COMP 8790-01-2020W Topics in Applied Artificial Intelligence

News Collection, Sentiment Analysis and Visualization

Jiaxiu Li, Yan Ding, and Rui Wang
School of Computer Science, University of Windsor
March 30th ,2020

***Abstract*—With the development of science and technology, as a promising technology, timeliness and relativity become essential in journalism when people keen to new things with their personal interest and limit time to read tremendous amount of news. That urge demand driven us to build a system to help readers get fast answers with informative and actionable visualizations and it also could help company find current and past events and articles that exactly match their criteria by using AI-driven filter. Designing and making a news analyzer project has broad prospects and great practical value. JavaScript and Ajax is eventually chosen to visualize the analyze result in web based front end, and some knowledge about convolutional neural network and python are also used in backend with natural language processing. What's more, a series of tools are used such as Anaconda, Bracket, Spyder, Flask, etc. Finally, the project can grab and analyze every relevant news average sentiment by typing keywords on the web interface, then markers could be pointed on the map with different colors to show the location and sentiment.**

***Index Terms*—sentiment analyze; python; JavaScript**

I. Introduction

Natural language processing is an important direction in the field of computer science and artificial intelligence. It studies various theories and methods that enable effective communication between humans and computers using natural language. Natural language processing is a science that integrates linguistics, computer science, and mathematics. Therefore, research in this field will involve natural language, that is, the language that people use every day, so it is closely related to the study of linguistics, but there are important differences. Natural language processing is not a general study of natural language, but rather the development of computer systems, especially software systems, that can effectively implement natural language communication. [1] The semantic analyze which belongs to NLP is the key to analyze the sentiment that the specific news present in average.

A. Architecture

Overall Architecture

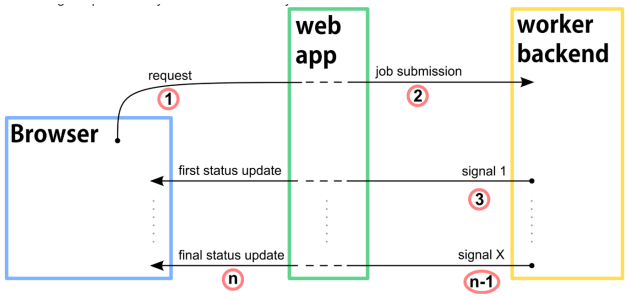

Fig. 1. Overall structure of the news analyzer

On top of that sit the Ajax and JavaScript frontends (with more to come). The Event Registry API provides a simpler interface for commonly used layers in web news scrawler. And finally, several commonly used models and Geocode API are ready to use out of the box, with more to come in future.

Front End Architecture

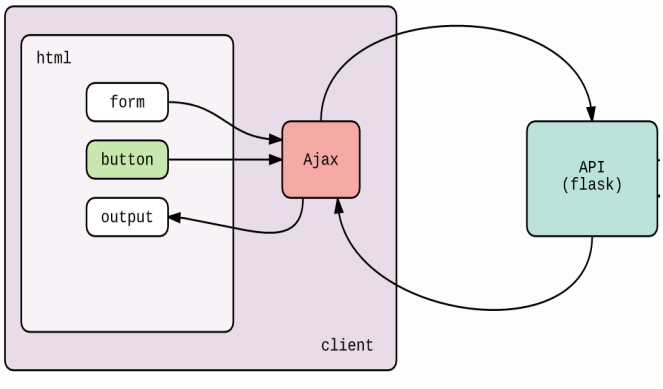

Fig. 2. Front end procedure

B. Platform

Operating System:

Windows 10 64bit

HTML

HTML is the standard markup language for creating Web pages. Its stands for Hyper Text Markup Language. Its describes the structure of a Web page. It consists of a series of elements. Its elements tell the browser how to display the content. Its elements are represented by tags
Its tags label pieces of content such as "heading", "paragraph", "table", and so on. Browsers do not display the Its tags, but use them to render the content of the page. [2]

python

Python is a high-level, interpreted and general-purpose dynamic programming language that focuses on code readability. It usually works well for imperative and object-oriented functional programming. It has a comprehensive and large standard library that has automatic memory management and dynamic features. [3]

Flask

Flask is a lightweight WSGI web application framework. It is designed to make getting started quick and easy, with the ability to scale up to complex applications. It began as a simple wrapper around Werkzeug and Jinja and has become one of the most popular Python web application frameworks.[4]

Flask offers suggestions but not enforce any dependencies or project layout. It is up to the developer to choose the tools and libraries they want to use. There are many extensions provided by the community that make adding new functionality easy.

Event registry

Event registry is a media intelligence platform that enables companies to analyze current and past news content in the world. [5]

JSON

JSON(JavaScript Object Notation) is an open-standard file format or data interchange format that uses text that is easy for human to read to transmit data objects that consist attribute-value paires and array data types.

Anaconda

Distribution of python language aiming at simplifying package managing the deployment.[6]

Echarts

ECharts, a pure JavaScript chart library, can run smoothly on PC and mobile devices, compatible with most current browsers (IE8/9/10/11, Chrome, Firefox, Safari, etc.), the underlying rely on lightweight Canvas class library ZRender, provide intuitive, vivid, interactive, highly customized data visualization chart.[7]

II. FRONTEND DESIGN

The front-end of the program is designed with JavaScript, Ajax and HTML, as well as Google map API to show the map.HTML is used to display the elements on the interface, including text “Enter Keyword”, “News Details”, the search bar, as well as the submit button. Ajax is used to connect front-end and back-end and transmit parameters.

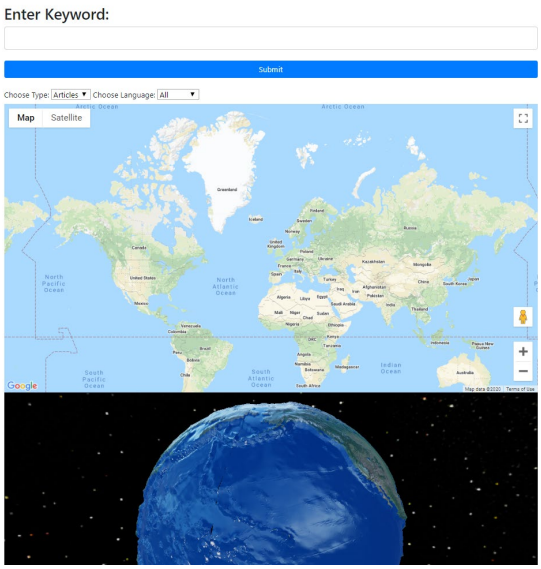

Fig. 3. Basic web interface with search bar, map and 3D earth

Except users can input the keyword they want to search, users can also choose the type of the news that they want to read, and the language users want. The condition of type and language will be sent to back-end with the keyword together.

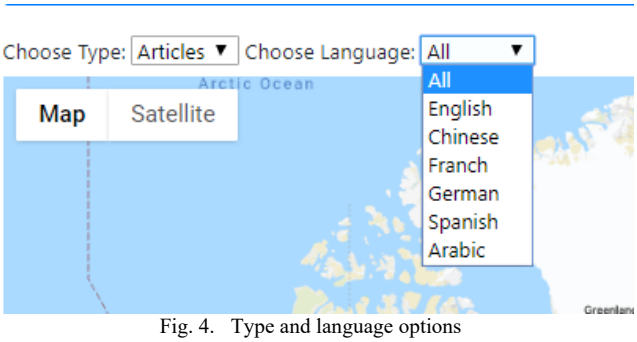

Fig. 4. Type and language options

Here is the code:

Ajax part:

```
    let keywordData = {
    keyword: keyword,
    language: lang,
    type : type
  }
  $.ajax({
     crossDomain: true,
    url: 'http://127.0.0.1:5000/',
    datatype: "json",
    data: keywordData,
    type: 'GET',

    success: function (data) {
      //成功后回调
        console.log(data);
      if(type_gl == "event")
            getJSON_event(data);
       else
            getJSON_article(data);
      showEarth();
    },
    error: function (jqXHR,textStatus,errorThrown) {
      //失败后回调
      console.log("You can not send Cross Domain AJAX requests: "+errorThrown);
    },
  })
}
```

Fig. 5. Get parameters and send to back-end

Google map initialization part:

```
var map, infoWindow;
function initMap() {
  map = new google.maps.Map(document.getElementById('map'), {
    center: {lat: 54.387267, lng: -7.823077},
    zoom: 2
  });
  infoWindow = new google.maps.InfoWindow;
}
```

Fig. 6. Initialize Google map

Earth construction part:

```
$.getJSON('http://127.0.0.1:51379/population.json', function (data) {

    data = data.filter(function (dataItem) {
        return dataItem[2] > 0;
    }).map(function (dataItem) {
        return [dataItem[0], dataItem[1], Math.sqrt(dataItem[2])];
    });

    myChart.setOption({
        visualMap: {
            show: false,
            min: 0,
            max: 60
//            inRange: {
//                symbolSize: [1.0, 100.0]
//            }
        },
        globe: {

            environment: 'http://127.0.0.1:51379/starfield.jpg',
            baseTexture: 'http://127.0.0.1:51379/earth.jpg',
            heightTexture: 'http://127.0.0.1:51379/bathymetry_bw_composite_4k.jpg',

            displacementScale: 0.1,
            displacementQuality: 'high',

            globeOuterRadius: 108,

          //  baseColor: '#000',

            shading: 'realistic',
            realisticMaterial: {
                roughness: 0.2,
                metalness: 0
            },

            postEffect: {
                enable: true,
                bloom : 300,
                intensity: 1
//                depthOfField: {
//                    focalRange: 20,
//                    enable: true,
//                    focalDistance: 160
//                }
            },
            temporalSuperSampling: {
```

Fig. 7. Show Earth function using Echarts

The JavaScript part is to send and receive data from users, and process the data in front-end. when the button is clicked, the keyword input by the users will be sent to the back-end, and receive the data returned from back-end. If this process was succeed, then the function getJSON(data) will be called, the parameter is the data returned from the back-end. However, the data returned from back-end is Json String, so we transformed Json String to Json List.

```
function getJSON_event(data){
 var obj = $.parseJSON(data);
   console.log(obj.events.results[0].title.eng);
     console.log(obj.events.results);
```

Fig. 8. Json String to Json List

Functions:

All of the functions are written in two version, event and article, to process different kinds of data respectively.

GetJSON( ):

This is the function getJSON(data), which is used to output the data to the interface. If the data type is events, then in a "for" loop, the title, the date of event and the event summary will be shown on the interface. If the data type is article, then the title, body, post time, and sentiment will be shown on the interface.

```
function getJSON_article(data){
     var obj = $.parseJSON(data);
   console.log(obj.articles.results);
  console.log(typeof(obj.articles.results[0].url));
  //  document.write(obj.articles.results[0].uri);
    let str = "";
    var i = 0;
     for (i=0; i<obj.articles.results.length; i++){
          str = str + "Title: " + "\n"
          str = str + obj.articles.results[i].title + "\n\n";
          str = str + "Body: " + "\n";
          str = str + obj.articles.results[i].body + "\n\n";
          str = str + "Post time: " + "\n";
          str = str + obj.articles.results[i].dateTime + "\n\n";
          str = str + "Sentiment: " + "\n";
          str = str + obj.articles.results[i].sentiment + "\n\n\n\n";

     }

   document.getElementById("demo").innerText=str;
   pointMap_article(obj);

}
```

Fig. 9. getJSON_article to show title, body, post time and sentiment

PointMap( ):

This is the function point map, which is used to point marks on the map. This function is provided by Google map API. Before point the map, we need to define the icon of the markrs and the information box, which is the box that show the link of the news when the marker is clicked. Point map function also point the different color of icon on the map by calculating the sentiment value, if the value is greater than 0, then the news is considered positive, smaller than 0 is considered negative.

```
function pointMap_article(obj){
    initMap();
    for (i=0; i < obj.articles.results.length; i++) {
      if(obj.articles.results[i].location != null){
        var position = new google.maps.LatLng(obj.articles.results[i].location.lat, obj.articles.results[i].location.long);

        if (obj.articles.results[i].sentiment > 0)
            var icon=icons['positiveP'].icon;
        else
            var icon=icons['negativeP'].icon;

      var marker = new google.maps.Marker({
        position: position,
        icon: icon,
        map:map,
        title: "click to open: "+obj.articles.results[i].title + '\n' + "url: " + obj.articles.results[i].url
      });
        url[i] = '<div id="content">'+
'<div id="siteNotice">'+
'</div>'+
'<div id="bodyContent">'+
'<p><b>'+obj.articles.results[i].title+'</b></p>'+'<a href="'+obj.articles.results[i].url+'" target="_blank" rel="noopener noreferrer">'+
obj.articles.results[i].url+'</a>'+
'</div>';

      }else
          continue;

      attachSecretMessage(marker, url[i]);
    }
```

Fig. 10. Point_map function point the location and fulfill the data in a information box

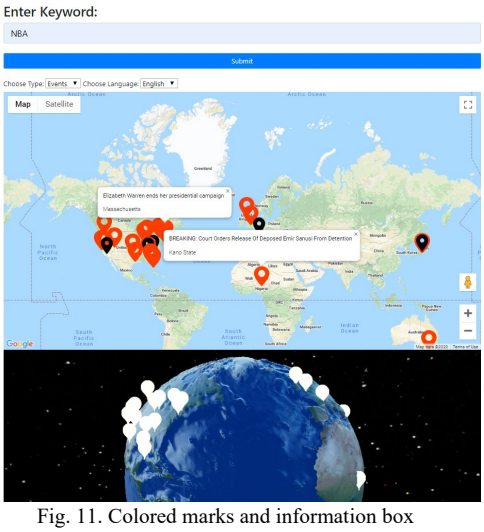

Fig. 11. Colored marks and information box

showEarth():

This is the function of 3D earth showing, the frame comes from Echarts, we only need to put the data on this earth.

```
function showEarth(){


    var myChart = echarts.init(document.getElementById('earth'));

// option = {
  $.getJSON('http://127.0.0.1:51379/population.json', function (data) {

    data = data.filter(function (dataItem) {
        return dataItem[2] > 0;
    }).map(function (dataItem) {
        return [dataItem[0], dataItem[1], Math.sqrt(dataItem[2])];
    });

    myChart.setOption({
        visualMap: {
            show: false,
            min: 0,
            max: 60
//            inRange: {
//                symbolSize: [1.0, 100.0]
//            }
        },
        globe: {

            environment: 'http://127.0.0.1:51379/starfield.jpg',
            baseTexture: 'http://127.0.0.1:51379/earth.jpg',
            heightTexture: 'http://127.0.0.1:51379/bathymetry_bw_composite_4k.jpg',

            displacementScale: 0.1,
            displacementQuality: 'high',

            globeOuterRadius: 108,

          //  baseColor: '#000',

            shading: 'realistic',
            realisticMaterial: {
                roughness: 0.2,
                metalness: 0
            },

            postEffect: {
                enable: true,
                bloom : 300,
```

Fig. 12. 3D earth showing code

Results show:

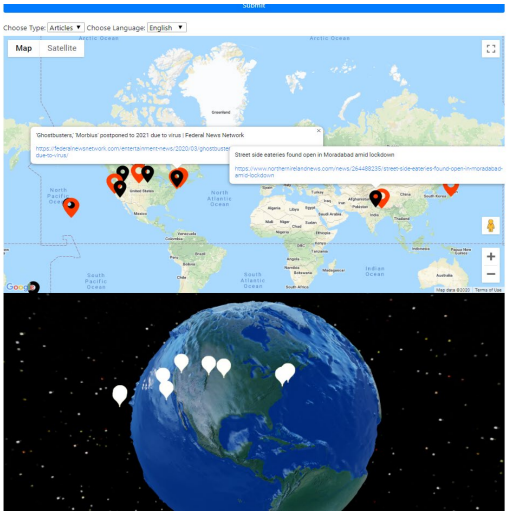

Fig. 13. Results show – visual part

Title:
Freight service still remains a challenge, no food on expressways | Gurgaon News - Times of India

Body:
Gurgaon: While the Centre's decision to allow inter-state movement of both essential and non-essential goods during the national lockdown seems to have come as a major relief to thousands of stranded truckers, transporters say keeping the continuity in the freight service still remains a major challenge.

In fact, transporters say truck drivers are still not too keen on hitting the roads anytime soon -- lack of food and medical services along the highway and harassment at the hands of cops on the ground being the major deterrents. This, transporters claimed, will further stress the industry that is already facing economic crunch.

Moreover, the villagers along the highway are apprehensive that many of these trucks are not being sanitised and drivers are not tested on borders, thus, stoking concerns about amplifying virus footprint in their community, according to the transporters.

"The information about the movement of essential and non-essential goods without discrimination has not yet reached to the officials at the ground-level and it's leading to harassment, extortion and beatings of truck drivers," said Kultaran Singh Atwal, president of All India Motor Transport Congress (AIMTC).

Other teething problem faced by the transporters is the scarcity of drivers. Humiliated and harassed by the local police and stranded on the highways for days bereft of food, water and safety has discouraged the truckers to take up new assignments.

"Thousands of vehicles are abandoned by the drivers at whatever places possible. This type of insensitivity has disheartened drives. There could be of scarcity of drivers as well vehicles for carrying essential goods unless immediate pragmatic solutions are not worked out,"noted Naveen Kumar Gupta of AIMTC.

Many truck drivers that TOI spoke to have pledged to not take the highways unless the government restores confidence among them about police support and availability of food supply. "If we all have to die only, then it's better we die with the family," said Mewat-based trucker Akhtar Khan.

Transporters claimed a pragmatic solution to support the industry and the drivers is much needed at this time when the cities can't afford to run out of essentials.

"Reorganising workforce is a major challenge. The driver no longer wish to continue. Further, national and state permits, fitness, road tax, and insurance are getting expired and movement without them during lockdown would be nearly im-possible. There should be rebate on these things," Ashok Sharma of Bhai Chara All India Truck Operator We-lfare Association (BAITOWA), said.

"Further, at every 200km, there should be provision for food and water supply. Without all this, the lost confidence among the drivers cannot be restored and they won't get down on the highways to keep the wheels moving," he added.

Residents of some villages along the highways are also apprehensive.

"Our district is basically truckers' district. And many truckers from other states and regions also at times take a halt over here. But, now we are in a dilemma. These trucks are not cleaned and drivers were saying there's no testing done on them either. We are worried about our safety too. If the virus spreads here ,it will take down all of us in one go," said Zubair, a Mewat resident.

Post time:
2020-03-31T02:14:00Z

Sentiment:
-0.3411764705882353

Fig. 14. Results show – text part

III. Backend Design

Parameters receiving

Firstly, back-end receive the parameter sent from front-end. There are three parameters, keyword, language and type.

```
def search():

    if request.method == 'GET':
            key = request.args.get("keyword")
            lang0 = request.args.get("language")
            typ = request.args.get("type")
```

Fig. 15. Parameter receiving

After received the parameter, if the type is event, then call the event searching API to grab event on the Internet based on the condition of language and keyword.

```
    if typ == "event":
        q = QueryEvents(lang = lang0,
                        keywords = key
                          #  sourceLocationUri  != None,
                        #dataType = ["news"])
                       )

        q.setRequestedResult(RequestEventsInfo(
returnInfo = ReturnInfo(
 #   articleInfo = ArticleInfoFlags(location = True),
    locationInfo = LocationInfoFlags(geoLocation = True)
    )))
        res = er.execQuery(q)
#        datalist = [[]*3]*100
#        i = 0
        k = 0
        str0 = "["
        for art in res['events']['results']:
           if res['events']['results'][k]['location'] :
# ================================================================
#                print(type(res['articles']['results'][k]['location']['lat']))
#                templist = [res['articles']['results'][k]['location']['lat'],res['articles']['resu
#                datalist[i] = templist
#                i = i + 1
# ================================================================
                str0 = str0 + "[" + str(res['events']['results'][k]['location']['long'])+","+str(r
           k = k + 1
        str0 = str0[:-1]
        str0 = str0 + "]"
        fileObject = open("C:/xampp/htdocs/map/population.json", 'w')
#   for num in datalist:
        fileObject.write(str0)
        fileObject.close()
    # print(str0)
```

Fig. 16. Event searching

If the type is article, then call the article searching API to search articles on the Internet with the condition of language and keyword.

```
    if typ == "article":

        q = QueryArticles(lang = lang0,
                        keywords = key
                          #  sourceLocationUri  != None,
                        #dataType = ["news"])
                       )

        q.setRequestedResult(RequestArticlesInfo(
returnInfo = ReturnInfo(
    articleInfo = ArticleInfoFlags(location = True),
    locationInfo = LocationInfoFlags(geoLocation = True)
    )))
        res = er.execQuery(q)
#        datalist = [[]*3]*100
#        i = 0
        k = 0
        str0 = "["
        for art in res['articles']['results']:
           if res['articles']['results'][k]['location'] :
# ================================================================
#                print(type(res['articles']['results'][k]['location']['lat']))
#                templist = [res['articles']['results'][k]['location']['lat'],res['articles']['resu
#                datalist[i] = templist
#                i = i + 1
# ================================================================
                str0 = str0 + "[" + str(res['articles']['results'][k]['location']['long'])+","+str
           k = k + 1
        str0 = str0[:-1]
        str0 = str0 + "]"
        fileObject = open("C:/xampp/htdocs/map/population.json", 'w')
#   for num in datalist:
        fileObject.write(str0)
        fileObject.close()
```

Fig. 17. Article searching

Finally, return the Json data that collected by the API from Internet to front-end. If we can see “success” in console, means that we have successfully get the news and returned to the front-end.

```
    print(key)
    print(lang0)
    print(typ)
    print("sucess!")
    return json.dumps(res,ensure_ascii=False)
    return "ok"
```

Fig. 18. Data returning

IV. Conclusion

This project has basically completed the expected objectives. We accomplished the functions of web news collection and sentiment analyze, and data visualization. With web interface, user could type any keyword they want to find the most relevant news, and choose the type they want to search, events or articles. Then, the location of the article publishing or the events happened will be marked on the map and 3D earth with sentiment analysis. If the marker is red, the news is positive, if the marker is black, the news is negative. By click the marker, users can see a information box with the title of the article or event, and the link of the news could be open. Also, users can read the content of articles or basic information in the area of news details.

There are still many parts that can be improved in this project:

1. We can improve UI interface to offer better user experience.
2. We should give user the limit news amount setting to let user decide how many news they want to see and set maximum amount with short response time.

Appendix

Appendixes are included in a separate zip file.

References

[1] “Modern Deep Learning Techniques Applied to Natural Language Processing,” Modern Deep Learning Techniques Applied to Natural Language Processing by Authors. [Online]. Available: https://nlpoverview.com/. [Accessed: 28-Feb-2020].M. Solutions, “Advantages and Disadvantages of Python Programming Language,” Medium, 24-Apr-2017. [Online]. Available: https://medium.com/@mindfiresolutions.usa/advantages-and-disadvantages-of-python-programming-language-fd0b394f2121. [Accessed: 01-Dec-2019].

[2] HTML Tutorial. [Online]. Available: https://www.w3schools.com/html/. [Accessed: 28-Feb-2020].Keras Tutorial: The Ultimate Beginner's Guide to Deep Learning in Python. (2019, January 25). Retrieved from https://elitedatascience.com/keras-tutorial-deep-learning-in-python.

[3] M. Solutions, “Advantages and Disadvantages of Python Programming Language,” Medium, 24-Apr-2017. [Online]. Available:

[4] “Welcome to Flask¶,” Welcome to Flask - Flask Documentation (1.1.x). [Online]. Available: https://flask.palletsprojects.com/en/1.1.x/. [Accessed: 28-Feb-2020].

[5] G. Leban, “Use the power of AI to turn raw news content into actionable insights,” Event Registry. [Online]. Available: https://eventregistry.org/. [Accessed: 28-Feb-2020].

[6] Installation. (n.d.). Retrieved from https://docs.anaconda.com/anaconda/install/.

[7] https://www.echartsjs.com/zh/index.htmlhttps://towardsdatascience.com/simple-introduction-to-convolutional-neural-networks-cdf8d3077bac.