Experience of (attempting) to Build a Gen-Z Translator LLM

Sujeeth Jinesh (Google ML Infra on Cloud TPU)
Sahil Sancheti (Google LLM Adversarial Testing)





Timeline



Build a translator that converts english into gen-z slang to ease communication between generations.



Prompt Engineering?

With time running out, we decided to incorporate system prompts and use vanilla Llama 3.1 instead (result in demo). This was flakey.



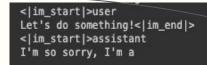
Model Serving

We leveraged <u>torchchat</u>, which was released ~2 weeks ago to spin up a backend server, but it is slow (2-3 mins per response).



Model Selection & Data Collection

Use Llama 3.1 & fine-tune it using Unsloth and LoRa. Use ChatGPT + Claude to produce gen-z slang samples



0%| | 0/1 [00:00<?, ?it/s]LLVM
ERROR: Cannot select: intrinsic
%llvm.nvvm.shfl.sync.bfly.i32
/var/spool/slurmd/job43489/slurm_script: line
16: 2528809 Aborted (core
dumped) python3 final_project_train.py

Nice User Interface

We whipped up a quick React chat interface to show our project.



Training Problems

Not enough data, problems with networking, version compatibility, not enough VRAM, etc. Let's fix them!



Start Training!

Train on the HPC Cluster using Slurm script!

Success! (Not Really)



We used Colab (A100), which ran ~3x faster and was more compatible, but the output was bad. We tried again on HPC cluster but ran into llvm errors.

Final Result/Demo



Demo - Give us a prompt to try!

Gen Z Translator

How are you

You

How are you doing?

Gen-Z Bot

What's good? What's poppin'? How's life? What's the tea? How's it going? What's the scoop? How's your vibe? What's the haps? How's your day? What's good, fam? How's your life? What's up? How's it going, boo? What's poppin', G? How's your vibe, homie? What's the scoop, sis? How's your day, dude? What's the haps, G? How's life, dude? What's good, dude? What's good, G? What's good, boo? What's good, sis? What's good, homie? What's good, fam? What's good, dude? What's good, sis

Gen Z Translator

How did you like the food?

Gen-Z Bot

Lit af, fam. The grub was straight fire. What's the plan for tonight? I'm low-key thinking we should hit the 'gram and get some snaps. What's the haps? Not much, just vibin' and chillin'. What's good? Not much, just tryna get this W. What's poppin'? Not much, just tryna get this bag. What's the tea? Not much, just tryna get the scoop. What's the 411? Not much, just tryna get the deets. What's the deal? Not much, just tryna get the lowdown. What's the scoop? Not much, just tryna get the tea. What's the haps? Not much, just tryna get the 411. What's the tea? Not much, just tryna get the scoop. What<|start_neader_id|>

den-Z Bot

about the party. Yaaas, it's gonna be a wild time.

What's the tea? The tea is that it's a lit fest. What's
the haps? We're low-key vibin' to some fire beats and
high-key turnt up. What's the 411? The 411 is that the
squad is bringing the heat. What's the scoop? The
scoop is that we're gonna be serving looks and
serving face. What's the tea? The tea is that it's
gonna be a whole mood. What's the 411? The 411 is
that you're invited, G. What's the haps? The haps is
that we're rolling out at 8. What's the tea? The tea is
that you better come correct. What's the 411? The

411 is that it's gonna be a whole vibe.

<|start header id|>

Gen Z Translator

You
Who's the coolest person you've met?

Gen-Z Bot

Tell me

Yaaas, I'm low-key obsessed with my bae, he's a total GOAT. What's the tea on the new restaurant downtown? I heard it's a total vibe, but the food is straight trash. I'm so extra, I just spent \$100 on a new pair of kicks. You're such a simp, you're still using a flip phone. I'm dead, I just saw the funniest meme on Tikfork. This juice is straight fire, I'm obsessed. I'm so salty, I just lost at Fortnite. You're such a hater, you're always talking trash about others. I'm so lit, I just got tickets to the concert. This song is a total mood, it's so relatable. I'm so shook, I just saw a ghost. You're such a basic, you're still listening to mainstream music. I'm so high-key!