

Week 12 • 소셜네트워크 데이터마이닝과 분석

Advanced Text Analysis Topics & Machine Learning 2

Joonhwan Lee
human-computer interaction + design lab.

오늘 다룰 내용

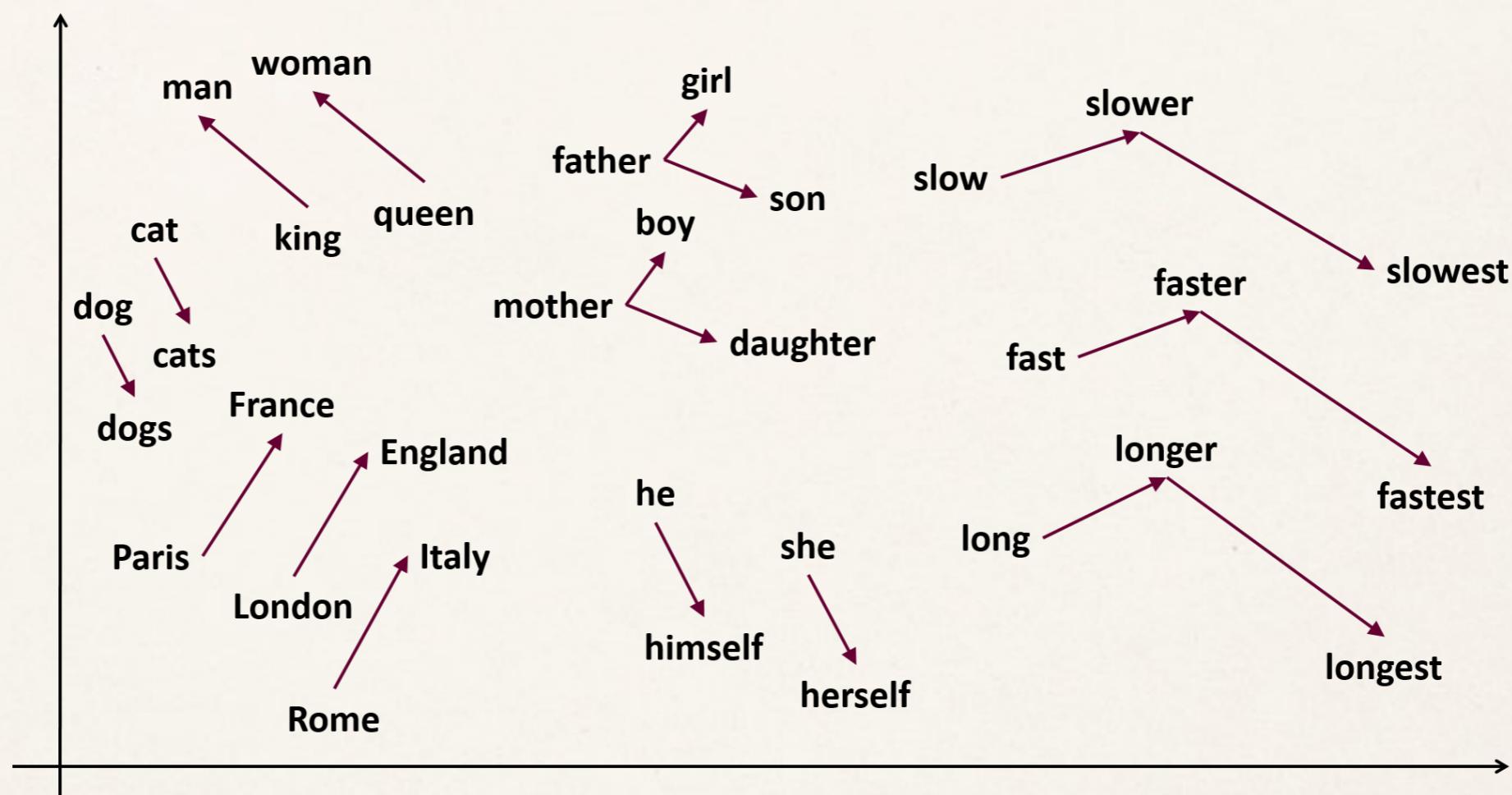
- Machine Learning의 기초
- Supervised Learning
- Unsupervised Learning

1. Word2Vec

Word2Vec

- ❖ Word2Vec 은 문장 내부의 단어를 벡터로 변환하는 도구
 - ❖ 단어의 연결구조를 기반으로 단어의 연관성을 벡터로 표현함
→ 단어의 의미를 파악하는데 도움이 됨
 - ❖ Co-occurrence 연결망 (one-hot encoding) → 도큐먼트 내에서 연결된 단어를 파악하는데 도움 (local structure)
 - ❖ 그러나 유사한 단어, 즉 “강아지”와 “댕댕이”는 서로 다른 벡터로 인식
 - ❖ 단어 간의 유사성을 파악하지 못함
 - ❖ 단어를 벡터로 구조화 하면 단어들 사이의 연관성을 파악할 수 있음

Word2Vec



Word2Vec

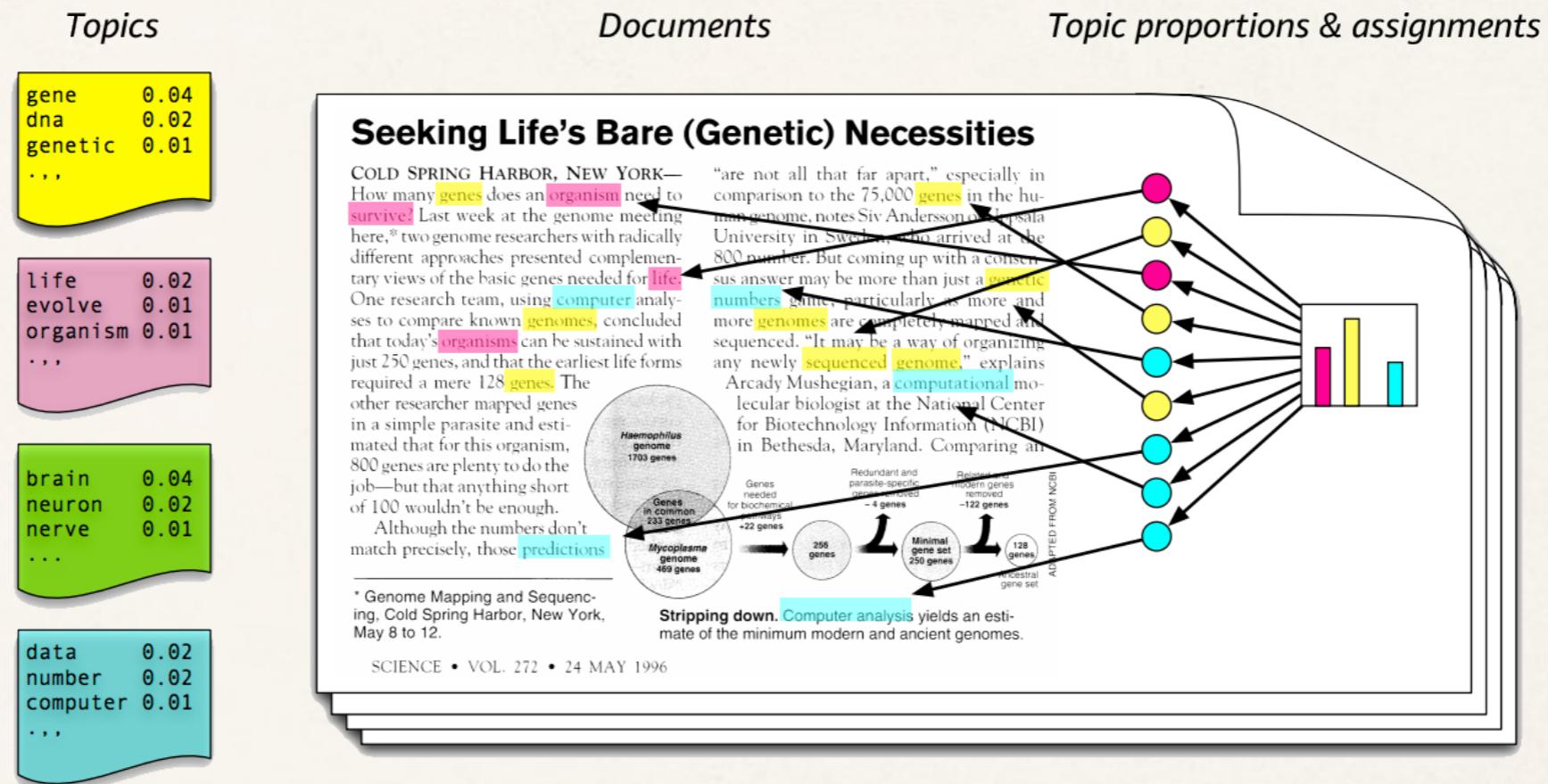
- ◆ 단어를 벡터화 하면 벡터의 연산이 가능
 - ◆ 아빠-남자+여자 = 엄마
 - ◆ 왕자-남성+여성 = 공주
 - ◆ <http://w.elnn.kr/search/> 참고

2. LDA (Latent Dirichlet Allocation)

LDA

- ❖ Document Model
 - ❖ Word
 - ❖ Document: collection of Words
 - ❖ Corpus: collection of documents
- ❖ Topic Model
 - ❖ Topic: collection of words
 - ❖ Document: representation of latent mixture of topics

LDA



- ◆ LDA는 특정 토픽에 특정 단어가 나타날 확률
- ◆ 노란색 토픽엔 gene이라는 단어가 등장할 확률이 0.04, dna는 0.02, genetic은 0.01 → 유전자 관련 주제임을 알 수 있음
- ◆ 문서 자체를 보면 빨강, 파랑 토픽 보다는 노란색의 비중이 많음
→ 따라서 이 문서의 메인 주제는 역시 유전자일 가능성이 높음

3. Google Cloud & AWS

Google Cloud Platform

The screenshot shows the Google Cloud Platform interface with a blue header bar. The header includes the 'Google Cloud Platform' logo, a dropdown menu, a search bar, and several icons. Below the header, a breadcrumb navigation shows '머신러닝' (Machine Learning) under a back arrow. A search bar contains the text 'API 및 서비스 검색'. On the left, there's a sidebar with '필터링 기준' (Filtering Criteria) and a '카테고리' (Category) section where '머신러닝' is selected. The main area displays eight search results cards:

Icon	Service Name	Provider	Description
	Cloud AutoML API	Google	Train high-quality custom machine learning models with minimum effort and machine learning...
	Cloud Machine Learning Engine	Google	An API to enable creating and using machine learning models.
	Cloud Natural Language API	Google	Provides natural language understanding technologies, such as sentiment analysis, entity...
	Cloud Speech-to-Text API	Google	Speech recognition
	Cloud Translation API	Google	Integrates text translation into your website or application.
	Cloud Video Intelligence API	Google	Detects objects, explicit content, and scene changes in videos. It also specifies the region for...
	Cloud Vision API	Google	Image Content Analysis
	Dialogflow API	Google	Builds conversational interfaces

- ♦ 구글 ML 서비스 API와 Storage를 제공

AWS

You can build custom deep learning models in the cloud using Amazon SageMaker, or use the collection of pre-trained models included with AWS DeepLens.



Object detection

Accurately detect and recognize objects.



Face recognition

Detect faces of people.



Activity recognition

Recognize more than 30 kinds of actions such as brushing teeth, applying lipstick, and playing guitar.



Artistic style transfer

Transfer the style from one image, such as painting, to an entire video sequence captured by your DeepLens in real-time.

You can explore the collection of AWS DeepLens projects contributed by the community of developers who have already been building with AWS DeepLens. You'll find a short video demo and in most cases a more detailed description and link to their project GitHub repo. See [Community projects](#).

AWS

You can build custom deep learning models in the cloud using Amazon SageMaker, or use the collection of pre-trained models included with AWS DeepLens.



Object detection

Accurately detect and recognize objects.



Activity recognition

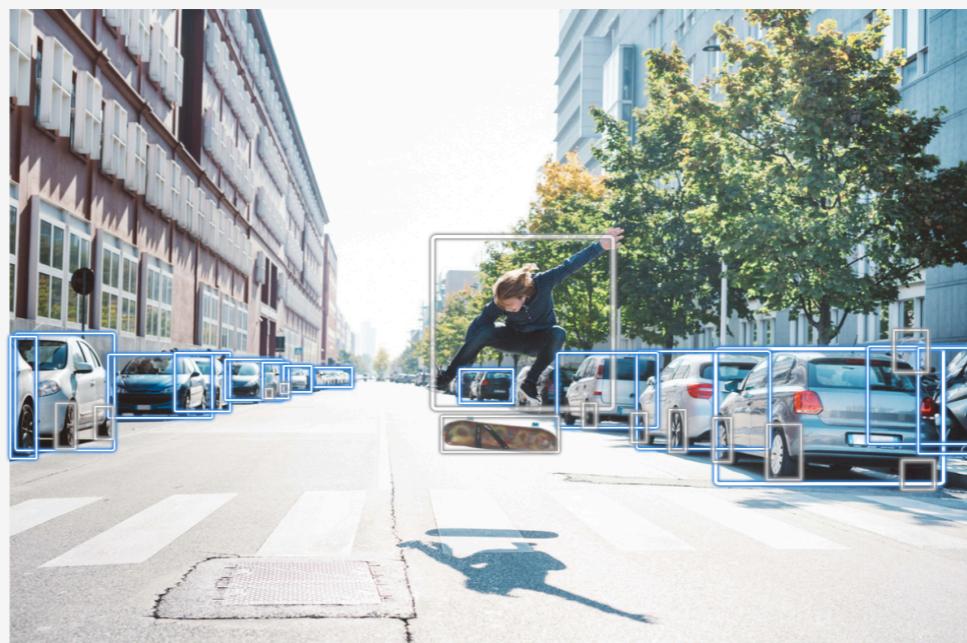
Recognize more than 30 kinds of actions such as brushing teeth, applying lipstick, and playing guitar.

You can explore the collection of AWS DeepLens projects built by developers who have already been building with DeepLens. You can view a live demo and in most cases a more detailed description and link to their project GitHub repo. See [Community projects](#).



객체 및 장면 감지

Rekognition은 이미지에 있는 피사체, 개념, 장면에 자동으로 레이블을 지정하고 신뢰도 점수를 제공합니다.



샘플 이미지 선택



자신만의 고유 이미지 사용

이미지는 .jpeg 또는 .png 형식이어야 하며 5MB 이하여야 합니다. 이미지가 저장되지 않았습니다.

업로드

또는 끌어서 놓기

이미지 URL 사용

결과

Transportation	98.8 %
Car	98.8 %
Automobile	98.8 %
Vehicle	98.8 %
Person	98.3 %
Human	98.3 %

자세히 표시

▶ 요청

▶ 응답

AWS Architecture



사용방법



Questions...?
