

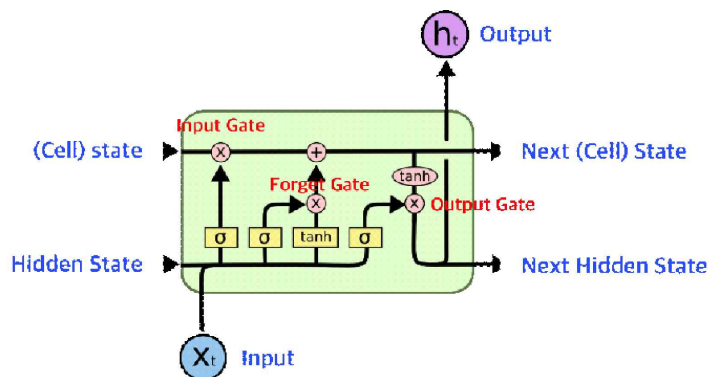
연구윤리 및 연구지도

(6주차 - RNN, LSTM)

경북대학교 전자공학부
2016113566 김남영

1) 책 6장의 내용 참고하여, RNN과 LSTM을 조사하고 차이점을 간단히 설명

RNN - 신경망을 시간 방향으로 펼치고 순환 경로를 포함하여 과거의 데이터를 기억할 수 있다. 따라서 시간의 흐름을 따르는 시계열 데이터가 입력으로 사용될 때 유용하게 사용할 수 있다. 모든 RNN 계층에 대한 상태를 기억한다면 메모리나 컴퓨팅 자원의 사용량이 증가하므로 여러 RNN 계층을 하나의 block으로 묶고, block 단위로 독립적으로 데이터의 전파를 수행한다. 하지만 단순한 구조의 RNN의 경우 역전파를 통해 기울기를 조정하는 단계를 거칠수록, 기울기가 작아져 소실되거나 혹은 너무 커져서 폭발하게 된다(tanh 함수에 의해 계속 값이 작아지거나, 기울기가 지수적으로 증가 또는 감소함) 즉, 시간적으로 멀리 떨어진 의존 관계를 가진 경우 잘 학습할 수 없다는 단점이 있다. 이런 단점을 보완하기 위해 게이트 구조가 더해진 RNN인 LSTM을 주로 사용한다.



LSTM - 위 그림과 같이 기존의 RNN과 달리 입출력 계층에 Cell state라는 경로가 추가적으로 존재한다. 이 계층은 LSTM 계층 내에서만 사용되고, 다른 계층으로 출력되지 않는다 (Output은 hidden state밖에 없음). 이 Cell state는 LSTM의 기억 메커니즘이며, 시각 t 에서의 기억이 저장되어 있다. 이 기억 셀이 게이트 역할을 하고, 이 게이트를 통해 이전 시각의 정보를 얼마나 수용할지(게이트를 얼마나 열지), 얼마나 잊을지(forget gate) 그리고 현재 시각의 정보를 얼마나 수용할지 등이 결정된다. 매 시각 다른 게이트 값을 가지므로 역전파시 곱셈의 효과가 누적되지 않아 기울기 소실이 일어나기 힘들다.

2) 본인이 관심있는 분야의 RNN(LSTM) 모델 사용 논문 2건 요약

1. 한국전자통신연구원, 휴먼케어 로봇과 소셜 상호작용 기술 동향, 2020.06.01.

현재까지 상용화된 휴먼케어 로봇은 간단한 상황에 단편적으로만 응답할 수 있는 수준이므로 아직까지 개인에게 맞춤형 서비스를 제공하기에는 한계가 있다. 따라서 이러한 로봇에 소셜 상호작용 기술을 적용하여 인간과 보다 자연스러운 유대관계를 형성하는 것은, 로봇 과학에서 중점적으로 연구되어야 할 분야이다. 소셜 지능은 크게 대인관계 지능과, 자기이해 지능으로

나뉜다. 휴먼케어 로봇의 경우에는 사람과의 유대 관계 형성이 중요하므로, 대인관계 지능이 필수적인 요소이다. 다양한 상황을 미리 예측하고 그에 따른 의사결정을 하기 위해서 로봇에 머신러닝, 딥러닝 기술이 활발히 도입되고 있다. 예를들어 LSTM 모델을 사용하여 과거의 장기적인 기억을 찾아 로봇 주도적인 발화와 행위를 할 수 있다. 현재까지는 입력과 출력까지 사람의 개입이 전혀 없지는 않고, 입력으로 어떤 데이터가 들어오면 출력으로는 사람이 지정해놓은 몇 개의 행위 중 하나를 선택하여 행동하는 방식이다. 최근에는 LSTM 기반으로 출력까지 사람의 개입없이 도출될 수 있도록 제스처 모델을 개발하고 있다.

(휴먼케어 로봇이나 언어지능에만 관심이 있다기보다, 다양한 로봇에 이런 딥러닝 기술이 활용되는 것에 초점을 두고 이 논문을 선정하였습니다.)

2. 한국전자통신연구원, 지능형 CCTV 기반 동적 범죄예측 기술 동향, 2020

지능형 CCTV의 영상을 분석하여 범죄 현장 상황에 보다 풍부한 정보를 획득하고, 경향성을 파악하여 범죄가 일어나기 전 선제적으로 예방하고자 하는 기술들이 활발히 연구 중이다. 본 논문에 의하면 날씨, 경제, 시간대, 총소리 등을 분석하여 범죄 발생 확률이 높은 지역을 예측하여 순찰을 강화하는 등의 기술이 활발히 개발 중이다. 또한, CCTV를 통해 안면인식, 사람 또는 차량 재식별 기술을 통해 범인 검거의 확률을 높이고 있다.

-> 이 논문은 제가 여름방학 때 아이디어 공모전을 준비하면서 읽어보았던 논문입니다. 이 논문을 보면서, 딥러닝 기술을 활용한 지능형 CCTV가 범인 검거에 탁월한 효과를 보이지만 위험상황을 미연에 방지하기에는 무리가 있다고 생각했습니다. 이 논문에 의하면, 이미지 인식을 통해 성별, 나이, 인상착의(모자나 선글라스 등을 착용했는지 등) 등으로 대상의 위험 정도를 추출하거나 범죄 전조 현상인 비명, 발자국 소리 등을 활용하는 모델이 활발히 연구 중이라고 합니다. 하지만 CCTV의 실시간 영상에서 대상의 이전 행동을 기억하고, 다음 행동을 예측할 수 있다면(RNN 모델을 적용하여) 단순 이미지나 음성 정보를 활용하는 것 보다 훨씬 더 예민하게 범죄 상황에 반응할 수 있을거란 생각에 이 논문을 선정하게 되었습니다.

3)고찰

1. 논문 (1)에서 휴먼케어 로봇을 입력에서 출력까지 전 과정을 사람의 개입없이 학습시키는 모델을 개발 중이라는 것을 알게 되었다. 하지만 모든 과정을 사용자에게 맡긴다면 잘못된 방향으로 학습되어 원하지 않은 서비스를 제공받는 경우가 생기지 않을까?

2. CCTV 영상은 단순 영상 데이터 이면서 시계열 데이터이므로, 누군가는 RNN 모델에 적용해봐야겠다고 생각을 했을텐데 왜 안했을까?