A Comprehensive Trial of Functions in Model

CHC Dec. 14, 2020

Candidate Functions

1. Activation function of Convolution layer(AC):

   relu, selu, elu, softplus, exponential

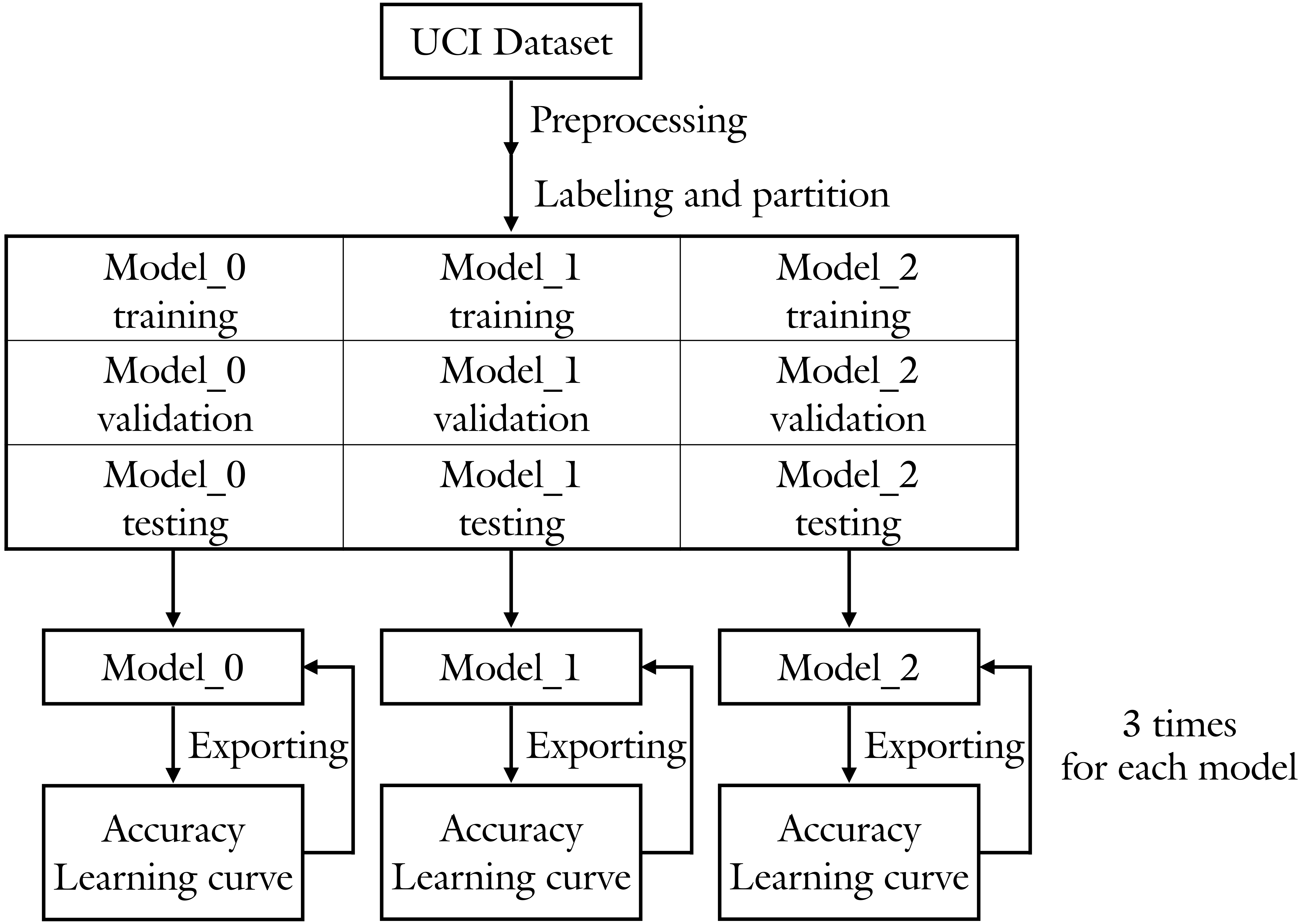
2. Activation function of Full connection layer(AF):

   softmax, sigmoid

3. Optimizer:

   sgd, adam, admax, adagrad, adadelta, nadam, rmsprop

Implementation of replacing functions



UCI Dataset

Preprocessing

Labeling and partition

Model_0 training	Model_1 training	Model_2 training
Model_0 validation	Model_1 validation	Model_2 validation
Model_0 testing	Model_1 testing	Model_2 testing

Model_0

Exporting

Accuracy Learning curve

Model_1

Exporting

Accuracy Learning curve

Model_2

Exporting

Accuracy Learning curve

3 times
for each model

Layer	#	size	activation
ZeroPadding		2, 2	
Conv2D	16	5, 1	i
ZeroPadding		2, 2	
MaxPooling		5, 1	
ZeroPadding		2, 2	
Conv2D	32	5, 1	i
ZeroPadding		2, 2	
MaxPooling		5, 1	
ZeroPadding		2, 2	
Conv2D	64	5, 1	i
ZeroPadding		2, 2	
MaxPooling		5, 1	
ZeroPadding		2, 2	
Conv2D	128	5, 1	i
Dropout	0.5		
Flatten			
Dense	2/7/6		j

Regularizer=l2

Loss=categorical cross entropy

Optimizer=k

Epochs=300

Batch size=64

$$\forall i \in \{relu, selu, elu, softplus, exponential\}$$

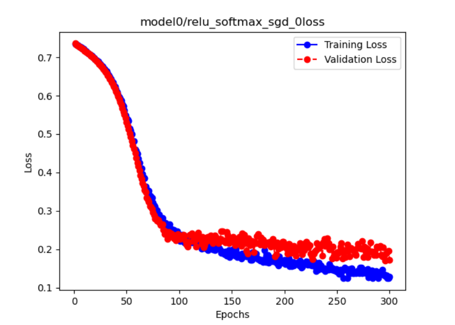
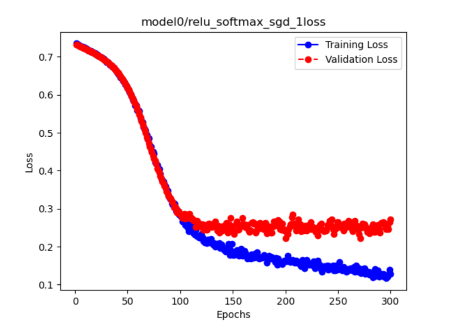
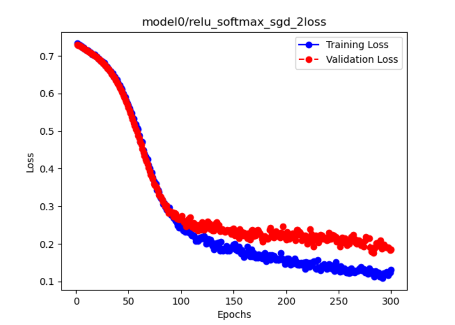
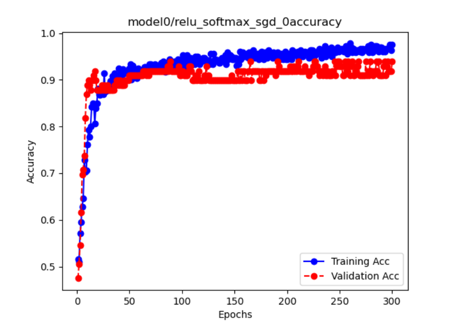
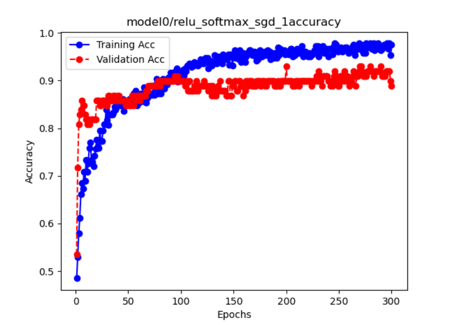
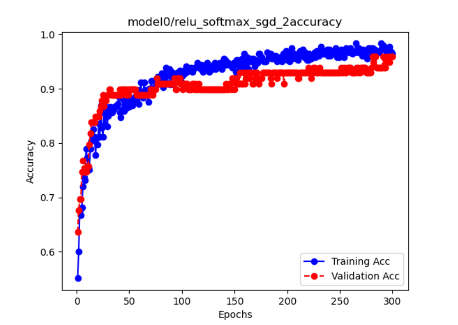
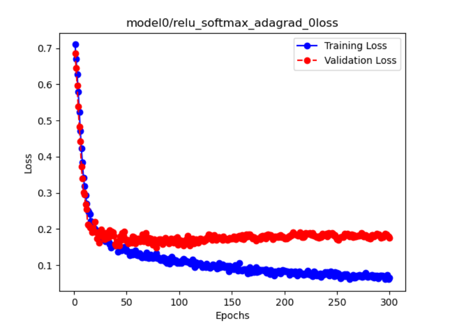
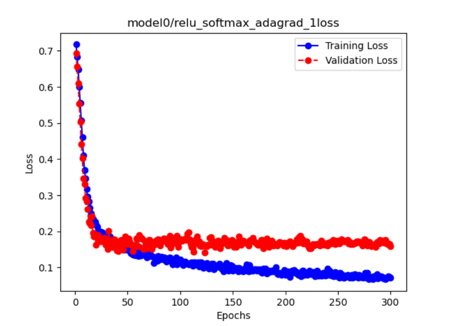
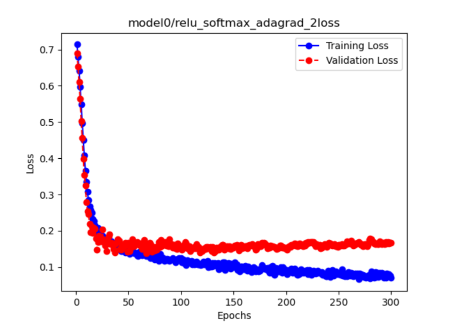
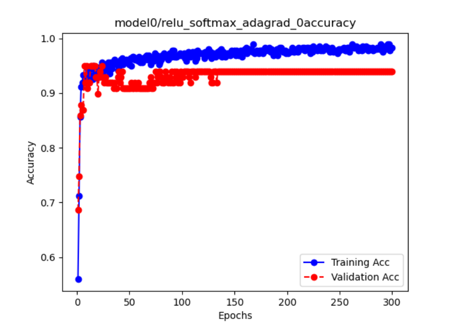
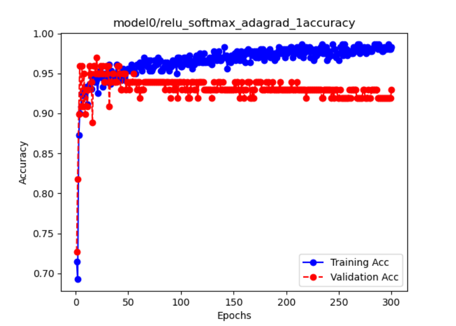
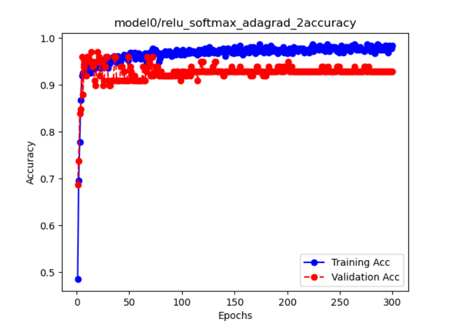
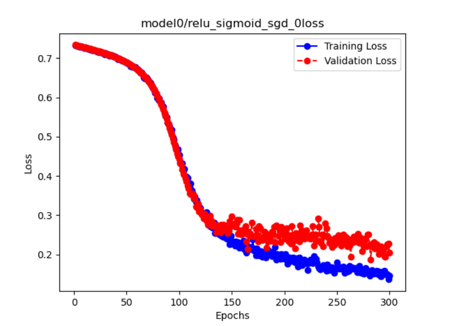
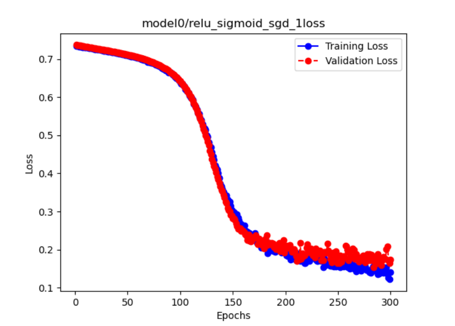
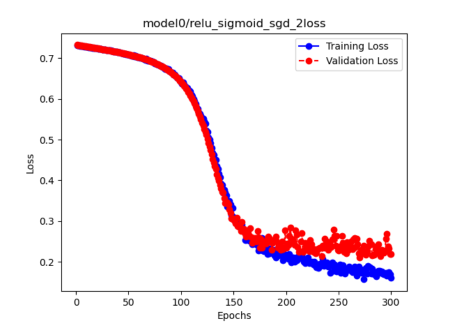
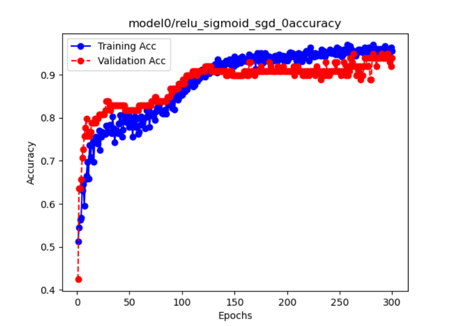
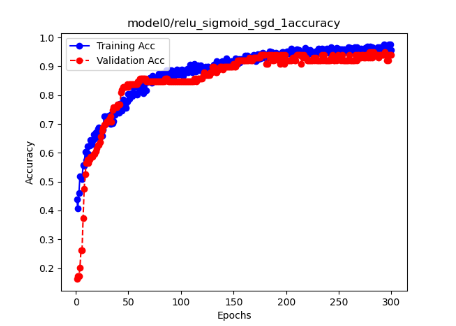
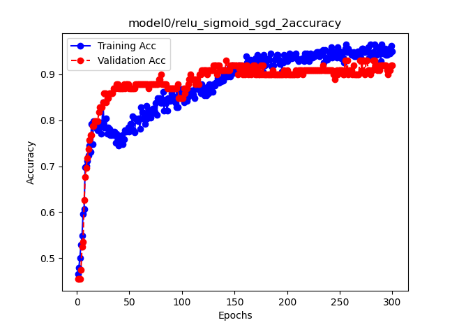
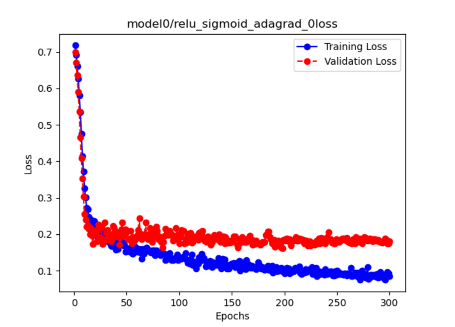
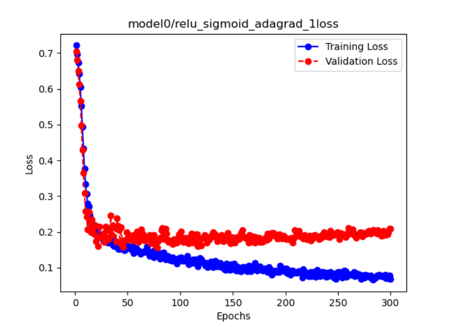
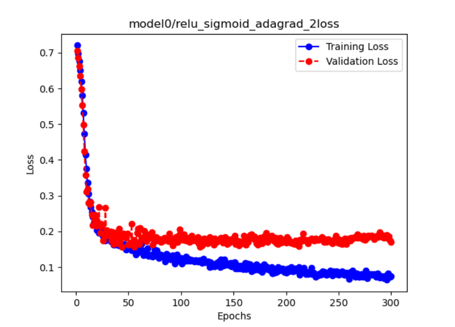
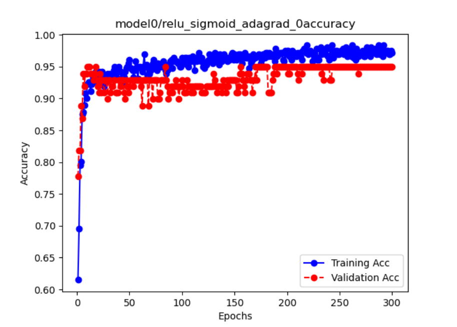
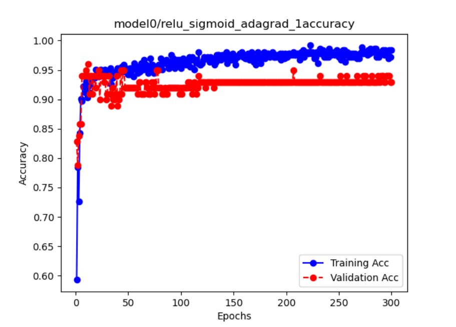
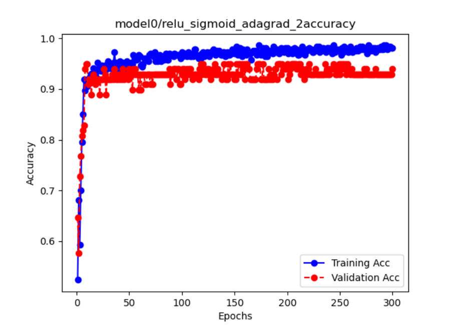
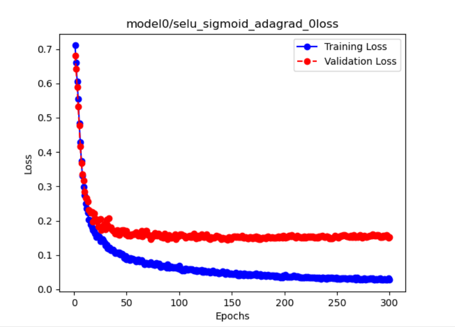
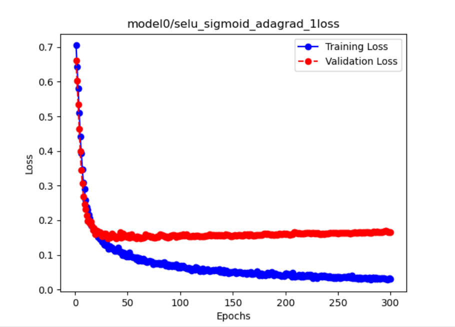
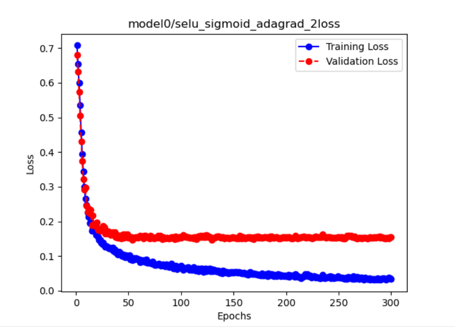
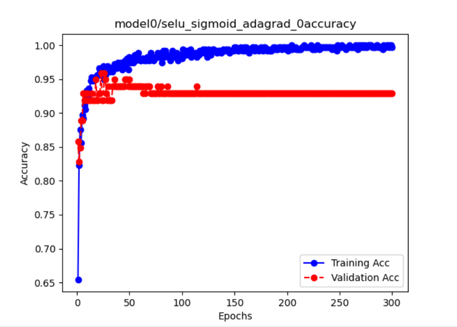
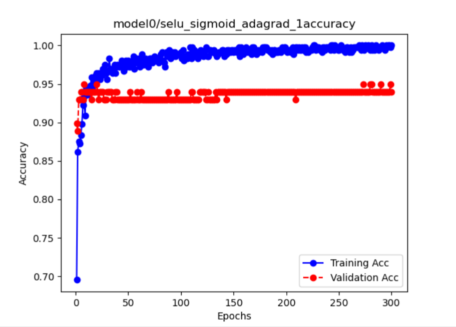
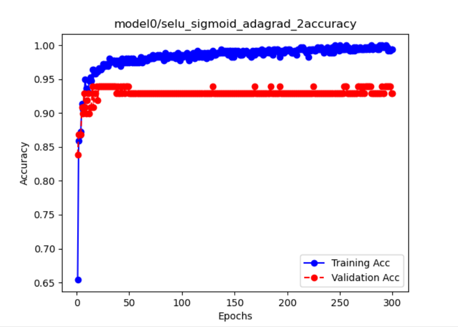
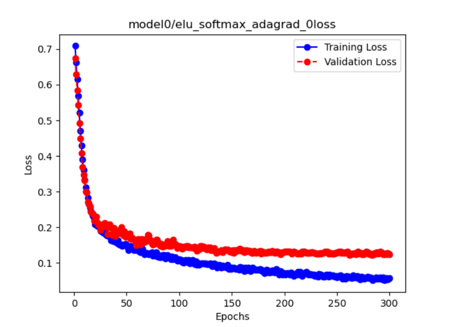
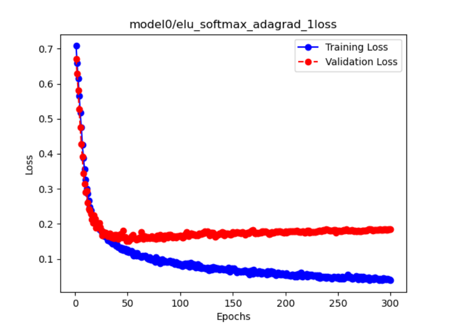
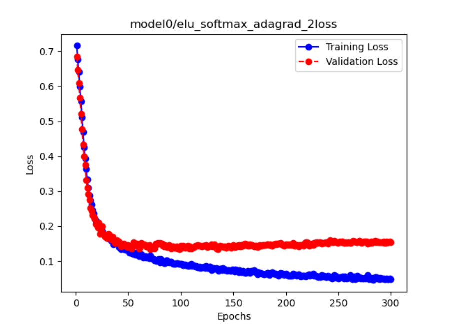
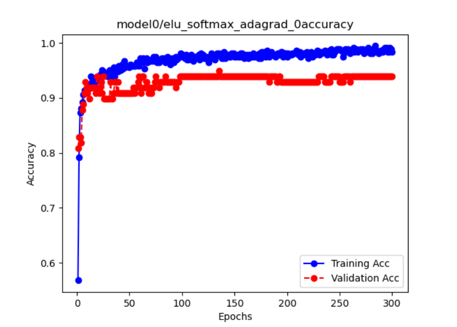
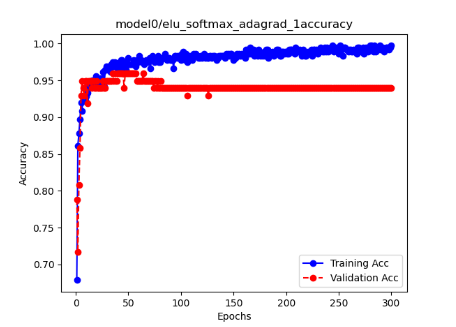
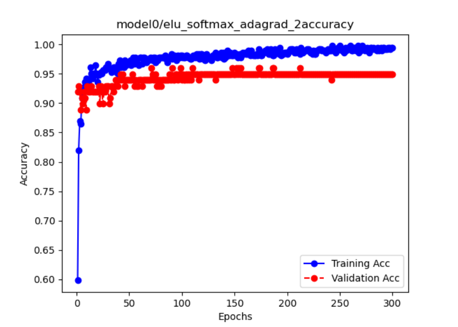
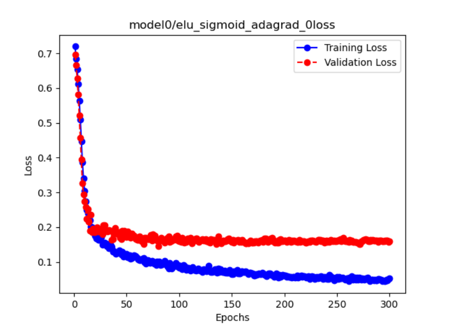
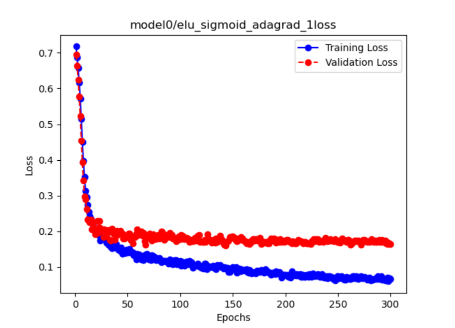
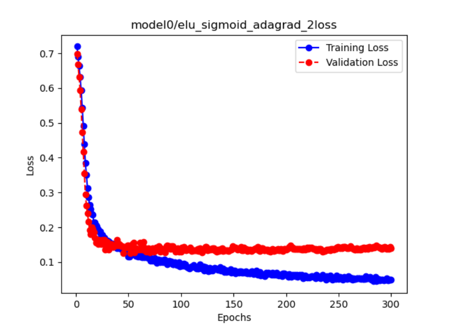
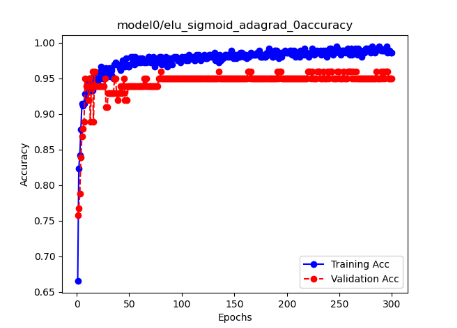
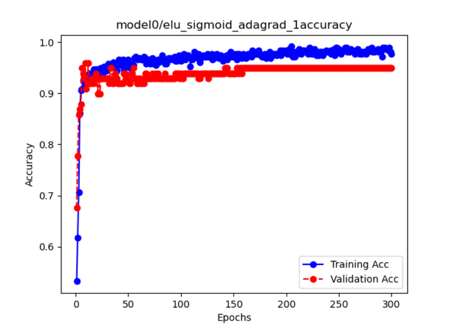
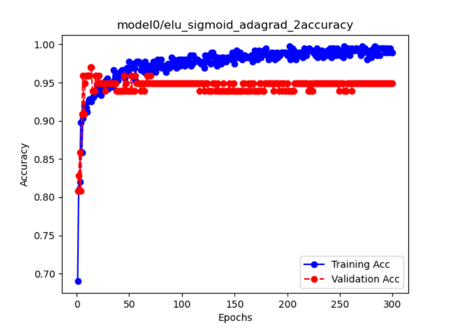
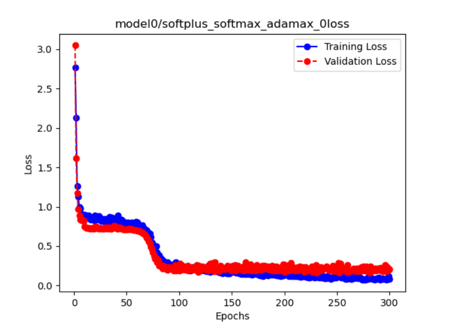
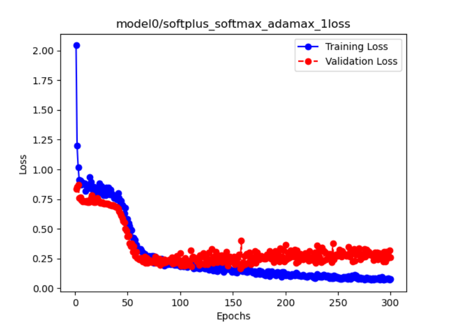
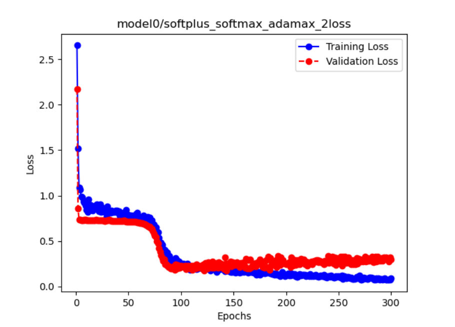
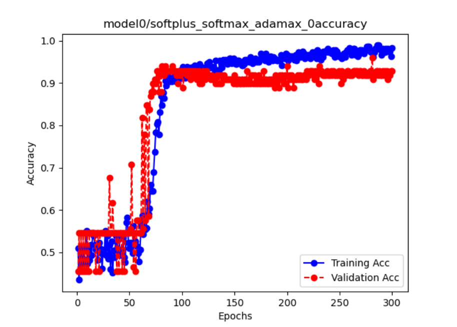
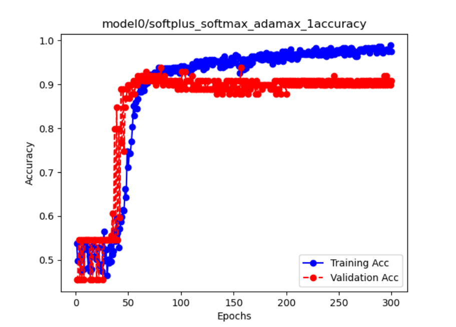
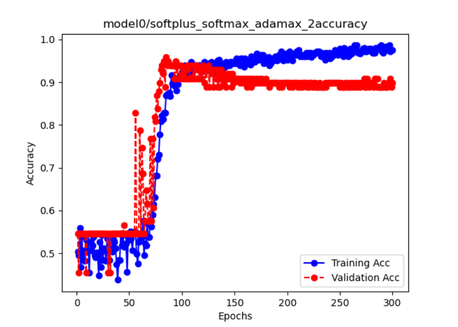
$$j \in \{softmax, sigmoid\}$$

$$k \in \{sgd, adam, admax, adagrad, adadelta, nadam, rmsprop\}$$

Results of Model_0

Testing Accuracy over 97.5% and without severe overfitting

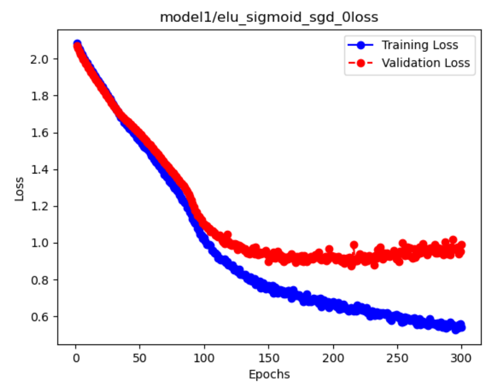
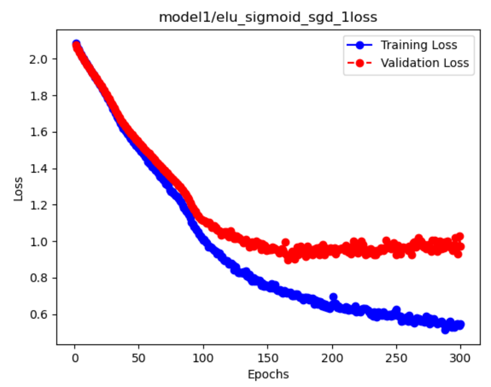
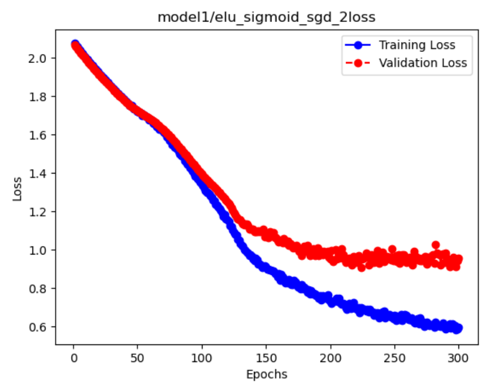
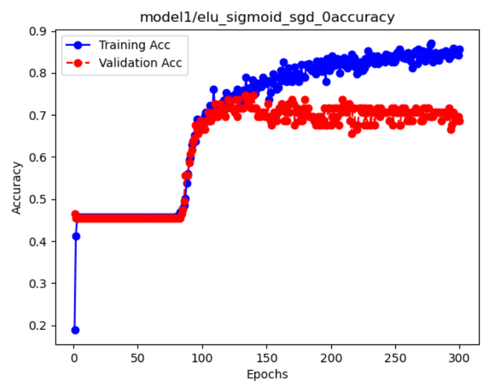
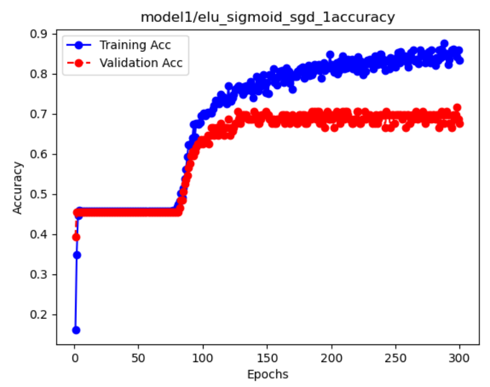
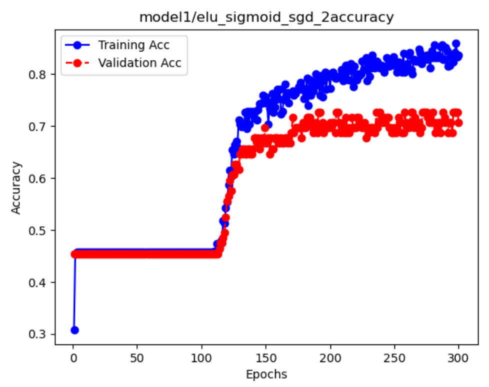
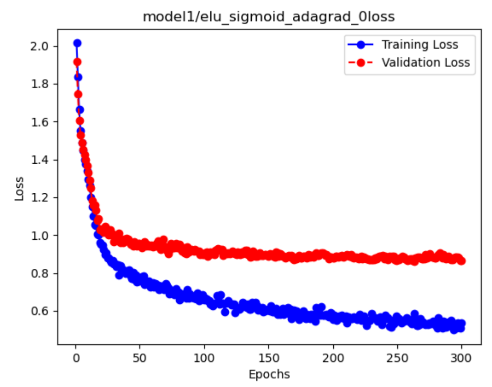
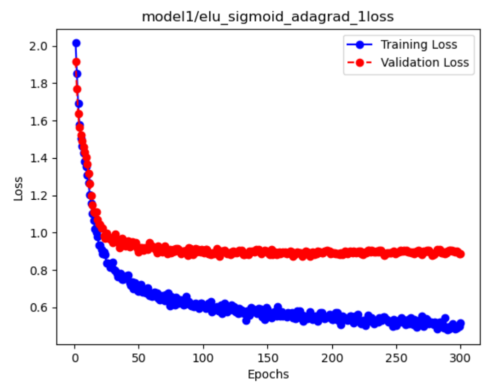
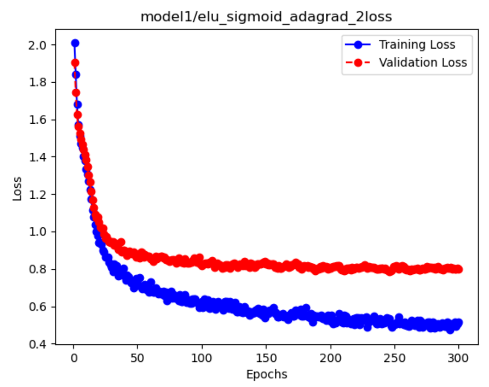
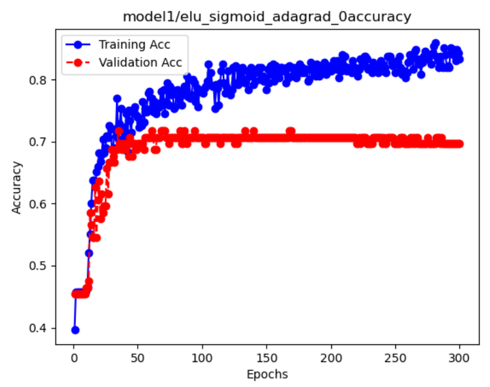
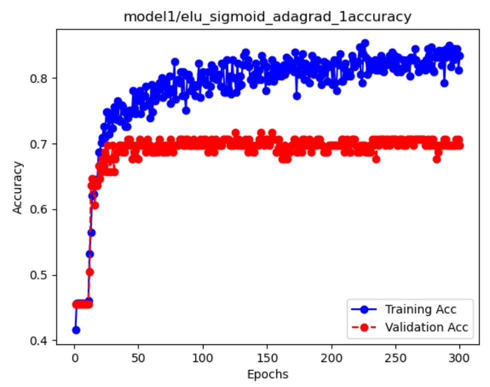
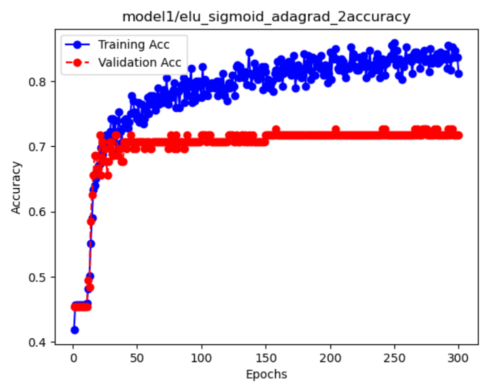
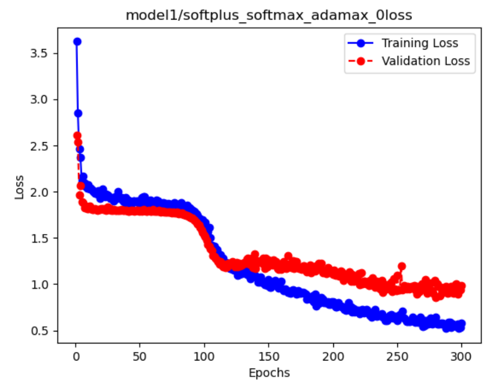
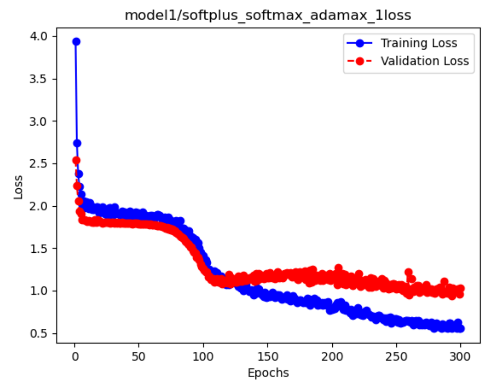
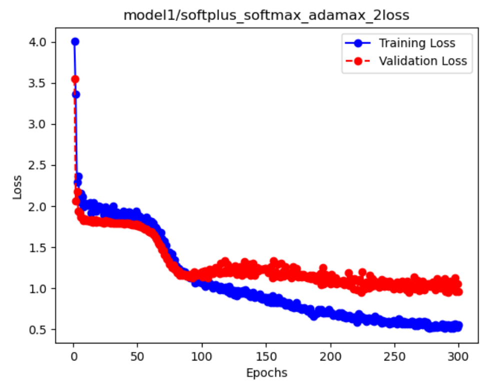
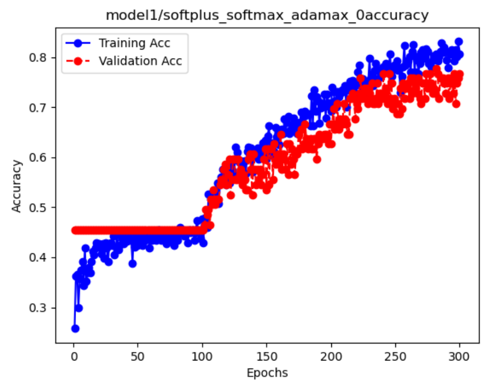
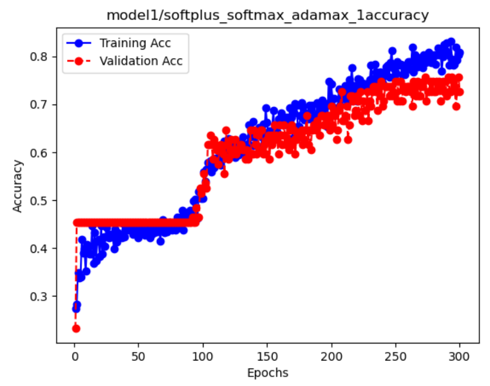
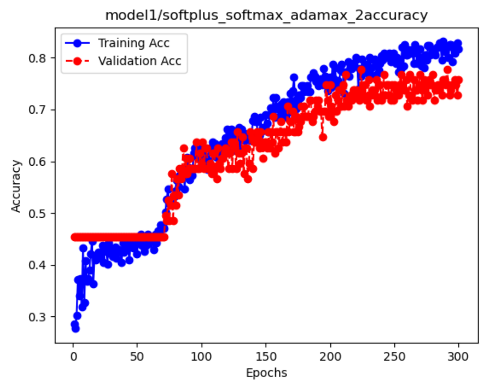
AC	AF	Optimizer	Training Accuracy	Testing Accuracy
relu	softmax	sgd	98%	98.7%
		adagrad	98.7%	98.3%
	sigmoid	sgd	96.7%	97.7%
		adagrad	98%	98.3%
selu	sigmoid	adagrad	100%	98%
elu	softmax	adagrad	99%	97.7%
	sigmoid	adagrad	99.7%	98%
softplus	softmax	adamax	99%	98%

relu	softmax	sgd	 model0/relu_softmax_sgd_0loss	 model0/relu_softmax_sgd_1loss	 model0/relu_softmax_sgd_2loss	 model0/relu_softmax_sgd_0accuracy	 model0/relu_softmax_sgd_1accuracy	 model0/relu_softmax_sgd_2accuracy
		adagrad	 model0/relu_softmax_adagrad_0loss	 model0/relu_softmax_adagrad_1loss	 model0/relu_softmax_adagrad_2loss	 model0/relu_softmax_adagrad_0accuracy	 model0/relu_softmax_adagrad_1accuracy	 model0/relu_softmax_adagrad_2accuracy
	sigmoid	sgd	 model0/relu_sigmoid_sgd_0loss	 model0/relu_sigmoid_sgd_1loss	 model0/relu_sigmoid_sgd_2loss	 model0/relu_sigmoid_sgd_0accuracy	 model0/relu_sigmoid_sgd_1accuracy	 model0/relu_sigmoid_sgd_2accuracy
		adagrad	 model0/relu_sigmoid_adagrad_0loss	 model0/relu_sigmoid_adagrad_1loss	 model0/relu_sigmoid_adagrad_2loss	 model0/relu_sigmoid_adagrad_0accuracy	 model0/relu_sigmoid_adagrad_1accuracy	 model0/relu_sigmoid_adagrad_2accuracy
selu	sigmoid	adagrad	 model0/selu_sigmoid_adagrad_0loss	 model0/selu_sigmoid_adagrad_1loss	 model0/selu_sigmoid_adagrad_2loss	 model0/selu_sigmoid_adagrad_0accuracy	 model0/selu_sigmoid_adagrad_1accuracy	 model0/selu_sigmoid_adagrad_2accuracy
elu	softmax	adagrad	 model0/elu_softmax_adagrad_0loss	 model0/elu_softmax_adagrad_1loss	 model0/elu_softmax_adagrad_2loss	 model0/elu_softmax_adagrad_0accuracy	 model0/elu_softmax_adagrad_1accuracy	 model0/elu_softmax_adagrad_2accuracy
	sigmoid	adagrad	 model0/elu_sigmoid_adagrad_0loss	 model0/elu_sigmoid_adagrad_1loss	 model0/elu_sigmoid_adagrad_2loss	 model0/elu_sigmoid_adagrad_0accuracy	 model0/elu_sigmoid_adagrad_1accuracy	 model0/elu_sigmoid_adagrad_2accuracy
softplus	softmax	adamax	 model0/softplus_softmax_adamax_0loss	 model0/softplus_softmax_adamax_1loss	 model0/softplus_softmax_adamax_2loss	 model0/softplus_softmax_adamax_0accuracy	 model0/softplus_softmax_adamax_1accuracy	 model0/softplus_softmax_adamax_2accuracy

Results of Model_1

Testing Accuracy over 80% and without severe overfitting

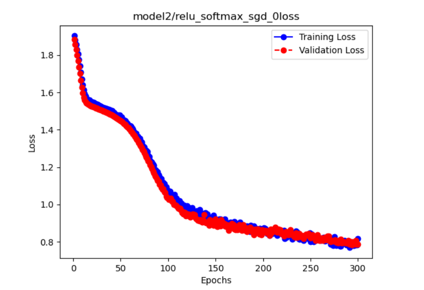
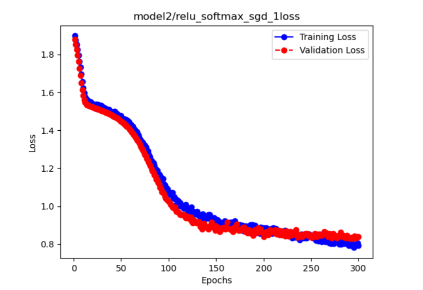
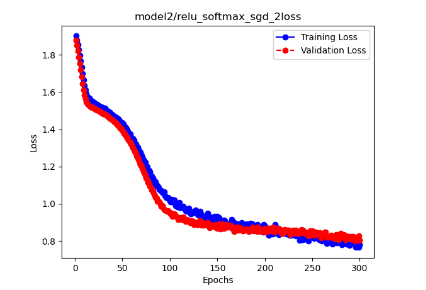
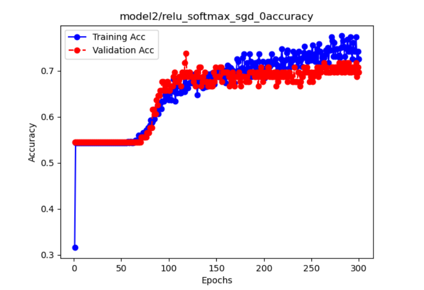
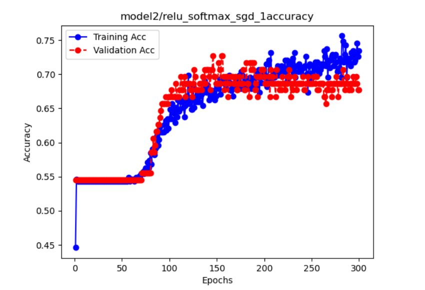
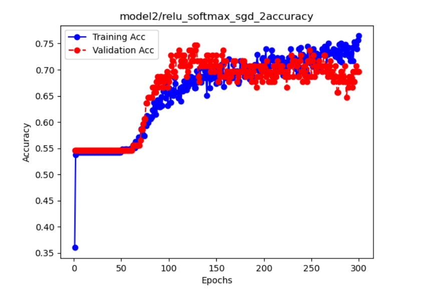
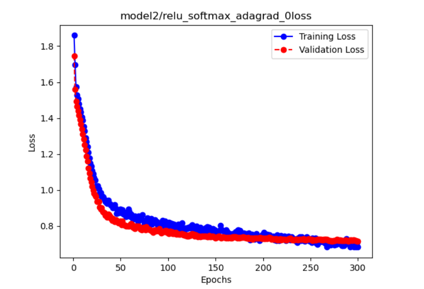
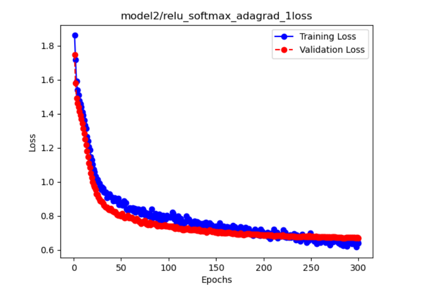
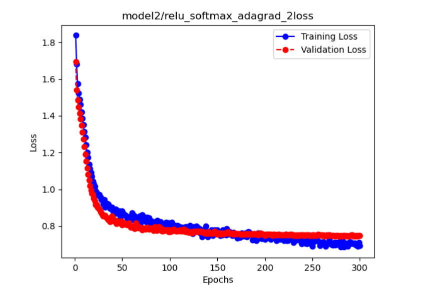
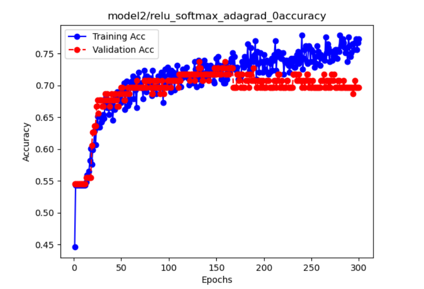
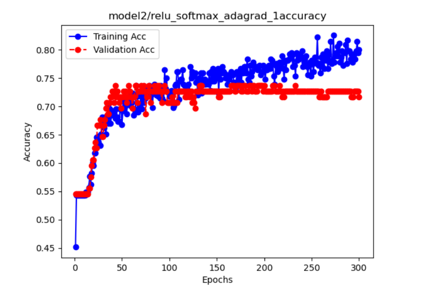
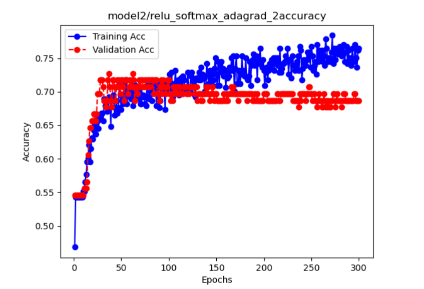
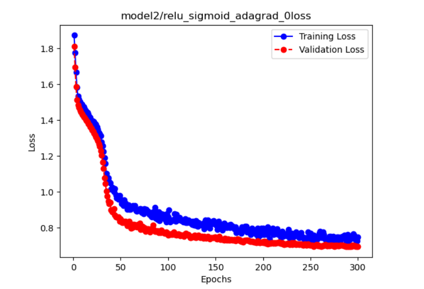
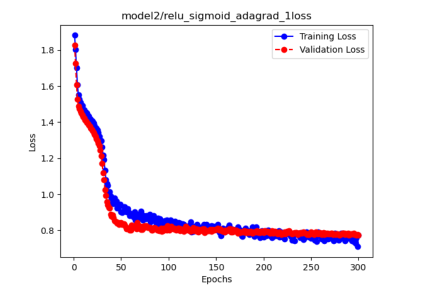
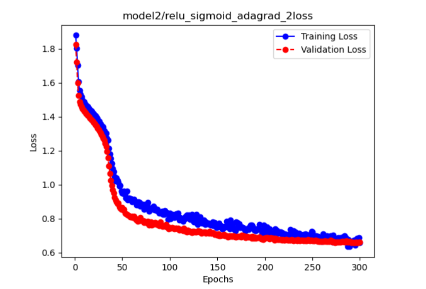
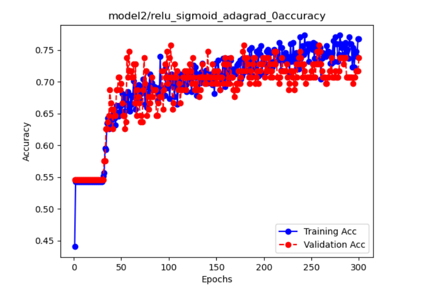
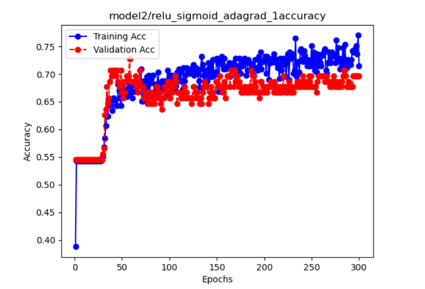
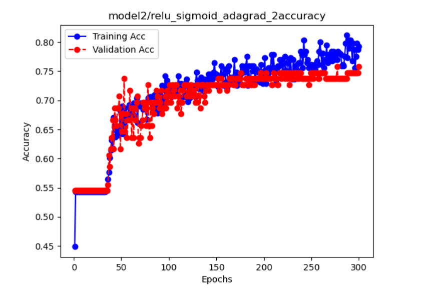
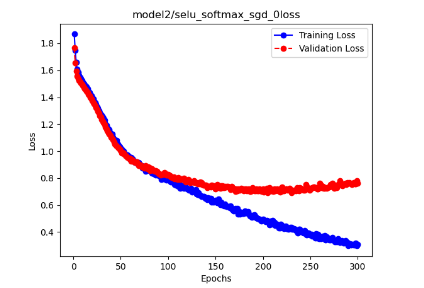
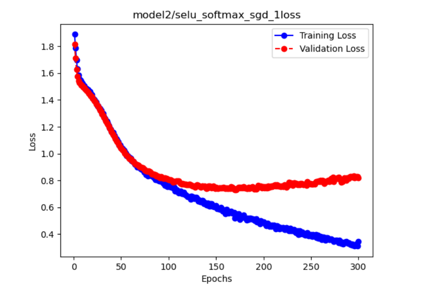
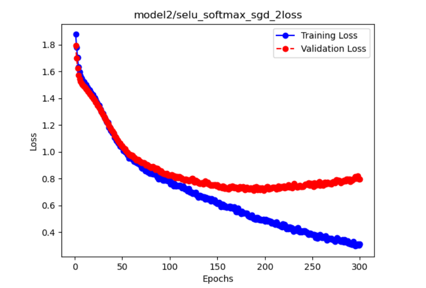
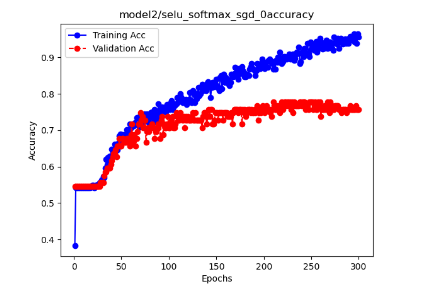
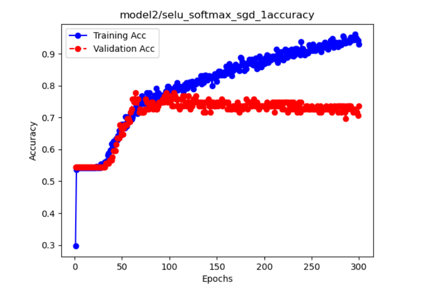
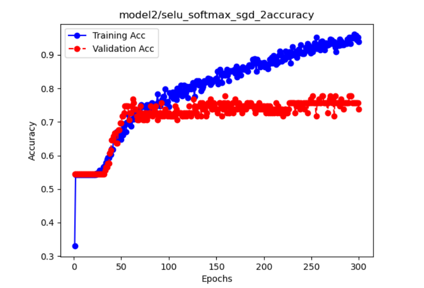
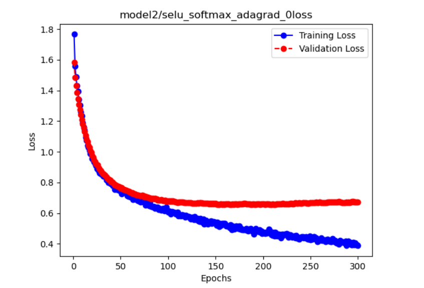
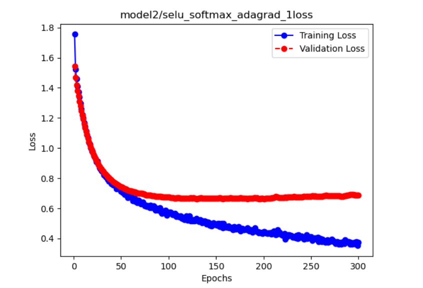
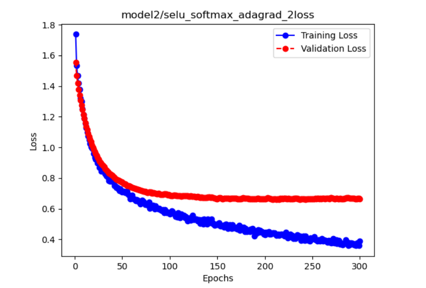
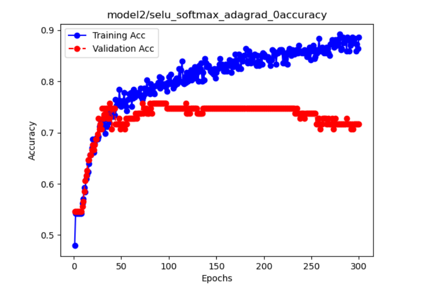
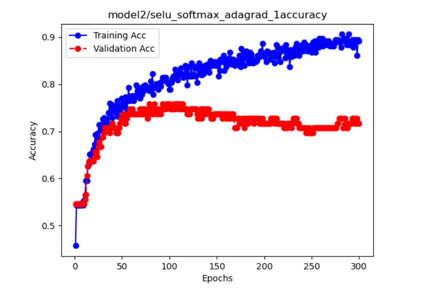
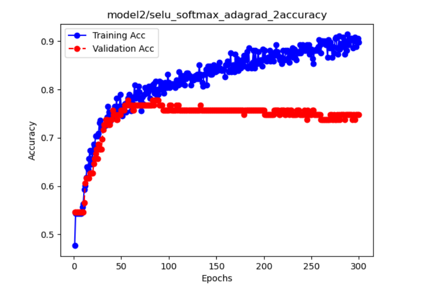
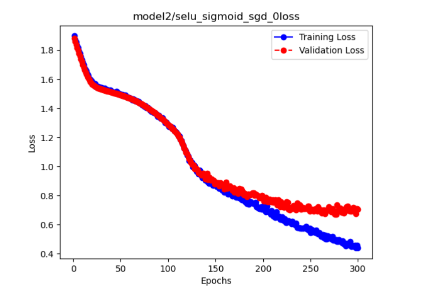
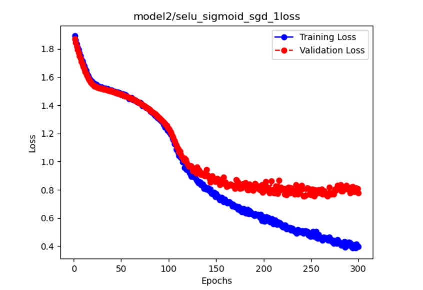
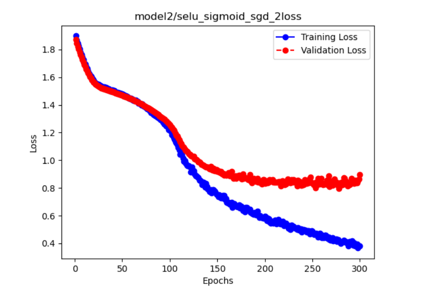
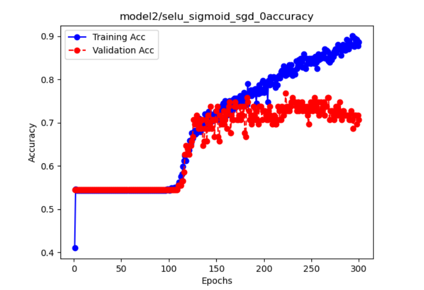
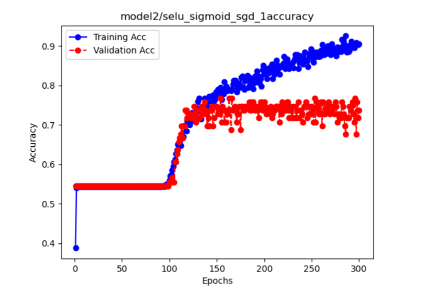
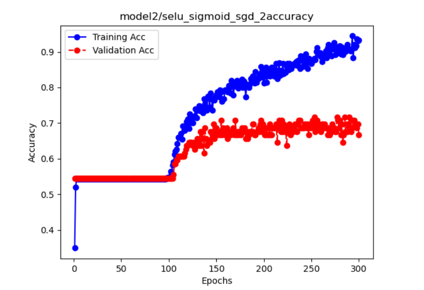
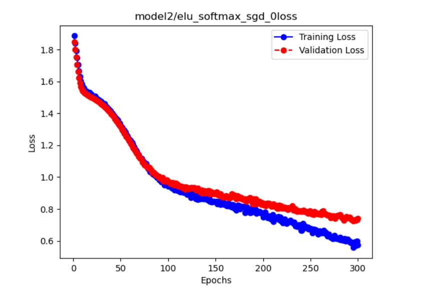
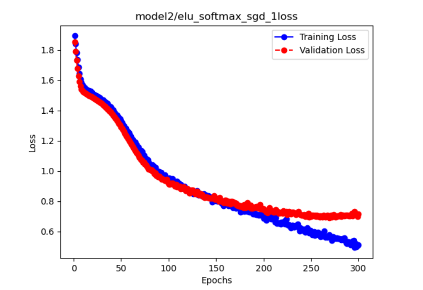
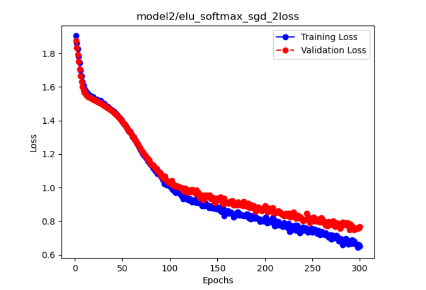
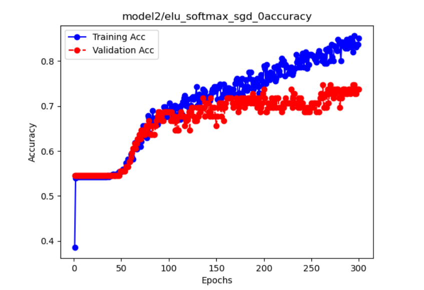
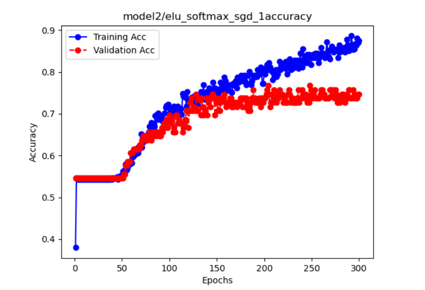
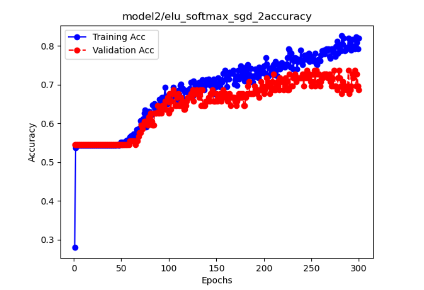
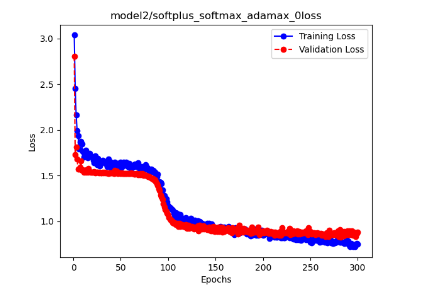
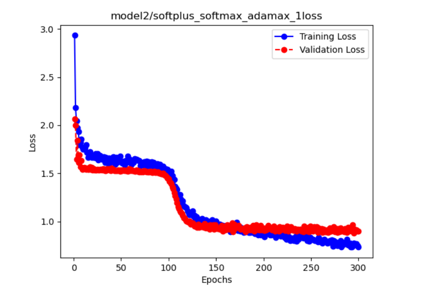
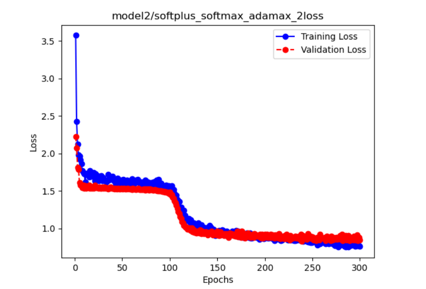
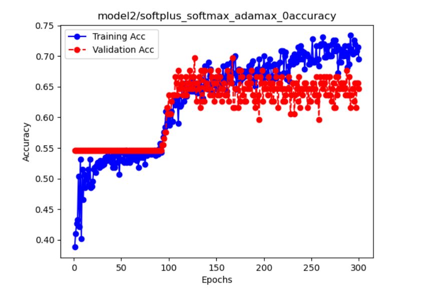
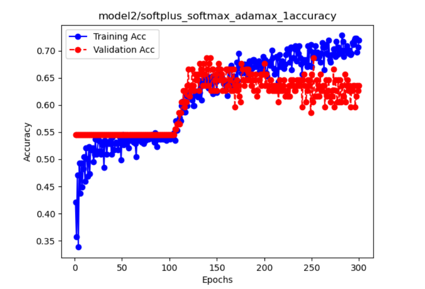
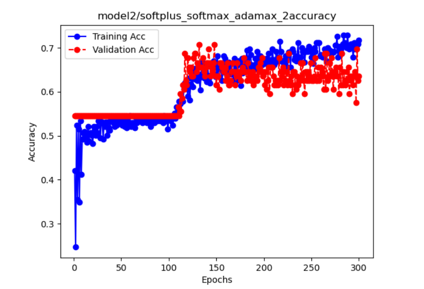
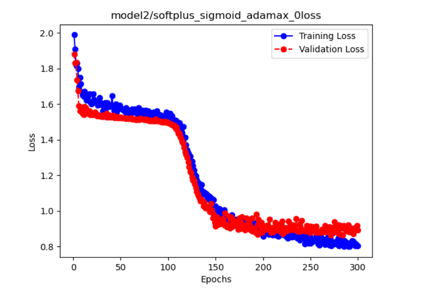
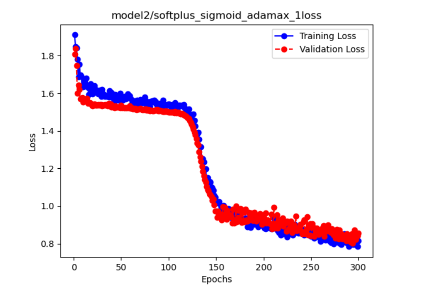
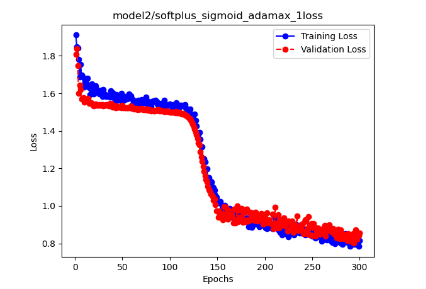
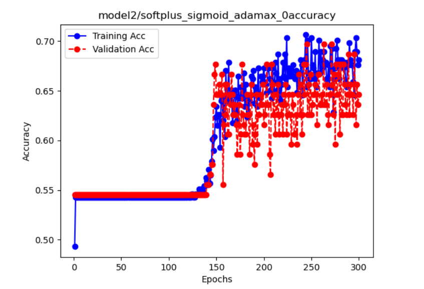
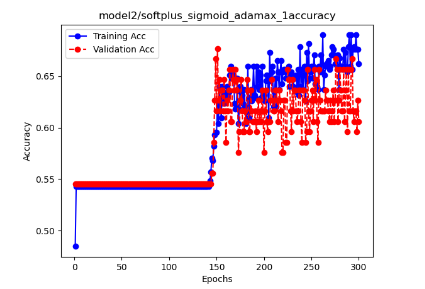
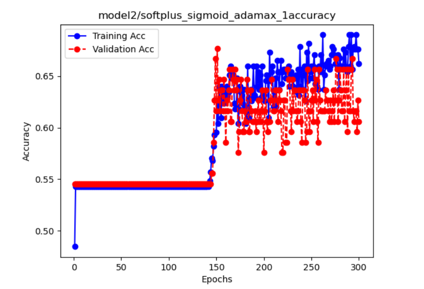
AC	AF	Optimizer	Training Accuracy	Testing Accuracy
elu	sigmoid	sgd	87.3%	81%
		adagrad	87.3%	85.3%
softplus	softmax	adamax	86.3%	84.7%

elu	sigmoid	sgd						
		adagrad						
softplus	softmax	adamax						

Results of Model_2

Testing Accuracy over 70% and without severe overfitting

AC	AF	Optimizer	Training Accuracy	Testing Accuracy
relu	softmax	sgd	76.3%	71.3%
		adagrad	81%	70%
	sigmoid	adagrad	80%	70.7%
selu	softmax	sgd	97.7%	73%
		adagrad	91.3%	71%
	sigmoid	sgd	93.7%	72.3%
elu	softmax	sgd	86.3%	70.6%
softplus	softmax	adamax	73.7%	69%
	sigmoid	adamax	72%	69%

relu	softmax	sgd	 model2/relu_softmax_sgd_0loss	 model2/relu_softmax_sgd_1loss	 model2/relu_softmax_sgd_2loss	 model2/relu_softmax_sgd_0accuracy	 model2/relu_softmax_sgd_1accuracy	 model2/relu_softmax_sgd_2accuracy
		adagrad	 model2/relu_softmax_adagrad_0loss	 model2/relu_softmax_adagrad_1loss	 model2/relu_softmax_adagrad_2loss	 model2/relu_softmax_adagrad_0accuracy	 model2/relu_softmax_adagrad_1accuracy	 model2/relu_softmax_adagrad_2accuracy
	sigmoid	adagrad	 model2/relu_sigmoid_adagrad_0loss	 model2/relu_sigmoid_adagrad_1loss	 model2/relu_sigmoid_adagrad_2loss	 model2/relu_sigmoid_adagrad_0accuracy	 model2/relu_sigmoid_adagrad_1accuracy	 model2/relu_sigmoid_adagrad_2accuracy
selu	softmax	sgd	 model2/selu_softmax_sgd_0loss	 model2/selu_softmax_sgd_1loss	 model2/selu_softmax_sgd_2loss	 model2/selu_softmax_sgd_0accuracy	 model2/selu_softmax_sgd_1accuracy	 model2/selu_softmax_sgd_2accuracy
		adagrad	 model2/selu_softmax_adagrad_0loss	 model2/selu_softmax_adagrad_1loss	 model2/selu_softmax_adagrad_2loss	 model2/selu_softmax_adagrad_0accuracy	 model2/selu_softmax_adagrad_1accuracy	 model2/selu_softmax_adagrad_2accuracy
	sigmoid	sgd	 model2/selu_sigmoid_sgd_0loss	 model2/selu_sigmoid_sgd_1loss	 model2/selu_sigmoid_sgd_2loss	 model2/selu_sigmoid_sgd_0accuracy	 model2/selu_sigmoid_sgd_1accuracy	 model2/selu_sigmoid_sgd_2accuracy
elu	softmax	sgd	 model2/elu_softmax_sgd_0loss	 model2/elu_softmax_sgd_1loss	 model2/elu_softmax_sgd_2loss	 model2/elu_softmax_sgd_0accuracy	 model2/elu_softmax_sgd_1accuracy	 model2/elu_softmax_sgd_2accuracy
softplus	softmax	adamax	 model2/softplus_softmax_adamax_0loss	 model2/softplus_softmax_adamax_1loss	 model2/softplus_softmax_adamax_2loss	 model2/softplus_softmax_adamax_0accuracy	 model2/softplus_softmax_adamax_1accuracy	 model2/softplus_softmax_adamax_2accuracy
	sigmoid	adamax	 model2/softplus_sigmoid_adamax_0loss	 model2/softplus_sigmoid_adamax_1loss	 model2/softplus_sigmoid_adamax_1loss	 model2/softplus_sigmoid_adamax_0accuracy	 model2/softplus_sigmoid_adamax_1accuracy	 model2/softplus_sigmoid_adamax_1accuracy

Future Work

1. Testing candidate model with 1000 epochs

   Model_0: relu+softmax+sgd, relu+sigmoid+sgd

   Model_1: elu+sigmoid+sgd, elu+sigmoid+adagrad, softplus+softmax+adamax

   Model_2: relu+softmax+sgd, relu+softmax+adagrad, relu+sigmoid+adagrad, elu+softmax+sgd, softplus+softmax+adamax, softplus+sigmoid+adamax

2. Adjusting number of layer, number of kernel, and ratio of Dropout layer

   # of layer: ±[ZeroPadding/Conv2D/ZeroPadding/MaxPooling]

   # of kernel: [32, 64, 128, 256] so on and so forth

   (:) of Dropout: $i \in \{0.3,0.35,0.4,0.45,0.5,0.55,0.6\}$