Markus Lähde

LLM enhanced Google search with Python

Setup

I set up Ollama into Python virtual environment and tested it with few models. To my surprise models up to 4B size worked reasonably well on my laptop with no Nvidia GPU. I used tinyllama, stablelm2 and phi3 (released previous night of writing this!).

Idea

I thought I'd try creating a search engine iterator that generates alternative search queries automatically to probe a search engine with multiple slightly different queries with the same theme.

Execution

The search

For searches I used a googlesearch library which uses beautifulsoup under the hood for google results scraping.

```python
from googlesearch import search     Import "googlesearch" could not be resolved

query = 'House for sale in London'

for item in search(query, advanced=True, num_results=1):
    print(item.title)
    print(item.url)
    print(item.description)
    print('\n')
```

The LLM

Getting local LLM running with ollama is very easy. After launching 'ollama serve', queries can be made with just a few lines of python.

```python
import ollama     Import "ollama" could not be resolved
response = ollama.chat(model='stablelm2', messages=[
  {
    'role': 'user',
    'content': f'Answer with one sentence only. In that sentence should be alternative google search to this query: {query}.',
  },
])
print(response['message']['content'])
```

Fine tuning the interaction

Then comes arguably the hardest part, how to instruct the LLM so that it outputs something that can used in code without much parsing. Usually, problem is one of the following:

Too verbose (model: stablelm2):

```
Here's a Python code snippet that returns a list of alternative Google searches when you search for "House for sale in London":

```python
google_searches = [
    "London apartment for sale",
    "Property for sale in London listed here",
    "Houses and properties for sale in London",
```

Model has been trained to answer like a customer service person so it cannot help talking too much all the time.

Too helpful (model: stablelm2):

```
{
  "location": {
    "city": "London",
    "country": "United Kingdom"
  },
  "query": "House for sale in London",
  "results": [
    {
      "title": "For Sale - 4 Bedroom House in London - Right Move",
      "description": "A spacious four-bedroom house available to purchase in London. Features, amenities and more details can be found on the Right Move website.",
      "url": "https://www.rightmove.co.uk/houses-for-sale/london-30815742.html"
    },
```

Model tries to "fix the problem" on its own instead of just doing what have been requested. Worst part is that these pages it supposedly found by running its own code are **all hallusinations**.

Inconsistency (model: stablelm2):

Model might make a really nice answer:

```python
house_sales_in_london = [
    "Search houses for sale in 'London' on Google now!",
    "Check out the latest listings of houses available for sale in London on Zillow!",
    "Explore unique house designs and styles available for sale in London on Redfin.",
    "Discover a variety of beautiful, well-maintained houses for sale in London on Rightmove.",
    "Get expert advice on buying and selling houses in London from local real estate agents on Homestore."
]
```

The next time the same code is run it might answer something completely different (model: stablelm2):

```python
To fulfill this task, I will create a Python list containing two items:

1) The first item is a simple description of a house for sale in London.
2) The second item is an alternative Google search sentence that matches the user's input.

Here's what I'm going to do:
- Create a dictionary where each key-value pair represents a property detail (e.g., price, size).
- Populate the list with these dictionaries.

Let me code this for you:

```python
# Define the structure of our objects. For simplicity, we'll have just 'Address' and 'Price'.
properties = [
    {"address": "1234 Elm Street, London", "price": 100000},
    {"address": "5678 Pine Lane, London", "price": 150000},
]

# Create a list that includes the properties.
house_list = [property for property in properties]
```

(This is not the full answer, but in general instead of returning a python list it decided to write me some code.)

Solution

I was quite happy with my prompt:

```python
f'Return your answer in a python list format. Your answer should only contain the list. For
items in the list, make alternative google search sentences to this query: {query}.'
```

So instead of trying to make the model obey me by tweaking the prompt more, I decided first to try another model... and look at this! Microsoft's tiny 3.8B model does exactly what I ask for and it does it 5 times in a row!

```
markuslahde@Elite:~/ENVIRONMENTS/py3.11env/CHAT$ ./chat.py
["Searching for available residential properties on the market in London", "Looking up houses currently listed for
purchase in London area", "Exploring options for buying a home in London city", "Investigating real estate listings
of homes for sale in London"]
markuslahde@Elite:~/ENVIRONMENTS/py3.11env/CHAT$ ./chat.py
["Searching alternatives for 'House for sale in London'", "Exploring options for buying property in London", "Looki
ng up real estate listings of houses available in London"]
markuslahde@Elite:~/ENVIRONMENTS/py3.11env/CHAT$ ./chat.py
["Searching alternatives for 'Houses available for purchase in London'", "Exploring other phrases like 'Buying resi
dential property in London', 'Real estate options on sale in London', 'London homes up for grabs']
markuslahde@Elite:~/ENVIRONMENTS/py3.11env/CHAT$ ./chat.py
["Search for real estate options available on the market in London", "Looking up properties listed for purchase in
London", "Exploring current listings of houses for sale in London area"]
markuslahde@Elite:~/ENVIRONMENTS/py3.11env/CHAT$ ./chat.py
["Searching for available properties for sale in London's real estate market", "Looking up current listings of hous
es on sale within London city limits", "Finding residential properties for purchase located in London area", "Explo
ring options for buying a home sold in the capital of England, London"]
```

Final product

For the final version I added user input and printed the results separately. For added user experience, processes could be multithreaded as all the results are now displayed only after the LLM-inference is ready.

Overall, I'm really happy how well this works!

```
Enter search query: House for sale in London

Search results:
Properties For Sale in Greater London
https://www.rightmove.co.uk/property-for-sale/London.html
Flats & Houses For Sale in London - Find properties with Rightmove - the UK's largest selection of properties.


Property for sale in London
https://www.zoopla.co.uk/for-sale/property/london/
Find houses and flats for sale in London with the UK's largest data-driven property portal. Discover properties for sale from the top estate agents and ...


Houses for sale in London
https://www.onthemarket.com/for-sale/property/london/
Find the latest properties available for sale in London with the UK's most user-friendly property portal. Search houses & flats to buy from leading estate ...


Alternative search results:
Properties for sale in London
https://www.foxtons.co.uk/properties-for-sale/london
Use our search function to find houses, apartments or studio flats in your desired location. You can narrow down your chosen properties in London even further ...


Property for sale in London | Savills
https://search.savills.com/list/property-for-sale/england/london/london/ec3n
2124 Properties for sale in London ; Kent Road · East Molesey, Surrey, KT8 9JZ. 1,487 sq ft(138.15 sq m) ; Eaton Place · London, SW1X 8AT. 10,704 sq ft(994.43 sq m).


Properties For Sale in Greater London
https://www.rightmove.co.uk/property-for-sale/London.html
Flats & Houses For Sale in London - Find properties with Rightmove - the UK's largest selection of properties.
```

```
London Local Guide to Buying a House
https://www.purplebricks.co.uk/moving-guides/london-local-house-guide
Looking for a house for sale in London to suit your budget? Here's how to view the right property, flat and apartment for sale in London and sell your ...


This is where to buy in London in 2023, according to a top ...
https://www.black-brick.com/insights/in-the-press/this-is-where-to-buy-in-london-in-2023-according-to-a-top-buying-agency/
9 Jan 2023 — North Kensington, Herne Hill, Mayfair, St John's Wood and Acton have been picked out as the top places to buy in London this year.


New Builds For Sale in London | New Homes
https://www.foxtons.co.uk/buy/buy_newhomes
Leading estate agents dealing with New Homes and new builds in London and the Home Counties. See our new homes on our website.


New Builds For Sale in London | New Homes
https://www.foxtons.co.uk/buy/buy_newhomes
Leading estate agents dealing with New Homes and new builds in London and the Home Counties. See our new homes on our website.


London's Leading Estate Agent Offers New Homes for Sale
https://www.foxtons.co.uk/newhomes
Looking for a New Home for sale? London's leading estate agency offers new houses with remarkable facilities.


Find New Homes & Developments For Sale in City Of London
https://www.rightmove.co.uk/new-homes-for-sale/City-Of-London.html
Find New Builds & Developments For Sale in City Of London with Rightmove. New build flats & houses for sale in City Of London.


Property for sale in London | Savills
https://search.savills.com/list/property-for-sale/england/london/london/ec3n
Discover our selection of properties and houses for sale in London. With over 32 boroughs and a diverse range of cultures, England's capital city is a hive ...


London Property Search | SiteSales
https://site-sales.co.uk/property-search
Search our extensive range of developments and find your next home.
```