

Contents

Chapter 1

Introduction

This documentation describes the “news intensity” dashboard and API. These are tools for gauging the amount of relevant media coverage for any asset. That is to say, our news intensities (NI) measure the volume of news stories that are realistically pertinent to an asset, regardless of whether they mention that asset by name. For example, rare earths are used in making smartphones, therefore a story about rare-earth mining is potentially of relevance to a company manufacturing smartphones, whether or not the company is mentioned in the story. Likewise, Peru is a leader in zinc mining, therefore a story about political turmoil in Peru may have an impact on the prices of zinc futures, whether or not the connection is made in the story itself.

The notion that asset prices are impacted by the nature and extent of media coverage (even when it does not contain new information) has found strong support in the finance literature (e.g. ?, ?, ?). By capturing the amount of related media coverage for any asset, NI are expected to have incremental explanatory power for asset price movements.

The calculation of NI relies on Cortical.io’s multi-lingual semantic fingerprinting technology, which ingests text and outputs its “semantic fingerprint” in the form of a sparse 128x128 matrix of 0s and 1s. Each position in the matrix represents a topic, so that the fingerprint of a text indicates which of the possible 16,384 topics the text is related to. Every day and for every supported language, we obtain the aggregate semantic fingerprint of online news stories from thousands of sources. The count of the relevant stories for each topic is then normalized so that the average normalized value across topics is equal to 1. We also obtain semantic fingerprints for traded assets based on their textual descriptions (?, ? show that such fingerprints have significant information content). The value of NI for any day/language/asset combination is the average of the normalized article counts across all topics related to the asset in question.

Chapter 2

Getting started

This section contains general information about querying companies and commodities for both the API and the dashboard.

If you have been given an API key, visit the API section. If you have been given access to the FinTxt dashboard, visit the dashboard section.

2.1 Languages

FinTxt news intensity values are based on news articles. These articles are written in either one of the following languages:

1. English
2. French
3. German
4. Arabic
5. Russian

When you retrieve news intensities, you can either select one of these languages or opt to use the aggregate of all languages ('total').

2.2 Selecting companies

When using the API or the dashboard, you should query companies by using their Reuters Instrument Code.

2.3 Selecting commodities

You can query the API and dashboard for either companies or commodities (but not a mix of both). The following commodities are available:

- corn
- oats
- rice

- soybeans
- rapeseed
- wheat
- milk
- cocoa
- coffee
- cotton
- sugar
- oranges
- ethanol
- propane
- copper
- lead
- zinc
- tin
- aluminium
- nickel
- cobalt
- molybdenum
- steel
- gold
- platinum
- palladium
- silver
- rubber
- wool
- amber

Chapter 3

News Intensity API

The FinTxt news intensity API is hosted at <https://api.fintxt.io/rest>. You can view the endpoints and their documentation [here](#).

The API contains four endpoints:

1. **languages**: returns a list of available language for which you can query news intensities
2. **live**: returns live news intensity metrics for a company or commodity and a language
3. **historic**: returns historic news intensity metrics for a company or commodity, a language and a date
4. **portfolio**: returns live or historic news intensity metrics for a weighted portfolio of stocks or commodities, a language and a date.

The *language* endpoint can be queried without an API key. The *historic* endpoint can be queried without an API key for dates that go back beyond 30 days on the date of today.

3.1 Basic usage

View the documentation for each endpoint for more information about headers and post bodies.

Essentially, you can query each endpoint by building the URL. For example, you want to query the news intensity value for the commodity `wool` on 04-06-2018 for texts written in arabic. Your query would then look as follows:

```
languages = 'arabic'
date = '04-06-2018'
endpoint = 'historic'
type = 'commodities'
q = 'wool'
```

You could construct this request using `curl`:

```
curl -X GET "https://api.fintxt.io/rest/historic/commodities/arabic/04-06-2018?q=wool" -H "accept: app
```

You could use `httr` in R:

```
languages <- 'arabic'
date <- '04-06-2018'
endpoint <- 'historic'
type <- 'commodities'
q <- 'wool'

url <- paste0("https://api.fintxt.io/rest/", endpoint, "/", type, "/", languages, "/", date, "?q=", q)

resp <- httr::GET(url)
httr::content(resp)
```

Or you could use `requests` in Python:

```
import json
import requests

languages = 'arabic'
date = '04-06-2018'
endpoint = 'historic'
type_ = 'commodities'
q = 'wool'

url = "https://api.fintxt.io/rest/{}/{}/{}/{}?q={}".format(endpoint, type_, languages, date, q)

# Send request
r = requests.get(url)

# Load response
c = json.loads(r.content)

print(c)
```

But it would be easier to use the clients described below.

3.2 R and Python clients

You can use the R and Python 3 clients to retrieve data from the API. These clients can be installed from their GitHub repositories.

3.2.1 R client

Install the R client by executing

```
devtools::install_github("FinTxt/FinTxtClient")
```

After installing the package, register your API token by calling the following code:

```
library(FinTxtClient)
Sys.setenv("FINTXT_CLIENT_TOKEN" = "<yourtoken>")
```


You can now access the various endpoints:

[illegible]

3.2.2 Python client

To install the python client, execute the following:

```
pip install git+https://github.com/FinTxt/FinTxtClient-Py.git --user
```

You can import the package as follows:

```
from FinTxtClient import FinTxtClient
```

Then, you can initiate the client using:

```
client = FinTxtClient() # Optionally, pass 'key = <your-key-here>'
```

Using the client is as simple as calling the following functions:

```
# Call the languages endpoint
client.languages()

# Call the live endpoint for a commodity
client.live_one("commodities", "english", "milk")

# Call the historic endpoint for a commodity and a date
client.historic_one("commodities", "english", "13-07-2018", "milk")

# Call the live portfolio endpoint
client.live_portfolio( _type = "companies", language = "english",
                      identifiers=["TRI.TO", "IBM.N", "RRD.N", "SPGI.N", "INTU.OQ", "RELN.AS", "WLSNc.AS"],
                      weights=[0.3, 0.1,0.05,0.05, 0.2,0.1,0.1,0.1])

# Call the historic portfolio endpoint
client.historic_portfolio( _type = "companies", language = "english", date="13-07-2018",
                          identifiers=["TRI.TO", "IBM.N", "RRD.N", "SPGI.N", "INTU.OQ", "RELN.AS", "WLSNc.AS"],
                          weights=[0.3, 0.1,0.05,0.05, 0.2,0.1,0.1,0.1])
```