Project 2: Predicting Sentiment

aws flair

OpenAI GPT-3

Project Members: Bryan LaSane, Rina Weiner, Andrew McCoy

Group Task Assignments

- Bryan - GPT-3 and Openai
- Andrew - sentiment prediction model
- Rina - email notification to end user

Roles and Roadmap

- Bryan to discuss GPT-3 and Openai
- Andrew to discuss and demonstrate the sentiment model
- Rina to demonstrate the notification email function based on sentiment

Initial Imports

```python
#Initial import

import openai
import os
import pandas as pd
from newsapi import NewsApiClient
from dotenv import load_dotenv
load_dotenv()
import json
from datetime import date, timedelta
import flair #<----- new NLP module not used in class
from flair.data import Sentence

from flair.models import SequenceTagger
```

[83]

Openai Modules

Classification
Classify items into categories via example.

Movie to Emoji
Convert movie titles into emoji.

Translate programming languages
Translate from one programming language ...

Explain code
Explain a complicated piece of code.

Factual answering
Guide the model towards factual answering ...

Product name generator
Create product names from examples word...

Python bug fixer
Find and fix bugs in source code.

Python to natural language
Explain a piece of Python code in human un...

Calculate Time Complexity
Find the time complexity of a function.

Advanced tweet classifier
Advanced sentiment detection for a piece o...

Keywords
Extract keywords from a block of text.

Ad from product description
Turn a product description into ad copy.

TL;DR summarization
Summarize text by adding a 'tl;dr:' to the en...

Spreadsheet generator
Create spreadsheets of various kinds of dat...

Openai

Examples

Explore what's possible with some example applications

🔍 Search...

All categories

Chat
Open ended conversation with an AI assist...

Q&A
Answer questions based on existing knowle...

Grammar correction
Corrects sentences into standard English.

Summarize for a 2nd grader
Translates difficult text into simpler concep...

Natural language to OpenAI API
Create code to call to the OpenAI API usin...

Text to command
Translate text into programmatic commands.

English to French
Translates English text into French.

Natural language to Stripe API
Create code to call the Stripe API using nat...

Openai code

View code

You can use the following code to start integrating your current prompt and settings into your application.

```python
POST /v1/engines/davinci/completions                    python ∨    ⧉ Copy
1   import os
2   import openai
3
4   openai.api_key = os.getenv("OPENAI_API_KEY")
5
6   response = openai.Completion.create(
7       engine="davinci",
8       prompt="My second grader asked me what this passage means:\n\"\"\"\n
9       temperature=0.5,
10      max_tokens=100,
11      top_p=1,
12      frequency_penalty=0.2,
13      presence_penalty=0,
14      stop=["\"\"\"\"\"\"\"]
15  )
```

Get the news

```python
#News api
api_key = os.getenv("NEWS_API")

# Create a newsapi client
newsapi = NewsApiClient(api_key=api_key)

# Fetch XRP Coin Headlines
xrp_news = newsapi.get_everything(q= 'xrp',
                                  from_param= str(date.today() - timedelta(5)),
                                  sort_by='relevancy',
                                  language='en',)

xrp_news
```

Pull news from the last 5 days only

Project team's OpenAi Code

```python
#Summerize XRP Data

openai.api_key = os.getenv("OPENAI_API_KEY")

xrp_summeries=[]

for row,data in xrp_article_content.iterrows():

    response = openai.Completion.create(
        engine="davinci",
        prompt=data["XRP Headlines"]+"\n\nI rephrased this for my daughter, in plain language a second grader can understand:",
        temperature=0.5,
        max_tokens=100,
        top_p=1,
        frequency_penalty=0.2,
        presence_penalty=0,
        stop=["\"\"\"\"\"\"\"]
    )
    xrp_summeries.append(response)

xrp_summeries
```
[95]

```python
#print btc article summeries

xrp_summeries
```
[96]

Comparison

Beginning of Original Article:

"The SEC likely did not expect pushback from its lawsuit against Ripple and the rise of the XRP Army. After more than a decade of dismissing the revolutionary potential of blockchain technology….

Summarized for a second grader:

"The SEC is a government agency that is supposed to protect people from scams and bad guys. They're like the police, but for the stock market. They think Ripple is a scam, and they don't like that. They don't think Ripple should be able to sell XRP. They sued Ripple to try to get a judge to make them stop selling XRP. But the people who use XRP didn't like that. So they hired a lawyer to help them fight back."

NLP Model: Flair

IMDb

flair 0.10

```
pip install flair
```

- Developed by Humboldt University of Berlin
- Offers a pre-trained NLP model for use on any language dataset based on IMDB movie reviews to gauge sentiment

```
# Initiate a Movie Sentiment Model

movie_sentiment_model = flair.models.TextClassifier.load('en-sentiment')
✓ 4.8s
```

Sentiment Classification and Confidence

```python
sample = flair.data.Sentence(sentence)
movie_sentiment_model.predict(sample)
movie_sentiment.append(sample.labels[0].value)
movie_confidence.append(sample.labels[0].score)
```

Value = either "POSITIVE" or "NEGATIVE"

Score = range from 0.00 - 1.0, represents the confidence level of the sentiment being positive or negative

Class Participation!

Train and Test a new model

The class will view a few movie quotes and the results of the Flair sentiment analysis.

Then they will predict the results of 2 quotes and observe the results

Braveheart

"This Wallace, he doesn't even have a knighthood, but he fights with passion and he inspires"

Sentiment: Positive

Confidence: 93%

The Matrix

"Neo, I'm not afraid anymore. The Oracle told me that I would fall in love and that that man... the man that I loved would be The One"

Sentiment: Positive

Confidence: 99%

Full Metal Jacket

"I am Gunnery Sergeant Hartman, your senior drill instructor. From now on you will speak only when spoken to, and the first and last words out of your filthy sewers will be "Sir". Do you maggots understand that?"

Sentiment: Positive

Confidence: 61%

Romeo and Juliet

"From ancient grudge break to new mutiny, where civil blood makes civil hands unclean"

Sentiment: Negative

Confidence: 98%

The Blair Witch Project

I'm scared to close my eyes; I'm scared to open them!

Sentiment: Positive

Confidence: 60%

Test 1

Now, you know it's up to you whether or not you want to just do the bare minimum. Or... well, like Brian, for example, has thirty seven pieces of flair, okay. And a terrific smile."

Class Agreed on Sentiment:
Class Agreed on Confidence:

Results from Test 1

Class Agreed on Sentiment:
Class Agreed on Confidence:

Flair Model Sentiment: Positive
Flair Model Confidence: 99%
Comments: phrase seems to be encouraging and the word "terrific" and "smile" are likely pushing up the confidence level

Now, you know it's up to you whether or not you want to just do the bare minimum. Or... well, like Brian, for example, has thirty seven pieces of flair, okay. And a terrific smile."

How might tone or sarcasm influence sentiment?

Integrating with a NLP that can process tone could potentially enhance sentiment related to the spoken word.

Test 2

Oh oh and I almost forgot. I'm also gonna need you to go ahead and come in on Sunday too Kay? We lost some people this week and we need to sort of play catch-up."

Class Agreed on Sentiment:
Class Agreed on Confidence:

Results from Test 2

Class Agreed on Sentiment:
Class Agreed on Confidence:

Flair Model Sentiment: Positive
Flair Model Confidence: 56%
Comments: Interesting how the model returns an essentially neutral result

Oh oh and I almost forgot. I'm also gonna need you to go ahead and come in on Sunday too Kay? We lost some people this week and we need to sort of play catch-up."

Once more, this quote has a heavy underlying tone.

Sample news data run through flair model

	description	sentiment	confidence
0	The Singapore affiliate of Binance, one of the...	NEGATIVE	0.999621
1	Yi He, a cofounder of the cryptocurrency excha...	NEGATIVE	0.995340
2	Cryptocurrency exchange Binance on Tuesday sai...	POSITIVE	0.536603
3	Binance will pick one winner from each team, r...	POSITIVE	0.527720

"The Singapore affiliate of Binance, one of the world's largest cryptocurrency exchanges, said in a statement on Monday it will withdraw its local license application and wind down its digital payment token business in the city-state"

Taking it further

Why not use get dummies function?

Get dummies seemed to assign "0" to what ever was in the first row, then "1" when the value changed.

For live news pulls we never know if the first row sentiment will be "positive" or "negative"

```python
# Assign numerical values based on sentiment value

sentiment_numerical = []

for row in binance_df["sentiment"]:
    i = ""
    if row == "POSITIVE":
        i = 1
        sentiment_numerical.append(i)
    else:
        i = -1
        sentiment_numerical.append(i)

binance_df["sentiment_numerical"] = sentiment_numerical

binance_df = binance_df.astype({"sentiment_numerical" : float})
```

Result

	description	sentiment	confidence	sentiment_numerical
0	The Singapore affiliate of Binance, one of the...	NEGATIVE	0.999621	-1.0
1	Yi He, a cofounder of the cryptocurrency excha...	NEGATIVE	0.995340	-1.0
2	Cryptocurrency exchange Binance on Tuesday sai...	POSITIVE	0.536603	1.0
3	Binance will pick one winner from each team, r...	POSITIVE	0.527720	1.0
4	Binance's relations with the FCA have improved...	POSITIVE	0.916237	1.0
5	Alethea AI and BeingAI are collaborating with ...	POSITIVE	0.987426	1.0

Taking it further

```
# Create variable to store total confidence value
total_confidence_value = binance_df['confidence'].sum().astype(float)
print(total_confidence_value)
```

18.059917211532593

There were 19 articles, but since some were lower confidence the total confidence value for this pool of articles is less than 19

Calculate Weighted Sentiment

```python
# Add weighted sentiment
## Weighted sentiment takes the confidence value and divides by the total of confidence values to determine
## how much weight that particular news article should influence the overal sentiment

weighted_sentiment = []

binance_df["weighted_sentiment"] = (binance_df["confidence"] * binance_df["sentiment_numerical"] / total_confidence_value)
```

Result

The articles with a higher confidence value will influence the aggregate sentiment whether it be strongly positive, negative or more neutral.

description	sentiment	confidence	sentiment_numerical	weighted_sentiment
The Singapore affiliate of Binance, one of the...	NEGATIVE	0.999621	-1.0	-0.056655
Yi He, a cofounder of the cryptocurrency excha...	NEGATIVE	0.995340	-1.0	-0.056413
Cryptocurrency exchange Binance on Tuesday sai...	POSITIVE	0.536603	1.0	0.030413
Binance will pick one winner from each team, r...	POSITIVE	0.527720	1.0	0.029910
Binance's relations with the FCA have improved...	POSITIVE	0.916237	1.0	0.051929

Next, we will sum this column

Resulting Weighted Sentiment

```
# Total of weighted sentiment

total_weighted_sentiment = binance_df['weighted_sentiment'].sum()

total_weighted_sentiment

-0.3639357743872927
```

Weighted Sentiment = -36.3%

This result indicates that overall sentiment for the given articles is 36.3% to the negative

Summarized news data run through flair model

```
# Add weighted sentiment
## Weighted sentiment takes the confidence value and divides by the total of confidence values to determine
## how much weight that particular news article should influence the overal sentiment

btc_weighted_sentiment = []

btc_summeries_df["weighted_sentiment"] = (btc_summeries_df["confidence"] * btc_summeries_df["sentiment_numerical"]) / btc_total_confidence_value)

btc_summeries_df
```

[12]:

	Article Summary	sentiment	confidence	sentiment_numerical	weighted_sentiment
0	"A long time ago, in a galaxy far, far away, t...	NEGATIVE	0.927638	-1.0	-0.049289
1	A representation of the virtual cryptocurrency...	NEGATIVE	0.999458	-1.0	-0.053105
2	"The government has given the power to steal f...	NEGATIVE	0.846402	-1.0	-0.044972
3	"What is the secret to making a lot of money?"...	POSITIVE	0.793606	1.0	0.042167
4	"A company called MicroStrategy bought a lot o...	NEGATIVE	0.997138	-1.0	-0.052981
5	"The people who buy and sell stuff on the inte...	NEGATIVE	0.964435	-1.0	-0.051244
6	https://www.reddit.com/r/CryptoCurrency/comme...	NEGATIVE	0.996223	-1.0	-0.052933
7	"The government is making it so that if you w...	NEGATIVE	0.999652	-1.0	-0.053115
8	"What is this, a stupid question or what?" r...	NEGATIVE	0.999916	-1.0	-0.053129
9	"The buzz around the cryptocurrency market is ...	NEGATIVE	0.909596	-1.0	-0.048330
10	The S&P 500 and Dow futures hit all-time ...	POSITIVE	0.595949	1.0	0.031665
11	" Don't trust everything you read on the inte...	NEGATIVE	0.999954	-1.0	-0.053131
12	For Immediate ReleaseChicago, IL December 28, ...	NEGATIVE	0.999772	-1.0	-0.053121
13	"Bitcoin is a digital currency. It's like a do...	POSITIVE	0.931952	1.0	0.049518
14	"The S&P 500 was on track to open at all-time ...	POSITIVE	0.977785	1.0	0.051953
15	"I don't like to spend money on things I don'...	NEGATIVE	0.999971	-1.0	-0.053132

...

"I always trusted code more than people anyway."

- Elsie Hughes, Westworld TV show, Season two, Episode four

Flair sentiment: Positive, Flair confidence: 99%