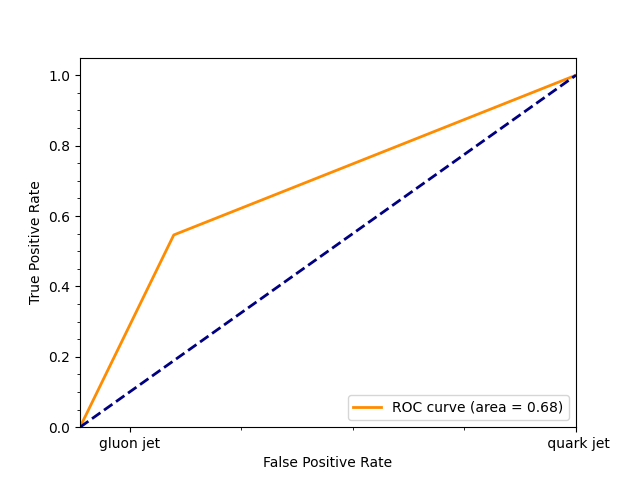
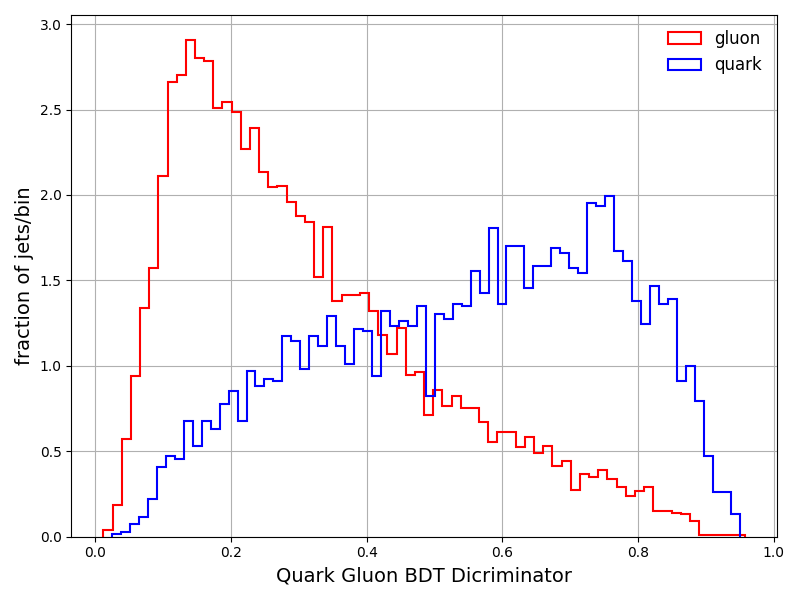
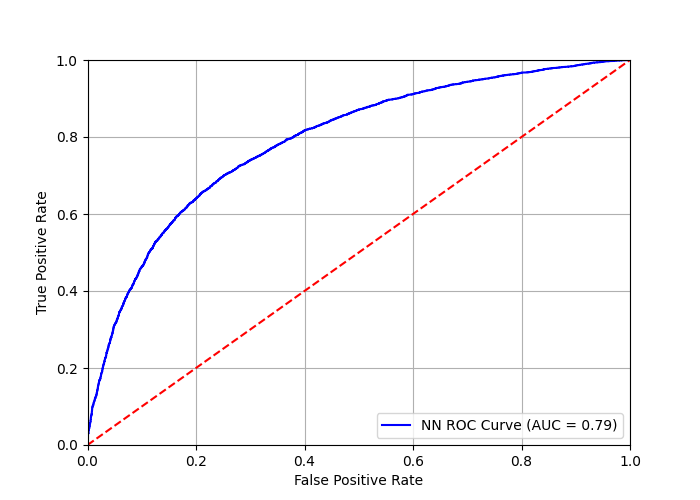
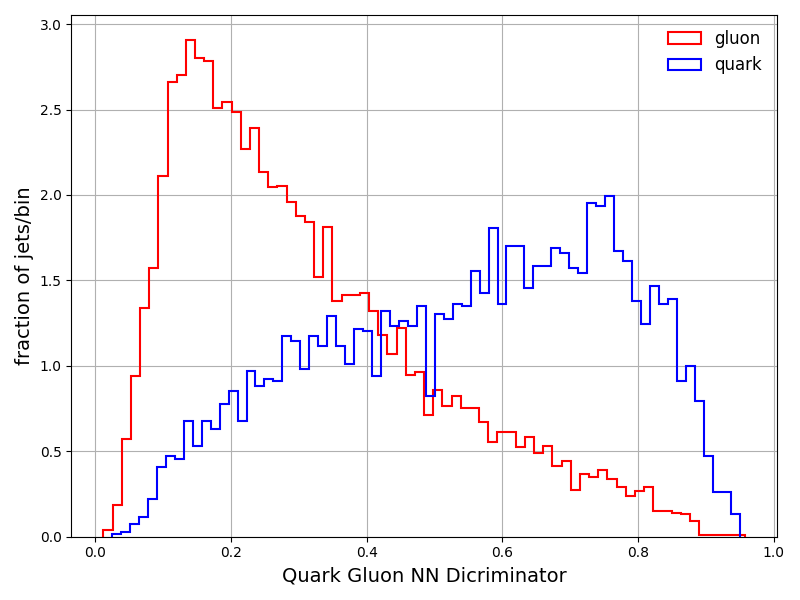
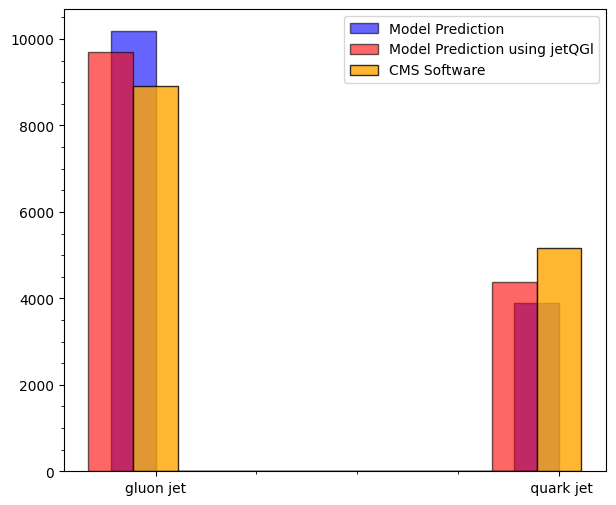
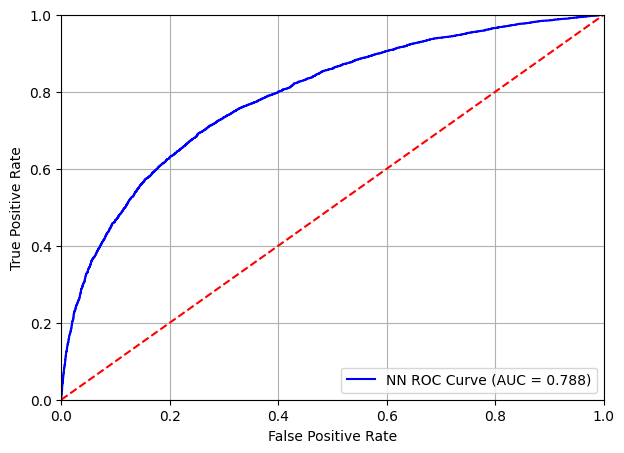
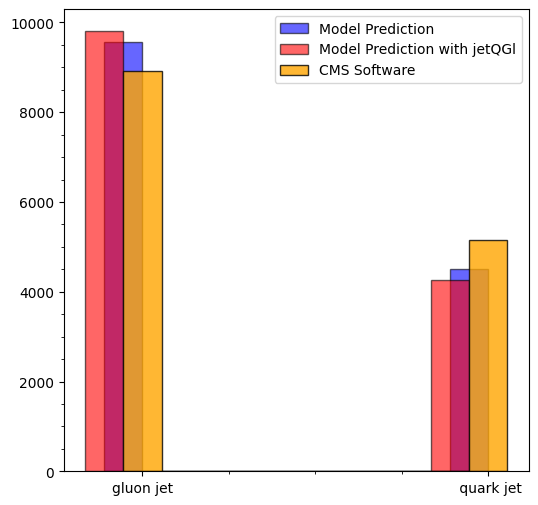
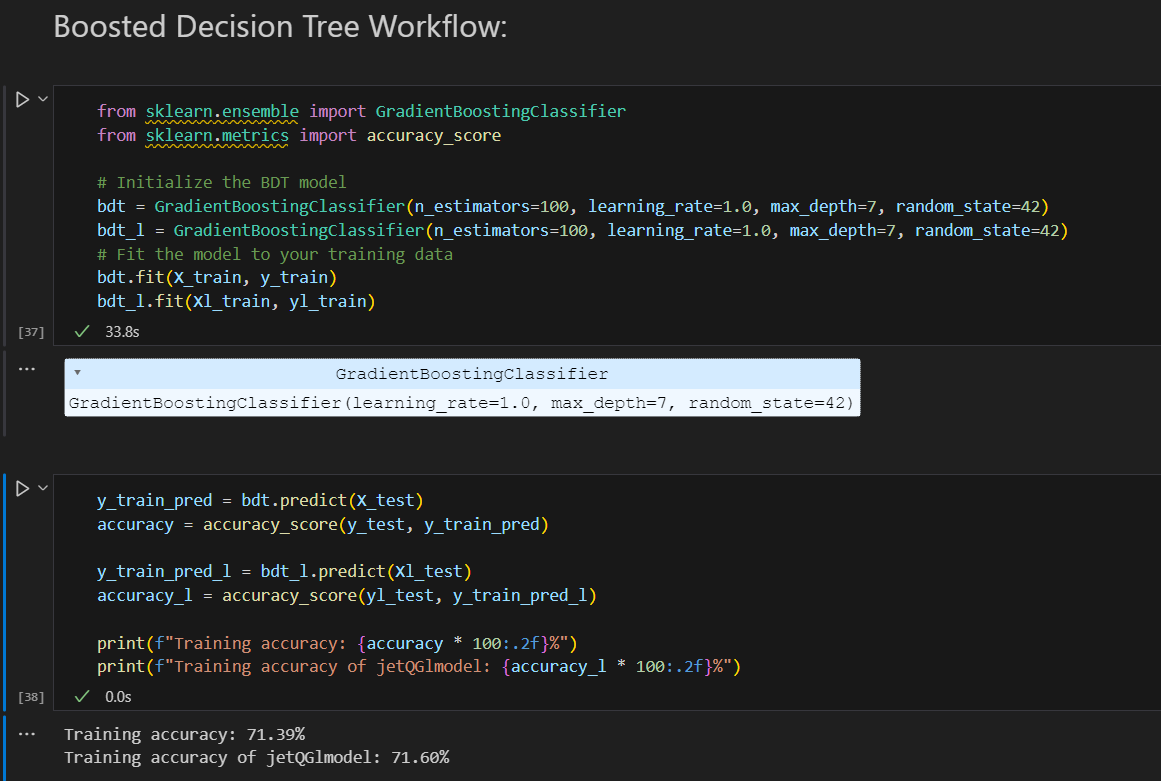
1. 3 variables
   1. BDT:
   2. NN
   3. Comparison with jetQGl as the discriminator variable: this is done by adding

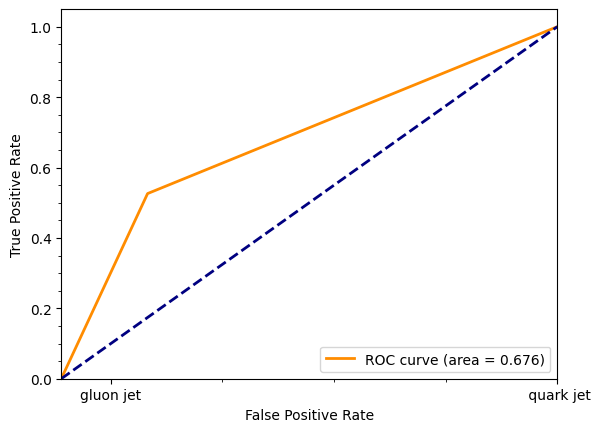
x\_l = df[['jetQGl']] and training the models on x\_l and y (isPhysUDS)

NN: We can see that the model trained on jetQGl is better at identifying a quark jet and does not misidentify a quark jet as a gluon jet as much as the model trained on 3 variablesWe see that the newer model has slightly lower AUC

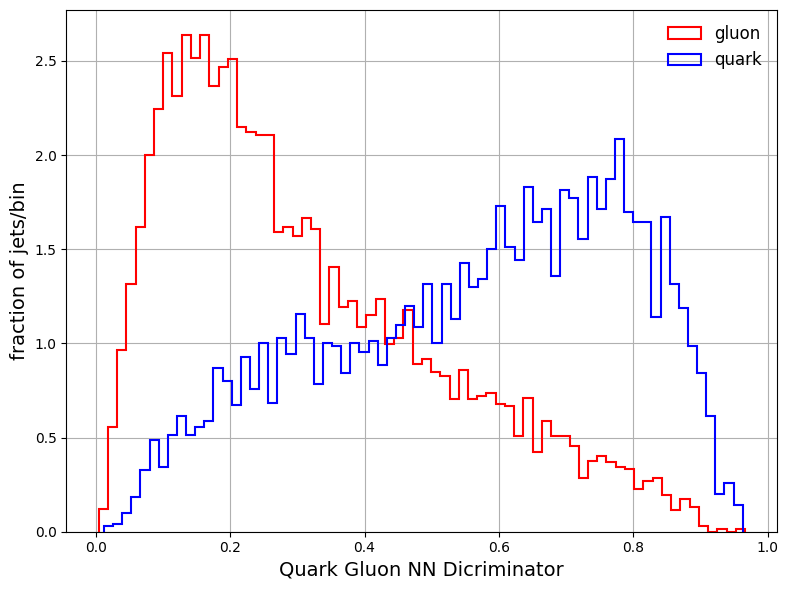


BDT: We see that BDT model trained with jetQGl is worse at identifying a jet as compared to the older model

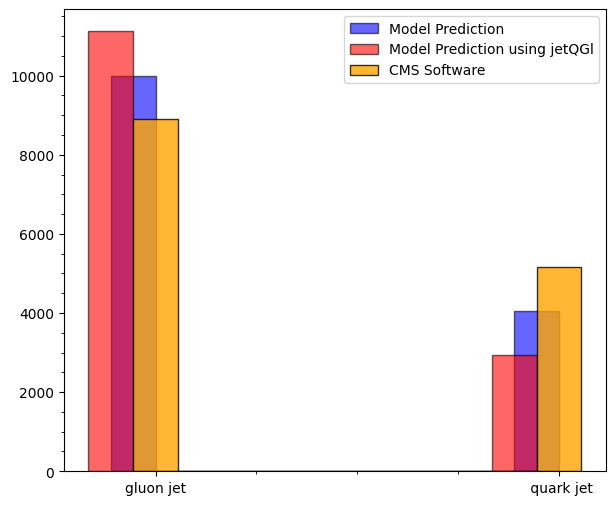
Similarly, the AUC is slightly lower than the AUC for the older model, implying lower effectiveness when classifying a jet.

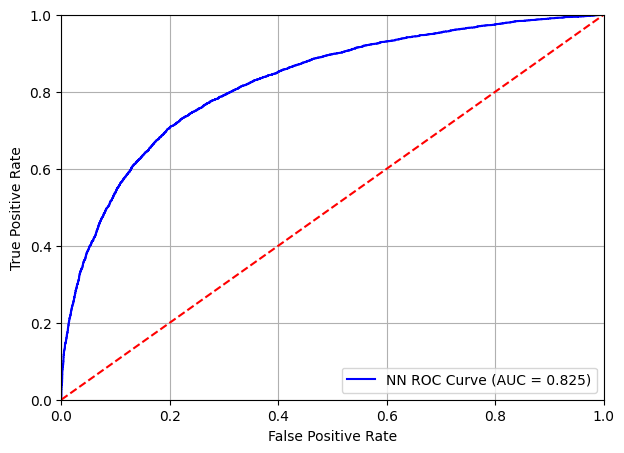


1. Six variables:
   1. NN:

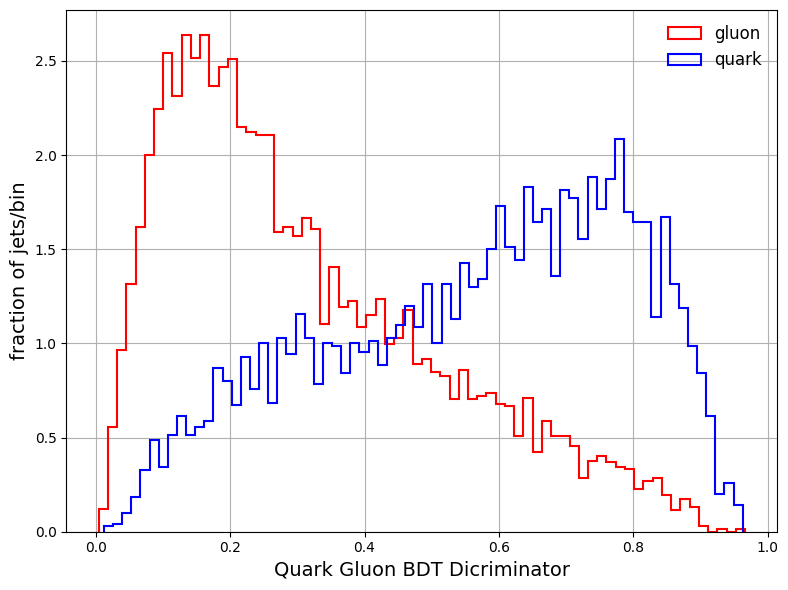
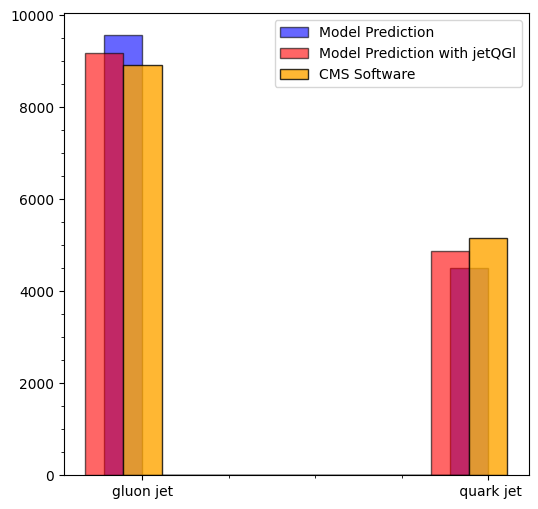


We see that the newer model (with 6 variables) makes more number of classification mistakes than the older model

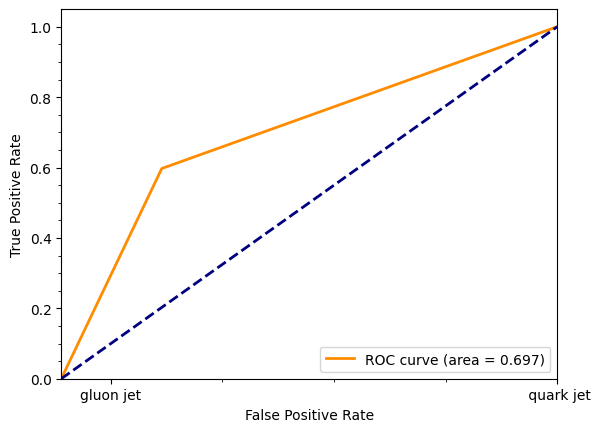


However, the AUC is much higher, implying the model has more likelihood of classifying a jet correctly:

* 1. BDT: We see the BDT has improved performance over the BDT trained using only 3 variables



The higher AUC implies better chance of the model to correctly classify a given jet



1. Input feature distributions: