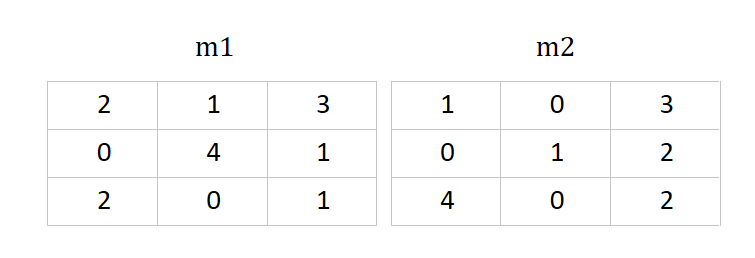
**Machine Learning & Data Mining, Spring 2020**

**Homework 8**

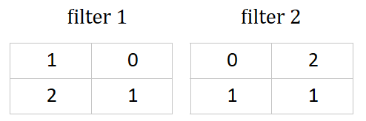
Due May 28

1. In Convolutional Neural Network (CNN), the -th output feature map at layer is given as follows.

Where is the input feature map and is the convolution filter (or kernel). Suppose a CNN layer has the following two feature maps at a certain layer:



1. Apply the convolution operation to the above feature maps using the following filters(2x2x2). Show the output feature maps (Assume stride = 1 and with zero-padding size 1) [5]

(i)m1 \* filter1 (ii)m2 \* filter2

|  |  |  |  |
| --- | --- | --- | --- |
| 2 | 5 | 5 | 6 |
| 0 | 6 | 10 | 5 |
| 2 | 4 | 5 | 3 |
| 0 | 2 | 0 | 1 |

|  |  |  |  |
| --- | --- | --- | --- |
| 1 | 1 | 3 | 3 |
| 2 | 1 | 9 | 2 |
| 4 | 6 | 6 | 2 |
| 8 | 0 | 4 | 0 |

* Output feature maps (m1\*filter1 + m2\*filter2)

|  |  |  |  |
| --- | --- | --- | --- |
| 3 | 6 | 8 | 9 |
| 2 | 7 | 19 | 7 |
| 6 | 10 | 11 | 5 |
| 8 | 2 | 4 | 1 |

1. Apply 2x2 max pooling operation to the output maps obtained in (a). Show the resulting layer. [5]

* Resulting layer

|  |  |
| --- | --- |
| 7 | 19 |
| 10 | 11 |

1. Unlike VGG-net or GoogleNet, ResNet can consist of more than 100 layers. Briefly explain what technology makes it possible and how. [10]

ResNet은 VGG-net 또는 GoogleNet과 다르게 residual connections 기술을 사용하여 very deep network가 가능하며, layer의 입력을 layer의 출력에 바로 연결시키는 skip connection을 사용하므로 100이 넘는 layer들로 구성이 가능하다.

1. Among deep learning-based methods, the CNN-based model in the computer vision field and the RNN-based model in the NLP field have good performance. For each field, describe at least one reason why the model has a good performance based on the characteristics of data. [10]

3-1. CNN-based model in the computer vision field

Computer vision 분야는 시각적 세계를 해석하고 이해하도록 컴퓨터를 학습시키는 인공지능 분야이다. 즉, 이미지에 대한 분석을 하는 분야이다. 이때, image의 size가 굉장히 큰 것도 존재한다.

CNN은 convolution 연산을 사용한다. 이때, 각 layer의 입출력 형상을 유지시키고 이미지의 공간 정보(데이터의 공간적 정보가 소실되지 않음)를 유지하면서 인접 이미지와의 특징을 효과적으로 인식한다. 즉, 이미지특화 모델이다. 또한 복수의 filter로 이미지의 특징을 추출 및 학습시킨다. 그리고 pooling을 통해서 data의 size를 줄이고 연산 속도를 향상시킨다. 이러한 이유로 CNN은 컴퓨터 비전 분야에서 좋은 성능을 보인다.

3-2. RNN-based model in the NLP field

NLP는 sequence 데이터이기때문에, 샘플들이 독립적이지 않고 순서가 중요하며, input이나 output이 가변적인 길이를 가질 수 있다. 이 때, RNN 기반 model을 사용하면 hidden layer 단계에서 이전 hidden layer 결과 값을 다음 hidden layer 입력 값으로 보내게 되면서, 이전 결과가 다음 결과 예측에 영향을 미친다. 따라서 종속적인 data이자 순서가 중요한 데이터에 적합한 모델이므로 NLP에서 좋은 성능을 보인다.