**REPORT**

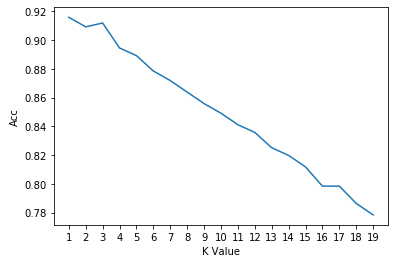
**User Identification Model:** This model is used to identify the user and not if the signature is valid or not. There are 100 users and hence the model predicts the user id based on input only genuine signatures are used to train the model. These Genuine signature have been split into training and testing data.

Training data shape:1747

Test data shape: 749

**K-Nearest Neighbor**

After training the model for various k values, these are the results recorded.



As observed in the graph, for k value 1 and 3 the accuracy is highest with values being 91.58% and 91.18% respectively.

**Neural Network**

**Train data set :1747 values**

**Validation Set : 249 values**

**Test Set : 500 values**

Model was trained for various hidden layers and other parameters. Model Specification can be found below:

**2 Hidden Layer:**

**1st hidden layer : 350 nodes , activation function relu.**

**2nd hidden layer: 150 nodes, activation function relu.**

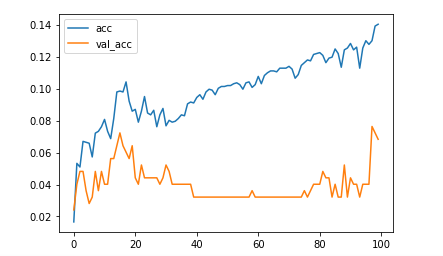
**Output layer: 100 nodes, activation function softmax.**

**Learning rate :** 0.0005 **, optimizer :** adam**, loss :** SparseCategoricalCrossentropy **,epoch :**100

Training acc : 14%

Validation acc: 7.5%

Test Acc : 6.4%



**3 Hidden Layers:**

**1st hidden layer : 350 nodes , activation function relu.**

**2nd hidden layer: 250 nodes, activation function relu.**

**3rd hidden layer:150nodes, activation function relu.**

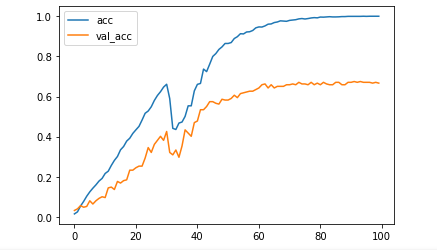
**Output layer: 100 nodes, activation function softmax.**

**Learning rate :** 0.0005 **, optimizer :** adam**, loss :** SparseCategoricalCrossentropy **,epoch :**100

Training acc : 100%

Validation acc: 67%

Test Acc : 66.2%



**4 Hidden Layers:**

**1st hidden layer : 350 nodes , activation function relu.**

**2nd hidden layer: 250 nodes, activation function relu.**

**3rd hidden layer:150 nodes, activation function relu.**

**4th hidden layer:125 nodes, activation function relu.**

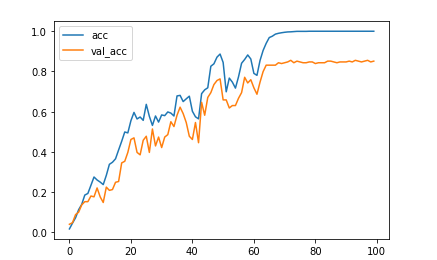
**Output layer: 100 nodes, activation function softmax.**

**Learning rate :** 0.0005 **, optimizer :** adam**, loss :** SparseCategoricalCrossentropy **,epoch :**100

Training acc : 100%

Validation acc: 85.5%

Test Acc : 85%



**4 Hidden Layers:**

**1st hidden layer : 350 nodes , activation function relu.**

**2nd hidden layer: 250 nodes, activation function relu.**

**3rd hidden layer:200 nodes, activation function relu.**

**4th hidden layer:150 nodes, activation function relu.**

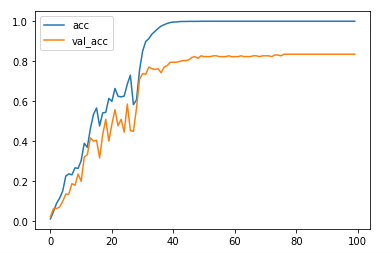
**Output layer: 100 nodes, activation function softmax.**

**Learning rate :** 0.0005 **, optimizer :** adam**, loss :** SparseCategoricalCrossentropy **,epoch :**100

Training acc : 100%

Validation acc: 83.5%

Test Acc : 81.2%



**5 Hidden Layers:**

**1st hidden layer : 350 nodes , activation function relu.**

**2nd hidden layer: 250 nodes, activation function relu.**

**3rd hidden layer:200 nodes, activation function relu.**

**4th hidden layer:150 nodes, activation function relu.**

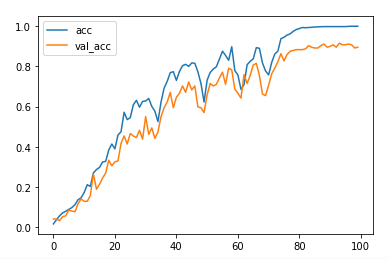
**5th hidden layer:125 nodes, activation function relu.**

**Output layer: 100 nodes, activation function softmax.**

**Learning rate :** 0.0005 **, optimizer :** adam**, loss :** SparseCategoricalCrossentropy **,epoch :**100

Training acc : 100%

Validation acc: 89.5%

Test Acc : 87%

**6 Hidden Layers:**

**1st hidden layer: 400 nodes, activation function relu.**

**2nd hidden layer: 350 nodes, activation function relu.**

**3rd hidden layer:250 nodes, activation function relu.**

**4th hidden layer:200 nodes, activation function relu.**

**5th hidden layer:150 nodes, activation function relu.**

**6th hidden layer:125 nodes, activation function relu.**

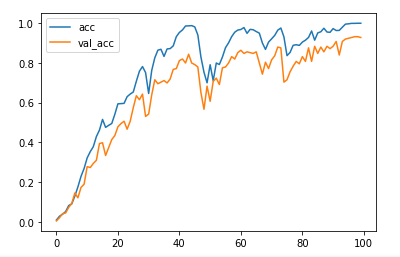
**Output layer: 100 nodes, activation function softmax.**

**Learning rate :** 0.0005 **, optimizer :** adam**, loss :** SparseCategoricalCrossentropy **,epoch :**100

Training acc : 99.6%

Validation acc: 92.3%

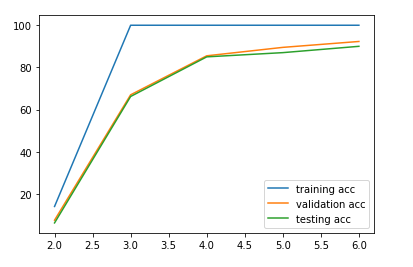
Test Acc : 90%



Summary:

User Accuracy Model

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Model | Specification | Training Acc | Validation Acc | Test Acc |
| KNN | 1 neighbor | - | - | 91.58 |
| KNN | 3 neighbor | - | - | 91.18 |
| DNN | 2 HL 350,150 | 14% | 7.5% | 6.4% |
| DNN | 3HL 350,250,150 | 100% | 67% | 66.2% |
| DNN | 4HL 350,250,150, 125 | 100% | 85.5% | 85% |
| DNN | 5HL 350, 250, 200, 150,125 | 100% | 89.5% | 87% |
| DNN | 6HL 400, 350 , 250, 200, 150,125 | 99.6% | 92.3% | 90% |



**SVM Model**

All the signatures were sampled to 100 points and each point of every parameter is taken as a feature. This gives a total of 513 features to train for the SVM model.

Features: user, x\_len, y\_len, avg\_ps, ratio, diff\_pen, diff\_az, pen\_up, sign\_time, x\_speed, y\_speed, x\_size, total\_length, x0 – x99, y0 – y99, pressure 0 – pressure 99, azimuth 0 – azimuth 99, pen elevation 0 – pen elevation 99

**Model 1:**

Kernel: linear

C (squared l2 penalty): 0.001

Gamma(kernel coefficient): 1/513

Tolerance: 0.001

Decision function: one vs rest

Test accuracy: 71.83%

**Model 2:**

Kernel: polynomial

C: 0.001

Gamma: 1/513

Degree: 3

Tolerance: 0.001

Decision function: one vs rest

Test accuracy: 83.78%