LLM sozlamalari

- Harorat, yuqori P va maksimal uzunlik haqida tushuncha

Kirish

Biz modelning turli jihatlarini, masalan, uning qanchalik "tasodifiy" ekanligini nazorat qilish uchun muayyan LLM[1] sozlamalaridan foydalanishimiz mumkin. Bu sozlamalar yanada ijodiy, xilma-xil va qiziqarli mahsulot ishlab chiqarish uchun sozlanishi mumkin. Harorat, Yuqori P va Maksimal uzunlik sozlamalari juda muhim, lekin biz OpenAI Playground sizga o‘zgartirish imkonini beradigan har bir sozlamani tavsiflaymiz.

Harorat

Harorat til modeli chiqishining oldindan aytib bo‘lmasligini tartibga soladi. Yuqori harorat sozlamalari bilan chiqishlar yanada ijodiyroq va kamroq bo‘ladi bashorat qilinishi mumkin, chunki u kamroq ehtimoliy tokenlar ehtimolini oshiradi, uni ehtimoli ko‘proq bo‘lganlar uchun kamaytirish. Aksincha, past haroratlarda hosil bo‘ladi. yanada konservativ va bashorat qilinadigan natijalar. Quyidagi misol shuni ko‘rsatadi ishlab chiqarishdagi ushbu farqlar:

Plyajda 10 ta g‘alati, noyob va qiziqarli narsa nima? Ro‘yxat tuzmoq tavsifsiz.

{’1. Qum qal’a qurmoq 2. Chig‘anoqlarni yig‘ish 3. Plyaj voleyboli o‘ynash 4. Fly a kite 5. Piknik qilmoq 6. Eshkak eshishga harakat qiling 7. Frizbi o‘ynash 8. Snorklingga borish 9. Qirg‘oq bo‘ylab uzoq sayr qiling. 10. Quyosh botishini tomosha qiling"}

{’1. Veydi sayoz suvda pufakchalarni puflayapti 2. Murakkab qum qal’a haykallarini yaratish 3. Qo‘lbola plyaj voleyboli matchida qatnashing 4. Yaqin-atrofdagi qoyalar bo‘ylab sayr qiling 5. Odamlar har birining plyaj bilan bog‘liq hikoyasini tomosha qiling va taxmin qiling 6. O‘z plyaj san’atingizni yaratish uchun dengiz chig‘anoqlarini to‘plang 7. Syorfing yoki boshqa suv sport turlarini sinab ko‘rishni o‘rganing 8. Spontan qum jangini boshlash 9. Mahalliy aholi kabi qirg‘oq bo‘ylab baliq ovlashga harakat qiling. 10. Qumli qal’a qurish bo‘yicha musobaqa tashkil etish orqali do‘stona raqobatga kirishish"}

Yuqori harorat o‘rnatilganda ishlab chiqarilgan mahsulot plyajda bajarilishi kerak bo‘lgan mashg‘ulotlarning yanada qiziqarli va xilma-xil ro‘yxatini taklif etadi. Bu ijod uchun juda foydali bo‘lishi mumkin.

Agar haroratni juda yuqori qilib sozlasangiz, quyidagi kabi ma’nosiz natijalarni olishingiz mumkin: "Beksmit-Shteyn-Men yaqinida beysbol uy yugurish musobaqasini boshlang. Beach`.

Top P

Top P[2] - bu til modellarida ularning chiqishidagi tasodifiylikni boshqarishga yordam beradigan sozlama. U ehtimollik chegarasini o‘rnatish va keyin umumiy ehtimolligi bu chegaradan oshib ketadigan tokenlarni tanlash orqali ishlaydi.

Masalan, model keyingi holatni bashorat qiladigan misolni ko‘rib chiqaylik. `Mushuk ___` kodiga ko‘tarildi. Eng sara beshta so‘z `tree` (ehtimollik 0,5) bo‘lishi mumkin, `roof` (probability 0.25), `wall` (probability 0.15), `window` (probability .07) and `carpet`, with probability of .03.

Agar Top P qiymatini ’90’ deb belgilasak, AI faqat umumiy yig‘indisi kamida ~90% bo‘lgan tokenlarni ko‘rib chiqadi. Bizning holatda:

- `tree` qo‘shilmoqda -> hozircha jami ’50%’.
- Keyin `tom` -> jami qo‘shilganda ’75%’ hosil bo‘ladi.
- Keyin `devor` keladi va endi yig‘indimiz ’90%’ga yetadi.

Shunday qilib, natijani yaratish uchun sun’iy intellekt ushbu uchta variant (`daraxt`, `tom` va `devor`) dan birini tasodifiy tanlaydi, chunki ular

barcha ehtimolliklarning ~90 foizini tashkil etadi. Bu usul butun lug‘at zaxirasidan tanlab oladigan an’anaviy usullarga qaraganda ko‘proq xilma-xil natijalar berishi mumkin, chunki u individual tokenlarga emas, balki kumulyativ ehtimolliklarga asoslangan tanlovlarni qisqartiradi.

Maksimal uzunlik

Maksimal uzunlik - AI yaratishi mumkin bo‘lgan jami # ta token. Bu sozlama foydali, chunki u foydalanuvchilarga uzunligini boshqarish imkonini beradi modelning javobi, haddan tashqari uzun yoki ahamiyatsiz javoblarning oldini oladi. Uzunlik Playground maydonchasidagi `USER` va {" "} o‘rtasida ulashiladi `ASSISTANT` generated response. Notice how with a limit of 256 tokens, our PirateGPT from earlier is forced to cut its story short mid-sentence.

Bu, shuningdek, agar modeldan foydalanish uchun orqali to‘layotgan bo‘lsangiz, narxni nazorat qilishga yordam beradi API, Playground ishlatish o‘rniga.

Boshqa LLM sozlamalari

Til modeli chiqishiga ta’sir qilishi mumkin bo‘lgan ko‘plab boshqa sozlamalar mavjud, masalan, to‘xtash ketma-ketligi, chastota va mavjudlik jazolari.

Ketma-ketlikni to‘xtatish

To‘xtatish ketma-ketliklari modelga chiqish generatsiyasini qachon to‘xtatish kerakligini aytadi, bu esa kontent uzunligi va tuzilishini boshqarishingiz mumkin. Agar sun’iy intellektni so‘rayotgan bo‘lsangiz to‘xtash ketma-ketligi sifatida "Eng yaxshi tilaklar" yoki "Hurmat bilan" deb belgilagan holda xat yozing emailni saqlab qoluvchi yakuniy salomlashuvdan oldin modelning to‘xtashini ta’minlaydi qisqa va aniq. To‘xtatish ketma-ketliklari siz kutgan natija uchun foydali email, raqamlangan ro‘yxat yoki dialogue.

Chastota jarimasi

Chastota jarimasi - bu yaratilganda takrorlanishni oldini oluvchi sozlama tokenlar qanchalik tez-tez paydo bo‘lishiga mutanosib ravishda ularni jazolash orqali matn yaratish. O‘sha matnda token qanchalik ko‘p ishlatilsa, AI undan foydalanish ehtimoli shunchalik kam bo‘ladi. again.

Mavjudlik uchun jarima

Mavjudlik jarimasi chastota jarimasiga o‘xshaydi, lekin qat’iy tokenlar sodir bo‘lgan yoki bo‘lmaganiga qarab jazolaydi, o‘rniga proporsional ravishda

Determinizm eslatmasi

Harorat va Top-P to‘liq nolga sozlanganida ham, AI har safar bir xil aniq natija bermasligi mumkin. Sababi, GPU (grafik protsessor) dagi hisob-kitoblar sun’iy intellektning "miyasi"da amalga oshirilayotgani tasodifiylikdir.

Xulosa

Xulosa qilib aytganda, til modellari bilan ishlashda harorat, top p, maksimal uzunlik va boshqalar kabi sozlamalarni o‘zlashtirish juda muhimdir. Ushbu parametrlar ma’lum vazifalar yoki ilovalarni qondirish uchun modelning chiqishini aniq boshqarish imkonini beradi. Ular javoblardagi tasodifiylik, javob uzunligi va takrorlanish chastotasi kabi jihatlarni boshqaradi - bularning barchasi AI bilan o‘zaro aloqangizni yaxshilashga yordam beradi.

Qisman jekdikenlar382 va evintunadorlar tomonidan yozilgan

---

1. Yanada texnikroq so‘z - "konfiguratsiya giperparametrlari" ↩
2. Yadro namunasi sifatida ham tanilgan ↩