Comparison between MLP, CNN, KAN, CNNKAN

Training strategy:

- Trained for 100 epochs. Interrupted training if no improvement in highest validation accuracy for 15 continuous epochs.
- Normal model trained on original MNIST samples, PGD adversarial model trained on PGD MNIST samples
- For adversarial training, PGD adversarial samples are used with parameters:
  alpha=8/255
  epsilon=0.2
  iter=20
- For KAN model, we have used Efficient KAN since the original KAN implementation was very slow for adversarial training.

Model specifications:

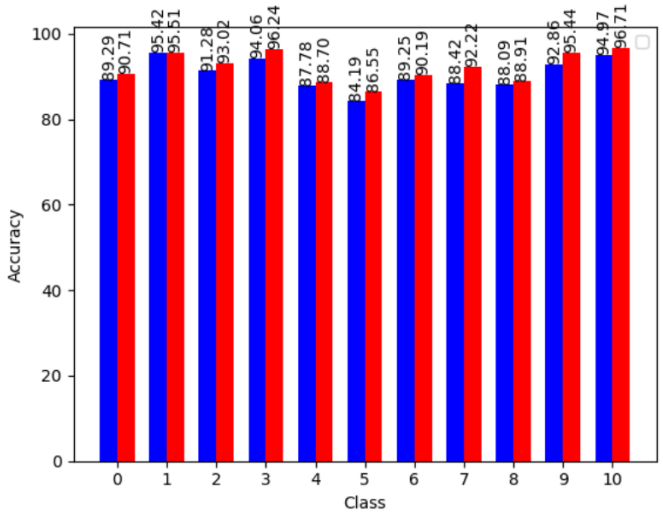
1) MLP
    a) Linear(28*28, 512) -> Linear(512, 256) -> Linear([256,10])
    b) Num parameters: 535,818
2) CNN
    a) Conv(16) -> Conv(32) -> ->Maxpool2d -> Dropout(0.25) -> Linear(4608, 110) -> Dropout(0.5) -> Linear(110, 10)
    b) Num parameters: 512,900
3) KAN
    a) KAN([28*28,64]) -> KAN([64,10])
    b) Num parameters: 508,160
4) CNNKAN
    a) Conv(16) -> Conv(32) -> ->Maxpool2d -> Dropout(0.25) -> KAN(4608, 110) -> KAN(110, 10)
    b) Num parameters: 5,084,600

Results:

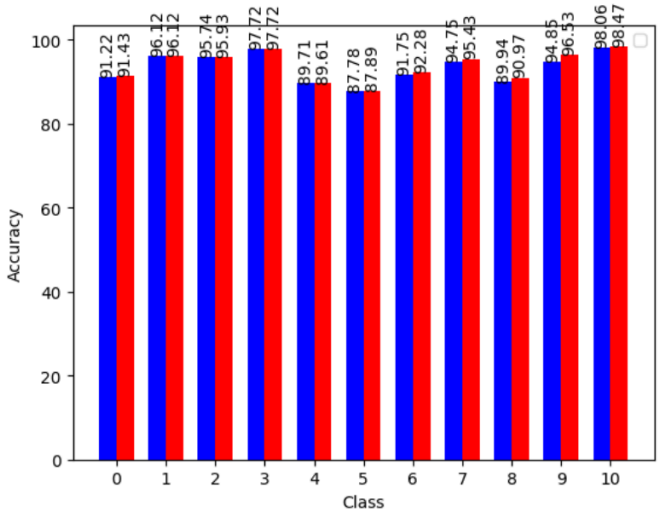
1) **Accuracy on PGD samples for Normal Model v/s PGD Adversarial Model**
   Testing PGD parameters: alpha=8/255; epsilon=0.2 ; iter=20
   0-9 show the MNIST digit classes; 10 shows the test on all classes with random sampling

   Observation: CNN and CNNKAN perform better than others.
   MLP, CNN and KAN have similar number of parameters

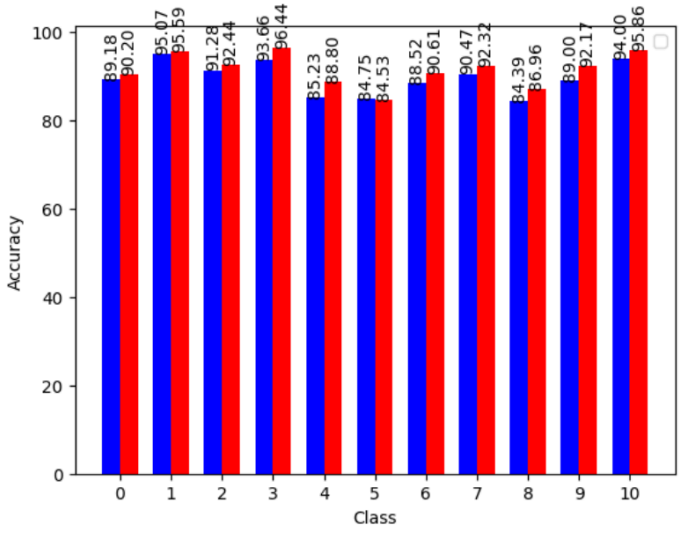
MLP



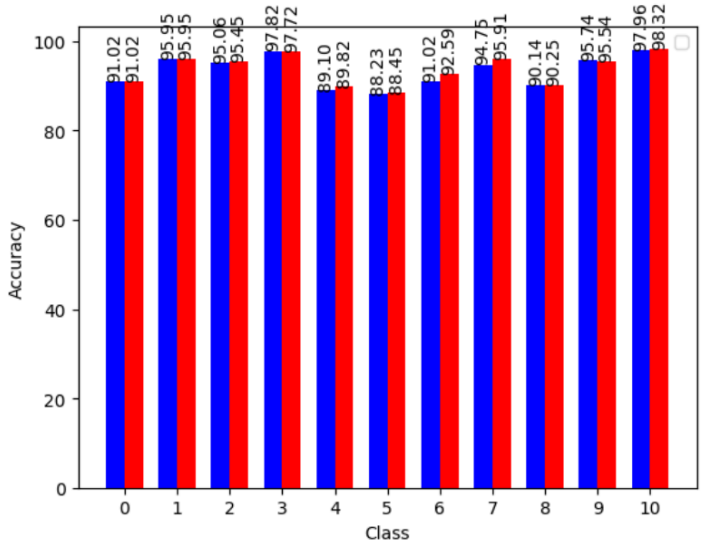
CNN



KAN



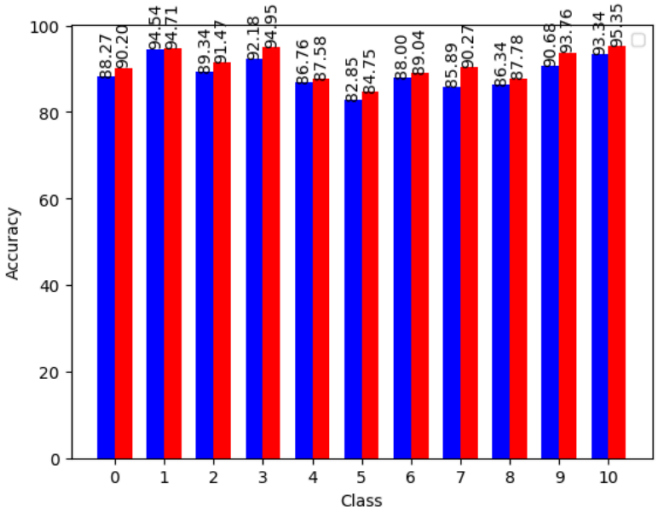
CNNKAN



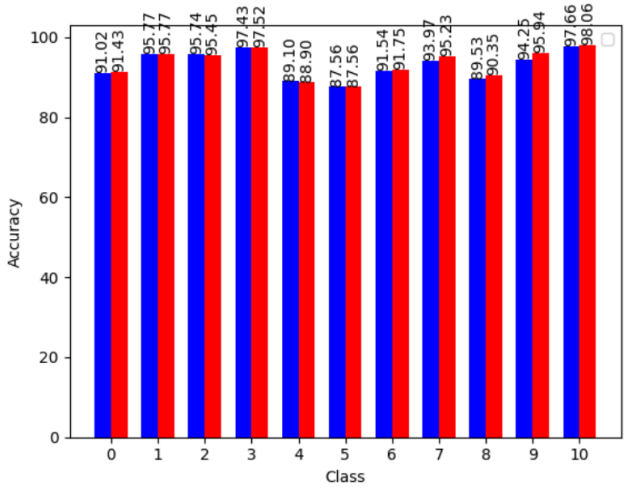
2) **Accuracy on PGD samples for Normal Model v/s PGD Adversarial Model**
Testing PGD parameters: alpha=8/255; epsilon=0.2 ; iter=40
0-9 show the MNIST digit classes; 10 shows the test on all classes with random sampling

Observation: CNN and CNNKAN perform better than others.
MLP, CNN and KAN have similar number of parameters

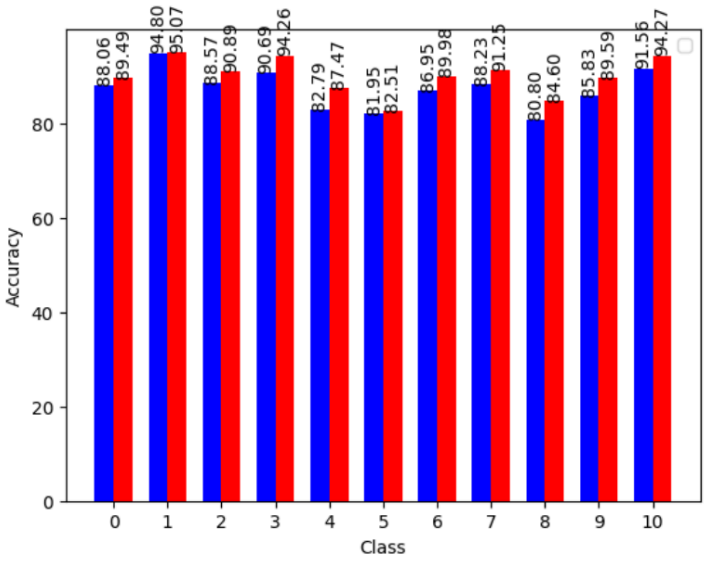
MLP



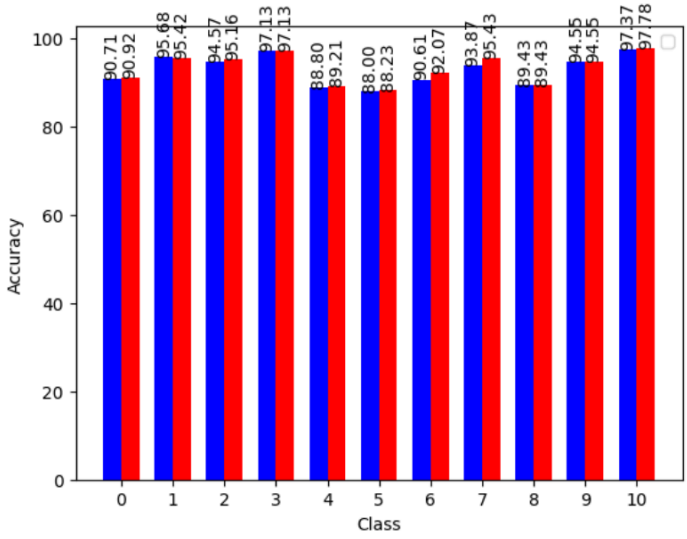
CNN



KAN



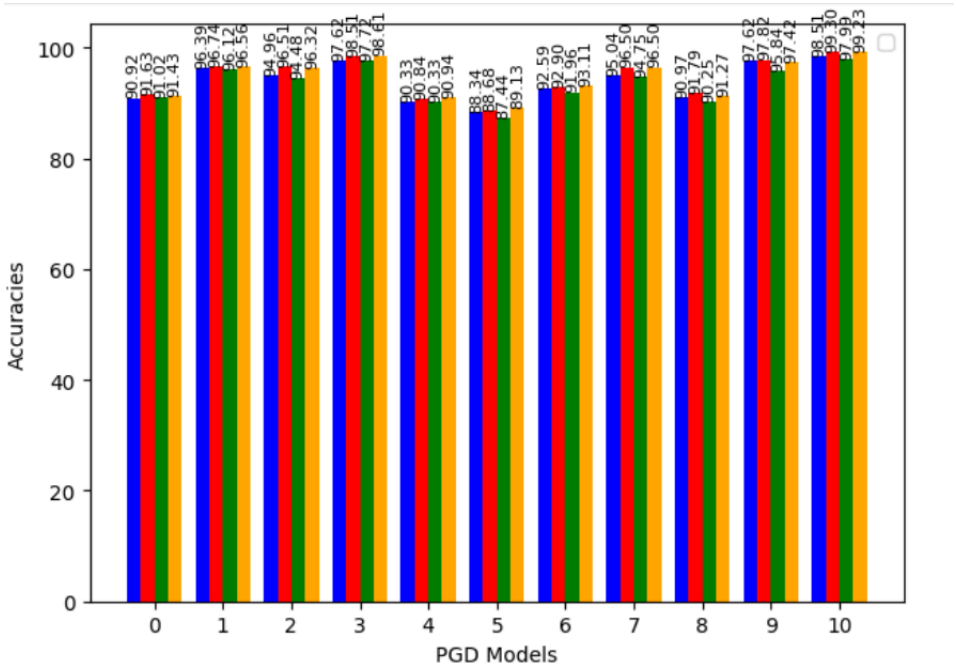
CNNKAN



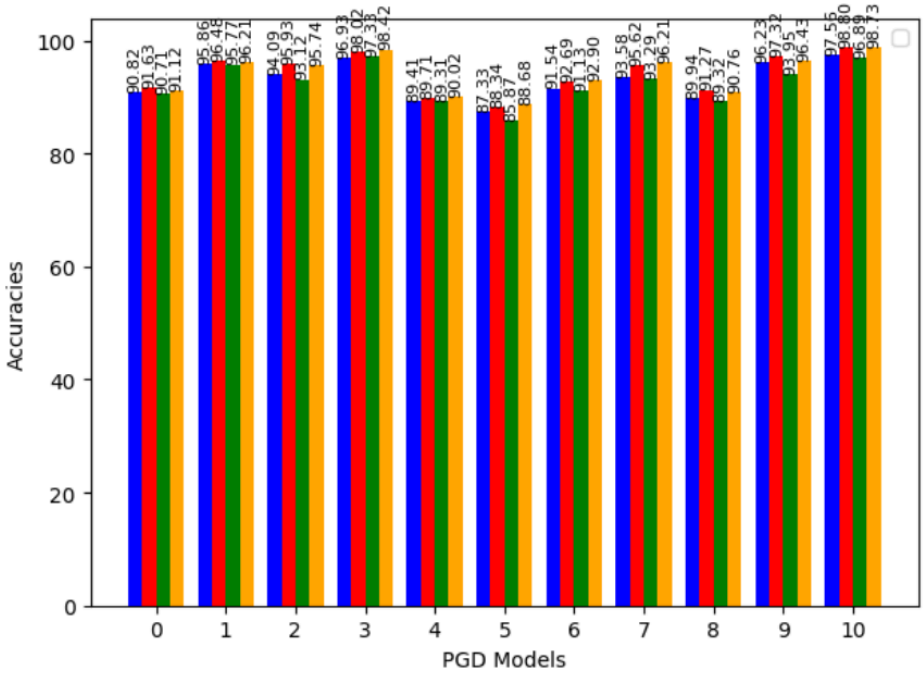
3) **Accuracy on PGD samples for Normal Model v/s PGD Adversarial Model**
Blue - MLP; Red - CNN; Green - KAN; Orange - CNNKAN
Testing PGD parameters: alpha=8/255; epsilon=0.2 ;

Observation: CNN and CNNKAN perform better than others.
KAN has slightly lower accuracy than other models

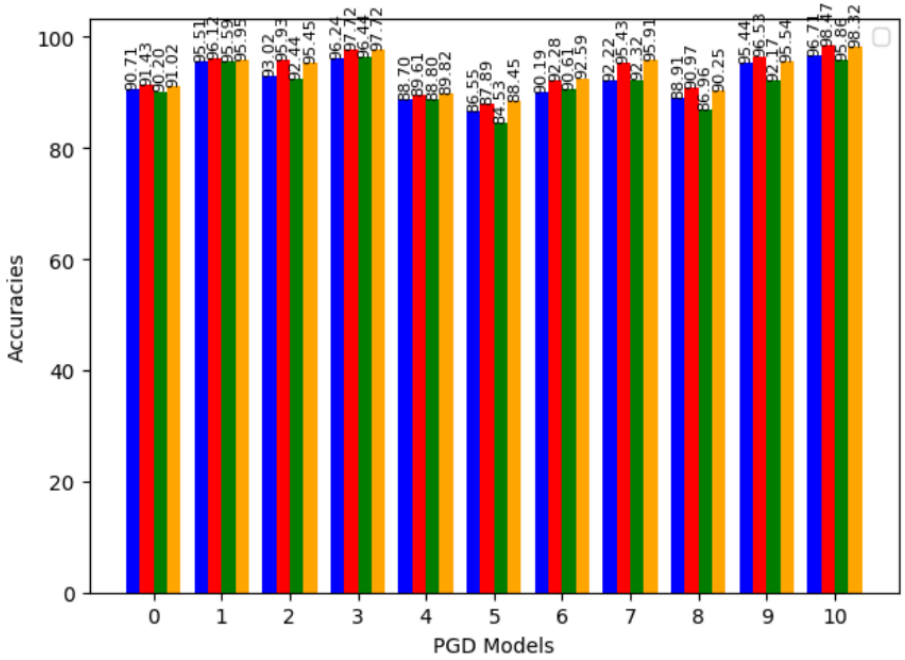
Iters = 0



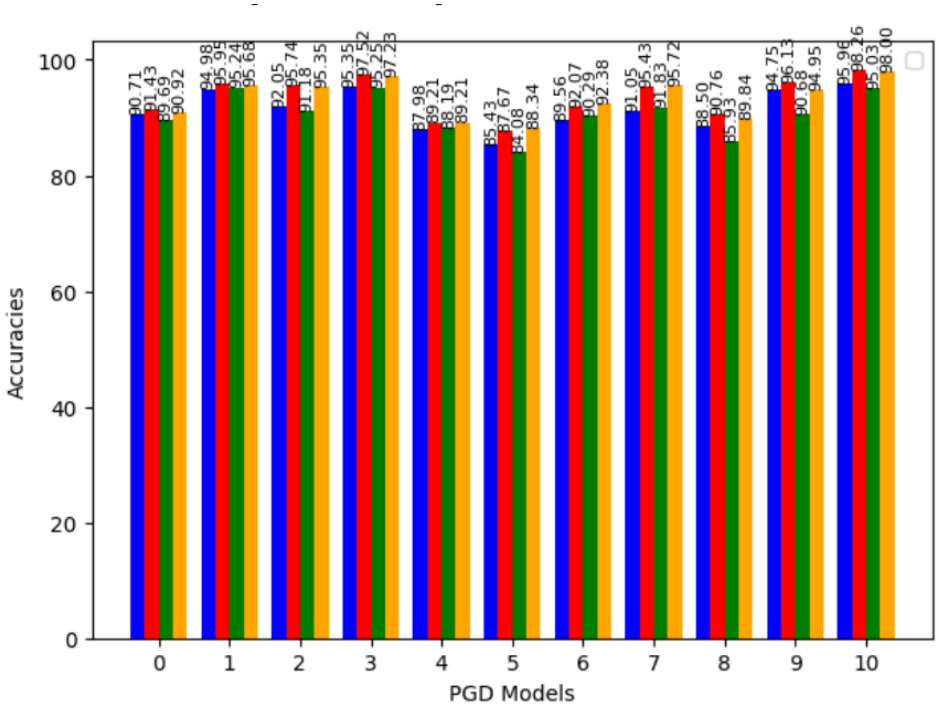
Iters = 10



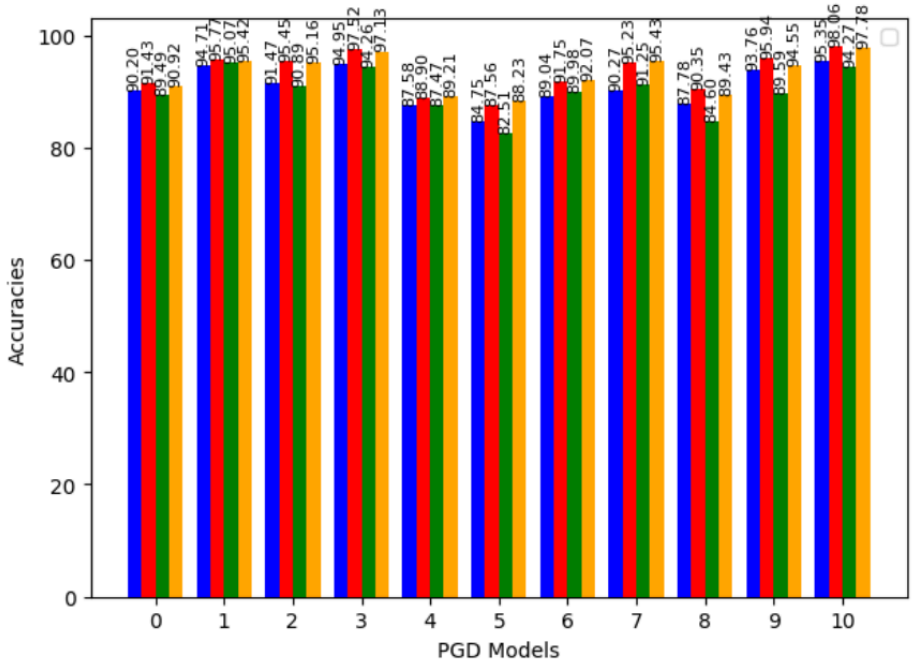
Iters = 20



Iters = 30



Iters = 40

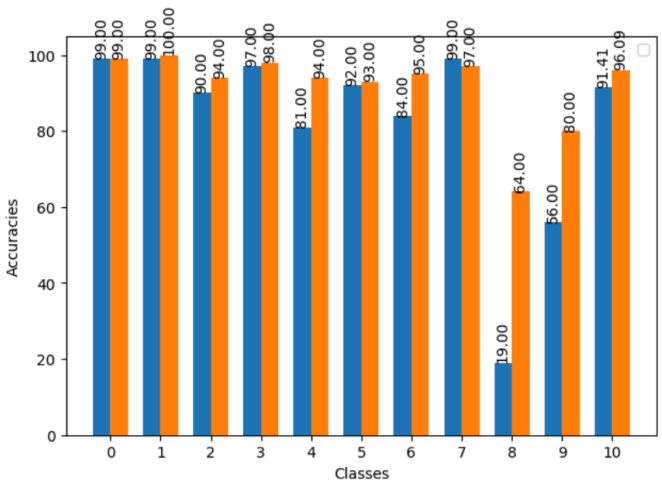


4) **Accuracy on APGD-CE samples for Normal Model v/s PGD Adversarial Model (Trained on PGD samples but tested on APGD-CE samples)**
Blue - Normal; Orange - PGD
Testing APGD-CE parameters: norm='Linf'; eps=8/255; version='standard'

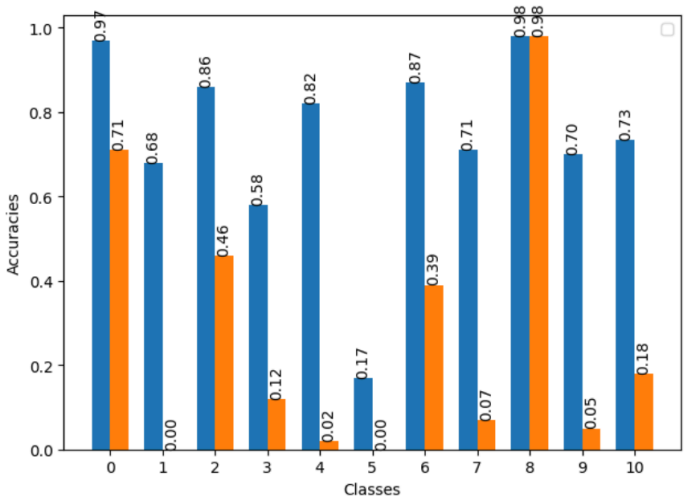
0-9 show the MNIST digit classes; 10 shows the test on all classes with random sampling

Observation: CNN accuracy decreases when testing PGD trained model on APGD-CE samples.
KAN performs better. KAN has the best accuracies on normal model tests.

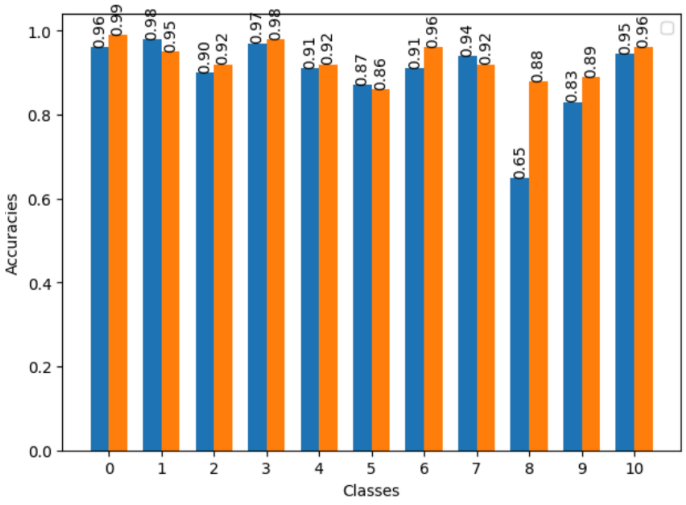
MLP



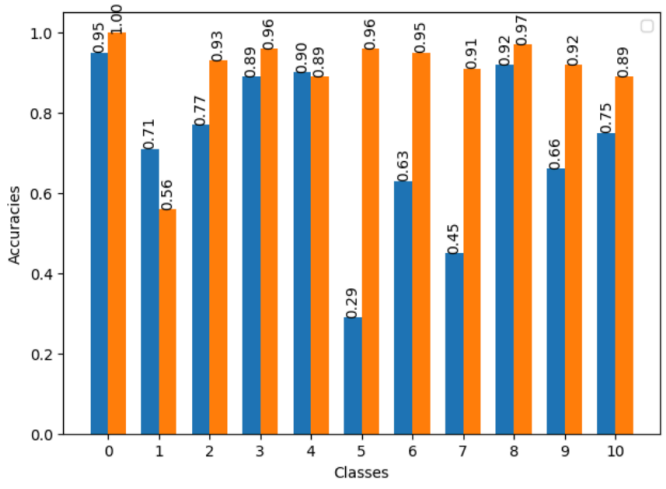
CNN



KAN



CNNKAN

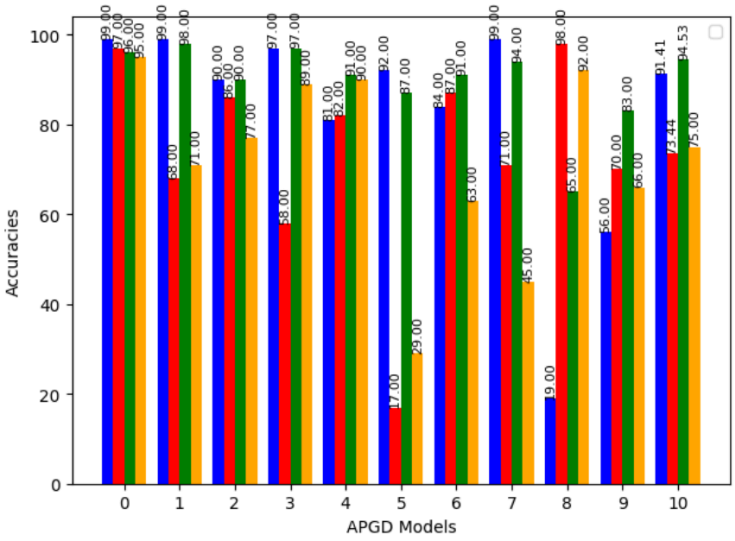


5) **Accuracy on APGD-CE samples for Normal Model v/s PGD Adversarial Model**
Comparision b/w MLPKAN1, MLPKAN2, MLPKAN3
(Trained on PGD samples but tested on APGD-CE samples)
Blue - MLP; Red - CNN; Green - KAN; Orange - CNNKAN

Testing APGD-CE parameters: norm='Linf'; eps=8/255; version='standard'

Observation: No model performs better on all classes. But KAN has best overall performance on APGD-CE samples, for both normal model and PGD trained model.

Normal models



PGD trained models