



Download Dataset  
from Hugging face

# Hugging face Text Datasets

```
from datasets import load_dataset
```

```
raw_datasets = load_dataset("glue", "mrpc")  
raw_datasets
```

```
DatasetDict({  
  train: Dataset({  
    features: ['sentence1', 'sentence2', 'label', 'idx'],  
    num_rows: 3668  
  })  
  validation: Dataset({  
    features: ['sentence1', 'sentence2', 'label', 'idx'],  
    num_rows: 408  
  })  
  test: Dataset({  
    features: ['sentence1', 'sentence2', 'label', 'idx'],  
    num_rows: 1725  
  })  
})
```

```
raw_datasets["train"]
```

```
Dataset({  
  features: ['sentence1', 'sentence2', 'label', 'idx'],  
  num_rows: 3668  
})
```

```
raw_datasets["train"][6]
```

```
{'idx': 6,  
 'label': 0,  
 'sentence1': 'The Nasdaq had a weekly gain of 17.27 , or 1.2 percent , closing at 1,520.15 on Friday .',  
 'sentence2': 'The tech-laced Nasdaq Composite .IXIC rallied 30.46 points , or 2.04 percent , to 1,520.15 .'}
```

```
[ ] 1 raw_datasets["train"][:5]
```

```
→ {'idx': [0, 1, 2, 3, 4],  
   'label': [1, 0, 1, 0, 1],  
   'sentence1': ['Amrozi accused his brother , whom he called " the witness " , of deliberately distorting his evidence .',  
                  "Yucaipa owned Dominick 's before selling the chain to Safeway in 1998 for $ 2.5 billion .",  
                  'They had published an advertisement on the Internet on June 10 , offering the cargo for sale , he added .',  
                  'Around 0335 GMT , Tab shares were up 19 cents , or 4.4 % , at A $ 4.56 , having earlier set a record high of A $ 4.57 .',  
                  'The stock rose $ 2.11 , or about 11 percent , to close Friday at $ 21.51 on the New York Stock Exchange .'],  
   'sentence2': ['Referring to him as only " the witness " , Amrozi accused his brother of deliberately distorting his evidence .',  
                  "Yucaipa bought Dominick 's in 1995 for $ 693 million and sold it to Safeway for $ 1.8 billion in 1998 .",  
                  "On June 10 , the ship 's owners had published an advertisement on the Internet , offering the explosives for sale .",  
                  'Tab shares jumped 20 cents , or 4.6 % , to set a record closing high at A $ 4.57 .',  
                  'PG & E Corp. shares jumped $ 1.63 or 8 percent to $ 21.03 on the New York Stock Exchange on Friday .']}]
```

# Hugging face Image Datasets

## Step 2: Load the Dataset

<https://huggingface.co/datasets/eurecom-ds/celeba-hq-small>



```
from datasets import load_dataset

# Load the CelebA dataset
dataset = load_dataset('eurecom-ds/celeba-hq-small')

# Print the dataset structure
print(dataset)
```

[3]

```
DatasetDict({
  train: Dataset({
    features: ['image', 'attributes'],
    num_rows: 246
  })
  val: Dataset({
    features: ['image', 'attributes'],
    num_rows: 150
  })
  test: Dataset({
    features: ['image', 'attributes'],
    num_rows: 371
  })
})
```



