**중간 보고서**

**T-SA: Twitter keyword Search API based tweet Analysis**

**(트위터 키워드 검색 API기반 트윗 분석)**



과 목: 산학캡스톤디자인1

분 반: 01

담 당 교 수: 정현숙 교수님

팀 명: 브이아이(VI)

이 름(학 번): 이석준(20165072)

이윤혁(20165062)

배인규(20165073)

서재익(20144773)

목차

서론 ----------

1. 개발 동기 및 필요성

2. 기존 연구

3. 개발 목표

4. 시스템 명세

4 - 1. 시스템 동작 과정

4 - 2. 원본 데이터 SET

본론 ----------

5. 객체지향적 설계

5 - 1. 환경 및 설치과정

5 - 2. T-SA 흐름도

6. 구현(Python)

6 - 1. TwitterAPI.py

6 - 2. DBModule.py

6 - 3. Visual.py

7. 구현(DB)

7 - 1. DB/Table 정의서

8. 구현(Hadoop)

8 - 1. Sqoop(DB/HDFS 연동)

8 - 2. Map/Reduce

결론 ----------

9. 실행 화면

10. 참고문헌 및 사이트

부록 ----------

깃허브

**1. 개발 동기 및 필요성**

**2. 기존 연구**

|  |  |
| --- | --- |
|  |  |
| URL: <http://tweetrend.com/>  Input: 기간, 단어  Ouput: 기간별 언급 횟수 그래프,  트윗정보 | URL: <https://foller.me/>  Input: 사용자 아이디  Output: 사용자의 정보 |

**3. 개발 목표**

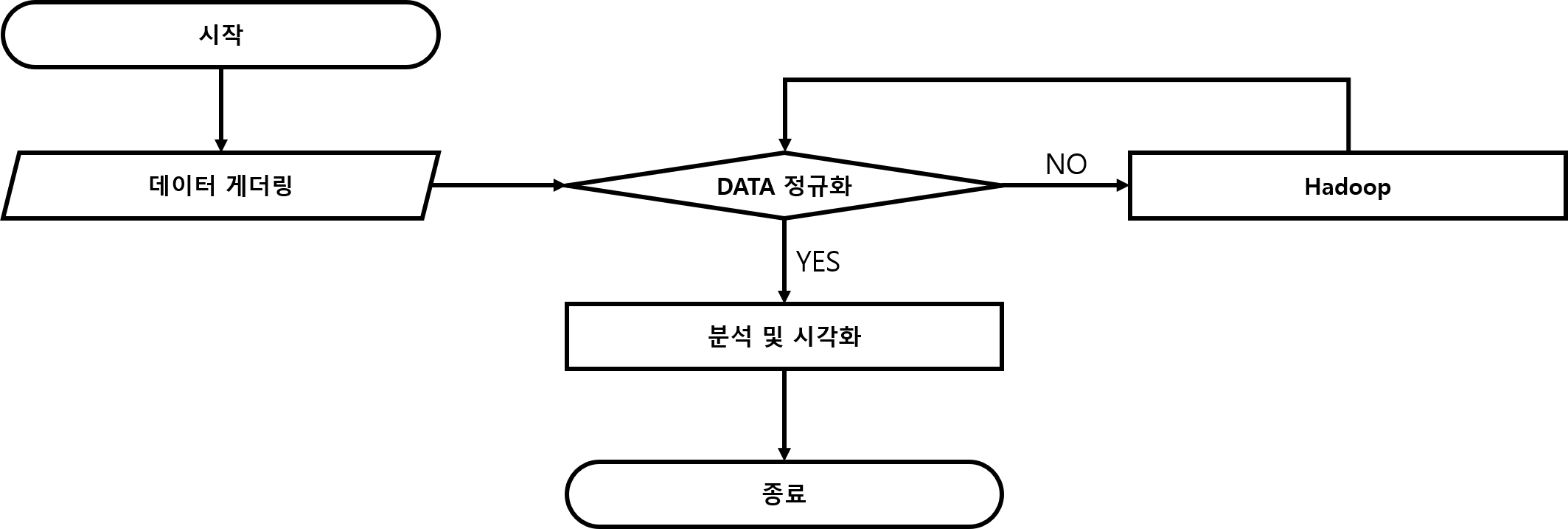
이 프로젝트에서는 개발 동기 및 필요성과 기존 연구를 토대로 2가지의 목표를 설정하였다.

첫 번째로는, 키워드와 기간을 설정하여 해당 기간동안 작성된 트윗을 가져와서 포함된 단어와 해시태그의 빈도수를 분석하고 시각화할 것이다.

두 번째로는, 사용자의 아이디혹은 이름을 통해 해당 사용자의 정보를 가져와서 분석하고 시각화할 것이다.

**4. 시스템 명세**

**4 - 1. 시스템 동작 과정**



**4 - 2. 원본 데이터 SET ()**

|  |
| --- |
| Status(\_api=<tweepy.api.API object at 0x7f6fd80f87b8>, \_json={'created\_at': 'Sun Apr 28 08:12:14 +0000 2019', 'id': 1122413198013911040, 'id\_str': '1122413198013911040', 'text': 'aaaaaaaaa 이윤혁', 'truncated': False, 'entities': {'hashtags': [], 'symbols': [], 'user\_mentions': [], 'urls': []}, 'metadata': {'result\_type': 'recent', 'iso\_language\_code': 'ko'}, 'source': '<a href="http://twitter.com" rel="nofollow">Twitter Web Client</a>', 'in\_reply\_to\_status\_id':None,'in\_reply\_to\_status\_id \_str':None,'in\_reply\_to\_user\_id':None,'in\_reply\_to\_user\_id\_str':None, 'in\_reply\_to\_screen\_name': None, 'user': {'id': 1107113829123555328, 'id\_str': '110711382912355 5328', 'name': 'Lee SeokJune', 'screen\_name': 'LSeokjune', 'location': '', 'description': '', 'url': None, 'entities': {'description': {'urls': []}}, 'protected': False, 'followers\_count ': 0, 'friends\_count': 0, 'listed\_count': 0, 'created\_at': 'Sun Mar 17 02:58:00 +0000 2019', 'favourites\_count': 0, 'utc\_offset': None, 'time\_zone': None, 'geo\_enabled': False, 'verified': False, 'statuses\_count': 8, 'lang': 'ko', 'contributors\_enabled': False, 'is\_translator': False, 'is\_translation\_enabled': False, 'profile\_background\_color': 'F5 F8FA', 'profile\_background\_image\_url': None, 'profile\_background\_image\_url\_https': None, 'profile\_background\_tile': False, 'profile\_image\_url': 'http://abs.twim g.com/sticky/default\_profile\_images/default\_profile\_normal.png','profile\_image\_url\_https': '[https://abs.twimg.com/sticky/default\_profile\_images/default\_profile\_nor mal.png',](https://abs.twimg.com/sticky/default_profile_images/default_profile_nor%20mal.png',) 'profile\_link\_color': '1DA1F2', 'profile\_sidebar\_border\_color': 'C0DEED', 'profile\_sidebar\_fill\_color': 'DDEEF6','profile\_text\_color':'333333','profile\_use\_backgro und\_image':True,'has\_extended\_profile':False,'default\_profile':True, 'default\_profile\_ima ge': True, 'following': False, 'follow\_request\_sent': False, 'notifications': False, 'translator\_type': 'none'}, 'geo': None, 'coordinates': None, 'place': None,'contributors':None,'is\_quote\_status':False,'retweet\_count':0,'favorite\_count':0,'favorited': Fals e,'retweeted':False,'lang':'ko'},created\_at=datetime.datetime(2019,4,28,8,12,14),id=1122413198013911040,id\_str='1122413198013911040',text='aaaaaaaaa이윤혁',tr uncated=False, entities ={'hashtags': [], 'symbols': [], 'user\_mentions': [], 'urls': []}, metadata={'result\_type': 'recent', 'iso\_language\_code':'ko'},source='TwitterWebClien t',source\_url='http://twitter.com',in\_reply\_to\_status\_id=None, in\_reply\_to\_status\_id\_str=None, in\_reply\_to\_user\_id=None, in\_reply\_to\_user\_id\_str=None,In\_reply\_to\_s creen\_name=None, author=User(\_api=<tweepy.api.API object at 0x7f6fd80f87b8>, \_json={'id': 1107113829123555328, 'id\_str': '1107113829123555328', 'name': 'L ee SeokJune', 'screen\_name': 'LSeokjune', 'location': '', 'description': '', 'url': None, 'entities': {'description': {'urls': []}}, 'protected': False, 'followers\_count': 0, 'friends\_co unt': 0, 'listed\_count': 0, 'created\_at': 'Sun Mar 17 02:58:00 +0000 2019', 'favourites\_count': 0, 'utc\_offset': None, 'time\_zone': None, 'geo\_enabled': False, 'verified': False, 'statuses\_count': 8, 'lang': 'ko', 'contributors\_enabled': False, 'is\_translator': False, 'is\_translation\_enabled': False, 'profile\_background\_color': 'F5F8FA', 'profile\_ba ckground\_imag e\_url': None, 'profile\_background\_image\_url\_https': None, 'profile\_background\_tile': False, 'profile\_image\_url': '[http://abs.twimg.com/sticky/defa ult\_profile\_images/default\_profile\_normal.png',](http://abs.twimg.com/sticky/defa%20ult_profile_images/default_profile_normal.png',)'profile\_image\_url\_https': '[https://abs.twimg.com/sticky/default\_profile\_images/default\_profile\_n ormal.png',](https://abs.twimg.com/sticky/default_profile_images/default_profile_n%20ormal.png',) 'profile\_ link\_color': '1DA1F2', 'profile\_sidebar\_border\_color': 'C0DEED', 'profile\_sidebar\_fill\_color': 'DDEEF6', 'profile\_text\_color': '333333', 'profile\_use\_background\_image': Tru e,'has\_extended\_profile': False, 'default\_profile': True, 'default\_profile\_image': True, 'following': False, 'follow\_request\_sent': False, 'notifications': False, 'translator\_type': 'none'}, id=1107113829123555328, id\_str='1107113829123555328', name='Lee SeokJune', screen\_name='LSeokjune', location='', description='', url=None, entities ={'description': {'urls': []}}, protected=False, followers\_count=0, friends\_count=0, listed\_count=0, created\_at=datetime.datetime(2019, 3, 17, 2, 58), favourites\_count= 0,utc\_offset=None,time\_zone=None,geo\_enabled=False,verified=False,statuses\_count=8,lang='ko',contributors\_enabled=False,is\_translator=False, is\_translation\_en abled=False, profile\_background\_color='F5F8FA', profile\_background\_image\_url=None, profile\_background\_image\_url\_https= None, profile\_background\_tile=False, profile\_image\_url='<http://abs.twimg.com/sticky/default_profile_images/default_profile_normal.png',>profile\_image\_url\_https='<https://abs.twimg.com/sticky/default_profile_images/default_profile_normal.png',> profile\_link\_color='1DA1F2', profile\_sideb ar\_border\_color='C0DEED',profile\_si debar\_fill\_color='DDEEF6', profile\_text\_co lor='333333',profile\_use\_background\_image=True,has\_extended\_profile=False,default\_profile=True,default\_profile\_image=True,following=False, follow\_request\_sen t=False,notifications=False,translator\_type='none'),user=User(\_api=<tweepy.api.API object at 0x7f6fd80f87b8>,\_json={'id':1107113829123555328,'id\_str': '1107113 829123555328', 'name': 'Lee SeokJune', 'screen\_name': 'LSeokjune', 'location': '', 'description': '', 'url': None, 'entities': {'description': {'urls': []}}, 'protected': False, 'followers\_count': 0, 'friends\_count': 0, 'listed\_count': 0, 'created\_at': 'Sun Mar 17 02:58:00 +0000 2019', 'favourites\_count': 0, 'utc\_offset': None, 'time\_zone': None, 'geo\_enabled': False, 'verified': False, 'statuses\_count': 8, 'lang': 'ko', 'contributors\_enabled': False, 'is\_translator': False, 'is\_translation\_enabled': False, 'profile\_backgro und\_color': 'F5F8FA', 'profile\_background\_image\_url': None, 'profile\_background\_image\_url\_https': None, 'profile\_background\_tile': False, 'profile\_image\_url': 'http://a bs.twimg.com/sticky/default\_profile\_images/default\_profile\_normal.png', 'profile\_image\_url\_https': '[https://abs.twimg.com/sticky/default\_profile\_images/default\_pr ofile\_normal.png',](https://abs.twimg.com/sticky/default_profile_images/default_pr%20ofile_normal.png',) 'profile\_link\_color': '1DA1F2', 'profile\_sidebar\_border\_color': 'C0DEED', 'profile\_sidebar\_fill\_color': 'DDEEF6', 'profile\_text\_color': '333333', 'profile\_us e\_background\_image':True,'has\_extended\_profile':False,'default\_profile':True,'default\_profile\_image': True, 'following': False, 'follow\_request\_sent': False, 'notifications': False,'translator\_type':'none'},id=1107113829123555328,id\_str='1107113829123555328',name='LeeSeokJune', screen\_name='LSeokjune', location='', description='', url=None,entities={'description': {'urls': []}}, protected=False, followers\_count=0, friends\_count=0, listed\_count=0, created\_at=datetime.datetime(2019, 3, 17, 2, 58), favourites\_count=0,utc\_offset=None,time\_zone=None,geo\_enabled=False,verified=False,statuses\_count=8,lang='ko',contributors\_enabled=False, is\_translator=Fals e,is\_translation\_enabled=False,profile\_background\_color='F5F8FA',profile\_background\_image\_url=None,profile\_background\_image\_url\_https=None, profile\_backgr ound\_tile=False,profile\_image\_url='<http://abs.twimg.com/sticky/default_profile_images/default_profile_normal.png',>profile\_image\_url\_https='<https://abs.twimg.com/sticky/default_profile_images/default_profile_normal.png',>profile\_link\_color='1DA1F2',profile\_sidebar\_border\_color='C0DEED',profile\_sidebar\_fill\_color='DDEEF6', profile\_text\_color='333333', profile\_use\_background\_image=True, has\_extended\_profile=False, default\_profile=True, default\_profile\_image=True, following=False, follow\_request\_sent=False, notifications=False, translator\_type='none'), geo=None, coordinates=None, place=None, contributors=None, is\_quote\_status=False, retweet\_count=0, favorite\_count=0, favorited=False, retweeted=False, lang='ko') |

- 하나의 row만 가져온 경우