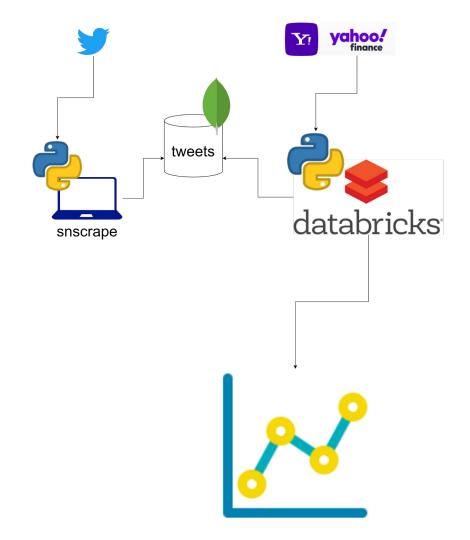
Crypto Price Visualization

Big Data Technology & Architecture

Overview



Twitter Scraper: tweets by username

```
def snscrape_ids(keyword_user_search_param, keywords_users_list, since, until, lang):
   dir_name = f"{since.replace('-', '')}_{until.replace('-', '')}"
   Path("./scraped_tweet").mkdir(exist_ok=True)
   os.chdir(os.path.join(ROOT_DIR, "scraped_tweet"))
   Path(dir_name).mkdir(exist_ok=True)
   temppath = "C:/twittertest/twits.json"
   for keyword in keywords_users_list:
       if len(keyword) > 0:
           output_name = f"{keyword.replace(' ', '_')}_{since.replace('-', '')}_{until.replace('-', '')}.txt"
           output_path = os.path.join(ROOT_DIR, 'scraped_tweet', dir_name, output_name)
           print(f'scraping tweets with keyword: "{keyword}" ...')
               os.system(f'snscrape twitter-user {keyword} > {temppath}')
               print('Searching for twitter-user:', keyword)
           except Exception as err:
               print(f"SNSCRAPE ERROR: {err}")
   shutil.move(temppath,output_path)
   print(f'Scraped all tweets in keywords list.')
```

Twitter Scraper: save to Mongo

```
#when saving to mongo, objects in list are changed and Json serialization fails later
tweets_batch_copy = [t.copy() for t in tweets_batch]
save_to_mongo.saveToMongo("twitter", "user_tweets", tweets_batch_copy)
tweets.append(tweets_batch)
```

```
def saveToMongo(database_name, collection_name, tweets_batch):
    db = client[database_name]
    mycollection = db[collection_name]
    #ensure unique tweet id
    mycollection.create_index('id', unique=True)
    for twit in tweets_batch:
        try:
        mycollection.insert_one(twit)
        except pymongo.errors.DuplicateKeyError as e:
        print(f"Duplicate tweet insert: {str(e)}")
```

DOGE vs @WhalePanda



BTC vs @elonmusk

 BTC-USD @elonmusk

50000

40000

35000

25000

@DocumentingBTC The dollar is shorting itself 30000

09-February



ETH vs @VitalikButerin



Doge @elon

