Discussion 4.1 – Model Architecture

The following questions are designed to guide critical reflection on the interplay between architectural choices, lookback window length, and data characteristics in deep learning for time series forecasting.

1. **Lookback Length vs. Signal Quality**  
   *How do different lookback periods (e.g., 14, 30, 90, 180, 270 days) influence the trade-off between capturing long-term seasonal patterns and avoiding the dilution of recent signals? In which domains might a very short or very long lookback be advantageous, and why?*
2. **Role of Convolutional Front-Ends in Sequence Models**  
   *Conv1D layers are used before recurrent layers to capture local motifs. How might this architectural decision improve performance for time series with volatility clusters or abrupt local events? Could replacing Conv1D with another feature extractor (e.g., WaveNet blocks or temporal convolutions) yield benefits in certain datasets?*
3. **Attention Mechanisms Across Horizons**  
   *Why might attention layers have different levels of utility for short vs. long lookbacks? How does attention help mitigate the risk of “uniform importance smearing” in long sequences, and what would you monitor in attention maps to evaluate model focus quality?*
4. **Regularization Strategy Selection**  
   *The notes propose varying dropout, recurrent dropout, and L2 weight decay depending on lookback length. How should regularization intensity be matched to the data regime? What are the risks of over-regularization for short lookbacks or under-regularization for long lookbacks?*

**Required:** Read all your peers' posts, then comment meaningfully on two or more.