Notes

- Select 3 events and try to get 0 loss (maybe with lr = 5e-3) → **started, get almost 0 loss (we get not 0 loss, maybe hint that using only 8 features is not correct model)**
- Use 12 features (8 + 4 delta features) → try to increase number of nodes. Does this improve loss?
- Create validation data → **done**
- Test which of the Δfeatures is the relevant one
- Add also mean features: E1+E1, theta1+theta2,ph1+phi2,time1+time2

  → does adding up angles make really sense? I doubt...
- Cleanup of code → work in progress
- Plot Energy reconstruction with “Edge method”, classic califa clustering, agglomerative clustering,