# LangChain: Evaluation

## Outline:

- Example generation
- Manual evaluation (and debuging)
- LLM-assisted evaluation

```
In [25]: import os

from dotenv import load_dotenv, find_dotenv
_ = load_dotenv(find_dotenv()) # read local .env file
```

## Create our QandA application

```
In [26]: from langchain.chains import RetrievalQA
from langchain.chat_models import ChatOpenAI
from langchain.document_loaders import CSVLoader
from langchain.indexes import VectorstoreIndexCreator
from langchain.vectorstores import DocArrayInMemorySearch
```

```
In [27]: file = 'OutdoorClothingCatalog_1000.csv'
loader = CSVLoader(file_path=file)
data = loader.load()
```

```
In [28]: index = VectorstoreIndexCreator(
    vectorstore_cls=DocArrayInMemorySearch
).from_loaders([loader])
```

```
In [29]: llm = ChatOpenAI(temperature = 0.0)
qa = RetrievalQA.from_chain_type(
    llm=llm,
    chain_type="stuff",
    retriever=index.vectorstore.as_retriever(),
    verbose=True,
    chain_type_kwargs = {
        "document_separator": "<<<<>>>>>"
    }
)
```

## Coming up with test datapoints

```
In [30]: data[10]
```

Document(page_content=": 10\nname: Cozy Comfort Pullover Set, Stripe\ndescription: Perfect for lounging, this striped knit set lives up to its name. We used ultrasoft fabric and an easy design that's as comfortable at bedtime as it is when we have to make a quick run out.\n\nSize & Fit\n- Pants are Favorite Fit: Sits lower on the waist.\n- Relaxed Fit: Our most generous fit sits farthest from the body.\nFabric & Care\n- In the softest blend of 63% polyester, 35% rayon and 2% spandex.\n\nAdditional Features\n- Relaxed fit top with raglan sleeves and rounded hem.\n- Pull-on pants have a wide elastic waistband and drawstring, side pockets and a modern slim leg.\nImported.", metadata={'source': 'OutdoorClothingCatalog_1000.csv', 'row': 10})

```
In [31]: data[11]
```

Document(page_content=': 11\nname: Ultra-Lofty 850 Stretch Down Hooded Jacket\ndescription: This technical stretch down jacket from our DownTek collection is sure to keep you warm and comfortable with its full-stretch construction providing exceptional range of motion. With a slightly fitted style that falls at the hip and best with a midweight layer, this jacket is suitable for light activity up to 20° and moderate activity up to -30°. The soft and durable 100% polyester shell offers complete windproof protection and is insulated with warm, lofty goose down. Other features include welded baffles for a no-stitch construction and excellent stretch, an adjustable hood, an interior media port and mesh stash pocket and a hem drawcord. Machine wash and dry. Imported.', metadata={'source': 'OutdoorClothingCatalog_1000.csv', 'row': 11})

## Hard-coded examples

```
In [32]: examples = [
    {
        "query": "Do the Cozy Comfort Pullover Set\
        have side pockets?",
        "answer": "Yes"
    },
    {

        "query": "What collection is the Ultra-Lofty \
        850 Stretch Down Hooded Jacket from?",
        "answer": "The DownTek collection"
    }
]
```

## LLM-Generated examples

```
In [33]: from langchain.evaluation.qa import QAGenerateChain
```

```
In [34]:  example_gen_chain = QAGenerateChain.from_llm(ChatOpenAI())
```

```
In [35]:  new_examples = example_gen_chain.apply_and_parse(
              [{"doc": t} for t in data[:5]]
          )
```

```
In [36]:  new_examples[0]
```

```
{'query': "What is the weight of one pair of Women's Campside Oxfords accordi
ng to the document?",
 'answer': "The approximate weight of one pair of Women's Campside Oxfords is
1 lb. 1 oz. per pair."}
```

```
In [37]:  data[0]
```

```
Document(page_content=": 0\nname: Women's Campside Oxfords\ndescription: This
ultracomfortable lace-to-toe Oxford boasts a super-soft canvas, thick cushion
ing, and quality construction for a broken-in feel from the first time you pu
t them on. \n\nSize & Fit: Order regular shoe size. For half sizes not offere
d, order up to next whole size. \n\nSpecs: Approx. weight: 1 lb.1 oz. per pai
r. \n\nConstruction: Soft canvas material for a broken-in feel and look. Comf
ortable EVA innersole with Cleansport NXT® antimicrobial odor control. Vintag
e hunt, fish and camping motif on innersole. Moderate arch contour of innerso
le. EVA foam midsole for cushioning and support. Chain-tread-inspired molded
rubber outsole with modified chain-tread pattern. Imported. \n\nQuestions? Pl
ease contact us for any inquiries.", metadata={'source': 'OutdoorClothingCata
log_1000.csv', 'row': 0})
```

## Combine examples

```
In [38]:  examples += new_examples
```

```
In [39]:  qa.run(examples[0]["query"])
```

```
> Entering new RetrievalQA chain...

> Finished chain.
```

```
'The Cozy Comfort Pullover Set, Stripe has side pockets on the pull-on pant
s.'
```

## Manual Evaluation

```
In [40]:  import langchain
          langchain.debug = True
```

```
In [41]: qa.run(examples[0]["query"])
```

```
[chain/start] [1:chain:RetrievalQA] Entering Chain run with input:
{
  "query": "Do the Cozy Comfort Pullover Set        have side pockets?"
}
[chain/start] [1:chain:RetrievalQA > 2:chain:StuffDocumentsChain] Entering Ch
ain run with input:
[inputs]
[chain/start] [1:chain:RetrievalQA > 2:chain:StuffDocumentsChain > 3:chain:LL
MChain] Entering Chain run with input:
{
  "question": "Do the Cozy Comfort Pullover Set        have side pockets?",
  "context": ": 10\nname: Cozy Comfort Pullover Set, Stripe\ndescription: Per
fect for lounging, this striped knit set lives up to its name. We used ultras
oft fabric and an easy design that's as comfortable at bedtime as it is when
we have to make a quick run out.\n\nSize & Fit\n- Pants are Favorite Fit: Sit
s lower on the waist.\n- Relaxed Fit: Our most generous fit sits farthest fro
m the body.\n\nFabric & Care\n- In the softest blend of 63% polyester, 35% ra
yon and 2% spandex.\n\nAdditional Features\n- Relaxed fit top with raglan sle
eves and rounded hem.\n- Pull-on pants have a wide elastic waistband and draw
string, side pockets and a modern slim leg.\nImported.<<<<>>>>: 73\nname:
Cozy Cuddles Knit Pullover Set\ndescription: Perfect for lounging, this knit
set lives up to its name. We used ultrasoft fabric and an easy design that's
as comfortable at bedtime as it is when we have to make a quick run out. \n\n
Size & Fit \nPants are Favorite Fit: Sits lower on the waist. \nRelaxed Fit:
Our most generous fit sits farthest from the body. \n\nFabric & Care \nIn the
softest blend of 63% polyester, 35% rayon and 2% spandex.\n\nAdditional Featu
res \nRelaxed fit top with raglan sleeves and rounded hem. \nPull-on pants ha
ve a wide elastic waistband and drawstring, side pockets and a modern slim le
g. \nImported.<<<<>>>>: 632\nname: Cozy Comfort Fleece Pullover\ndescriptio
n: The ultimate sweater fleece \u2013 made from superior fabric and offered a
t an unbeatable price. \n\nSize & Fit\nSlightly Fitted: Softly shapes the bod
y. Falls at hip. \n\nWhy We Love It\nOur customers (and employees) love the r
ugged construction and heritage-inspired styling of our popular Sweater Fleec
e Pullover and wear it for absolutely everything. From high-intensity activit
ies to everyday tasks, you'll find yourself reaching for it every time.\n\nFa
bric & Care\nRugged sweater-knit exterior and soft brushed interior for excep
tional warmth and comfort. Made from soft, 100% polyester. Machine wash and d
ry.\n\nAdditional Features\nFeatures our classic Mount Katahdin logo. Snap pl
acket. Front princess seams create a feminine shape. Kangaroo handwarmer pock
ets. Cuffs and hem reinforced with jersey binding. Imported.\n\n \u2013 Offic
ial Supplier to the U.S. Ski Team\nTHEIR WILL TO WIN, WOVEN RIGHT IN. LEARN M
ORE<<<<>>>>: 151\nname: Cozy Quilted Sweatshirt\ndescription: Our sweatshirt
is an instant classic with its great quilted texture and versatile weight tha
t easily transitions between seasons. With a traditional fit that is relaxed
through the chest, sleeve, and waist, this pullover is lightweight enough to
be worn most months of the year. The cotton blend fabric is super soft and co
mfortable, making it the perfect casual layer. To make dressing easy, this sw
eatshirt also features a snap placket and a heritage-inspired Mt. Katahdin lo
go patch. For care, machine wash and dry. Imported."
}
[llm/start] [1:chain:RetrievalQA > 2:chain:StuffDocumentsChain > 3:chain:LLMC
hain > 4:llm:ChatOpenAI] Entering LLM run with input:
{
  "prompts": [
    "System: Use the following pieces of context to answer the users questio
n. \nIf you don't know the answer, just say that you don't know, don't try to
make up an answer.\n---------------\n: 10\nname: Cozy Comfort Pullover Set,
```

Stripe\ndescription: Perfect for lounging, this striped knit set lives up to its name. We used ultrasoft fabric and an easy design that's as comfortable at bedtime as it is when we have to make a quick run out.\n\nSize & Fit\n- Pants are Favorite Fit: Sits lower on the waist.\n- Relaxed Fit: Our most generous fit sits farthest from the body.\n\nFabric & Care\n- In the softest blend of 63% polyester, 35% rayon and 2% spandex.\n\nAdditional Features\n- Relaxed fit top with raglan sleeves and rounded hem.\n- Pull-on pants have a wide elastic waistband and drawstring, side pockets and a modern slim leg.\n\nImported.<<<<>>>>: 73\nname: Cozy Cuddles Knit Pullover Set\ndescription: Perfect for lounging, this knit set lives up to its name. We used ultrasoft fabric and an easy design that's as comfortable at bedtime as it is when we have to make a quick run out. \n\nSize & Fit \nPants are Favorite Fit: Sits lower on the waist. \nRelaxed Fit: Our most generous fit sits farthest from the body. \n\nFabric & Care \nIn the softest blend of 63% polyester, 35% rayon and 2% spandex.\n\nAdditional Features \nRelaxed fit top with raglan sleeves and rounded hem. \nPull-on pants have a wide elastic waistband and drawstring, side pockets and a modern slim leg. \nImported.<<<<>>>>: 632\nname: Cozy Comfort Fleece Pullover\ndescription: The ultimate sweater fleece \u2013 made from superior fabric and offered at an unbeatable price. \n\nSize & Fit\nSlightly Fitted: Softly shapes the body. Falls at hip. \n\nWhy We Love It\nOur customers (and employees) love the rugged construction and heritage-inspired styling of our popular Sweater Fleece Pullover and wear it for absolutely everything. From high-intensity activities to everyday tasks, you'll find yourself reaching for it every time.\n\nFabric & Care\nRugged sweater-knit exterior and soft brushed interior for exceptional warmth and comfort. Made from soft, 100% polyester. Machine wash and dry.\n\nAdditional Features\nFeatures our classic Mount Katahdin logo. Snap placket. Front princess seams create a feminine shape. Kangaroo handwarmer pockets. Cuffs and hem reinforced with jersey binding. Imported.\n\n \u2013 Official Supplier to the U.S. Ski Team\nTHEIR WILL TO WIN, WOVEN RIGHT IN. LEARN MORE<<<<>>>>: 151\nname: Cozy Quilted Sweatshirt\ndescription: Our sweatshirt is an instant classic with its great quilted texture and versatile weight that easily transitions between seasons. With a traditional fit that is relaxed through the chest, sleeve, and waist, this pullover is lightweight enough to be worn most months of the year. The cotton blend fabric is super soft and comfortable, making it the perfect casual layer. To make dressing easy, this sweatshirt also features a snap placket and a heritage-inspired Mt. Katahdin logo patch. For care, machine wash and dry. Imported.\nHuman: Do the Cozy Comfort Pullover Set        have side pockets?"
    ]
}

[llm/end] [1:chain:RetrievalQA > 2:chain:StuffDocumentsChain > 3:chain:LLMChain > 4:llm:ChatOpenAI] [896.399ms] Exiting LLM run with output:
{
  "generations": [
    [
      {
        "text": "The Cozy Comfort Pullover Set, Stripe has side pockets on the pull-on pants.",
        "generation_info": null,
        "message": {
          "content": "The Cozy Comfort Pullover Set, Stripe has side pockets on the pull-on pants.",
          "additional_kwargs": {},
          "example": false
        }
      }
    ]

```
      ],
      "llm_output": {
        "token_usage": {
          "prompt_tokens": 734,
          "completion_tokens": 18,
          "total_tokens": 752
        },
        "model_name": "gpt-3.5-turbo"
      }
    }
```

[chain/end] [1:chain:RetrievalQA > 2:chain:StuffDocumentsChain > 3:chain:LLMC
hain] [897.2620000000001ms] Exiting Chain run with output:
```
{
  "text": "The Cozy Comfort Pullover Set, Stripe has side pockets on the pull
-on pants."
}
```
[chain/end] [1:chain:RetrievalQA > 2:chain:StuffDocumentsChain] [897.952ms] E
xiting Chain run with output:
```
{
  "output_text": "The Cozy Comfort Pullover Set, Stripe has side pockets on t
he pull-on pants."
}
```
[chain/end] [1:chain:RetrievalQA] [1.07s] Exiting Chain run with output:
```
{
  "result": "The Cozy Comfort Pullover Set, Stripe has side pockets on the pu
ll-on pants."
}
```

'The Cozy Comfort Pullover Set, Stripe has side pockets on the pull-on pant
s.'

In [42]:
```python
# Turn off the debug mode
langchain.debug = False
```

# LLM assisted evaluation

```
In [43]: predictions = qa.apply(examples)
```

> Entering new RetrievalQA chain...

> Finished chain.

> Entering new RetrievalQA chain...

> Finished chain.

> Entering new RetrievalQA chain...

> Finished chain.

> Entering new RetrievalQA chain...

> Finished chain.

> Entering new RetrievalQA chain...

> Finished chain.

> Entering new RetrievalQA chain...

> Finished chain.

> Entering new RetrievalQA chain...

> Finished chain.

```
In [44]: from langchain.evaluation.qa import QAEvalChain
```

```
In [45]: llm = ChatOpenAI(temperature=0)
         eval_chain = QAEvalChain.from_llm(llm)
```

```
In [46]: graded_outputs = eval_chain.evaluate(examples, predictions)
```

```
In [47]: for i, eg in enumerate(examples):
    print(f"Example {i}:")
    print("Question: " + predictions[i]['query'])
    print("Real Answer: " + predictions[i]['answer'])
    print("Predicted Answer: " + predictions[i]['result'])
    print("Predicted Grade: " + graded_outputs[i]['text'])
    print()
```

Example 0:
Question: Do the Cozy Comfort Pullover Set      have side pockets?
Real Answer: Yes
Predicted Answer: The Cozy Comfort Pullover Set, Stripe has side pockets on the pull-on pants.
Predicted Grade: CORRECT

Example 1:
Question: What collection is the Ultra-Lofty      850 Stretch Down Hooded Jacket from?
Real Answer: The DownTek collection
Predicted Answer: The Ultra-Lofty 850 Stretch Down Hooded Jacket is from the DownTek collection.
Predicted Grade: CORRECT

Example 2:
Question: What is the weight of one pair of Women's Campside Oxfords according to the document?
Real Answer: The approximate weight of one pair of Women's Campside Oxfords is 1 lb. 1 oz. per pair.
Predicted Answer: The weight of one pair of Women's Campside Oxfords is approximately 1 lb. 1 oz. according to the document.
Predicted Grade: CORRECT

Example 3:
Question: What are the dimensions of the small and medium Recycled Waterhog dog mats?
Real Answer: The small Recycled Waterhog dog mat has dimensions of 18" x 28" and the medium has dimensions of 22.5" x 34.5".
Predicted Answer: The small Recycled Waterhog dog mat has dimensions of 18" x 28" and the medium size has dimensions of 22.5" x 34.5".
Predicted Grade: CORRECT

Example 4:
Question: What are some features of the Infant and Toddler Girls' Coastal Chill Swimsuit, Two-Piece?
Real Answer: The Infant and Toddler Girls' Coastal Chill Swimsuit, Two-Piece features bright colors, ruffles, and exclusive whimsical prints. It is made of four-way-stretch and chlorine-resistant fabric that keeps its shape and resists snags. The swimsuit has a UPF 50+ rated fabric that provides the highest rated sun protection possible, blocking 98% of the sun's harmful rays. The crossover no-slip straps and fully lined bottom ensure a secure fit and maximum coverage. It can be machine washed and line dried for best results, and is imported.
Predicted Answer: The Infant and Toddler Girls' Coastal Chill Swimsuit, Two-Piece features bright colors, ruffles, and exclusive whimsical prints. The four-way-stretch and chlorine-resistant fabric keeps its shape and resists snags. The UPF 50+ rated fabric provides the highest rated sun protection possible, blocking 98% of the sun's harmful rays. The crossover no-slip straps and fully lined bottom ensure a secure fit and maximum coverage. It is machine washable and should be line dried for best results.
Predicted Grade: CORRECT

Example 5:
Question: What is the fabric composition of the Refresh Swimwear, V-Neck Tankini Contrasts?
Real Answer: The Refresh Swimwear, V-Neck Tankini Contrasts is made of 82% re

cycled nylon with 18% Lycra® spandex for the body and 90% recycled nylon with
10% Lycra® spandex for the lining.
Predicted Answer: The Refresh Swimwear, V-Neck Tankini Contrasts is made of 8
2% recycled nylon with 18% Lycra® spandex for the body and 90% recycled nylon
with 10% Lycra® spandex for the lining.
Predicted Grade: CORRECT

Example 6:
Question: What is the technology used in EcoFlex 3L Storm Pants and what is i
ts main benefit?
Real Answer: The technology used in EcoFlex 3L Storm Pants is TEK O2 technolo
gy which offers the most breathability ever tested. Its main benefit is keepi
ng the wearer dry and comfortable during any outdoor activity and in any weat
her.
Predicted Answer: The EcoFlex 3L Storm Pants use TEK O2 technology, which is
designed to make the pants more breathable. The main benefit of this technolo
gy is that it keeps you dry and comfortable, no matter the activity or weathe
r.
Predicted Grade: CORRECT

In [ ]:

In [ ]:

In [ ]:

In [ ]:

In [ ]:

In [ ]:

In [ ]:

In [ ]:

In [ ]:

In [ ]: