Hands-on: Implementing Novel Models

MoE Implementation

- DeepSpeed-MoE
- Fairseq Switch Transformer
- PyTorch custom routing

Long-Context Tools

- Flash Attention library
- LongFormer
- Rope positional encoding

Graph Libraries

- PyTorch Geometric
- DGL (Deep Graph Library)
- Graph Transformer networks

Experimentation

- Start with open-source models
- Medical dataset adaptation
- Benchmark on standard tasks

💻 Practical Implementation

Modern libraries make novel architectures accessible. Start with pretrained models, fine-tune on medical data, and evaluate rigorously on clinical benchmarks.