Model:

Activation Function	ReLU
Loss Function	Cross Entropy Loss
Optimizer	SGD, weight decay = 0.001
Training Epochs	10000
Batch Size	32
Learning Rate	0.007
Number of Hidden Layers	1
Hidden Layers: Inputs and Outputs	1: (512) => (256)
Batch Normalization	-
Early Stopping Patience	20
Drop Out Rate	0.1
Last Epoch Information	Epoch 296: | Train Loss: 1.04244 | Val Loss: 1.12140 | Train Acc: 80.000| Val Acc: 48.077
Test Accuracy	52%

Output Prediction:

[0, 1, 0, 0, 0, 3, 0, 3, 0, 0, 4, 2, 1, 0, 0, 0, 0, 0, 0, 0, 2, 3, 1, 0, 0, 4, 0, 1, 0, 0, 3, 0, 0, 1, 0, 2, 0, 0, 2, 1, 0, 2, 0, 2, 2, 0, 1, 2, 2, 3, 0, 2, 0, 2, 0, 1, 0, 1, 4, 3, 0, 2, 4, 0, 0, 0, 2, 3, 0, 2, 0, 1, 0, 0, 1, 0, 4, 0, 0, 2, 0, 0, 0, 2, 1, 2, 0, 2, 0, 2, 1, 3, 0, 2, 0, 4, 2, 0, 0, 0, 1, 0, 0, 0, 0, 0, 0, 2, 0, 0, 0, 0, 4, 2, 0, 0, 2, 0, 0, 2, 0, 0, 0, 1, 0, 0, 2, 2, 0, 0]

Actual Output:

[0, 2, 0, 2, 0, 0, 0, 3, 0, 2, 4, 0, 3, 1, 2, 0, 0, 0, 0, 0, 3, 3, 0, 0, 0, 4, 0, 1, 0, 1, 1, 0, 1, 2, 3, 2, 1, 0, 1, 1, 0, 0, 0, 2, 1, 0, 0, 3, 1, 3, 1, 3, 1, 3, 2, 1, 0, 1, 4, 4, 0, 3, 4, 1, 1, 0, 0, 2, 0, 3, 2, 1, 0, 2, 1, 0, 4, 0, 3, 2, 1, 1, 0, 3, 1, 0, 0, 1, 0, 2, 4, 1, 0, 2, 4, 3, 1, 0, 0, 0, 2, 0, 2, 0, 2, 0, 0, 3, 1, 1, 1, 2, 4, 3, 0, 2, 2, 2, 0, 2, 0, 0, 1, 1, 2, 1, 2, 3, 0, 0]

Results table:

```
              precision    recall  f1-score   support

           0       0.59      0.86      0.70        50
           1       0.53      0.27      0.36        30
           2       0.30      0.33      0.31        24
           3       0.38      0.18      0.24        17
           4       0.86      0.67      0.75         9

    accuracy                           0.52       130
   macro avg       0.53      0.46      0.47       130
weighted avg       0.51      0.52      0.49       130
```

Train-Val Graphs:

Train-Val Accuracy/Epoch

Train-Val Loss/Epoch