Parameter Tuning: Real Examples & Effects

See How Parameters Shape LLM Responses

Experiment 1: Temperature Effects

```python
# temperature_effects.py
from litellm import completion

def test_temperature(prompt, temperatures=[0.0, 0.5, 1.0,
    results = {}

    for temp in temperatures:
        response = completion(
            model="gpt-3.5-turbo",
            messages=[{"role": "user", "content": prompt}]
            temperature=temp,
            max_tokens=50
        )
        results[temp] = response.choices[0].message.conten

    return results

# Test with creative writing prompt
test_prompt = "Write the opening line of a short story abo
results = test_temperature(test_prompt)
```

Results: From Deterministic to Creative

Deterministic (temp=0.0)

"Sarah had always walked past the old wooden door at the end of the hallway, but today something made her stop and stare."

Balanced (temp=0.5)

"The brass handle of the peculiar door seemed to glow faintly in the dim corridor light, beckoning Emma closer."

Creative (temp=1.0)

"Nobody remembered when the shimmering door appeared in Mrs. Chen's basement, or why it hummed softly at midnight."

Highly Creative (temp=1.5)

"Seventeen purple butterflies danced around the singing doorframe that definitely hadn't existed yesterday morning."

Experiment 2: Top-p Sampling Effects

```python
# top_p_effects.py
def test_top_p(prompt, top_p_values=[0.1, 0.9, 1.0]):
    results = {}
    for top_p in top_p_values:
        response = completion(
            model="gpt-3.5-turbo",
            messages=[{"role": "user", "content": prompt}],
            temperature=0.8, # Fixed temperature
            top_p=top_p, max_tokens=40
        )
        results[top_p] = response.choices[0].message.content
    return results
# Test with explanation prompt
explain_prompt = "Explain quantum computing in simple terms."
top_p_results = test_top_p(explain_prompt)
```

Focused (top_p=0.1)

"Quantum computing uses quantum bits that can exist in multiple states simultaneously, allowing for incredibly fast calculations."

Balanced (top_p=0.9)

"Think of quantum computing like having a magical computer that can explore many different solutions at once, making it super

🎯 Parameter Combination Strategy

- 📚 **Factual content:** Low temp (0.1-0.3) + top_p (0.8-0.9)
- ✍️ **Creative writing:** Med-high temp (0.7-1.2) + top_p (0.9-1.0)
- 📄 **Technical docs:** Very low temp (0.0-0.2) + top_p (0.7-0.9)
- 💡 **Brainstorming:** High temp (1.0-1.5) + top_p (1.0)

powerful."

Full Diversity (top_p=1.0)

"Quantum computing harnesses weird physics phenomena where particles dance between possibilities, creating computational superpowers!"

⚙️ **Consistent output:** temp=0.0 (deterministic mode)