Regression: Steps Beyond Baseline

Dominick Olivito (UCSD)

Overview

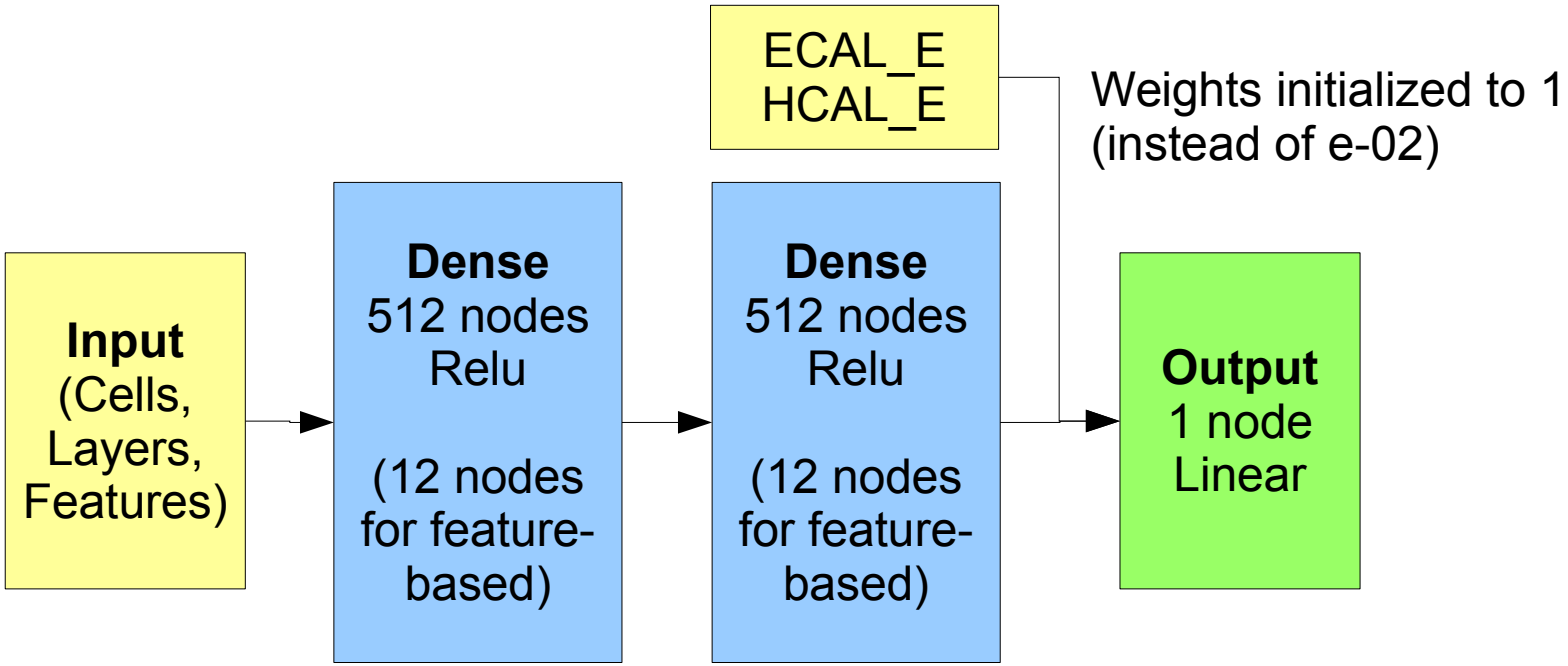
- Try going beyond feature-based regression XGBoost baseline
  - Feature-based NN with same features as XGBoost
    - ECAL_E, HCAL_E, ECALMomentZ1
  - Cell-based NN with original (smaller) window size
    - 25x25x25 for ECAL
    - 5x5x60 for HCAL
    - Larger window size takes twice as long to load data, haven't checked to see if it improves performance. Doubtful for electrons
  - Layer-based NN (running now)
    - since depth in Z was the main feature to improve baseline BDT, try summing in X,Y and only using energy in layers of Z

- Also tried "ResNet" like skip connection for ECAL_E, HCAL_E variables → seems to speed convergence

- All this on fixed angle Electon samples
  - No further news beyond skype thread on variable angle samples

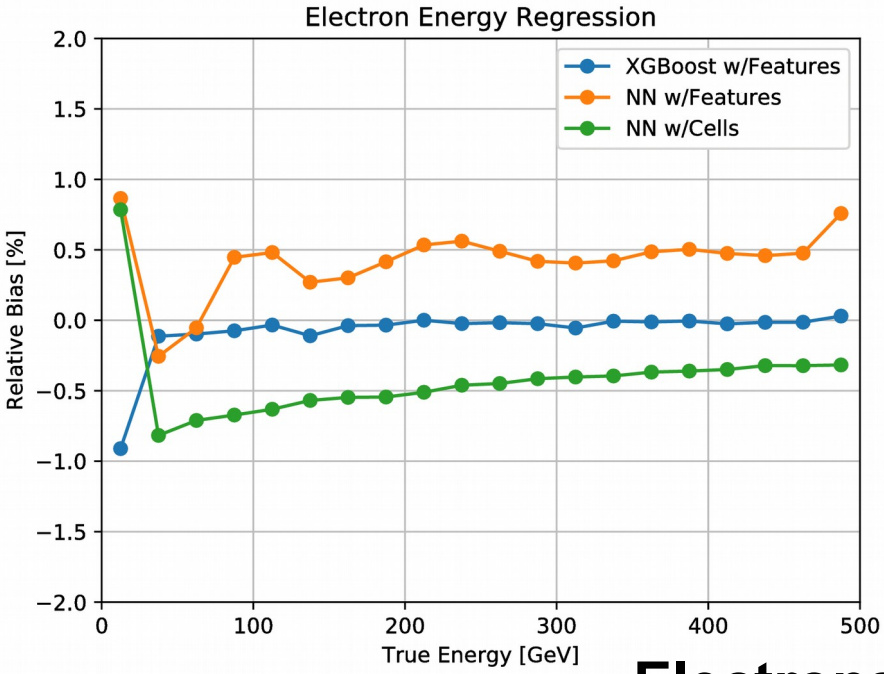
Skip Connections

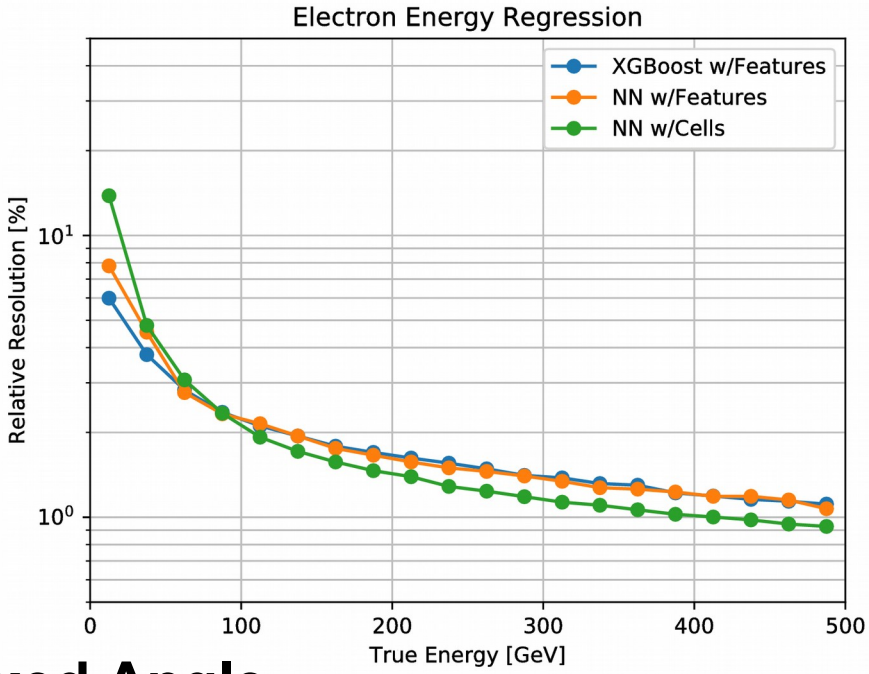
- Basic idea: hardcode **Identity function** into network, to make other layers learn residual correction to identity
- Appropriate for our case: we know linear regression in ECAL_E, HCAL_E gets close to the right answer
- Performance is similar, training converges faster
  - For Feature Based NN, in 10-20 epochs instead of 40



Performance

- NN with features similar to XGBoost
  - Worse resolution below ~50 GeV, slightly worse bias
- NN with cells has better resolution at high energy
  - Better by up to ~20% relative at high energy
  - At low energy, resolution 2x worse than XGBoost
    - Also worse than simpler feature-based NN..
    - Trying again sigmoid activation, to see if this helps with cell-based





Electrons, Fixed Angle

Other stuff / To Do

- Tried varying some hyperparameters
  - Some improvement in convergence from tuning learning rate in Adam, settled on 0.001
  - For cell-based NN, no improvement from making network deeper (4 layers) or wider (1024 nodes/layer)

- Trying:
  - Layer-based NN
  - Sigmoid activation
- Want to also find minimal window size needed
  - Likely depends a bit on Energy

Bonus Slides

Samples / Details

- Samples: new larger window samples, fixed angle, with features
  - On culture-plate at caltech:
    - /data/shared/LCDLargeWindow/fixedangle/*Escan/*.h5
    - /data/shared/LCDLargeWindow/varangle/*Escan/*.h5
  - Slimmed versions with only features (no images):
    - /data/shared/LCDLargeWindow/fixedangle/*Escan/merged_featuresonly/
    - /data/shared/LCDLargeWindow/varangle/*Escan/merged_featuresonly/
  - ~800k events, 70% train, 30% test

- Running XGBoost in python with:
  - maxdepth 3, up to 1000 rounds
  - Early stopping if test loss doesn't improve for 10 rounds
-
- Running NNs in pytorch, python3 using Triforce