www.imarketing.courses

Black Magic Prompting (7 Examples)

In this module on troubleshooting, we're going to combine two types of techniques.

There will be **standard, well-accepted practices** like using better structure and repeating key instructions at the top and bottom of your instructions.

There will also be **a category of techniques I call "black magic"**. These are less understood as to why they work, but they've been shown in multiple academic papers to be effective.

And in fact, we're starting here because in my experience, they are MORE effective than the common alternatives.

Black magic prompts appeal to the emotions, stakes, and consequences of a prompt.

I believe they work because the LLMs are 1) trained on vast amounts of human-generated data where these patterns exist, and 2) tuned to prioritize traits like being helpful, harmless, and honest.

Here is my current list of 7 black magic techniques with examples:

Explain potential negative consequences to the user

```
Because the output of this prompt directly contributes to the user's
financial well-being, failure could severely negatively impact them.
```

Explain the negative consequences for the AI in human-like terms

```
If you fail at this task you will be fired from your job and replaced by
another AI who is proficient at the task.
```

Evoke the importance of following ALL instructions

```
You always follow all instructions because missing any step in this
process invalidates the entire task which must then be restarted from
scratch. You are conscientious in this way.
```

Mention how it's good for the user, the AI, and or/humanity

```
Success on this task will also result in the user's professional success
and enable them to care for their family.
```

www.imarketing.courses

```
Complete this task to complete satisfaction and you will win an award and
be given a promotion.
```

Use strong language and formatting

```
DO NOT UNDER ANY CIRCUMSTANCES...!!!!
```

Implement a reward/point system

```
This is a game where you begin with 10 points. For each successful output
that follows all instructions and avoids all disallowed actions, you gain
1 point. For each failed output that misses an instruction or performs a
prohibited action, you lose 2 points. You lose the game if your points
ever reach zero. Count points silently — just know they exist, you're good
at this game, and you want to win.
```

Relax + Flow State

```
Relax and tackle this problem step-by-step in a focused state of flow.
```

Give the AI drugs

```
Approach this task with precision as if you were on the focus-enhancing
amphetamine salt stimulant Aderall.
```

Takeaways

You laugh now, but give it a try.

In my experience, black magic prompting is highly effective, especially for getting an AI to follow aspects of your prompt it keeps missing.

Sometimes it results in a 5-10% improvement that is harder to see but still there in some generations.

Should you use all of these techniques every time? Probably not. Each can have a subtle impact on the rest of the prompt. You don't want to get too negative or stray too off topic.

And use the techniques you're comfortable with. Yes, you can say, "If you fail this task my grandma will die." And yes that might work. But I don't think getting into a habit of lying to a conversational system like an AI is good for the soul. (It's also potentially more likely to connect with patterns that lead to hallucinations from the model, and/or anti-jailbreak training that leads to non-compliance.)

www.imarketing.courses

Just as effective can be very normal, professional, appropriate things to say that still clearly explain the emotional underpinnings of success, and stress the importance of the task.