



Observations:

- RNN Performance:**
 - RNN (10k) has the highest BLEU (~11.87).
 - RNN (10k) uses approximately 2.5M parameters.
 - RNN (10k) has a hidden dimension of 10k.
- Transformer Model Size:**
 - Transformer models (Base, Medium, Optimized) show higher BLEU for smaller parameter counts.
 - Transformer models (Base, Medium, Optimized) use 100k parameters.
 - Transformer models (Base, Medium, Optimized) have hidden dimensions of 10k.
- Resource Efficiency:**
 - BLEU / $d_{model}/hidden_dim$ is highest for Transformer Small (~2.15) and Transformer Optimized (~0.50).
 - RNN (10k) and Transformer Base have the lowest values (~0.25 and ~0.05 respectively).
- RNN vs Transformer:**
 - RNN (10k) has a BLEU / parameters ratio of ~4.75.
 - Transformer models have a BLEU / parameters ratio of ~0.5-1.5.