EVALUATE YOUR DOCUMENT RETRIEVAL CHAINS & AGENTS

LANGCHAIN DATASETS, QAGENRATION AND QAEVALUATION CHAIN

WHAT KIND OF EVALUATION

RETRIEVAL CHAINS WITH DOCUMENTS STORES HAVE TO BE ACCURATE

1) HOW MANY QUESTIONS THE AGENT ANSWERS CORRECTLY

2) HOW MANY QUESTIONS THE RETRIEVER GETS CORRECT DOCUMENTS

3) WHICH PART OF THE TEXT CREATES AMBIGUOUS QUESTIONS & ANSWERS

4) WHY SOME ANSWERS ARE ALWAYS WRONG

5) WHY DIFFERENT QUESTIONS RETURNS SAME ANSWERS

6) WITH DIFFFERENT LLMS THE EITHER RESULTS DIFFER OR ERRORS OUT

SOLUTION OFFERED :
1) TRACING SERVER
2) LANGCHAIN DATASETS
3) EVAL CHAINS

HTTPS://GITHUB.COM/INSIGHTBUILDER

INTRODUCING DATASETS

LANGCHAIN PROVIDES DATASETS & CLASSES TO CREATE DATASETS

1) 5 DATASETS CURATED BY LANGCHAIN TEAM IN HUGGINGFACE HUB

2) ADDITIONAL DATASETS CAN BE CREATED YOURSELF

3) DATASET WILL HAVE THE BELOW DATA STRUCTURE

[{'QUESTION': 'WHAT IS THE VIM PLAY BOOK?',

'ANSWER': 'THE VIM PLAY BOOK IS A FEATURE INSIDE THE SPACEVIM FLAVOR.'}]

EVAL CHAINS

LANGCHAIN PROVIDES 3 DIFFERENT EVAL CLASSES

1) QA GENERATION: QAGENERATIONCHAIN

2) QA EVALUATION : QAEVALCHAIN

3) MULTIPLE TEMPLATES TO WORK WITH AGENTS / TOOLKITS AND CHAINS

HOW EVALUATION WORKS

1) PIP -Q INSTALL TRANSFORMERS DATASETS

2) IMPORT LOAD_DATASET

DATASET = LOAD_DATASET("QUESTION-ANSWERING-

STATE-OF-THE-UNION")

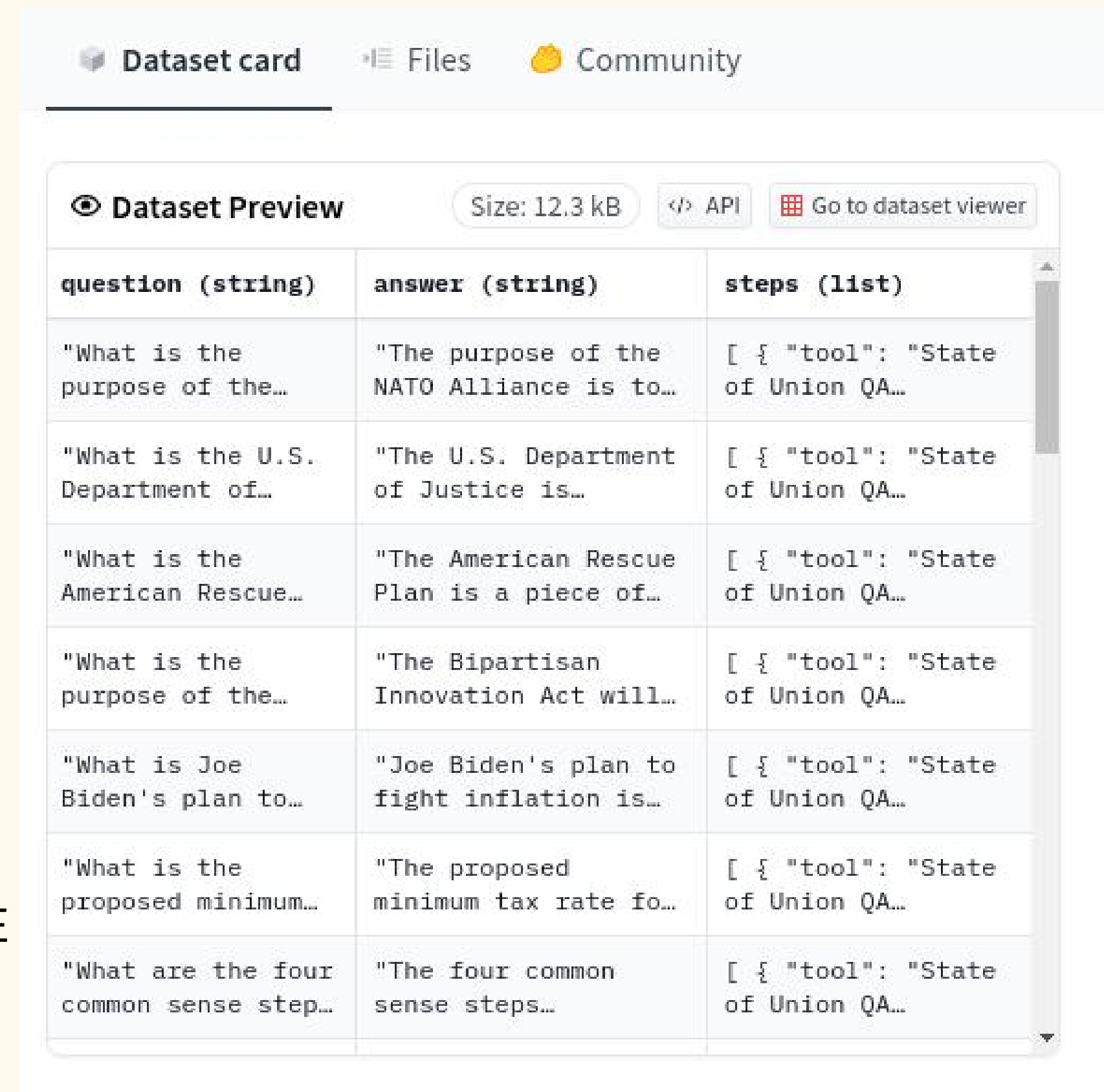HTTPS://HUGGINGFACE.CO/LANGCHAINDATASETS

3) DATASET IS ARRAY OF QUESTION & ANSWERS

4) DATASET IS EVALUATED USING QA EVAL CHAIN

5) DATASET IS CREATED USING QA GENERATION

CHAIN

6) BOTH THE ABOVE CHAINS USE LLMS TO GENERATE

AND EVALUATE THE QUESTIONS / ANSWERS

7) CAN INCLUDE TRACING TOO

HTTPS://GITHUB.COM/INSIGHTBUILDER

Dataset card · Files · Community

Dataset Preview — Size: 12.3 kB — API — Go to dataset viewer

question (string)	answer (string)	steps (list)
"What is the purpose of the…	"The purpose of the NATO Alliance is to…	[{ "tool": "State of Union QA…
"What is the U.S. Department of…	"The U.S. Department of Justice is…	[{ "tool": "State of Union QA…
"What is the American Rescue…	"The American Rescue Plan is a piece of…	[{ "tool": "State of Union QA…
"What is the purpose of the…	"The Bipartisan Innovation Act will…	[{ "tool": "State of Union QA…
"What is Joe Biden's plan to…	"Joe Biden's plan to fight inflation is…	[{ "tool": "State of Union QA…
"What is the proposed minimum…	"The proposed minimum tax rate fo…	[{ "tool": "State of Union QA…
"What are the four common sense step…	"The four common sense steps…	[{ "tool": "State of Union QA…

LETS HEAD TO COLAB NOTEBOOK

DATASTRUCTURES ARE BEST UNDERSTOOD VISUALLY

THE QAEVALUATION +

QAGENERATION WILL USE

TOKENS. REVIEW THE

DATASET SIZES BEFORE

EXECUTING

insightbuilder/
python_de_learners_data

Repo contains the code, data and supporting
documents including presentations, playbooks and
additional documents to support learning

👥 2   ⊙ 0   ☆ 25   ⑂ 10

Contributors   Issues   Stars   Forks

**python_de_learners_data/langchainDatasets_EvalChains.ipynb at main ·
insightbuilder/python_de_learners_data**

Repo contains the code, data and supporting documents including presentations, playbooks and additional
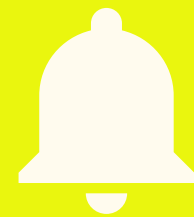documents to support learning - python_de_learners_data/langchainDatasets_EvalChains.ipynb at...

GitHub

THANKS FOR WATCHING

LIKE

SHARE

SUBSCRIBE