Synthesize

October 3, 2024

```
[3]: %pip install -U langchain-ollama
```

```
[1]: import os
     from   langchain_ollama      import ChatOllama
     from   langchain_core.prompts import ChatPromptTemplate
```

```
[ ]:
```

```
[7]: # ENVIRONMENT
     # host = os.getenv('OLLAMA_HOST')
     # port = os.getenv('OLLAMA_PORT')

     host = "212.189.145.27"
     port = 11435

     # MODEL
     llm = ChatOllama(model="llama3.1", base_url="http://{ip}:{port}".
       ↪format(ip=host, port=port), temperature=0)
```

```
[8]: # PROMPT DEFINITION
     system = """
     You are an expert in text synthesize.
     You need to maintain the right contexts related to each sentence.
     In addition, you have to return a short text starting from a list of sentences.
     You do not have to use bulleted lists.
     The returned text has to be very clear.
     """

     prompt = ChatPromptTemplate.from_messages(
         [
             ("system", system),
             ("human" , "{sentences}"),
         ]
     )
```

```
[9]: # SENTENCES
     sentences = [
```

```
    "Basigin (BSG, CD147) is a multifunctional protein involved in cancer cell␣
↪survival, mostly by controlling lactate transport through its interaction␣
↪with monocarboxylate transporters (MCTs) such as MCT1.",
    "CD147 (basigin, BSG) is a membrane-bound glycoprotein involved in energy␣
↪metabolism that plays a role in cancer cell survival. Its soluble form is a␣
↪promising marker of some diseases, but it is otherwise poorly studied.",
    "Basigin (BSG) is an essential factor for the infection and progression of␣
↪COVID-19 and tumorigenesis of multiple tumors, which may serve as a novel␣
↪target for the effective treatment against COVID-19 and multiple human␣
↪cancers.",
    "SARS-CoV-2 host receptors ACE2 and CD147 are present on the membrane of␣
↪trophectoderm, epiblast and hypoblast cells in human blastocysts. CD147 is␣
↪also present on the oolemma."
]
```

[10]:
```python
# CHAIN
chain = prompt | llm
```

[11]:
```python
chain.invoke({"sentences": sentences})
```

[11]: AIMessage(content='Basigin (BSG), also known as CD147, plays a crucial role in
cancer cell survival by regulating lactate transport through its interaction
with monocarboxylate transporters like MCT1. This multifunctional protein is
involved in energy metabolism and has been implicated in the progression of
COVID-19 and various human cancers. Interestingly, CD147 is also present on the
membrane of trophectoderm, epiblast, and hypoblast cells in human blastocysts,
suggesting its potential role in early embryonic development. Furthermore, the
soluble form of CD147 has been identified as a promising marker for certain
diseases, highlighting its importance in disease diagnosis and treatment.',
response_metadata={'model': 'llama3.1', 'created_at':
'2024-10-03T12:53:31.673202042Z', 'message': {'role': 'assistant', 'content':
''}, 'done_reason': 'stop', 'done': True, 'total_duration': 66308818851,
'load_duration': 5480931550, 'prompt_eval_count': 271, 'prompt_eval_duration':
19506371000, 'eval_count': 138, 'eval_duration': 41266731000},
id='run-b6a028c1-52b0-4afb-8580-12c6afd3fe6a-0', usage_metadata={'input_tokens':
271, 'output_tokens': 138, 'total_tokens': 409})

[ ]: