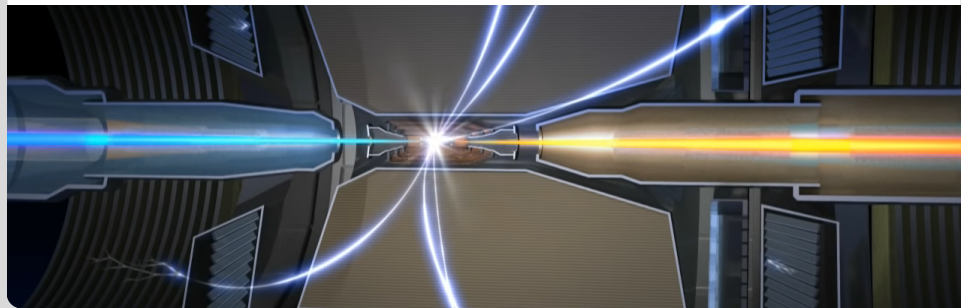


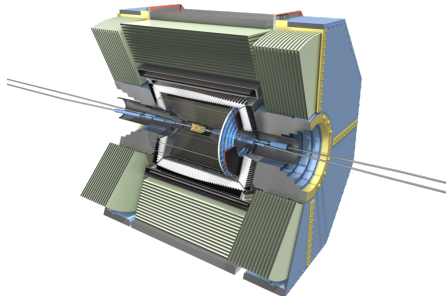
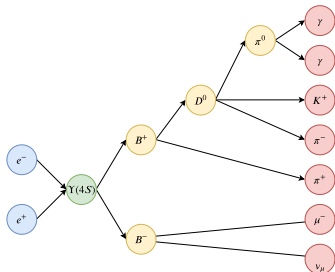
Deep Full Event Interpretation

Deep Learning Hackathon Dresden

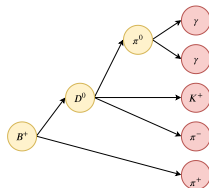
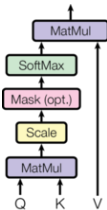
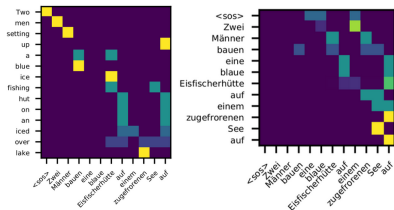
William Sutcliffe, Jochen Gemmler, Moritz Bauer, Tobias Böckh | 09.09.2019



- $e^+ e^-$ collider \rightarrow production of two B mesons
- measurement of final state particles with Belle II detector
- Full Event Interpretation: reconstruction of B mesons
predecessor FEI¹, low efficiency $\mathcal{O}(\epsilon) < 1\%$



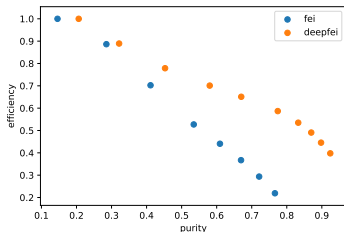
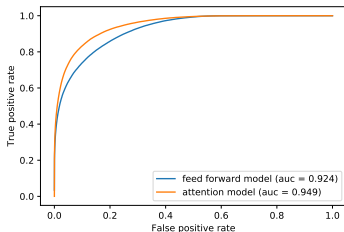
¹T. Keck et al.; 2018; The Full Event Interpretation – An exclusive tagging algorithm for the Belle II experiment; <https://arxiv.org/pdf/1807.08680.pdf>



$$\text{Attention}(Q, K, V) = \text{softmax} \left(\frac{Q \cdot K^T}{\sqrt{d_k}} \right) V = \text{softmax} \left(\frac{XW_Q \cdot (XW_K)^T}{\sqrt{d_k}} \right) XW_V$$

- output is weighted sum of the values
- weights are compatibility function of the query and keys
- Self-attention: query, key, value are linear projections of same input

Reconstruction of $\pi^0 \rightarrow \gamma\gamma$ usign self-attention



Goals for the Hackathon

- reconstruct D^0 for the decays $B^+ \rightarrow \bar{D}^0 \left(\rightarrow K^+ \pi^- \pi^0 \right) \pi^+$ and $B^+ \rightarrow \bar{D}^0 \left(\rightarrow K^+ \pi^- \right) \pi^+$
- multiple attention layers
- Open question: can we interpret attention-maps as intermediate particles?

