Conditional Neural Processes (CNPs) meta-learn a mapping from context sets $D_C$ to predictive distributions at target locations $x_T$, $p_\theta(y_T|x_T; D_C)$, using neural networks.

### Desirable Properties
1. Data-efficient (using meta-learning)
2. Fast predictions at test time: $O(n + m)$ for predicting at $m$ target locations with $n$ context observations
3. Good uncertainty representation (by modelling stochastic processes)
4. Data-driven expressivity (using deep learning)

### Model Architecture
![Model Architecture Diagram]

### Maximum Likelihood Training
Minimize using gradient descent:

$$
\mathcal{L}(\theta) = -\mathbb{E}_{D_C, D_T \sim D} \left[ \sum_{(x_T, y_T) \in D_T} \log p_\theta(y_T|x_T; D_C) \right]
$$

### References