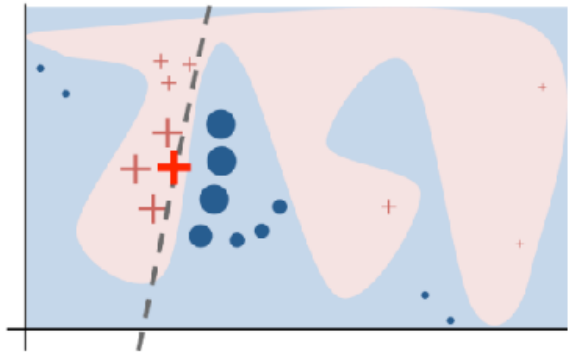
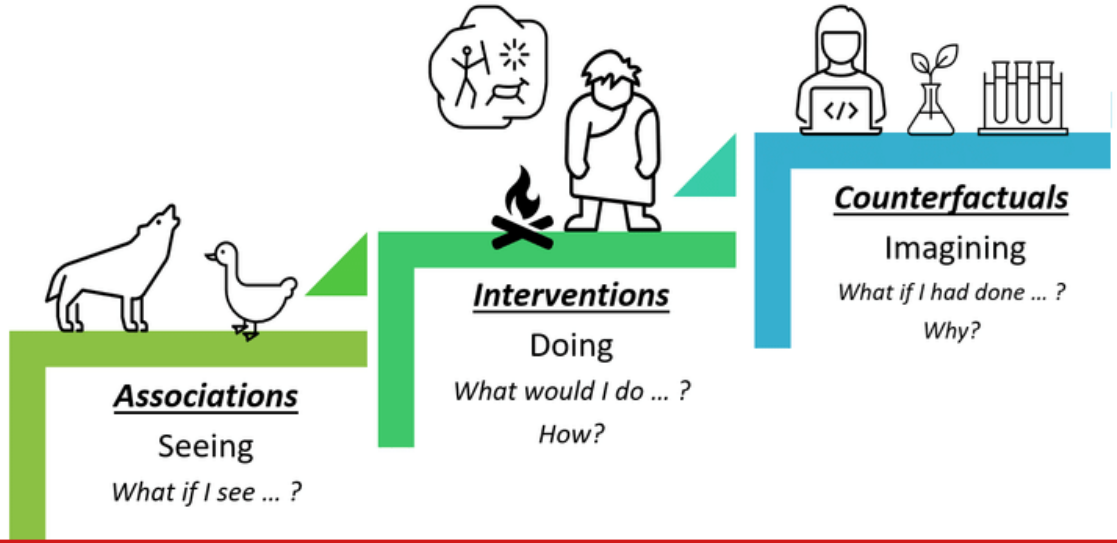
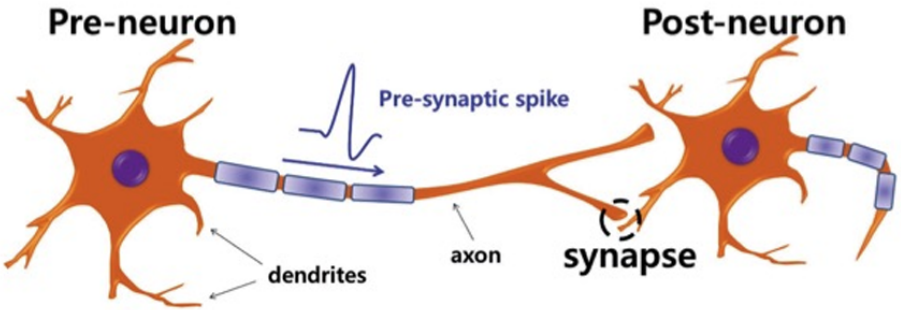


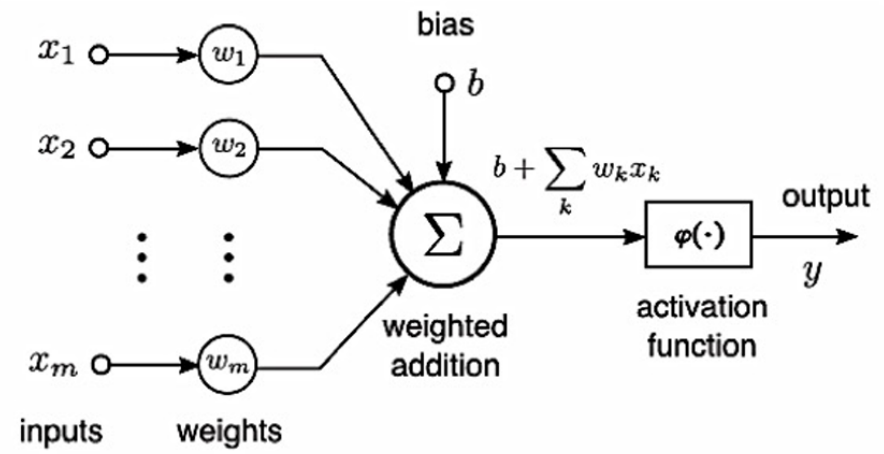


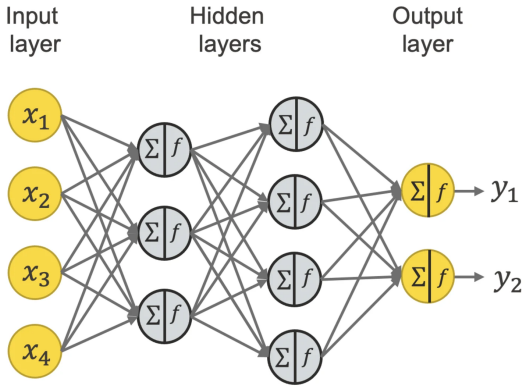


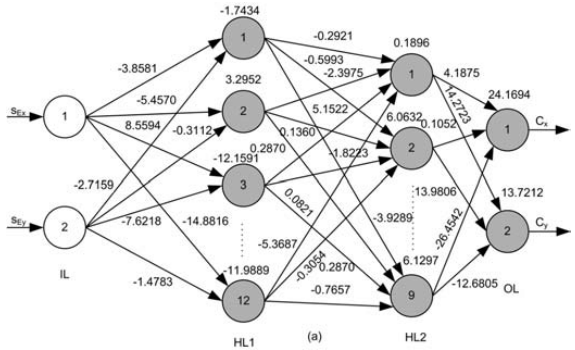
IA intepretabile

Reti neurali [1, 7]

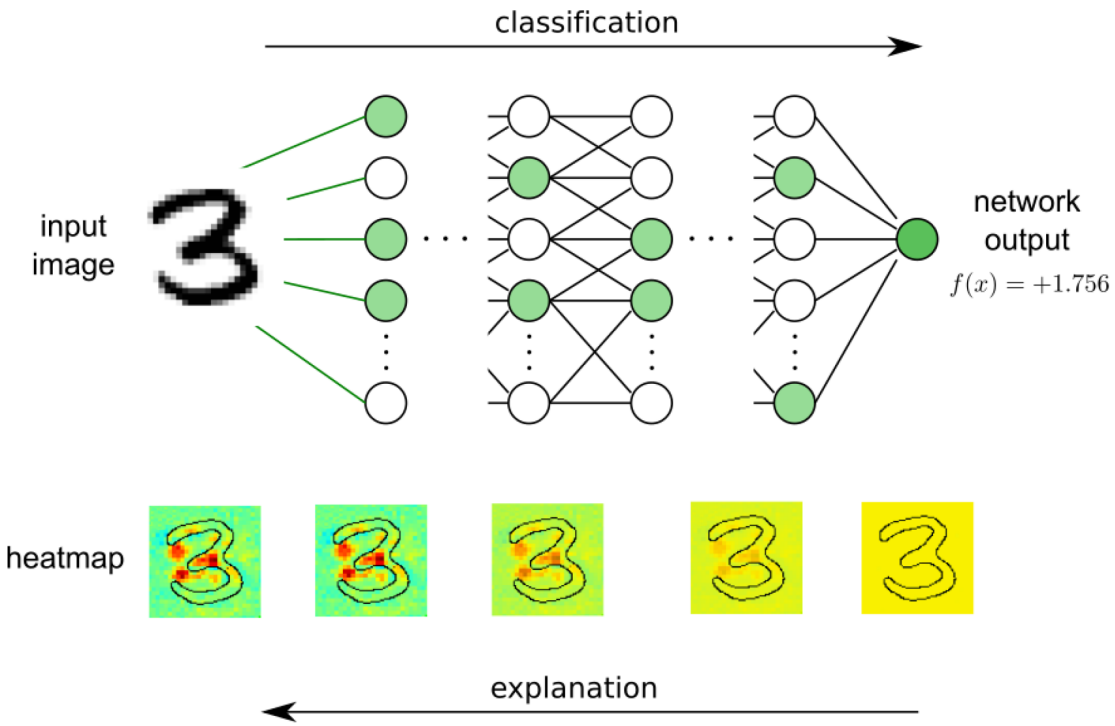








Interpretabilità per le reti neurali [2]

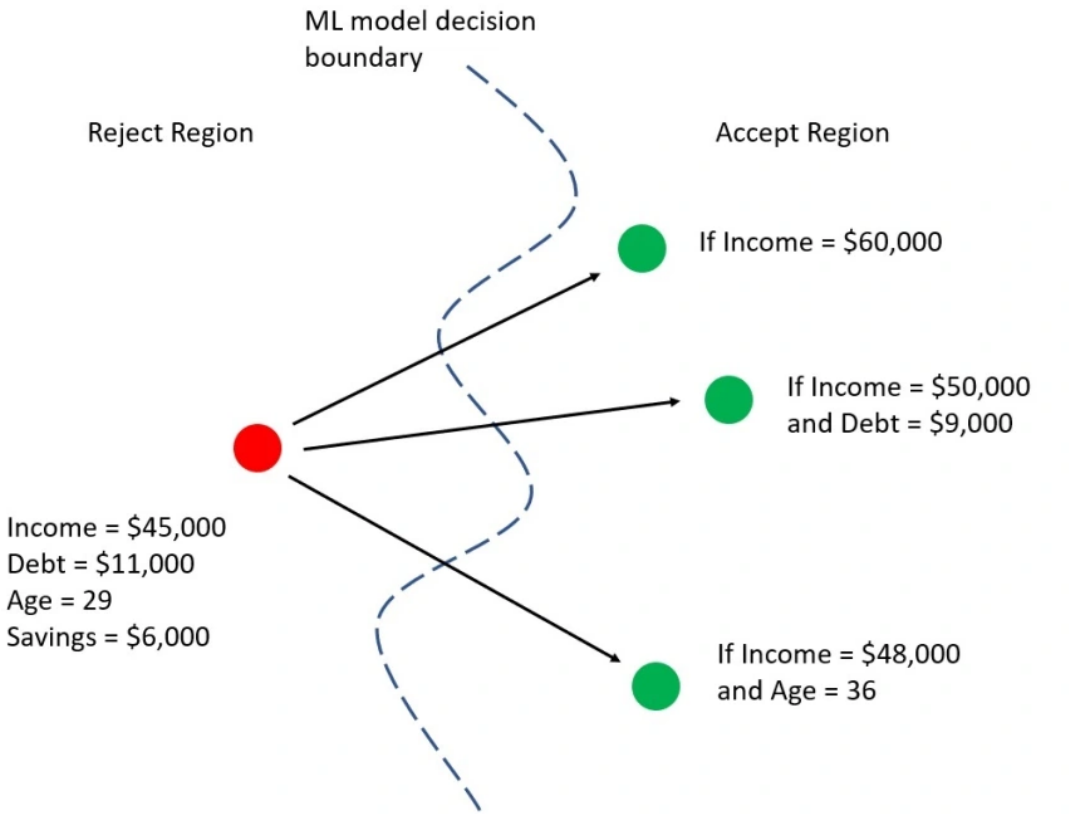


Saliency map [6]

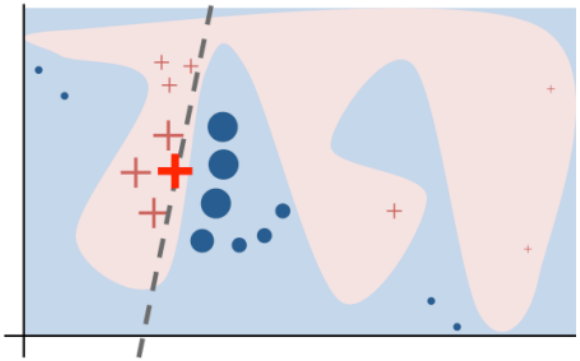
NLP Born-Rule [5]

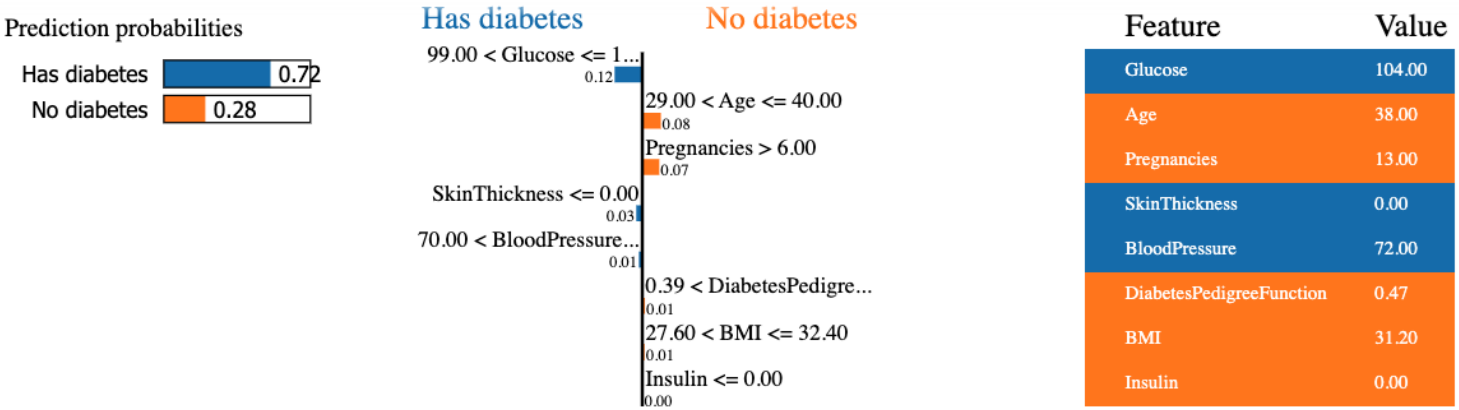
#	Baseball	Hockey	Autos	Graphics	Macintosh	Windows	Cryptography
1	Phillies	NHL	car	polygon	Centris	'AX	encryption
2	Braves	hockey	cars	TIFF	Quadra	Windows	Clipper
3	pitching	Leafs	eliot	graphics	Apple	3.1	clipper
4	Alomar	team	SHO	3D	Mac	windows	crypto
5	Baseball	Devils	automotive	3DO	Duo	W4WG	NSA
6	Players	ESPN	Callison	CView	LCIII	cica	escrow
7	Mets	Wings	Dumbest	POV	LC	font	key
8	Sox	Pens	rmt6r	cview	C650	BJ-200	DES
9	Cubs	playoffs	Thigpen	tdawson	BMUG	NDIS	Amanda
10	baseball	playoff	Toyota	MPEG	IIsi	Win	wiretap

XAI counterfactual [3]



XAI LIME [8, 4]




Prediction probabilities
Has diabetes
0.72
No diabetes
0.28
Has diabetes
No diabetes
99.00 < Glucose <= 1...
0.12
29.00 < Age <= 40.00
0.08
Pregnancies > 6.00
0.07
SkinThickness <= 0.00
0.03
70.00 < BloodPressure...
0.01
0.39 < DiabetesPedigre...
0.01
27.60 < BMI <= 32.40
0.01
Insulin <= 0.00
0.00
Feature
Value
Glucose
104.00
Age
38.00
Pregnancies
13.00
SkinThickness
0.00
BloodPressure
72.00
DiabetesPedigreeFunction
0.47
BMI
31.20
Insulin
0.00

Bibliography I

[1] https://www.knime.com/blog/a-friendly-introduction-to-deep-neural-networks.

[2] https://www.hhi.fraunhofer.de/en/departments/vca/research-groups/machine-learning/research-topics/interpretable-machine-learning.html.

[3] https://jamesmccaffrey.wordpress.com/2020/03/23/researchers-release-open-source-counterfactual-machine-learning-library/loan_application_counterfactuals/.

[4] https://towardsdatascience.com/explainable-ai-xai-lime-shap-two-great-candidates-to-help-you-explain-your-machine-learning-a95536a46c4e.

Bibliography II

[5] Emanuele Guidotti e Alfio Ferrara. "Text classification with born's rule". In: *Advances in Neural Information Processing Systems* 35 (2022), pp. 30990–31001.

[6] Mamuku Mokuwe, Michael Burke e Anna Sergeevna Bosman. "Black-box saliency map generation using bayesian optimisation". In: *2020 International Joint Conference on Neural Networks (IJCNN)*. IEEE. 2020, pp. 1–8.

[7] A Arockia Bazil Raj e S Padmavathi. "Quality metrics and reliability analysis of laser communication system". In: *Defence Science Journal* 66.2 (2016), pp. 175–185.

[8] Marco Tulio Ribeiro, Sameer Singh e Carlos Guestrin. "" Why should i trust you?" Explaining the predictions of any classifier". In: *Proceedings of the 22nd ACM SIGKDD international conference on knowledge discovery and data mining*. 2016, pp. 1135–1144.