

A APPENDIX

A.1 Convergence Curve of Fitness

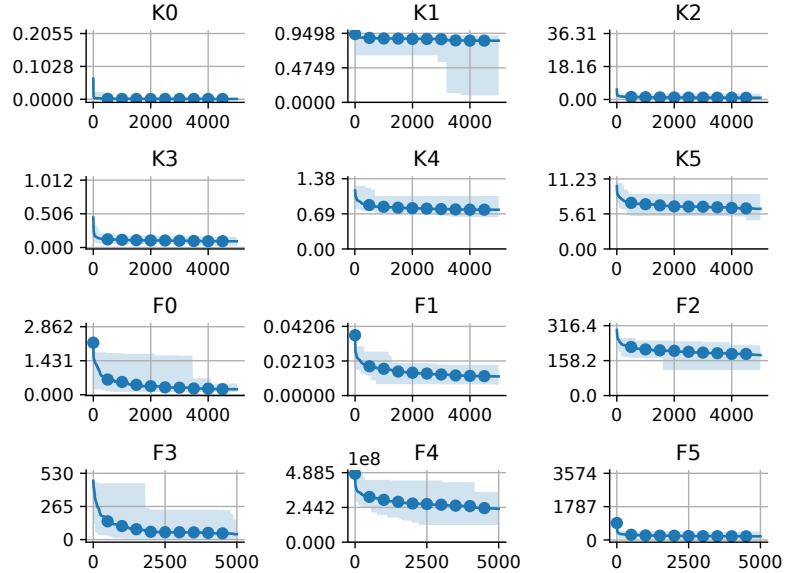


Figure 1: The convergence curve of fitness of MNNGP-ES that run 30 times on each MLP of $K_0 - K_5, F_0 - F_5$. The blue line represents the average fitness. The light blue range is the fitness range (maximum and minimum values) of the 30 results.

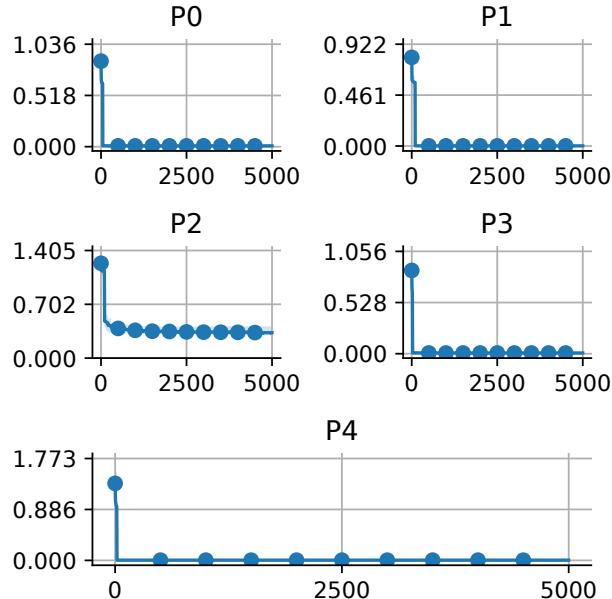


Figure 2: The convergence curve of fitness of MNNGP-ES that run 30 times on each MLP for the classification tasks $P_0 - P_5$.

A.2 Mathematical Expressions of Neural Networks

Dataset	$h_0^s(x)$	$h_1^s(h_0^s)$	$h_2^s(h_1^s)$	y^s
K_0	$\sin(0.26x + \sin(\sin(x)) - 0.068)$ (6.74e-04)	$0.88 - \cos(h_0^s)_0$ (7.54e-05)	—	$-\sin(\sin((h_0^s)_0 - 0.36))$ (2.37e-04)
K_1	$5.15e-05x_1 - 0.0072\sin(x_0) - 5.15e-05$ (5.34e-02)	$-(h_0^s)_0 + (h_0^s)_2$ (1.39e-03)	—	$-(h_1^s)_0 + (h_1^s)_2 + 0.0011$ (6.88e-02)
K_2	$\frac{\sqrt{x^2 \log(x^2)}}{x+1.20}$ (3.26e-02)	$-\tan(((h_0^s)_2 + (h_0^s)_3^2 - (h_0^s)_3))$ (1.04e-02)	—	$\sin\left(\frac{(h_1^s)_0}{(h_1^s)_2^2}\right)$ (2.00e-01)
K_3	$-\sin((0.24x_0 + \sin(0.23x_1)))$ (3.86e-02)	$(h_0^s)_1(h_0^s)_2^2 \sin((h_0^s)_1)$ (3.64e-03)	$(h_1^s)_0(-(h_1^s)_2 + (h_1^s)_3)$ (2.87e-04)	$-0.99 + \frac{(h_1^s)_3}{(h_1^s)_0}$ (1.32e-02)
K_4	$-\sin(x_1) + \sin(\frac{\sin(x_0)}{x_2})$ (1.17e-02)	$((h_0^s)_0 + (h_0^s)_3 - \tan((h_0^s)_1))^2$ (1.31e-02)	—	$-\frac{\sin((h_1^s)_1 - (h_1^s)_2)}{(h_1^s)_0^2}$ (6.14e-01)
K_5	$x_1 - 0.097$ (9.01e-02)	$\frac{\sin((h_0^s)_3 - 0.48)}{(h_0^s)_2}$ (2.44e-02)	—	$\frac{\tan((h_1^s)_0 + 0.51)}{(h_1^s)_2 - 0.26}$ (4.56)
F_0	$\cos((\frac{0.67x_1}{x_2} + \log(x_0)))$ (6.35e-03)	$-(h_0^s)_1 + \cos((h_0^s)_2)$ (1.04e-02)	—	$\tan((\tan(((h_1^s)_0 - 0.33)) + 0.11))$ (9.63e-02)
F_1	$\log\left(\frac{x_0 x_1}{x_2 x_3^2}\right)$ (9.59e-03)	$(h_0^s)_1^4$ (2.46e-04)	—	$(h_1^s)_0 + \sin((h_1^s)_0)$ (1.45e-03)
F_2	$-\frac{x_4}{x_3} + \frac{x_4}{x_2}$ (1.42e-02)	$(h_0^s)_2((h_0^s)_1 + 1)$ (1.36e-02)	—	$\frac{0.082 - (h_1^s)_2}{(h_1^s)_0}$ (1.15e02)
F_3	$\sin((\sqrt[4]{2}\sqrt[4]{x_0} + \sqrt{x_1}))$ (1.22e-02)	$\tan((h_0^s)_1) - 0.30$ (1.64e-02)	—	$(h_1^s)_1 - \cos(4(h_1^s)_2^2)$ (5.71)
F_4	$x_1 x_4 - \sqrt{x_2 + x_3}$ (7.49e-02)	$(h_0^s)_4 \cos^2((h_0^s)_1)$ (3.61e-02)	—	$(h_1^s)_0^2 (h_1^s)_4^2$ (1.21e08)
F_5	$\sqrt{x_1} + \log\left(\frac{x_1}{x_0 x_2}\right)$ (3.84e-03)	$\sin^2(0.51(h_0^s)_0)$ (1.20e-03)	$\frac{(2(h_1^s)_1 - \tan((h_1^s)_1))^4}{(h_1^s)_3^2}$ (5.12e-03)	$\tan(\tan((h_1^s)_0))$ (5.75e01)
P_0	$-x_6 + x_7 + x_9$ (1.58e-02)	$(h_0^s)_{85}$ (3.09e-04)	—	$\frac{(h_1^s)_{41}}{(h_1^s)_{31}}$ (0.0)
P_1	x_3 (3.76e-03)	$(h_0^s)_{14} - (h_0^s)_{67}$ (1.10e-04)	$(h_1^s)_2$ (3.53e-05)	$\sin\sqrt{(h_2^s)_0 + (h_2^s)_{19}}$ (0.0)
P_2	$-x_{10} - x_{20} + x_8 + x_9$ (2.08e-02)	$(h_0^s)_{63}$ (5.58e-03)	—	$(h_1^s)_{33}$ (2.91e-01)
P_3	$x_2 + x_4 + x_6 + x_9$ (1.40e-02)	$(h_0^s)_{17}$ (2.76e-04)	—	$\sqrt{(h_1^s)_{19}}$ (0.0)
P_4	$x_0 - x_1 + x_3 + x_4$ (1.09e-02)	$(h_0^s)_{72}$ (2.73e-04)	$(h_1^s)_{34}$ (7.14e-06)	$(h_2^s)_{40}$ (0.0)

Table 1: The best fitted mathematical expression (f_i) for explaining the hidden semantics of each layer h_i^s in NN. y^s is the output layer in NN. The number below each f_i is its fitness.

Dataset	$O^s(x)$
K_0	$0.29 - 4.01 \sin((\sin((2.36 \cos((0.41 \sin((0.26x + \sin(\sin(x)) - 0.068)) + 0.49)) - 2.03)))$ (6.11e-04)
K_1	$-0.01x_1 + 1.69 \sin(x_0) + 0.0021$ (9.62e-02)
K_2	$4.49 \sin\left(\frac{6.70 \left(0.51 \tan\left(0.061 \left(1 - \frac{0.19 \sqrt{x^2} \log(x^2)}{x+1.20}\right)^2 - 0.044 + \frac{0.084 \sqrt{x^2} \log(x^2)}{x+1.20}\right) + 0.27\right)}{\left(\tan\left(0.061 \left(1 - \frac{0.19 \sqrt{x^2} \log(x^2)}{x+1.20}\right)^2 - 0.044 + \frac{0.084 \sqrt{x^2} \log(x^2)}{x+1.20}\right) + 0.63\right)^2}\right)$ (2.22e-01)
K_3	$17.08(-0.81((-1.28(0.57 - 0.55 \sin(0.24x_0 + \sin(0.23x_1))))(0.93 \sin(0.24x_0 + \sin(0.23x_1)) + 1)^2 \sin(0.55 \sin(0.24x_0 + \sin(0.23x_1)) - 0.57) + 0.39))$ $((-0.88(0.57 - 0.55 \sin(0.24x_0 + \sin(0.23x_1))))(0.93 \sin(0.24x_0 + \sin(0.23x_1)) + 1)^2 \sin(0.55 \sin(0.24x_0 + \sin(0.23x_1)) - 0.57) - 0.12) + 0.43)$ $\div (-0.51(-1.28(0.57 - 0.55 \sin(0.24x_0 + \sin(0.23x_1))))(0.93 \sin(0.24x_0 + \sin(0.23x_1)) + 1)^2 \sin(0.55 \sin(0.24x_0 + \sin(0.23x_1)) - 0.57) - 0.12) + 0.44) - 15.36$ (2.73e-02)
K_4	$-0.0080 + 1.78 \sin((2.43(0.19 \sin(x_1) - 0.19 \sin(\frac{\sin(x_0)}{x_2})) - \tan(0.020 \sin(x_1) - 0.020 \sin(\frac{\sin(x_0)}{x_2}) + 0.73) + 0.90)^2 - 0.026))$ $\div (0.35 - (0.19 \sin(x_1) - 0.19 \sin(\frac{\sin(x_0)}{x_2})) - \tan(0.20 \sin(x_1) - 0.020 \sin(\frac{\sin(x_0)}{x_2}) + 0.73) + 0.90)^2)^2$ (6.27e-01)
K_5	$-0.14 + \frac{0.011 \tan(1.42 - \frac{0.051 \sin(0.25x_1 - 0.14)}{0.18 - 0.020x_1})}{0.16 - \frac{0.0414421632885933 \sin(0.25x_1 - 0.14)}{0.18 - 0.020x_1}}$ (4.62)
F_0	$3.05 - 7.00 \tan((\tan((0.58 \cos(\log(m_0) + \frac{0.67v}{c}) - 0.76 \cos(0.22 \cos(\log(m_0) + \frac{0.67v}{c}) - 1.01) + 0.35)) - 0.11))$ (1.05e-01)
F_1	$0.022 \left(0.46 \log\left(\frac{q_1 q_2}{e \pi^{\frac{3}{2}}}\right) + 1\right)^4 + 3.22 \sin\left(0.0068 \left(0.46 \log\left(\frac{q_1 q_2}{e \pi^{\frac{3}{2}}}\right) + 1\right)^4 - 0.00072\right) + 0.0059$ (6.37e-03)
F_2	$0.66 + \frac{173.88 \left(1.74 \left(0.0077 + \frac{0.014r_2}{r_1} - \frac{0.014r_2}{m_2}\right) \left(0.93 - \frac{0.054r_2}{r_1} + \frac{0.054r_2}{m_2}\right) - 0.023 + \frac{0.025r_2}{r_1} - \frac{0.025r_2}{m_2}\right)}{-0.72 \left(0.0077 + \frac{0.014r_2}{r_1} - \frac{0.014r_2}{m_2}\right) \left(0.93 - \frac{0.054r_2}{r_1} + \frac{0.054r_2}{m_2}\right) + 0.97 - \frac{0.010r_2}{r_1} + \frac{0.010r_2}{m_2}}$ (1.15e02)
F_3	$-31.43 \cos\left(4.40 \left(-\tan(0.46 \sin(\sqrt[4]{2} \sqrt[4]{k} + \sqrt{x}) - 0.27) - 0.091\right)^2\right) - 31.98 \tan\left(0.46 \sin(\sqrt[4]{2} \sqrt[4]{k} + \sqrt{x}) - 0.27\right) + 31.62$ (5.72)
F_4	$-1.15e14((-0.0026cr + 0.0026\sqrt{m_1 + m_2} + 0.00035)\cos^2(0.37cr - 0.37\sqrt{m_1 + m_2} + 0.58) - 0.00061)^2$ $((-0.0026cr + 0.0026\sqrt{m_1 + m_2} + 0.00035)\cos^2(0.37cr - 0.37\sqrt{m_1 + m_2} + 0.58) - 0.00050)^2 - 93.80$ (1.21e08)
F_5	$12.10 - 735.10 \tan((\tan(0.016 - (6.39e06(\sin^2(0.0072\sqrt{y} + 0.0072\log(\frac{y}{Vq}) - 0.0083) - 0.0010 \tan((489.69 \sin^2(0.0072\sqrt{y} + 0.0072\log(\frac{y}{Vq}) - 0.0084)$ $+ 0.10)) + 0.00021)^4) \div (\sin^2(0.0072\sqrt{y} + 0.0072\log(\frac{y}{Vq}) - 0.0084) + 0.00033)^2)))$ (57.47)
P_0	$Pr0 = -14.16 - \frac{0.00043x_5 - 0.00043x_7 - 0.00043x_9 + 0.42}{-0.0025x_0 + 0.0025x_7 + 0.0025x_9 + 0.489}$ $Pr1 = 14.57 + \frac{0.00047x_5 - 0.00047x_7 - 0.00047x_9 + 0.46}{-0.0025x_6 + 0.0025x_7 + 0.0025x_9 + 0.49}$ (8.05e-03)
P_1	$Pr0 = -5.10 \sin(0.00033x_3 + 0.67\sqrt{1 - 0.00044x_3} + 0.45) - 6.51$ $Pr1 = 5.42 \sin(0.00033x_3 + 0.67\sqrt{1 - 0.00044x_3} + 0.45) + 6.97$ (8.05e-03)
P_2	$Pr0 = 0.66x_{10} + 0.66x_{20} - 0.66x_8 - 0.66x_9 - 1.57$ $Pr1 = -0.53x_{10} - 0.53x_{20} + 0.53x_8 + 0.53x_9 + 1.17$ (3.04e-01)
P_3	$Pr0 = \sqrt{-0.15x_2 - 0.15x_4 - 0.15x_6 - 0.15x_9 + 19.27 + 14.87}$ $Pr1 = \sqrt{-0.18x_2 - 0.18x_4 - 0.18x_6 - 0.18x_9 + 22.28 - 15.34}$ (7.15e-03)
P_4	$Pr0 = -2.06e - 5x_0 + 2.06e - 5x_1 - 2.06e - 5x_3 - 2.06e - 5x_4 - 5.00$ $Pr1 = -1.64e - 5x_0 + 1.64e - 5x_1 - 1.64e - 5x_3 - 1.64e - 5x_4 - 5.18$ $Pr2 = 3.84e - 5x_0 - 3.84e - 5x_1 + 3.84e - 5x_3 + 3.84e - 5x_4 + 12.80$ $Pr3 = -1.35e - 5x_0 + 1.35e - 5x_1 - 1.35e - 5x_3 - 1.35e - 5x_4 - 4.57$ (3.73e-03)

Table 2: The mathematical expressions of explaining whole NN.

A.3 Comparison

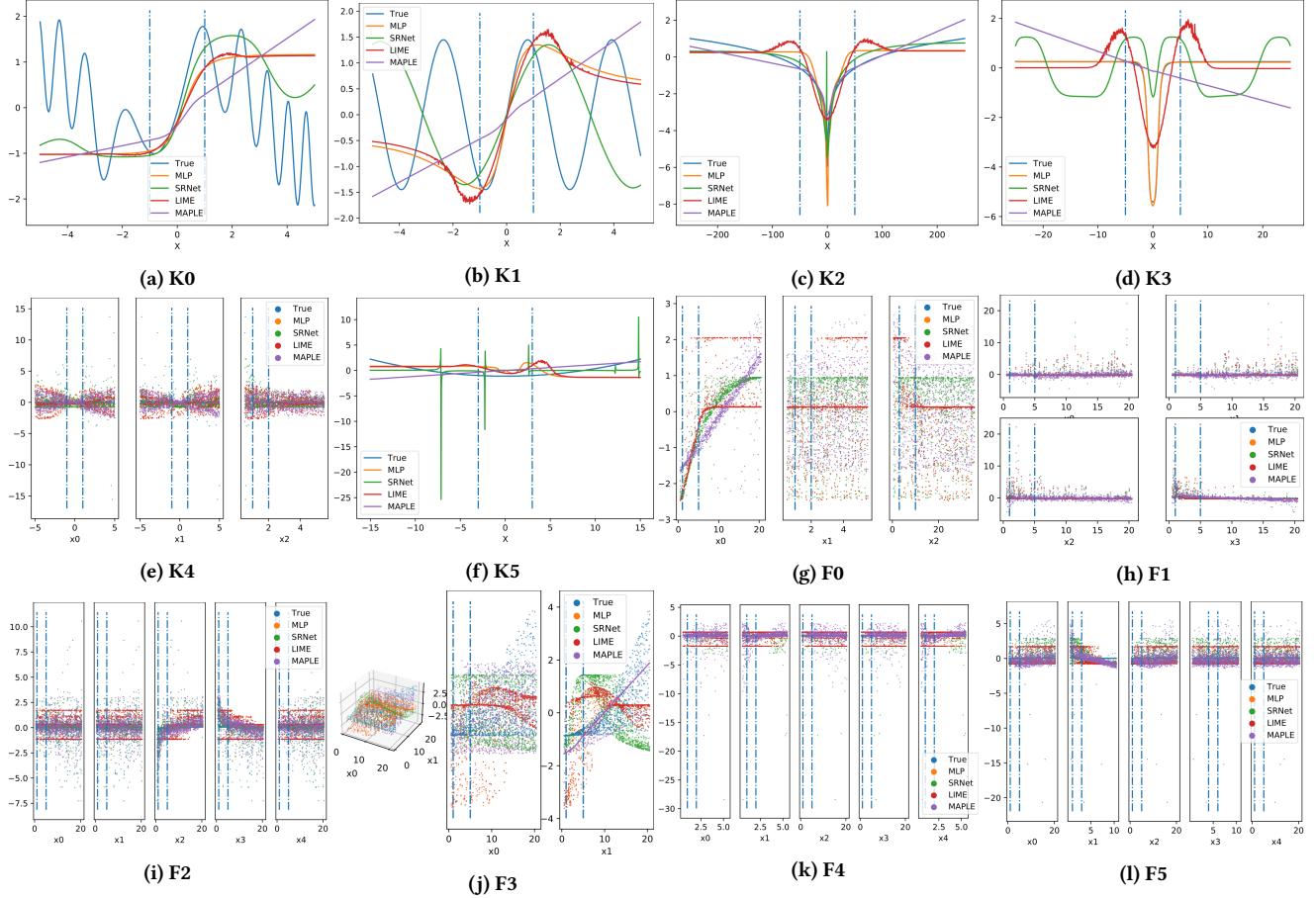


Figure 3: SRNet vs LIME vs MAPLE on the interpolation and extrapolation domain of 12 SR benchmarks. The area between two blue vertical lines is the interpolation domain. The other area is the extrapolation domain

A.4 Fitting of Hidden Neurons

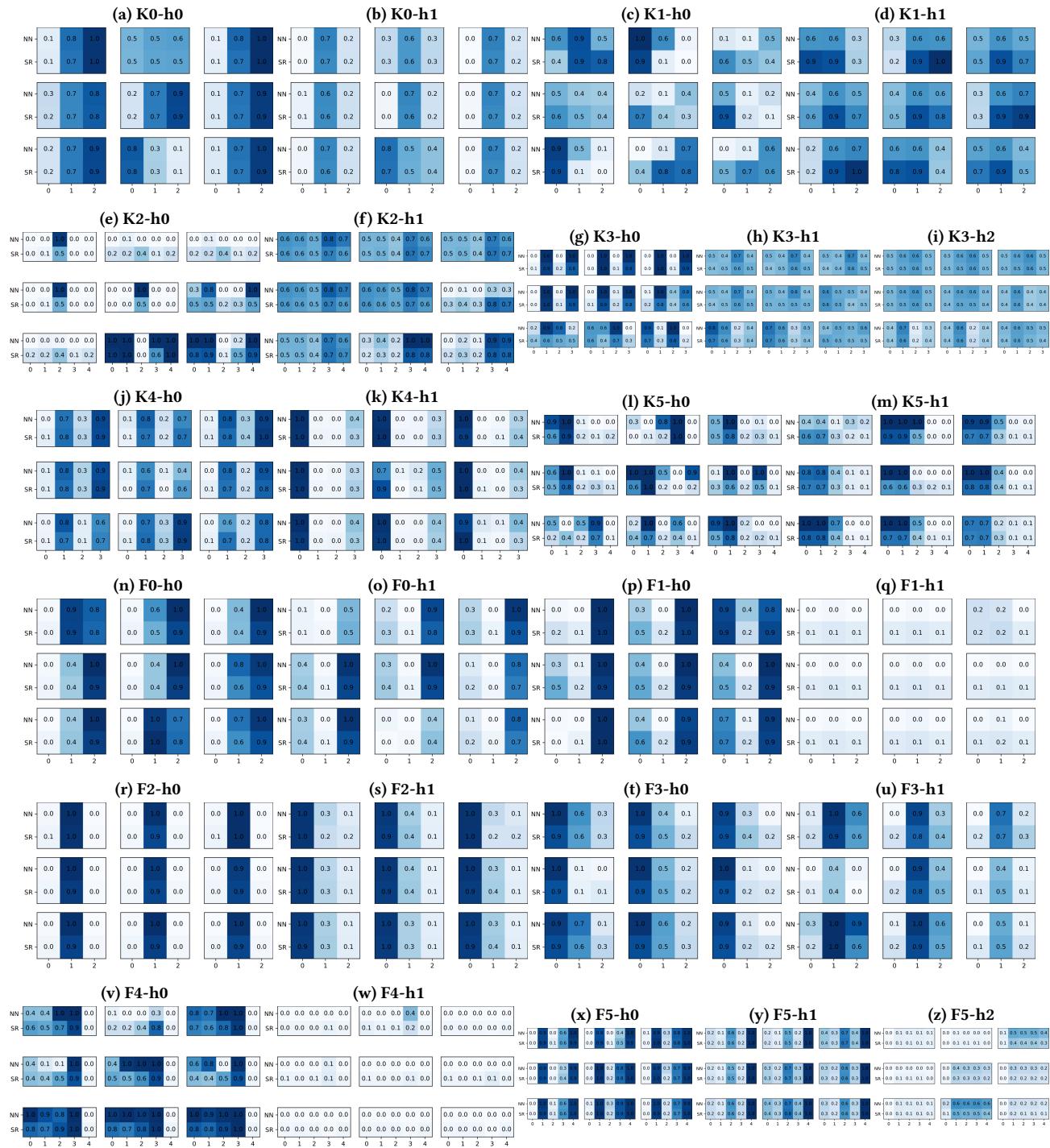


Figure 4: Each group of 9 heat maps represents the comparison of outputs of the SRNet layer vs the NN layer with 9 random input values 12 SR benchmarks.