

**Clustering Algorithm** 

구름 도시공학과 일반대학원

한양대학교

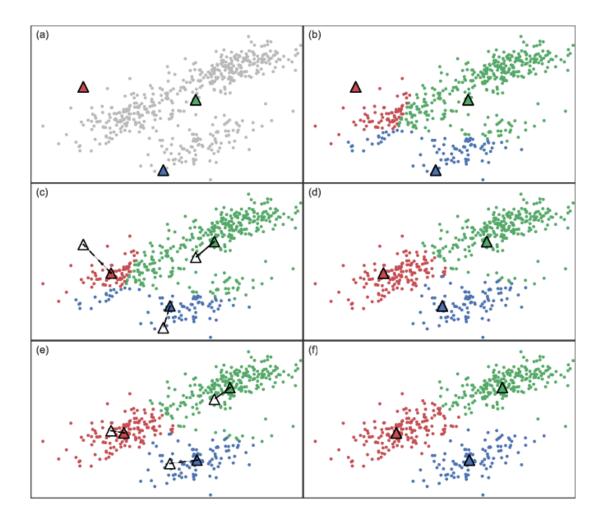
- 1. 클러스터링 알고리즘
- 2. 이미지 클러스터링
- 3. NLP 클러스터링

### Clustering (군집분석)

$$X(x_1, x_2, \dots, x_n) \to y$$

$$X(x_1, x_2, \dots, x_n)$$

## K-means 알고리즘



#### K-means 알고리즘

```
Input:

D= {t1, t2, .... Tn } // Set of elements

K // Number of desired clusters

Output:

K // Set of clusters

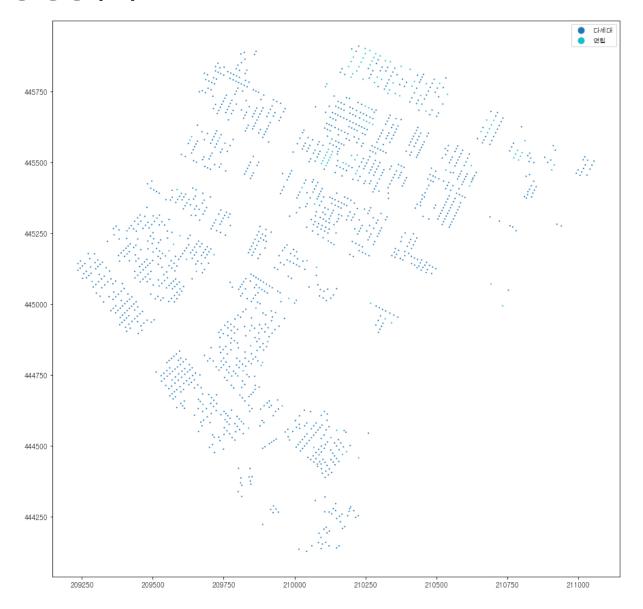
K-Means algorithm:

Assign initial values for m1, m2,.... mk

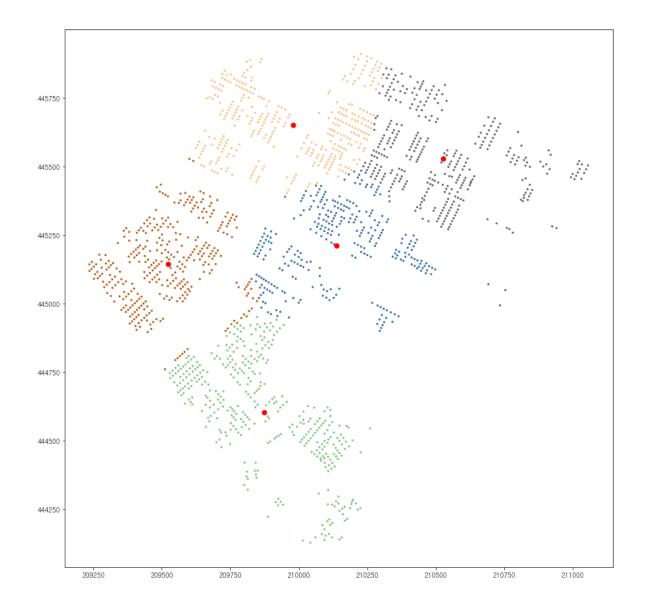
repeat

assign each item ti to the clusters which has the closest mean;
calculate new mean for each cluster;
until convergence criteria is met;
```

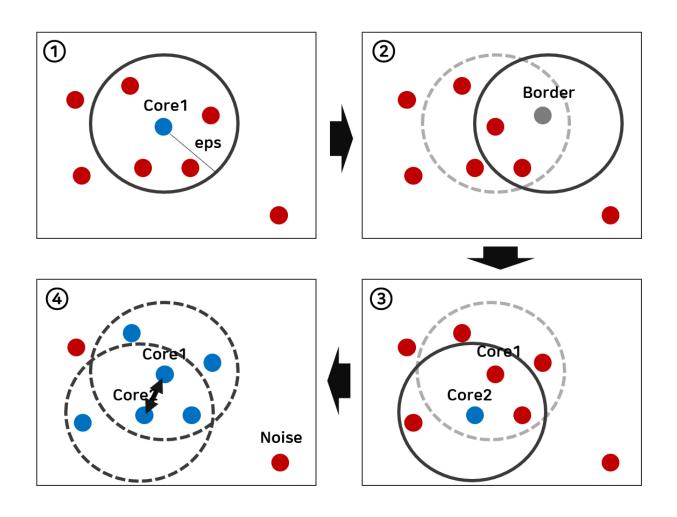
### 송파구 소형 공동주택



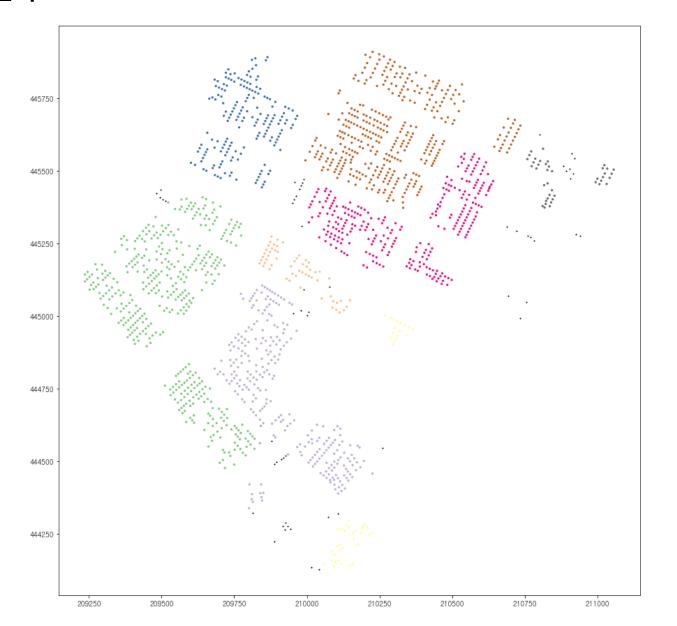
### K-means 알고리즘 결과



### DBSCAN(Density-based spatial clustering of applications with noise)



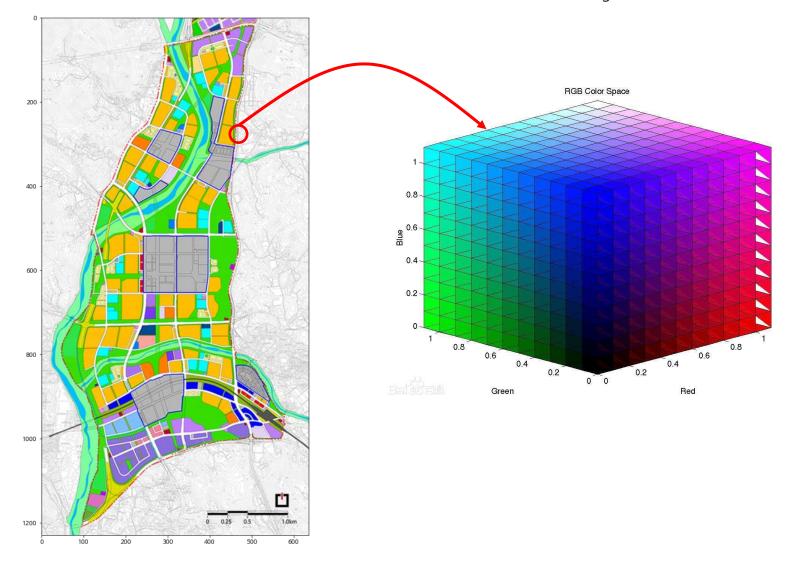
### DBSCAN 결과



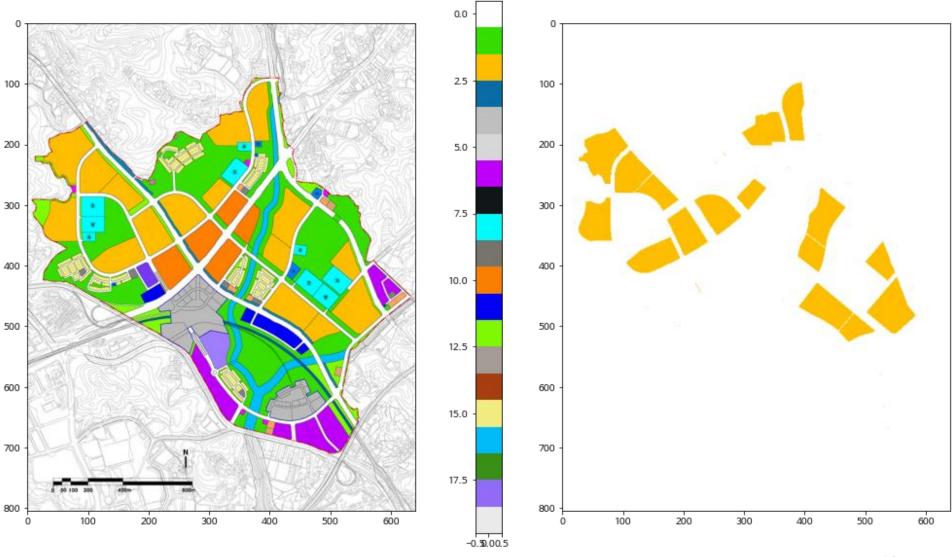
- 1. 클러스터링 알고리즘
- 2. 이미지 클러스터링
- 3. NLP 클러스터링

### K-means 알고리즘을 이용한 색깔 검출

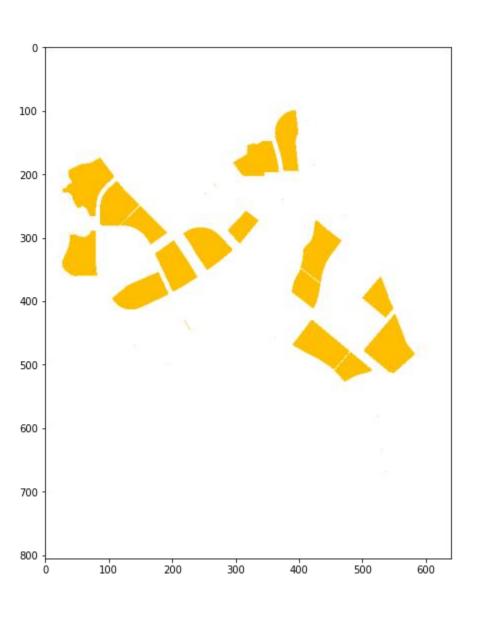
이미지의 각 픽셀은 RGB 3가지 색깔 공간의 특정 위치 좌표를 의미, 이를 이용하여 Clustering을 수행

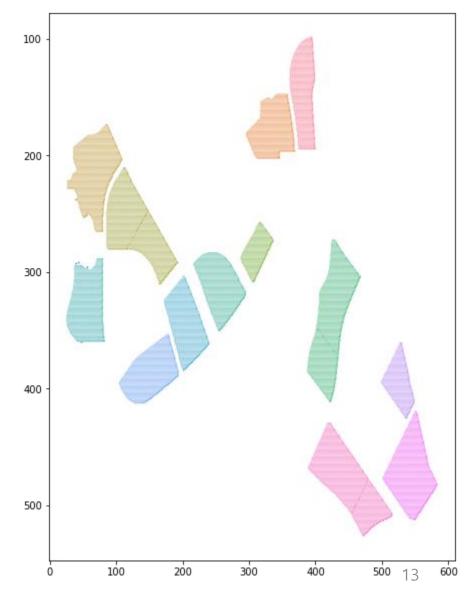


## Centroind 좌표의 색깔을 표시

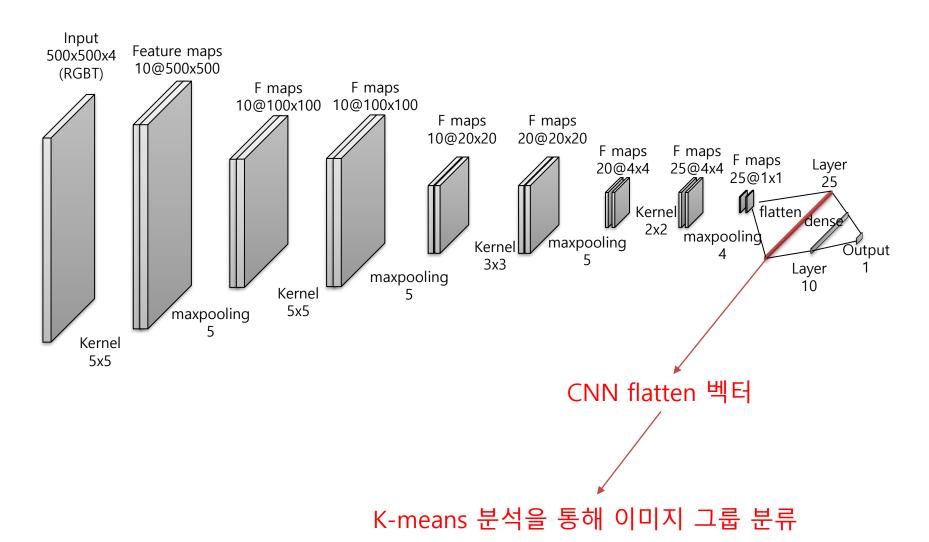


### DBScan으로 각 블록을 재 구분





#### CNN 학습 결과와 Kmeans 활용



## Flatten Layer 분석

25차원으로 데이터 축소 효과 유사한 벡터를 군집화 하면 이미지 차이 확인 가능

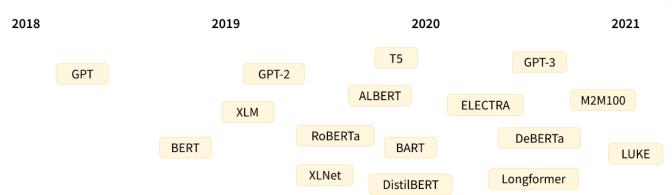
클러스터 1 클러스터 2 클러스터 3 클러스터 4 

- 1. 클러스터링 알고리즘
- 2. 이미지 클러스터링
- 3. NLP 클러스터링

## Transformers (신경망 언어모델 라이브러리) 강좌

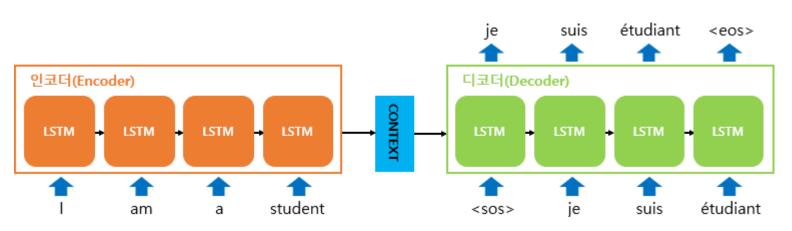


https://wikidocs.net/book/8056



## 딥 러닝을 이용한 자연어 처리 입문

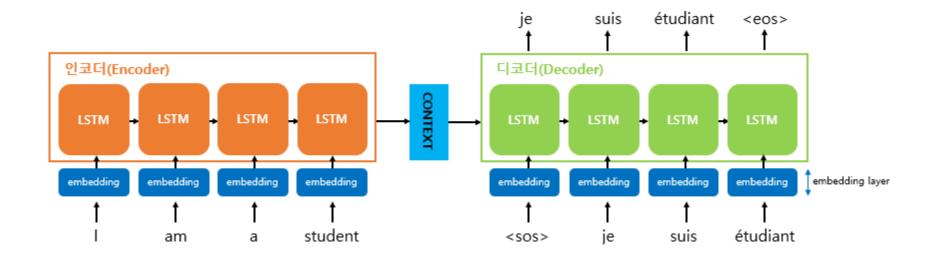
https://wikidocs.net/book/2155





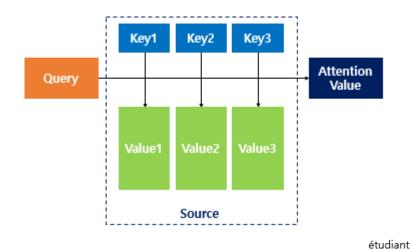
## LSTM Seq2Seq 모델

인코더와 디코더 형태



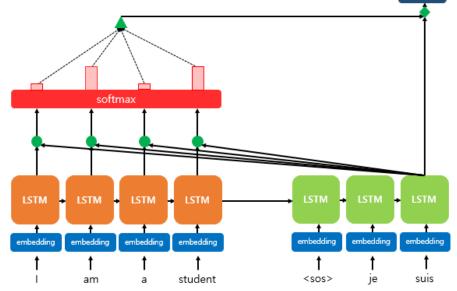
첫째, 하나의 고정된 크기의 벡터에 모든 정보를 압축하려고 하니까 정보 손실이 발생합니다. 둘째, RNN의 고질적인 문제인 기울기 소실(vanishing gradient) 문제가 존재합니다.

## 어텐션 함수(Attention Function)



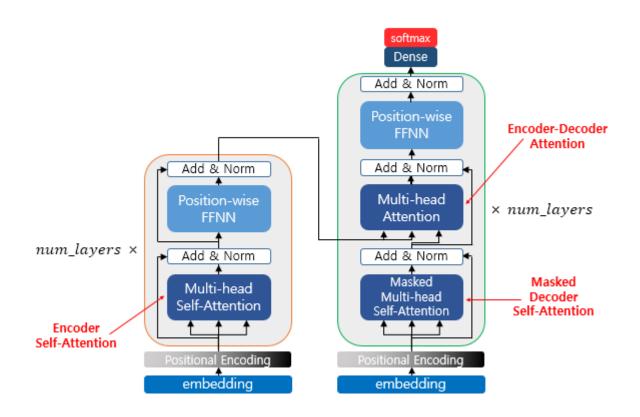
Dense

## 닷-프로덕트 어텐션(Dot-Product Attention)



출처: 딥 러닝을 이용한 자연어 처리 입문

# 어텐션(Attention)



## 인코더와 셀프 어텐션

