

Generative AI parameter

Language Tokens

DoiT is a great company to work for.

36 character are represented as
11 token

Tokenized



Temperature



- Deterministic output
- More reliable
- Less creativity



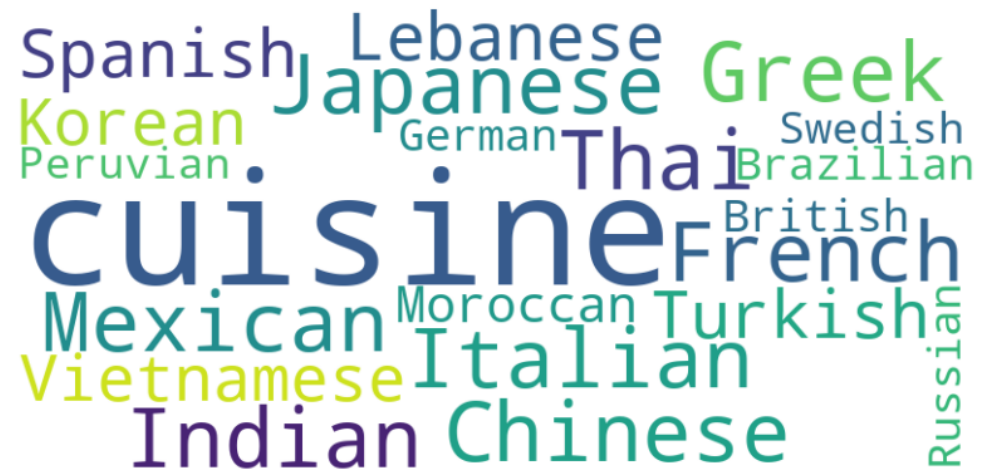
- More creativity
- Unpredictable



Top P



Dinner



Cuisine Type

Situation: Planning a special dinner menu and wants to include a variety of dishes. However, they want to ensure the menu is balanced and doesn't heavily favor one type of cuisine.

Top P



Chef

- top P = 0.9 (90%)



Italy

- pleasing: 0.4



Japan

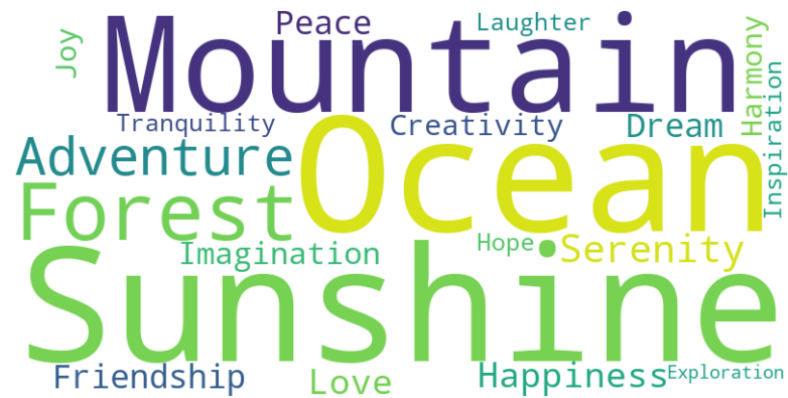
- pleasing: 0.3



Thailand

- pleasing: 0.2

Top K



K = 4

- Mountain
- Ocean
- Sunshine

K = 5

- Mountain
- Ocean
- Sunshine
- Adventure

K = 2

- Mountain
- Ocean

Summary

Parameter	Objective	Low	High
Max token	จำกัดจำนวน token ของผลที่ได้ออกมาจาก generative ai	-	-
Temperature	ควบคุมความ random ของการเลือกหยิบ token	ผลลัพธ์เกี่ยวข้องกับ prompt จริงๆ แต่อาจจะขาดความแปลกใหม่หรือสร้างสรรค์	ผลลัพธ์จะมีความสร้างสรรค์ หรือแปลกใหม่มากกว่า เหมาะกับการ generate งานที่ต้องใช้ความคิดสร้างสรรค์
Top P	ควบคุมความหลากหลายของคำที่ generate ออกมา	ผลลัพธ์ที่ได้อาจจะได้ text ที่มีคำที่ไม่หลากหลาย	ผลลัพธ์ที่ได้ จะได้ text ที่มีคำที่แปลกใหม่หรือหลากหลายมากกว่า
Top K	ควบคุมจำนวนของ token ที่ model จะเลือก generate ออกมา	ผลลัพธ์ที่ได้อาจจะได้ text ที่มีคำที่ไม่หลากหลาย	ผลลัพธ์ที่ได้ จะได้ text ที่มีคำที่แปลกใหม่หรือหลากหลายมากกว่า

Credit

Generative AI - Mastering the Language Model Parameters for Better Outputs

Parameters in large language models are crucial because they help control the model's behavior.



Sascha Heyer · Follow

Published in Google Cloud - Community · 6 min read · Jun 12, 2023



81



1



PaLM 2 and GPT provide parameters for temperature, token limit, top-k, and top-p. This Article explains those parameters with everyday language and relatable examples.

Before we dive into the parameters, it's crucial to understand the concept of tokens.

<https://medium.com/google-cloud/generative-ai-mastering-the-language-model-parameters-for-better-outputs-a82b07b4e383>