

▼ Automatic Text Summarization

- Given that time is scarce, one of the main goals when dealing with large volumes of information is to increase the efficiency with which it can be processed.
- This paper will look at five of the major Automatic Text Summarization Techniques (ATS) methods used for processing linguistic textual data to determine which method yields the results for the DoQA dataset.
- An assumption will be made that given the nature of this dataset; the best result will be yielded through the use of a Query-Based approach.
- Methods will be compared numerically based on the ROGUE metric and F1 score amongst others.
- Furthermore, the authors present the Cooking dataset in three separate forms consisting of training, development and testing formats an attempt will be made to see if there is any statistical significance distinguishing these three possible sets of the same subject domain.

▼ Import Necessary Libraries

```
import warnings
warnings.filterwarnings("ignore")

pip install transformers

Requirement already satisfied: transformers in /usr/local/lib/python3.10/dist-packages (4.35.2)
Requirement already satisfied: filelock in /usr/local/lib/python3.10/dist-packages (from transformers) (3.13.1)
Requirement already satisfied: huggingface-hub<1.0,>=0.16.4 in /usr/local/lib/python3.10/dist-packages (from tra
Requirement already satisfied: numpy>=1.17 in /usr/local/lib/python3.10/dist-packages (from transformers) (1.23.
Requirement already satisfied: packaging>=20.0 in /usr/local/lib/python3.10/dist-packages (from transformers) (2
Requirement already satisfied: pyyaml>=5.1 in /usr/local/lib/python3.10/dist-packages (from transformers) (6.0.1
Requirement already satisfied: regex!=2019.12.17 in /usr/local/lib/python3.10/dist-packages (from transformers)
Requirement already satisfied: requests in /usr/local/lib/python3.10/dist-packages (from transformers) (2.31.0)
Requirement already satisfied: tokenizers<0.19,>=0.14 in /usr/local/lib/python3.10/dist-packages (from transform
Requirement already satisfied: safetensors>=0.3.1 in /usr/local/lib/python3.10/dist-packages (from transformers)
Requirement already satisfied: tqdm>=4.27 in /usr/local/lib/python3.10/dist-packages (from transformers) (4.66.1
Requirement already satisfied: fsspec>=2023.5.0 in /usr/local/lib/python3.10/dist-packages (from huggingface-hub
Requirement already satisfied: typing-extensions>=3.7.4.3 in /usr/local/lib/python3.10/dist-packages (from huggi
Requirement already satisfied: charset-normalizer<4,>=2 in /usr/local/lib/python3.10/dist-packages (from request
Requirement already satisfied: idna<4,>=2.5 in /usr/local/lib/python3.10/dist-packages (from requests->transform
Requirement already satisfied: urllib3<3,>=1.21.1 in /usr/local/lib/python3.10/dist-packages (from requests->tra
Requirement already satisfied: certifi>=2017.4.17 in /usr/local/lib/python3.10/dist-packages (from requests->tra
```

```
import transformers
```

```
import pandas as pd
import numpy as np
from numpy import hstack
import matplotlib.pyplot as plt
#from keras.utils.np_utils import to_categorical
from keras.preprocessing.text import Tokenizer
from keras.preprocessing import sequence
#from keras.preprocessing.sequence import pad_sequences # from keras_preprocessing.sequence import pad_sequences
#from keras_preprocessing.sequence import pad_sequences
from keras.models import Sequential
from keras.layers import Dense, LSTM, GRU, Embedding
from keras import optimizers
from sklearn.metrics import r2_score, mean_absolute_error
from sklearn.ensemble import RandomForestClassifier
from sklearn.ensemble import GradientBoostingRegressor, RandomForestRegressor
from sklearn.model_selection import train_test_split
from sklearn.metrics import r2_score
from fractions import Fraction
import seaborn as sns
import matplotlib.pyplot as plt
import json
import pandas as pd
from pathlib import Path
import re
import math
import random
import sys
import os# os.environ["CUDA_VISIBLE_DEVICES"] = "-1" # Force TF to use only the CPU
from keras.utils import to_categorical
from keras import layers
from keras import models
from matplotlib.ticker import StrMethodFormatter

import time
import datetime
print(sys.version)
import numpy as np
print(np.__version__)

3.10.12 (main, Nov 20 2023, 15:14:05) [GCC 11.4.0]
1.23.5
```

```
pip install rouge-score
```

```
Requirement already satisfied: rouge-score in /usr/local/lib/python3.10/dist-packages (0.1.2)
Requirement already satisfied: absl-py in /usr/local/lib/python3.10/dist-packages (from rouge-score) (1.4.0)
Requirement already satisfied: nltk in /usr/local/lib/python3.10/dist-packages (from rouge-score) (3.8.1)
Requirement already satisfied: numpy in /usr/local/lib/python3.10/dist-packages (from rouge-score) (1.23.5)
Requirement already satisfied: six>=1.14.0 in /usr/local/lib/python3.10/dist-packages (from rouge-score) (1.16.0)
Requirement already satisfied: click in /usr/local/lib/python3.10/dist-packages (from nltk->rouge-score) (8.1.7)
Requirement already satisfied: joblib in /usr/local/lib/python3.10/dist-packages (from nltk->rouge-score) (1.3.2)
Requirement already satisfied: regex>=2021.8.3 in /usr/local/lib/python3.10/dist-packages (from nltk->rouge-score) (2021.8.3)
Requirement already satisfied: tqdm in /usr/local/lib/python3.10/dist-packages (from nltk->rouge-score) (4.66.1)
```

```
!pip install --upgrade tensorflow
```

```
Requirement already satisfied: tensorflow in /usr/local/lib/python3.10/dist-packages (2.15.0)
Requirement already satisfied: absl-py>=1.0.0 in /usr/local/lib/python3.10/dist-packages (from tensorflow) (1.4.0)
Requirement already satisfied: astunparse>=1.6.0 in /usr/local/lib/python3.10/dist-packages (from tensorflow) (1.16.1)
Requirement already satisfied: flatbuffers>=23.5.26 in /usr/local/lib/python3.10/dist-packages (from tensorflow) (23.5.26)
Requirement already satisfied: gast!=0.5.0,!0.5.1,!0.5.2,>=0.2.1 in /usr/local/lib/python3.10/dist-packages (from tensorflow) (0.5.3)
Requirement already satisfied: google-pasta>=0.1.1 in /usr/local/lib/python3.10/dist-packages (from tensorflow) (0.1.2)
Requirement already satisfied: h5py>=2.9.0 in /usr/local/lib/python3.10/dist-packages (from tensorflow) (3.9.0)
Requirement already satisfied: libclang>=13.0.0 in /usr/local/lib/python3.10/dist-packages (from tensorflow) (16.0.0)
```

```
Requirement already satisfied: ml-dtypes~==0.2.0 in /usr/local/lib/python3.10/dist-packages (from tensorflow) (0.
Requirement already satisfied: numpy<2.0.0,>=1.23.5 in /usr/local/lib/python3.10/dist-packages (from tensorflow)
Requirement already satisfied: opt-einsum>=2.3.2 in /usr/local/lib/python3.10/dist-packages (from tensorflow) (3
Requirement already satisfied: packaging in /usr/local/lib/python3.10/dist-packages (from tensorflow) (23.2)
Requirement already satisfied: protobuf!=4.21.0,!!=4.21.1,!!=4.21.2,!!=4.21.3,!!=4.21.4,!!=4.21.5,<5.0.0dev,>=3.20.3
Requirement already satisfied: setuptools in /usr/local/lib/python3.10/dist-packages (from tensorflow) (67.7.2)
Requirement already satisfied: six>=1.12.0 in /usr/local/lib/python3.10/dist-packages (from tensorflow) (1.16.0)
Requirement already satisfied: termcolor>=1.1.0 in /usr/local/lib/python3.10/dist-packages (from tensorflow) (2.
Requirement already satisfied: typing-extensions>=3.6.6 in /usr/local/lib/python3.10/dist-packages (from tensorflow)
Requirement already satisfied: wrapt<1.15,>=1.11.0 in /usr/local/lib/python3.10/dist-packages (from tensorflow)
Requirement already satisfied: tensorflow-io-gcs-filesystem>=0.23.1 in /usr/local/lib/python3.10/dist-packages (
Requirement already satisfied: grpcio<2.0,>=1.24.3 in /usr/local/lib/python3.10/dist-packages (from tensorflow)
Requirement already satisfied: tensorboard<2.16,>=2.15 in /usr/local/lib/python3.10/dist-packages (from tensorflow)
Requirement already satisfied: tensorflow-estimator<2.16,>=2.15.0 in /usr/local/lib/python3.10/dist-packages (fr
Requirement already satisfied: keras<2.16,>=2.15.0 in /usr/local/lib/python3.10/dist-packages (from tensorflow)
Requirement already satisfied: wheel<1.0,>=0.23.0 in /usr/local/lib/python3.10/dist-packages (from astunparse)>=1
Requirement already satisfied: google-auth<3,>=1.6.3 in /usr/local/lib/python3.10/dist-packages (from tensorbo
Requirement already satisfied: google-auth-oauthlib<2,>=0.5 in /usr/local/lib/python3.10/dist-packages (from ten
Requirement already satisfied: markdown>=2.6.8 in /usr/local/lib/python3.10/dist-packages (from tensorboard<2.16
Requirement already satisfied: requests<3,>=2.21.0 in /usr/local/lib/python3.10/dist-packages (from tensorboard<
Requirement already satisfied: tensorboard-data-server<0.8.0,>=0.7.0 in /usr/local/lib/python3.10/dist-packages
Requirement already satisfied: werkzeug>=1.0.1 in /usr/local/lib/python3.10/dist-packages (from tensorboard<2.16
Requirement already satisfied: cachetools<6.0,>=2.0.0 in /usr/local/lib/python3.10/dist-packages (from google-au
Requirement already satisfied: pyasn1-modules>=0.2.1 in /usr/local/lib/python3.10/dist-packages (from google-aut
Requirement already satisfied: rsa<5,>=3.1.4 in /usr/local/lib/python3.10/dist-packages (from google-auth<3,>=1.
Requirement already satisfied: requests-oauthlib>=0.7.0 in /usr/local/lib/python3.10/dist-packages (from google-
Requirement already satisfied: charset-normalizer<4,>=2 in /usr/local/lib/python3.10/dist-packages (from request
Requirement already satisfied: idna<4,>=2.5 in /usr/local/lib/python3.10/dist-packages (from requests<3,>=2.21.0
Requirement already satisfied: urllib3<3,>=1.21.1 in /usr/local/lib/python3.10/dist-packages (from requests<3,>=
Requirement already satisfied: certifi>=2017.4.17 in /usr/local/lib/python3.10/dist-packages (from requests<3,>=
Requirement already satisfied: MarkupSafe>=2.1.1 in /usr/local/lib/python3.10/dist-packages (from werkzeug>=1.0.
Requirement already satisfied: pyasn1<0.6.0,>=0.4.6 in /usr/local/lib/python3.10/dist-packages (from pyasn1-modu
Requirement already satisfied: oauthlib>=3.0.0 in /usr/local/lib/python3.10/dist-packages (from requests-oauthli
```

```
import tensorflow
```

```
! pip install transformers
```

```
Requirement already satisfied: transformers in /usr/local/lib/python3.10/dist-packages (4.35.2)
Requirement already satisfied: filelock in /usr/local/lib/python3.10/dist-packages (from transformers) (3.13.1)
Requirement already satisfied: huggingface-hub<1.0,>=0.16.4 in /usr/local/lib/python3.10/dist-packages (from tra
Requirement already satisfied: numpy>=1.17 in /usr/local/lib/python3.10/dist-packages (from transformers) (1.23.
Requirement already satisfied: packaging>=20.0 in /usr/local/lib/python3.10/dist-packages (from transformers) (2
Requirement already satisfied: pyyaml>=5.1 in /usr/local/lib/python3.10/dist-packages (from transformers) (6.0.1
Requirement already satisfied: regex!=2019.12.17 in /usr/local/lib/python3.10/dist-packages (from transformers)
Requirement already satisfied: requests in /usr/local/lib/python3.10/dist-packages (from transformers) (2.31.0)
Requirement already satisfied: tokenizers<0.19,>=0.14 in /usr/local/lib/python3.10/dist-packages (from transform
Requirement already satisfied: safetensors>=0.3.1 in /usr/local/lib/python3.10/dist-packages (from transformers)
Requirement already satisfied: tqdm>=4.27 in /usr/local/lib/python3.10/dist-packages (from transformers) (4.66.1
Requirement already satisfied: fsspec>=2023.5.0 in /usr/local/lib/python3.10/dist-packages (from huggingface-hub
Requirement already satisfied: typing-extensions>=3.7.4.3 in /usr/local/lib/python3.10/dist-packages (from huggi
Requirement already satisfied: charset-normalizer<4,>=2 in /usr/local/lib/python3.10/dist-packages (from request
Requirement already satisfied: idna<4,>=2.5 in /usr/local/lib/python3.10/dist-packages (from requests->transform
Requirement already satisfied: urllib3<3,>=1.21.1 in /usr/local/lib/python3.10/dist-packages (from requests->tra
Requirement already satisfied: certifi>=2017.4.17 in /usr/local/lib/python3.10/dist-packages (from requests->tra
```

Extract Data From Json Files

```
data = pd.read_json("/content/doqa-movies-test-v2.1.json")
data.head()
```

```

data   [grid icon]
0 {"title": "What's the meaning behind the phras... [info icon]
1 {"title": "Is Franz Oberhauser related to the ... [info icon]
2 {"title": "What was Margaery's motivation to c... [info icon]
3 {"title": "Why does Wolverine say these words ... [info icon]

# Function to extract data
def extract_data(data):
    cols = ["Context", "Question", "Answer", "OrigAnswer"]
    comp_list = []
    for entry in data["data"]:
        for paragraph in entry["paragraphs"]:
            for qa in paragraph["qas"]:
                # Extracting all answers for a question
                for answer in qa["answers"]:
                    temp_list = [
                        paragraph["context"],
                        qa["question"],
                        answer["text"],
                        qa["orig_answer"]["text"]
                    ]
                    comp_list.append(temp_list)
    # Creating a pandas DataFrame
    data_df = pd.DataFrame(comp_list, columns=cols)
    return data_df

```

▼ Calculate F1 Score

The F1 score ranges from 0 to 1, with a score of 1 indicating perfect precision and recall and 0 indicating poor performance. A high F1 score means that the model has high precision and high recall, which indicates that it is a good model for the binary classification task.

```
df1 = extract_data(data)
```

```
df1.head()
```

	Context	Question	Answer	OrigAnswer
0	I just did a bit of digging on your behalf reg...	What's the meaning behind the phrase "The bloo...	"You cannot wash away your sins."	there seems to be no concrete definition of it...
1	I just did a bit of digging on your behalf reg...	In Gangs of New York, a character is told not ...	1) "You cannot wash away your sins.", 2) The b...	The blood remains on the blade and not on your...
2	I just did a bit of digging on your behalf reg...	What about another meaning?	"You cannot wash away your sins."	You cannot wash away your sins
3	I just did a bit of digging on your behalf reg...	Are there any others?	The blood is worn as a badge of merit on the b...	The blood is worn as a badge of merit on the b...
4	I just did a bit of digging on your behalf reg...	Wait, in what way would there be merit?	CANNOTANSWER	CANNOTANSWER

```
df1.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 1884 entries, 0 to 1883
```

```
Data columns (total 4 columns):
 #   Column      Non-Null Count  Dtype  
 --- 
 0   Context     1884 non-null    object  
 1   Question    1884 non-null    object  
 2   Answer      1884 non-null    object  
 3   OrigAnswer  1884 non-null    object  
 dtypes: object(4)
 memory usage: 59.0+ KB
```

▼ Calculation Of F1 Metric Through Precision And Recall

```
def calculate_f1_score(true_answer, predicted_answer):
    true_set = set(true_answer.lower().split())
    pred_set = set(predicted_answer.lower().split())
    # Calculate precision and recall
    precision = len(true_set.intersection(pred_set)) / len(pred_set) if len(pred_set) > 0 else 0
    recall = len(true_set.intersection(pred_set)) / len(true_set) if len(true_set) > 0 else 0
    # Calculate F1 score
    f1 = 2 * (precision * recall) / (precision + recall) if (precision + recall) > 0 else 0
    return f1

display(df1.Context.values.tolist()[0])

'I just did a bit of digging on your behalf regarding this line, and there seems to be no concrete definition of it from any official source. That said, I found several interpretations along the way, including: 1) "You can not wash away your sins."; 2) The blood is worn as a badge of merit on the blade; 3) The blood remains on the blade and not on your hands (meaning you are relatively guilt-free). Personally, considering the overtly religious themes within the film, I would side with option 1, especially since Priest Vallon is saying the line. CANNOT ANSWER!'

trueAnswer = df1.OrigAnswer.values.tolist()
PredictAnswer = df1.Answer.values.tolist()

f1_scores = []

for ta in range(len(trueAnswer)):
    true_answer = trueAnswer[ta].lower()
    predicted_answer = PredictAnswer[ta].lower()
    f1_score = calculate_f1_score(true_answer, predicted_answer)
    f1_scores.append(f1_score)

# Create a DataFrame from the results
F1Score_df = pd.DataFrame({"F1 Score": f1_scores})

F1Score_df.info()

<class 'pandas.core.frame.DataFrame'>
RangeIndex: 1884 entries, 0 to 1883
Data columns (total 1 columns):
 #   Column      Non-Null Count  Dtype  
 --- 
 0   F1 Score   1884 non-null    float64
 dtypes: float64(1)
memory usage: 14.8 KB
```

▼ Calculation Of Exact Match

- Each Question Has One Correct Or "Gold" Answer
- Each Question Is Asked 3 Separate Times From The Same Context

- "Exact Match" Occurs When The Question Answer Is Identical To The Gold Answer
- Each Question Is Asked 3 Times And Is Either Correctly(Exact Matched) atleast once Thus It Will Have A Matching Average Of 1/3, 2/3, or 1 but not 0
- The Datasets Was Created Using two crowdworkers in Amazon Mechanical Turk(AMT) in accordance with the original source paper (Campos et al., 2020).

```
df2 = extract_data(data)
```

```
df2['Match'] = (df2['Answer'] == df2['OrigAnswer']).astype(float)
```

```
df2['EM3QAVG'] = df2.groupby('Question')['Match'].transform('mean')
```

```
df2.head(12)
```

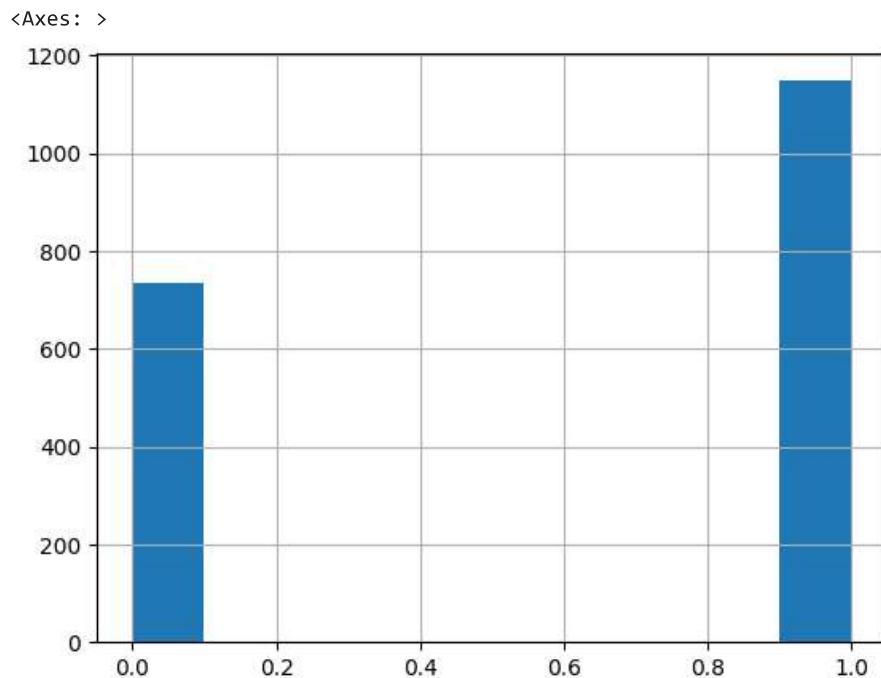
	Context	Question	Answer	OrigAnswer	Match	EM3QAVG	grid icon
0	I just did a bit of digging on your behalf reg...	What's the meaning behind the phrase "The bloo...	"You cannot wash away your sins."	there seems to be no concrete definition of it...	0.0	0.0	bar chart icon
1	I just did a bit of digging on your behalf reg...	In Gangs of New York, a character is told not ...	1) "You cannot wash away your sins."; 2) The b...	The blood remains on the blade and not on your...	0.0	0.0	
2	I just did a bit of digging on your behalf reg...	What about another meaning?	"You cannot wash away your sins."	You cannot wash away your sins	0.0	0.0	
3	I just did a bit of digging on your behalf reg...	Are there any others?	The blood is worn as a badge of merit on the b...	The blood is worn as a badge of merit on the b...	0.0	0.0	
4	I just did a bit of digging on your behalf reg...	Wait, in what way would there be merit?	CANNOTANSWER	CANNOTANSWER	1.0	1.0	
5	Yes. He's supposed to be the same character. T...	Is Franz Oberhauser related to the original Bl...	He's supposed to be the same character.	He's supposed to be the same character	0.0	0.0	
6	Yes. He's supposed to be the same character. T...	what is the role? is anything curious thing is...	The timeline of the story arc is all over the ...	The timeline of the story arc is all over the ...	1.0	1.0	
7	Yes. He's supposed to be the same character. T...	oh! then what about the main theme of his movie?	Some say a reboot but apparently its all suppo...	Some say a reboot but apparently its all suppo...	1.0	1.0	
8	Yes. He's supposed to be the same character. T...	oh! then what about the end of the story?	This falls apart because of Judi Denchs M, Bon...	This falls apart because of Judi Denchs M, Bon...	1.0	1.0	
9	Yes. He's supposed to be the same character. T...	do you personally know about bond?	CANNOTANSWER	CANNOTANSWER	1.0	1.0	
10	The balance of power had shifted considerably ...	What was Margaery's motivation to call Cersei ...	Whatever power Cersei held over Margaery has n...	Whatever power Cersei held over Margaery has n...	1.0	1.0	

```
df2.describe()
```

	Match	EM3QAVG	
count	1884.000000	1884.000000	
mean	0.609342	0.609342	
std	0.488027	0.485208	
min	0.000000	0.000000	
25%	0.000000	0.000000	
50%	1.000000	1.000000	
<hr/>			

▼ Match Visualization

```
df2["Match"].hist()
```



```
df2.head()
```

	Context	Question	Answer	OrigAnswer	Match	EM3QAVG	
0	I just did a bit of digging on your behalf reg...	What's the meaning behind the phrase "The bloo...	"You cannot wash away your sins."	there seems to be no concrete definition of it...	0.0	0.0	
1	I just did a bit of digging on your behalf reg...	In Gangs of New York, a character is told not ...	1) "You cannot wash away your sins."; 2) The b...	The blood remains on the blade and not on your...	0.0	0.0	
2	I just did a bit of digging on your behalf reg...	What about another meaning?	"You cannot wash away your sins."	You cannot wash away your sins	0.0	0.0	
3	I just did a bit of digging on your behalf reg...	Are there any others?	The blood is worn as a badge of merit on the b...	The blood is worn as a badge of merit on the b...	0.0	0.0	
4	I just did a bit of digging on your behalf reg...	Wait, in what way would there be merit?	CANNOTANSWER	CANNOTANSWER	1.0	1.0	

✓ Exact Match Visualizatioin For Each Question

- Each Question Will Have An Exact Match Average Of 1/3, 2/3 or 1

```
#Take Every Third Row Of The
df3 = df2[df2.index % 3 == 0]

#Create A New DataFrame With Context, Question, and ExactMatch
#df4rows=df3/3
df4 = df3[['Context', 'Question', 'EM3QAVG']].copy()
```

```
df4.info()
```

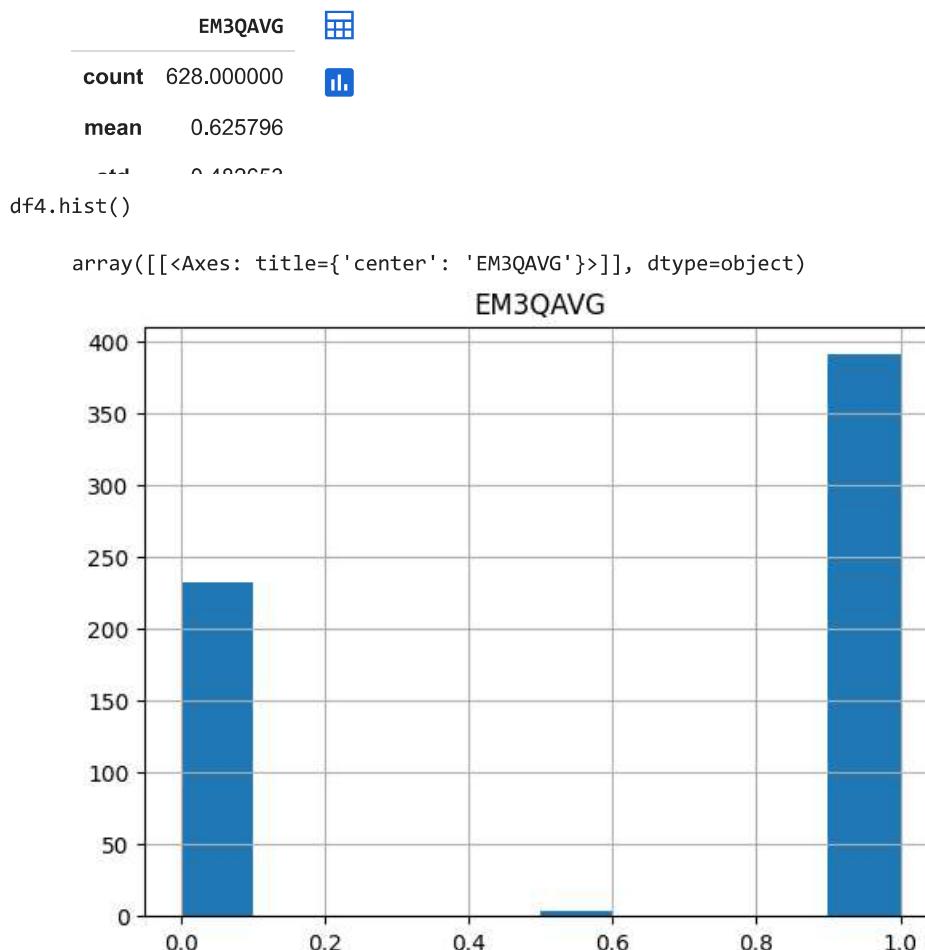
```
<class 'pandas.core.frame.DataFrame'>
Int64Index: 628 entries, 0 to 1881
Data columns (total 3 columns):
 #   Column      Non-Null Count  Dtype  
---  --          --          --      
 0   Context     628 non-null    object  
 1   Question    628 non-null    object  
 2   EM3QAVG    628 non-null    float64 
dtypes: float64(1), object(2)
memory usage: 19.6+ KB
```

```
df4=df4.reset_index(drop=True)
```

```
df4.head(10)
```

	Context	Question	EM3QAVG	
0	I just did a bit of digging on your behalf reg...	What's the meaning behind the phrase "The bloo...	0.0	
1	I just did a bit of digging on your behalf reg...	Are there any others?	0.0	
2	Yes. He's supposed to be the same character. T...	what is the role? is anything curious thing is...	1.0	
3	Yes. He's supposed to be the same character. T...	do you personally know about bond?	1.0	
4	The balance of power had shifted considerably ...	Who did Margaery marry Joffrey or Tommen?	0.0	
5	Logan's plan earlier in the movie was to buy a...	Why does wolverine say "It's got water, and it...	1.0	
6	The plot by MI 6 and the CIA is to give Le Chi...	Why didn't Felix Leiter catch Le Chiffre after...	1.0	
7	The plot by MI 6 and the CIA is to give Le Chi...	Did the CIA catch Le Chiffre?	1.0	
8	In the books it is yet to be discovered if he ...	What was the purpose of bringing back Jon?	0.0	
9	A Song of Ice and Fire really sounds quite fan...	Why is the TV show titled "Game of Thrones" an...	0.5	

```
df4.describe()
```



▼ F1 Score Analysis

```
df5=df2
```

```
df5.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 1884 entries, 0 to 1883
Data columns (total 6 columns):
 #   Column      Non-Null Count  Dtype  
 ---  --          --          --      
 0   Context     1884 non-null   object 
 1   Question    1884 non-null   object 
 2   Answer      1884 non-null   object 
 3   OrigAnswer  1884 non-null   object 
 4   Match       1884 non-null   float64
 5   EM3QAVG    1884 non-null   float64
dtypes: float64(2), object(4)
memory usage: 88.4+ KB
```

```
df6 = F1Score_df
```

```
df5.reset_index(drop=True, inplace=True)
df6.reset_index(drop=True, inplace=True)
```

```
df7 = pd.concat([df5, df6], axis=1)
```

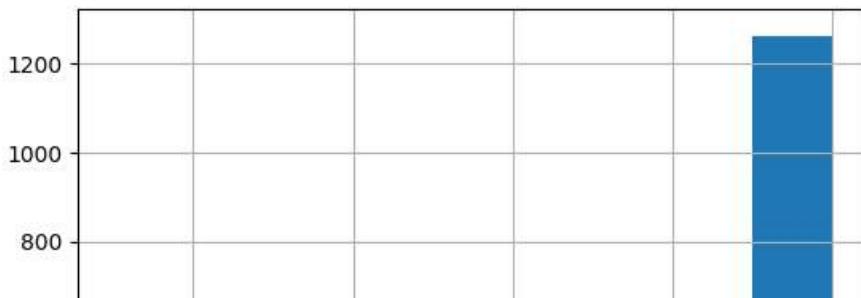
df7.head(12)

	Context	Question	Answer	OrigAnswer	Match	EM3QAVG	F1 Score
0	I just did a bit of digging on your behalf reg...	What's the meaning behind the phrase "The bloo...	"You cannot wash away your sins."	there seems to be no concrete definition of it...	0.0	0.0	0.000000
1	I just did a bit of digging on your behalf reg...	In Gangs of New York, a character is told not ...	1) "You cannot wash away your sins."; 2) The b...	The blood remains on the blade and not on your...	0.0	0.0	0.590909
2	I just did a bit of digging on your behalf reg...	What about another meaning?	"You cannot wash away your sins."	You cannot wash away your sins	0.0	0.0	0.666667
3	I just did a bit of digging on your behalf reg...	Are there any others?	The blood is worn as a badge of merit on the b...	The blood is worn as a badge of merit on the b...	0.0	0.0	0.909091
4	I just did a bit of digging on your behalf reg...	Wait, in what way would there be merit?	CANNOTANSWER	CANNOTANSWER	1.0	1.0	1.000000
5	Yes. He's supposed to be the same character. T...	Is Franz Oberhauser related to the original Bl...	He's supposed to be the same character.	He's supposed to be the same character	0.0	0.0	0.857143
6	Yes. He's supposed to be the same character. T...	what is the role? is anything curious thing is...	The timeline of the story arc is all over the ...	The timeline of the story arc is all over the ...	1.0	1.0	1.000000
7	Yes. He's supposed to be the same character. T...	oh! then what about the main theme of his movie?	Some say a reboot but apparently its all suppo...	Some say a reboot but apparently its all suppo...	1.0	1.0	1.000000
8	Yes. He's supposed to be the same character. T...	oh! then what about the end of the story?	This falls apart because of Judi Denchs M, Bon...	This falls apart because of Judi Denchs M, Bon...	1.0	1.0	1.000000
9	Yes. He's supposed to be the same character. T...	do you personally know about bond?	CANNOTANSWER	CANNOTANSWER	1.0	1.0	1.000000
10	The balance of power had shifted	What was Margaery's motivation to call	Whatever power Cersei held over Margaery has	Whatever power Cersei held over Margaery has	1.0	1.0	1.000000

#F1Score_df.hist(bins=10)

df7["F1 Score"].hist(bins=[0.10, 0.20, 0.30, 0.40, 0.50, 0.60, 0.70, 0.80, 0.90, 1.00])

<Axes: >



df7["F1 Score"].describe()

```

count    1884.000000
mean      0.771091
std       0.356426
min      0.000000
25%      0.569597
50%      1.000000
75%      1.000000
max      1.000000
Name: F1 Score, dtype: float64

```

✓ Use The Huggingface Transforms QA Model For Generation

From The Wikipedia Pace Of The AI Firm HuggingFace With Respect To Their Transformers Libraries The Transformers library is a Python package that contains open-source implementations of transformer models for text, image, and audio tasks. It is compatible with the PyTorch, TensorFlow and JAX deep learning libraries and includes implementations of notable models like BERT and GPT-2.[12] The library was originally called "pytorch-pretrained-bert"[13] which was then renamed to "pytorch-transformers" and finally "transformers."

Part 1 - Import The DataFrame

DF7 = df4

DF7.info()

```

<class 'pandas.core.frame.DataFrame'>
RangeIndex: 628 entries, 0 to 627
Data columns (total 3 columns):
 #   Column   Non-Null Count  Dtype  
--- 
 0   Context   628 non-null    object 
 1   Question  628 non-null    object 
 2   EM3QAVG  628 non-null    float64
dtypes: float64(1), object(2)
memory usage: 14.8+ KB

```

Step 2 - Take A Random Sample Of 50 Of The Origan Question-Answer Pairs

DF8 = DF7.sample(50)

! pip install pipeline

```
Requirement already satisfied: pipeline in /usr/local/lib/python3.10/dist-packages (0.1.0)
```

```
from transformers import AutoModelForQuestionAnswering, AutoTokenizer, pipeline
model_name = "deepset/roberta-base-squad2"
# a) Get predictions
nlp = pipeline('question-answering', model=model_name, tokenizer=model_name)
qa_output = []
# Iterate through the DataFrame and print the output
for index, row in DF8.iterrows():
    context = row['Context']
    question = row['Question']
    output = nlp(context=context, question=question)
    qa_output.append(output)
# Create a DataFrame from the results
Output_df = pd.DataFrame({"QA_Output": qa_output})
Output_df.head(10)
```

```
Output_df
```

QA_Output

```

0  {'score': 0.0021448724437505007, 'start': 0, '...
1  {'score': 0.0029164133593440056, 'start': 264, ...
2  {'score': 0.10357394814491272, 'start': 28, 'e...
3  {'score': 0.0013937543844804168, 'start': 354, ...
4  {'score': 0.0027404669672250748, 'start': 299, ...
5  {'score': 0.017981598153710365, 'start': 327, ...
6  {'score': 0.00032730898237787187, 'start': 362...
7  {'score': 0.011116974987089634, 'start': 262, ...
8  {'score': 0.002028859220445156, 'start': 263, ...
9  {'score': 0.002028907649219036, 'start': 277, ...
10  {'score': 0.06268451362848282, 'start': 106, '...
11  {'score': 0.3493128716945648, 'start': 784, 'e...
12  {'score': 0.005450374446809292, 'start': 405, ...
13  {'score': 0.0013724957825616002, 'start': 50, ...
14  {'score': 0.002362099476158619, 'start': 136, ...
15  {'score': 0.03062998689711094, 'start': 30, 'e...
16  {'score': 0.02064436674118042, 'start': 0, 'en...
17  {'score': 0.00046564466902054846, 'start': 78, ...
18  {'score': 0.0030710934661328793, 'start': 760, ...
19  {'score': 0.03519187122583389, 'start': 314, '...
20  {'score': 2.5030478809640044e-06, 'start': 773...
21  {'score': 0.0006155967130325735, 'start': 238, ...

```

❖ Compute The ROUGE Metric

```
24  %%capture <ipython> --output=rouge_out --error=rouge_err
```

ROUGE (Recall-Oriented Understudy for Gisting Evaluation), is a set of metrics and a software package specifically designed for evaluating automatic summarization, but that can be also used for machine translation. The metrics compare an automatically produced summary or translation against reference (high-quality and human-produced) summaries or translations.

```
28  {'score': 0.0038544582203030586, 'start': 320, ...}
```

ROUGE-N ROUGE-N measures the number of matching n-grams between the model-generated text and a human-produced reference. Then, ROUGE-1 F1-score can be directly obtained from the ROUGE-1 precision and recall using the standard F1-score formula.

$$\text{ROUGE-1 F1-score} = \frac{2 * (\text{precision} * \text{recall})}{(\text{precision} + \text{recall})}$$

```
dr = df8 = extract_data(data)
33  {'score': 0.1053428351891199, 'start': 192, '...
dr.head()
```



	Context	Question	Answer	OrigAnswer
0	I just did a bit of digging on your behalf reg...	What's the meaning behind the phrase "The bloo...	"You cannot wash away your sins."	there seems to be no concrete definition of it...
1	I just did a bit of digging on your behalf reg...	In Gangs of New York, a character is told not ...	1) "You cannot wash away your sins."; 2) The b...	The blood remains on the blade and not on your...
2	I just did a bit of digging on your behalf reg...	What about another meaning?	"You cannot wash away your sins."	You cannot wash away your sins
44	I just did a bit of digging on your behalf reg...	The blood is worn as a badge of	The blood is worn as a badge of	The blood is worn as a badge of

```
! pip install rouge
```

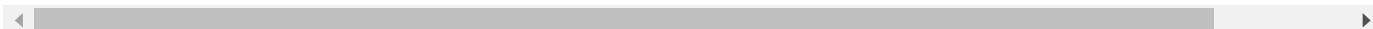
```
Requirement already satisfied: rouge in /usr/local/lib/python3.10/dist-packages (1.0.1)
Requirement already satisfied: six in /usr/local/lib/python3.10/dist-packages (from rouge) (1.16.0)
```

```
import rouge
```

```
49 {'score': 3.3587273264856776e-06, 'start': 291...}
```

```
! pip install rouge-score
```

```
Requirement already satisfied: rouge-score in /usr/local/lib/python3.10/dist-packages (0.1.2)
Requirement already satisfied: absl-py in /usr/local/lib/python3.10/dist-packages (from rouge-score) (1.4.0)
Requirement already satisfied: nltk in /usr/local/lib/python3.10/dist-packages (from rouge-score) (3.8.1)
Requirement already satisfied: numpy in /usr/local/lib/python3.10/dist-packages (from rouge-score) (1.23.5)
Requirement already satisfied: six>=1.14.0 in /usr/local/lib/python3.10/dist-packages (from rouge-score) (1.16.0)
Requirement already satisfied: click in /usr/local/lib/python3.10/dist-packages (from nltk->rouge-score) (8.1.7)
Requirement already satisfied: joblib in /usr/local/lib/python3.10/dist-packages (from nltk->rouge-score) (1.3.2)
Requirement already satisfied: regex>=2021.8.3 in /usr/local/lib/python3.10/dist-packages (from nltk->rouge-score) (4.6.1)
Requirement already satisfied: tqdm in /usr/local/lib/python3.10/dist-packages (from nltk->rouge-score) (4.66.1)
```



```
from rouge_score import rouge_scorer
```

```
scorer = rouge_scorer.RougeScorer(['rouge1', 'rouge2', 'rougeL'], use_stemmer=True)
```

```
sc_output = []
```

```
for index, row in dr.iterrows():
```

```
    hyps = row['Answer']
```

```
    refs = row['OrigAnswer']
```

```
# Calculate Rouge scores
```

```
scores = scorer.score(hyps, refs)
```

```
# Append scores to the list
```

```
sc_output.append(scores)
```

```
# Create a DataFrame from the results
```

```
SCOutput_df = pd.DataFrame({"RougeScore_Output": sc_output})
```

```
SCOutput_df.head(10)
```

index	RougeScore_Output
0	{'rouge1': Score(precision=0.0, recall=0.0, fmeasure=0.0), 'rouge2': Score(precision=0.0, recall=0.0, fmeasure=0.0), 'rougeL': Score(precision=0.0, recall=0.0, fmeasure=0.0)}
1	{'rouge1': Score(precision=1.0, recall=0.4473684210526316, fmeasure=0.61818181818182), 'rouge2': Score(precision=1.0, recall=0.43243243243243246, fmeasure=0.6037735849056604), 'rougeL': Score(precision=1.0, recall=0.4473684210526316, fmeasure=0.61818181818182)}
2	{'rouge1': Score(precision=1.0, recall=1.0, fmeasure=1.0), 'rouge2': Score(precision=1.0, recall=1.0, fmeasure=1.0), 'rougeL': Score(precision=1.0, recall=1.0, fmeasure=1.0)}
3	{'rouge1': Score(precision=1.0, recall=1.0, fmeasure=1.0), 'rouge2': Score(precision=1.0, recall=1.0, fmeasure=1.0), 'rougeL': Score(precision=1.0, recall=1.0, fmeasure=1.0)}
4	{'rouge1': Score(precision=1.0, recall=1.0, fmeasure=1.0), 'rouge2': Score(precision=0.0, recall=0.0, fmeasure=0.0), 'rougeL': Score(precision=1.0, recall=1.0, fmeasure=1.0)}
5	{'rouge1': Score(precision=1.0, recall=1.0, fmeasure=1.0), 'rouge2': Score(precision=1.0, recall=1.0, fmeasure=1.0), 'rougeL': Score(precision=1.0, recall=1.0, fmeasure=1.0)}
6	{'rouge1': Score(precision=1.0, recall=1.0, fmeasure=1.0), 'rouge2': Score(precision=1.0, recall=1.0, fmeasure=1.0), 'rougeL': Score(precision=1.0, recall=1.0, fmeasure=1.0)}
7	{'rouge1': Score(precision=1.0, recall=1.0, fmeasure=1.0), 'rouge2': Score(precision=1.0, recall=1.0, fmeasure=1.0), 'rougeL': Score(precision=1.0, recall=1.0, fmeasure=1.0)}
9	{'rouge1': Score(precision=1.0, recall=1.0, fmeasure=1.0), 'rouge2': Score(precision=1.0, recall=1.0, fmeasure=1.0), 'rougeL': Score(precision=1.0, recall=1.0, fmeasure=1.0)}

Show 100 per page

Like what you see? Visit the [data table notebook](#) to learn more about interactive tables.WARNING:root:Quickchart encountered unexpected dtypes in columns: "(['RougeScore_Output'],)"
No charts were generated by quickchart