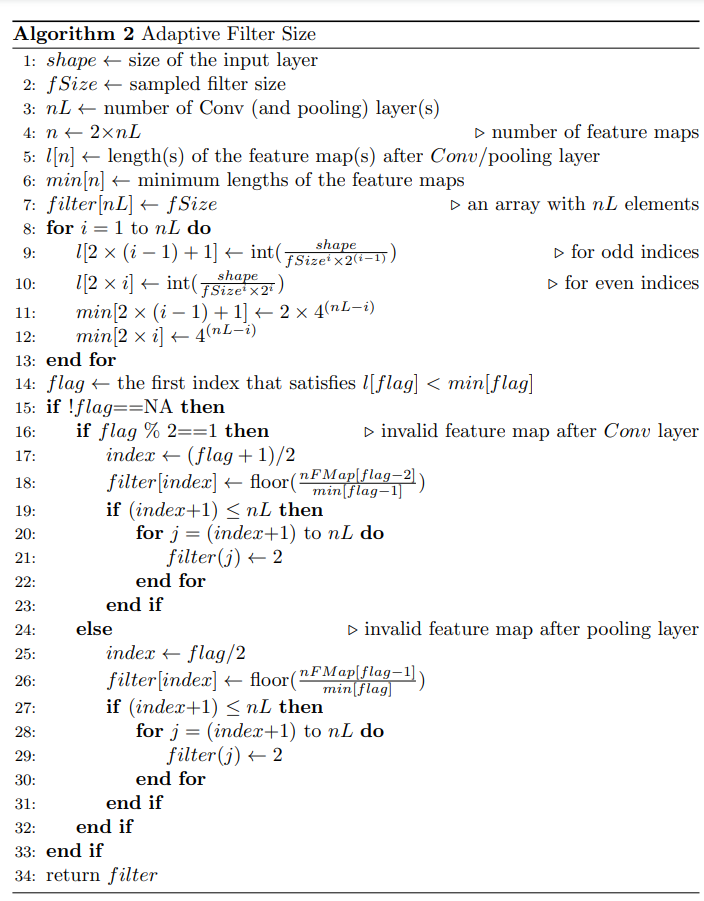
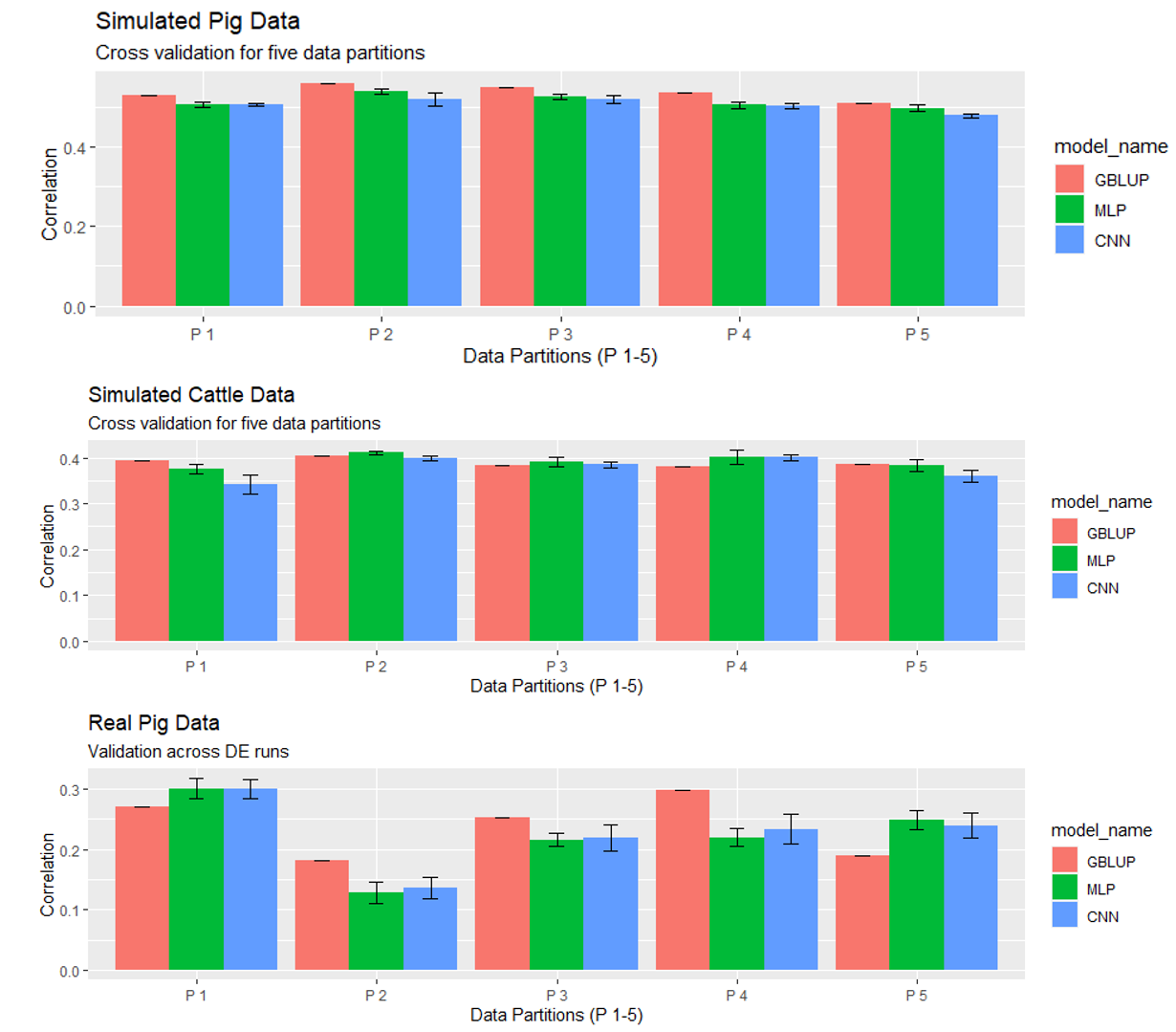
**Supplemental Material**



**Figure S1 Pseudocode for adaptive filter size.** Conv, convolutional layer; int(x), convert x into the nearest integer; floor(x), get the largest integer that is smaller or equal to x.



**Figure S2 Mean predictive performance and error bars across datasets and data partitions.** The error bar represents the mean ± standard deviation of cross validation by fitting the same model 30 times. Pink, green, and blue bars correspond to GBLUP, MLP, and CNN models, respectively. MLP, multilayer perceptron; CNN, convolutional neural network; GBLUP, genomic best linear unbiased prediction.

**Table S1** **Adaptive hyperparameter space for the number of neurons**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Layer** | **One HL** | **Two HLs** | **Three HLs** | **Four HLs** | **Five HLs** |
| **1** | [4-512] | [259-512] | [344-512] | [386-512] | [412-512] |
| **2** |  | [4-258] | [175-343] | [259-385] | [311-411] |
| **3** |  |  | [4-174] | [132-258] | [210-310] |
| **4** |  |  |  | [4-131] | [109-209] |
| **5** |  |  |  |  | [4-108] |

Number of neurons (nodes) given the depth of network (number of hidden layers, HL) in multilayer perceptron models.

**Table S2** **Adaptive hyperparameter space for number of filters**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Layer** | **One layer** | **Two layers** | **Three layers** | **Four layers** | **Five layers** |
| **1** | [4-128] | [4-65] | [4-44] | [4-34] | [4-28] |
| **2** | -- | [66-128] | [45-85] | [35-65] | [29-53] |
| **3** | -- | -- | [86-128] | [66-96] | [54-78] |
| **4** | -- | -- | -- | [97-128] | [79-103] |
| **5** | -- | -- | -- | -- | [104-128] |

Number of filters (kernels) given the depth (number of convolutional layers) of convolutional neural network.

**Table S3 Minimum length of feature maps applied to each layer of convolutional neural network.**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Layer** | **One layer** | **Two layers** | **Three layers** | **Four layers** | **Five layers** |
| **Conv 1** | 4 | 16 | 64 | 256 | 1024 |
| **Pooling 1** | 2 | 8 | 32 | 128 | 512 |
| **Conv 2** | -- | 4 | 16 | 64 | 256 |
| **Pooling 2** | -- | 2 | 8 | 32 | 128 |
| **Conv 3** | -- | -- | 4 | 16 | 64 |
| **Pooling 3** | -- | -- | 2 | 8 | 32 |
| **Conv 4** | -- | -- | -- | 4 | 16 |
| **Pooling 4** | -- | -- | -- | 2 | 8 |
| **Conv 5** | -- | -- | -- | -- | 4 |
| **Pooling 5** | -- | -- | -- | -- | 2 |

Conv: Convolutional layer.

**Table S4** **Distributions of** **optimized hyperparameters related to multilayer perceptron architectures for simulated pig data**

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | **Activation function** | | | | |  | **Number of layers** | | | |
|  | **elu** | **linear** | **selu** | **relu** | **softplus** |  | **One** | **Two** | **Three** | **Other** |
| **Pop1** | 6 | 38 | 4 | 0 | 2 |  | 3 | 18 | 29 | 0 |
| **Pop2** | 7 | 11 | 10 | 9 | 13 |  | 3 | 37 | 6 | 4 |
| **Pop3** | 11 | 20 | 10 | 6 | 3 |  | 8 | 33 | 7 | 2 |
| **Pop4** | 13 | 20 | 8 | 5 | 4 |  | 8 | 15 | 23 | 4 |
| **Pop5** | 11 | 1 | 13 | 15 | 10 |  | 8 | 30 | 7 | 5 |

Pop 1-5: MLP solutions to five data partitions (five differential evolution runs).

**Table S5** **Distributions of optimized hyperparameters related to CNN architectures for simulated pig data**

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | **Activation function** | | | |  | **Number of layers** | | | |  | **Filter size** | | |  | **Size of fully connected layer** | | |
|  | **elu** | **linear** | **selu** | **tanh** |  | **One** | **Two** | **Three** | **Other** |  | **Q0.05** | **Median** | **Q0.95** |  | **Q0.05** | **Median** | **Q0.95** |
| **Pop1** | 13 | 18 | 14 | 5 |  | 19 | 26 | 5 | 0 |  | 10 | 16 | 19 |  | 19 | 73 | 380 |
| **Pop2** | 20 | 14 | 16 | 0 |  | 31 | 19 | 0 | 0 |  | 5 | 10 | 20 |  | 22 | 149 | 477 |
| **Pop3** | 13 | 19 | 18 | 0 |  | 8 | 37 | 4 | 1 |  | 2 | 8 | 20 |  | 12 | 53 | 452 |
| **Pop4** | 16 | 27 | 7 | 0 |  | 35 | 15 | 0 | 0 |  | 6 | 11 | 20 |  | 9 | 36 | 308 |
| **Pop5** | 8 | 26 | 15 | 1 |  | 32 | 10 | 8 | 0 |  | 8 | 13 | 20 |  | 20 | 39 | 481 |

Pop 1-5: CNN solution populations of five differential evolutions runs. Size of fully connected layer: the number of neurons applied in the fully connected layer (after flatten layer). Q0.05, 5% quantile; Q0.95, 95% quantile.

**Table S6** **Distributions of optimized hyperparameters related to multilayer perceptron architectures for the simulated cattle data**

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | **Activation function** | | | | |  | **Number of layers** | | | |
|  | **linear** | **relu** | **elu** | **selu** | **softplus** |  | **One** | **Two** | **Three** | **Four** |
| **Pop1** | 43 | 0 | 4 | 2 | 1 |  | 17 | 17 | 13 | 3 |
| **Pop2** | 30 | 9 | 1 | 4 | 6 |  | 25 | 10 | 12 | 3 |
| **Pop3** | 41 | 1 | 6 | 2 | 0 |  | 6 | 28 | 14 | 2 |
| **Pop4** | 44 | 2 | 1 | 1 | 2 |  | 7 | 33 | 10 | 0 |
| **Pop5** | 49 | 1 | 0 | 0 | 0 |  | 19 | 20 | 10 | 1 |

Pop 1-5: MLP solutions to five data partitions (five differential evolution runs).

**Table S7** **Distributions of optimized hyperparameters related to CNN architectures for simulated cattle data**

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | **Activation function** | | | | |  | **Number of layers** | | | |  | **Filter size** | | |  | **Size of fully connected layer** | | |
|  | **elu** | **linear** | **selu** | **tanh** | **other** |  | **One** | **Two** | **Three** | **Other** |  | **Q0.05** | **Median** | **Q0.95** |  | **Q0.05** | **Median** | **Q0.95** |
| **Pop1** | 3 | 45 | 0 | 2 | 0 |  | 34 | 10 | 6 | 0 |  | 8 | 18 | 20 |  | 16 | 46 | 354 |
| **Pop2** | 11 | 27 | 5 | 7 | 0 |  | 8 | 9 | 33 | 0 |  | 10 | 17 | 20 |  | 12 | 195 | 386 |
| **Pop3** | 0 | 49 | 1 | 0 | 0 |  | 40 | 8 | 2 | 0 |  | 6 | 15 | 20 |  | 24 | 26 | 396 |
| **Pop4** | 13 | 17 | 19 | 0 | 1 |  | 9 | 9 | 32 | 0 |  | 10 | 18 | 18 |  | 27 | 221 | 485 |
| **Pop5** | 7 | 37 | 2 | 4 | 0 |  | 12 | 16 | 21 | 1 |  | 10 | 18 | 20 |  | 26 | 148 | 416 |

Pop 1-5: CNN populations of five differential evolutions runs. Size of fully connected layer: the number of neurons applied in the fully connected layer (after flatten layer). Q0.05, 5% quantile; Q0.95, 95% quantile.

**Table S8** **Distributions of optimized hyperparameters related to multilayer perceptron architectures for the real pig data**

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | **Activation function** | |  | **Number of layers** | | |
|  | **sigmoid** | **Other** |  | **Two** | **Three** | **Four** |
| **Pop1** | 47 | 3 |  | 6 | 44 | 0 |
| **Pop2** | 50 | 0 |  | 0 | 46 | 4 |
| **Pop3** | 50 | 0 |  | 0 | 44 | 6 |
| **Pop4** | 50 | 0 |  | 2 | 38 | 10 |
| **Pop5** | 50 | 0 |  | 1 | 46 | 3 |

Pop 1-5: MLP solutions to five data partitions (five differential evolution runs).

**Table S9** **Distributions of optimized hyperparameters related to CNN architectures for real pig data**

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | **Activation function** | | | |  | **Number of layers** | | | |  | **Filter size** | | | |  | | **Size of fully connected layer** | | | |
|  | **elu** | **linear** | **tanh** | **other** |  | **Two** | **Three** | **Four** | **Other** |  | **Q0.05** | **Median** | **Q0.95** |  | | **Q0.05** | | **Median** | **Q0.95** |
| **Pop1** | 19 | 6 | 25 | 0 |  | 16 | 24 | 10 | 0 |  | 12 | 13 | 18 |  | | 22 | | 367 | 506 |
| **Pop2** | 38 | 1 | 5 | 6 |  | 7 | 30 | 12 | 1 |  | 12 | 16 | 17 |  | | 50 | | 197 | 463 |
| **Pop3** | 3 | 4 | 40 | 3 |  | 16 | 26 | 8 | 0 |  | 8 | 13 | 18 |  | | 60 | | 150 | 463 |
| **Pop4** | 33 | 4 | 1 | 12 |  | 26 | 11 | 10 | 3 |  | 9 | 15 | 17 |  | | 50 | | 158 | 426 |
| **Pop5** | 20 | 14 | 16 | 0 |  | 2 | 25 | 22 | 1 |  | 5 | 11 | 18 |  | | 12 | | 195 | 416 |

Pop 1-5: CNN populations of five differential evolutions runs. Size of fully connected layer: the number of neurons applied in the fully connected layer (after flatten layer). Q0.05, 5% quantile; Q0.95, 95% quantile.

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | **Optimizer** | | |  | **Epochs** | | |  | **Batch size** | | |  | **Dropout rate** | | |  | **L2** | | |
|  | **adam** | **adamax** | **other** |  | **Q0.05** | **median** | **Q0.95** |  | **Q0.05** | **median** | **Q0.95** |  | **Q0.05** | **median** | **Q0.95** |  | **Q0.05** | **median** | **Q0.95** |
| **Pop1** | 48 | 0 | 2 |  | 27 | 35 | 47 |  | 14 | 30 | 53 |  | 0.03 | 0.16 | 0.45 |  | 0.03 | 0.18 | 0.67 |
| **Pop2** | 43 | 6 | 1 |  | 23 | 43 | 49 |  | 12 | 36 | 56 |  | 0.01 | 0.05 | 0.41 |  | 0.01 | 0.16 | 0.75 |
| **Pop3** | 43 | 6 | 1 |  | 23 | 32 | 48 |  | 6 | 28 | 52 |  | 0.01 | 0.05 | 0.57 |  | 0.01 | 0.17 | 0.69 |
| **Pop4** | 41 | 8 | 1 |  | 23 | 36 | 48 |  | 15 | 40 | 58 |  | 0.01 | 0.12 | 0.50 |  | 0.02 | 0.23 | 0.76 |
| **Pop5** | 35 | 15 | 0 |  | 22 | 35 | 50 |  | 19 | 38 | 57 |  | 0 | 0.04 | 0.14 |  | 0.01 | 0.14 | 0.87 |

**Table S10** **Distributions of optimized hyperparameters related to MLP model compilation and fitting for simulated pig data**

Pop 1-5: MLP solution populations of five differential evolution runs. Q0.05, 5% quantile; Q0.95, 95% quantile.

**Table S11** **Distributions of optimized hyperparameters related to CNN model compilation and fitting for simulated pig data**

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | **Optimizer** | | | | |  | **Epochs** | | |  | **Dropout rate** | | |  | **L2** | | | **Pooling** | |
|  | **adam** | **adamax** | **adadelta** | **nadam** | **rmsprop** |  | **Q0.05** | **median** | **Q0.95** |  | **Q0.05** | **median** | **Q0.95** |  | **Q0.05** | **median** | **Q0.95** | **Max** | **Average** |
| **Pop1** | 4 | 11 | 8 | 8 | 19 |  | 24 | 31 | 48 |  | 0.02 | 0.36 | 0.79 |  | 0.01 | 0.10 | 0.49 | 27 | 23 |
| **Pop2** | 10 | 10 | 20 | 0 | 10 |  | 21 | 34 | 49 |  | 0.04 | 0.42 | 0.79 |  | <0.01 | 0.04 | 0.25 | 22 | 28 |
| **Pop3** | 13 | 13 | 11 | 7 | 6 |  | 22 | 28 | 44 |  | 0.11 | 0.56 | 0.77 |  | 0.01 | 0.11 | 0.49 | 42 | 8 |
| **Pop4** | 10 | 29 | 5 | 3 | 3 |  | 23 | 31 | 46 |  | 0.05 | 0.39 | 0.78 |  | 0.01 | 0.12 | 0.64 | 22 | 28 |
| **Pop5** | 22 | 1 | 4 | 10 | 13 |  | 22 | 31 | 49 |  | 0.02 | 0.39 | 0.79 |  | 0.01 | 0.07 | 0.44 | 29 | 21 |

Pop 1-5: CNN solution populations of five differential evolution runs. Q0.05, 5% quantile; Q0.95, 95% quantile.

**Table S12** **Distributions of optimized hyperparameters related to MLP model compilation and fitting for simulated cattle data**

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | **Optimizer** | | |  | **Epochs** | | |  | **Batch size** | | |  | **Dropout rate** | | |  | **L2** | | |
|  | **adam** | **adamax** | **nadam** |  | **Q0.05** | **median** | **Q0.95** |  | **Q0.05** | **median** | **Q0.95** |  | **Q0.05** | **median** | **Q0.95** |  | **Q0.05** | **median** | **Q0.95** |
| **Pop1** | 47 | 1 | 2 |  | 23 | 41 | 47 |  | 11 | 31 | 57 |  | 0.01 | 0.15 | 0.52 |  | 0.04 | 0.27 | 0.81 |
| **Pop2** | 39 | 10 | 1 |  | 26 | 38 | 48 |  | 18 | 40 | 57 |  | <0.01 | 0.09 | 0.45 |  | 0.06 | 0.26 | 0.85 |
| **Pop3** | 44 | 4 | 2 |  | 26 | 33 | 44 |  | 12 | 26 | 52 |  | 0.01 | 0.18 | 0.53 |  | 0.03 | 0.22 | 0.87 |
| **Pop4** | 40 | 0 | 10 |  | 21 | 39 | 50 |  | 10 | 23 | 47 |  | 0.03 | 0.21 | 0.54 |  | 0.03 | 0.39 | 0.86 |
| **Pop5** | 41 | 2 | 7 |  | 22 | 29 | 48 |  | 9 | 24 | 52 |  | 0.02 | 0.26 | 0.72 |  | 0.06 | 0.27 | 0.91 |

Pop 1-5: MLP solution populations of five differential evolution runs. Q0.05, 5% quantile; Q0.95, 95% quantile.

**Table S13** **Distributions of optimized hyperparameters related to CNN model compilation and fitting for simulated cattle data**

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | **Optimizer** | | | | |  | **Epochs** | | |  | **Dropout rate** | | |  | **L2** | | | **Pooling** | |
|  | **adadelta** | **adam** | **adamax** | **nadam** | **rmsprop** |  | **Q0.05** | **median** | **Q0.95** |  | **Q0.05** | **median** | **Q0.95** |  | **Q0.05** | **median** | **Q0.95** | **Max** | **Average** |
| **Pop1** | 3 | 28 | 11 | 1 | 7 |  | 23 | 35 | 50 |  | 0.01 | 0.30 | 0.65 |  | 0.03 | 0.24 | 0.72 | 20 | 30 |
| **Pop2** | 17 | 16 | 7 | 2 | 8 |  | 22 | 33 | 50 |  | 0.06 | 0.28 | 0.68 |  | 0.02 | 0.29 | 0.59 | 10 | 40 |
| **Pop3** | 3 | 21 | 22 | 0 | 4 |  | 22 | 40 | 44 |  | 0.04 | 0.29 | 0.78 |  | 0.02 | 0.16 | 0.50 | 12 | 38 |
| **Pop4** | 11 | 3 | 29 | 2 | 5 |  | 23 | 29 | 46 |  | 0.05 | 0.35 | 0.67 |  | 0.01 | 0.23 | 0.75 | 17 | 33 |
| **Pop5** | 4 | 19 | 5 | 1 | 21 |  | 21 | 34 | 46 |  | 0.02 | 0.24 | 0.60 |  | <0.01 | 0.11 | 0.52 | 27 | 23 |

Pop 1-5: CNN solution populations of five differential evolution runs. Q0.05, 5% quantile; Q0.95, 95% quantile.

**Table S14** **Distributions of optimized hyperparameters related to MLP model compilation and fitting for real pig data**

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | **Optimizer** | | |  | **Epochs** | | |  | **Batch size** | | |  | **Dropout rate** | | |  | **L2** | | | |
|  | **adam** | **adamax** | **sgd** |  | **Q0.05** | **median** | **Q0.95** |  | **Q0.05** | **median** | **Q0.95** |  | **Q0.05** | **median** | **Q0.95** |  | **Q0.05** | **median** | **Q0.95** |
| **Pop1** | 1 | 9 | 40 |  | 25 | 40 | 45 |  | 7 | 18 | 64 |  | 0.03 | 0.38 | 0.85 |  | 0.12 | 0.66 | 0.95 |
| **Pop2** | 19 | 30 | 1 |  | 22 | 37 | 45 |  | 28 | 55 | 64 |  | 0.04 | 0.41 | 0.82 |  | 0.09 | 0.60 | 0.92 |
| **Pop3** | 0 | 47 | 3 |  | 30 | 44 | 49 |  | 21 | 44 | 68 |  | 0.05 | 0.33 | 0.79 |  | 0.05 | 0.55 | 0.94 |
| **Pop4** | 2 | 47 | 1 |  | 31 | 33 | 47 |  | 30 | 31 | 32 |  | 0.15 | 0.56 | 0.86 |  | 0.04 | 0.37 | 0.86 |
| **Pop5** | 1 | 40 | 9 |  | 21 | 24 | 41 |  | 8 | 47 | 63 |  | 0.05 | 0.38 | 0.84 |  | 0.05 | 0.48 | 0.90 |

Pop 1-5: MLP solution populations of five differential evolution runs. Q0.05, 5% quantile; Q0.95, 95% quantile.

**Table S15** **Distributions of optimized hyperparameters related to CNN model compilation and fitting for real pig data**

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | **Optimizer** | | |  | **Epochs** | | |  | **Dropout rate** | | |  | **L2** | | |  | **Pooling** | |
|  | **adam** | **adamax** | **other** |  | **Q0.05** | **median** | **Q0.95** |  | **Q0.05** | **median** | **Q0.95** |  | **Q0.05** | **median** | **Q0.95** |  | **Max** | **Average** |
| **Pop1** | 36 | 9 | 5 |  | 22 | 31 | 41 |  | 0.02 | 0.32 | 0.71 |  | 0.03 | 0.54 | 0.96 |  | 1 | 49 |
| **Pop2** | 34 | 13 | 3 |  | 25 | 36 | 50 |  | 0.08 | 0.41 | 0.78 |  | 0.04 | 0.51 | 0.97 |  | 2 | 48 |
| **Pop3** | 45 | 4 | 1 |  | 22 | 34 | 45 |  | 0.02 | 0.41 | 0.88 |  | 0.16 | 0.49 | 0.93 |  | 13 | 37 |
| **Pop4** | 40 | 7 | 3 |  | 29 | 31 | 49 |  | 0.03 | 0.33 | 0.75 |  | 0.05 | 0.60 | 0.98 |  | 8 | 42 |
| **Pop5** | 44 | 5 | 1 |  | 24 | 44 | 46 |  | 0.03 | 0.38 | 0.77 |  | 0.05 | 0.55 | 0.95 |  | 0 | 50 |

Pop 1-5: CNN solution populations of five differential evolution runs. Q0.05, 5% quantile; Q0.95, 95% quantile.

**Table S16 Selected MLP and CNN architecture derived from other studies**

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Study** | **Model** | **Activation** | **No. layers** | **No. neurons (filters)** | **Dropout** | **Filter size** |
| Bellot et al. (2018) | MLP | elu | 1 | 32 | 0.0100 | NA |
| Pérez-Enciso and Zingaretti (2019) | MLP | relu | 4 | [64,64,64,64] | 0.0005 | NA |
| Bellot et al. (2018) | CNN | linear | 1+1 | [16,32] | 0.0100 | 3 |
| Pérez-Enciso and Zingaretti (2019) | CNN | relu | 4 | [64,64,64,64] | 0.0005 | 3 |

No. layers, the number of fully connected layers or convolutional layers; No. neurons (filters), the number of neurons or filters adaptive based on the number of layers. In the No. layers column, 1+1 means one convolutional layer plus one fully connected layer.