인공지능을 이용한 운동 상태 예측   
생성 모델 개발에 대한 연구

저자 30826 조한흠

연구 동기

자율주행 자동차의 수요 증가가 시사하는 바와 같이, 가까운 미래에 로켓, 드론, 휴모노이드 등 완전 자율화된, 고도의 엔지니어링을 필요로 하는 제품들의 수요가 증가할 것이라고 볼 수 있다. 하지만 현재 사용되는 로봇, 로켓과 같은 환경과의 물리적 상호작용을 통해 작동하는 물건들은 모두 복잡한 물리학적 연산에 의존하고 있다. 이러한 점은 제품이 예상 외의 상황에 대응하기 어렵게 만든다. 특히 로켓의 경우, 운전자의 요구에 따라 실시간으로 목적지를 재설정 하고, 궤도와 방향에 관련된 복잡한 물리학적 연산을 순식간에 처리해내는 기술은 아직 없다. 그러나, 인공지능 기술을 사용하여 짧은 시간 내에 제품과 환경의 상호작용으로 운동 상태가 어떻게 변할지 예측할 수 있다면, 곧 일어날 사건을 예측할 수 있기 때문에, 복잡한 구조를 가진 제품이(자동차, 로켓, 비행기, 로봇 등) 사용자의 요구에 실시간으로 반응하는 것이 가능해진다. 또한 이 기술을 응용하면 특정 장비나 설비의 수명을 예측하거나, 주행 물체가 현 상태를 유지하며 나아갈 때 사고가 날 확률을 계산하는 등 기존에 어렵게 하던 연산작업들을 모델 하나로 전부 대체할 수 있다.

본 연구에서 개발할 인공지능 알고리즘 모델 속엔 운동상태를 예측하는데 필요한 물리 규칙이사람이 이해 할 수 없는 복잡한 행렬 속 성분들의 형태로 생성된다. 이때, 그 내재적으로 생성된 충돌 규칙들 중 사람이 발견하지 못한 규칙도 포함될 가능성이 있다. 예를 들면, 무중력 공간에 질량을 가진 물체 3개가 있을 때, 그 물체들이 운동하는 모습을 기록한 데이터만 충분히 있다면, 인공지능이 그 물체들이 어떻게 운동할 지 예측해낼 수 있는 것이다. 이는 계산 및 해석은 가능하나, 매우 복잡한 프로세스를 가진 연산이다. 이 모델을 더욱 발전시키면 이와 같은 사람이 해석하지 못했거나, 복잡하고 어렵게 해석하던 물리 현상들도 쉽고 빠르게 해석해낼 수 있는 모델이 될 것이라고 기대한다.

이론적 배경:

본 연구에서 개발할 모델은 물리현상을 기록한 시퀀스를 예측된 시퀀스로 변환하는 seq2seq모델이다. Seq2seq 작업에서 가장 좋은 성능을 보이는 Transformer(변환기) 모델을 기반으로 삼아 응용한 모델을 사용할 것이다.

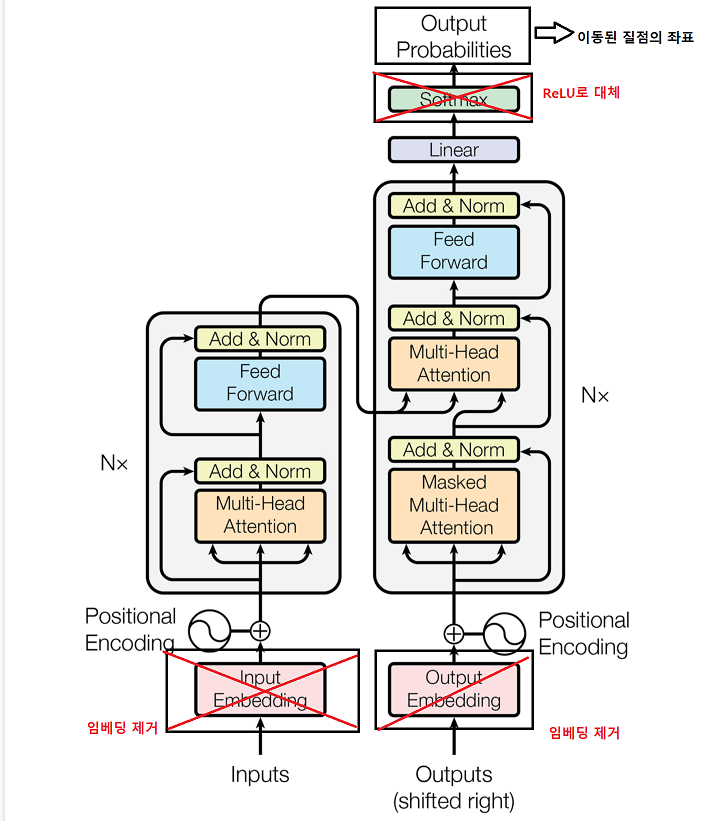
Seq2seq 모델은 여러가지가 있는데, 그 중에서 Transformer을 사용할 때의 장점은 다음과 같다:

1. 빠른 처리 속도 (병렬 가능 여부)  
   seq2seq은 시퀀스를 다루는 것이기에 각 성분의 순서가 중요하게 작용한다. 시퀀스의 순서는 말의 어순과 같은 것이다. 기존에 seq2seq작업을 위해 사용하던 RNN(Recurrent Neural Network, 순환신경망)이나 LSTM(Long Short Term Memory, 장단기 메모리)은 시퀀스의 각 성분의 순서를 반영하기 위해 성분을 순서대로 처리했다. 그러한 방법은 컴퓨터에서의 병렬연산을 불가능하게 만든다. Transformer 모델을 사용할 경우, Positional Encoding이라는 방법을 적용하므로 병렬 연산이 가능하다.   
     
   Positional Encoding은 다양한 방법으로 적용될 수 있는데, 주로 입력벡터의 성분에 미세한 순서를 나타내는 값을 더하는 방법으로 적용된다. 이때, Positional Encoding 값이 지나치게 크면 입력 벡터의 방향을 너무 크게 바꿔버린다는 문제가 있다. 고로 이러한 점 때문에 Positional Encoding은 1과 -1 사이의 값을 갖는 sin함수와 cos함수로 구현한다.
2. 시퀀스 길이 제한의 부재 (입력 시퀀스의 길이가 무제한)  
   기존의 seq2seq 작업을 구현할 때 사용하던 RNN과 LSTM은 한 번에 입력 받을 수 있는 시퀀스의 길이에 제한이 명확히 있었다. 입력 시퀀스 길이의 제한이 있다는 말은 곧, 한 번에 적용 가능한 맥락에 한계가 있다는 의미이다. 한 번에 처리할 수 있는 시퀀스의 길이를 window라고 한다. RNN의 window가 가장 작았고, 이를 조금 개선한 것이 LSTM이다. 허나 LSTM도 window의 크기가 그렇게 크지는 않았다.   
     
   Transformer는 window라는 개념을 사용하지 않는다. Transformer은 입력 시퀀스의 길이를 Positional Encoding에 사용된 sin함수와 cos함수의 주기를 이용해 결정한다. 주기가 길수록 입력 시퀀스의 길이가 늘어나는 것이다. 고로 Transformer 모델을 사용할 경우, 컴퓨팅 파워만 충분하다면 입력 시퀀스의 길이를 무한대로 늘릴 수 있다.
3. Self Attention을 사용한 상황 맥락 파악  
   Self Attention이란 입력으로 들어온 시퀀스 자체의 맥락을 파악하는 기법을 말한다. Self Attention은 그 기저에 다른 Attention기법들이 들어있다. Self attention의 수행 단계 중 제일 처음엔 Scaled Dot Product Attention이 있다. 이는 가장 기본적인 Attention인 Dot Product Attention(세개의 입력 시퀀스 사이의 관계를 나타내는 기법)의 결과물을 모델 고유의 차원 수(상수)의 양의 제곱근으로 나눈 것이다. 그렇게 하는 이유는, scaled dot product attention의 수행 직후 softmax를 이용해 normalize 하는데, 이때 dot product attention의 결과물이 너무 크면 softmax가 포화상태가 되어 제대로 값 산출이 불가능해지기 때문이다.   
     
   그 이후, Multihead Attention을 정의한다. 이는 단순히 Scaled Dot Product Attention을 여러 개 겹쳐 둔 것이다. ‘Attention Is All You Need’ 논문에 따르면 Multihead Attention을 사용함으로써 입력 값 사이의 관계를 더욱 잘 표현할 수 있었다고 한다.   
     
   최종적으로 Self Attention은 Multihead Attention의 입력 값 3개를 모두 하나의 시퀀스로 넣는 것을 말한다. Attention이 입력 값 3개의 관계를 찾는 과정이므로, Self Attention을 통해 한 시퀀스 자체의 맥락을 파악할 수 있다.

가설 및 연구 방법

가설: 물리현상을 기록한 자료가 있을 때, Transformer 모델을 응용한 모델을 사용하여 해당 자료를 분석하면 곧 일어날 현상을 예측할 수 있을 것이다.

연구 방법:

1. 데이터 확보  
   모델 학습에 사용될 데이터를 실제 실험상황에서 모으는 것은, 물리현상 예측 머신러닝 모델 구현이라는 주제에 견주어 볼 때 비효율적이다. 데이터 확보에 지나치게 시간을 많이 쓰게 되는 방법이기 때문이다. 데이터 확보의 시간적 효율성을 위해 물리엔진을 제작하여 데이터를 만들어 낼 것이다. 생성할 데이터의 세부사항은 다음과 같다  
   -2차원이고(이후 3차원으로 확장), 평면에 수직으로 들어가는 방향으로 중력 작용  
   -물체는 질점으로 설정, 총 2개의 물체가 평면에 있음  
   -물체는 질점이기에 충돌하기 어렵지만, 충돌 할 경우 일정한 반발계수로 탄성충돌 함 (추후 반발 계수를 변경할 수 있도록 후속 연구를 진행해볼 예정임)  
   -물체에 작용하는 마찰력은 무시하지 않으며, 바닥과 물체 모두 유리재질임. 따라서 마찰계수는 0.4로 설정) (반발계수와 마찬가지로 추후 후속 연구를 통해 변경 가능하도록 수정할 계획이 있음)  
   -데이터 생성 과정에서 두 물체(질점)은 서로 다른 두 지점에서 출발하여 동일한 점까지 이동함. 충돌하기까지 두 물체가 이동한 시간은 서로 같음 (특정 좌표에서 충돌함을 가정하기 위해 이동시간이 서로 같음)  
   -두 물체의 질량은 데이터 생성 절차에서 무작위로 설정됨  
   -충돌 후 물체가 분리되지 않는 것을 가정으로 함(쪼개지지 않음)
2. 모델 구성  
   가장 기본적인 Transformer 모델의 경우, 주된 목적이 NLP(Natural Language Procesing, 자연어 처리) 였기에 텍스트 토크나이징과 임베딩(Embedding)이라는 절차를 통해 자연어를 벡터로 변환하였다. 허나 본 연구에서 사용하는 데이터는 이미 벡터의 형태이므로 두 절차가 필요하지 않다.   
   해당 데이터 분석에서 결과물은 회귀방법(Regression)으로 도출 할 수 있다. 분류방법(Classification)은 연구를 진행하는 과정에서 사용하는 리소스와 뉴런 수에 비해 저조한 성능을 나타냄을 확인하였다. 기존의 transformer 모델은 분류를 사용하나, 해당 모델의 결과물의 해상도(Resolution, 얼마나 디테일 한 자료인가?)가 높아야 한다는 점을 고려한다면 회귀모델의 효율과 성능이 더 좋다는 점이 이상하지 않다.  
   Regression의 경우 모델 아키텍처
3. 학습 및 평가  
   Regression의 경우 모든 activation은 ReLU를 사용하고, Classification의 경우 마지막 activation만 softmax를 사용하고 나머지는 ReLU를 사용한다. Regression은 MSE(Mean Squared Error)를 사용해 평가하고, Classification은 CCE(Categorical Cross Entropy)를 사용해 평가한다. Optimizer은 두 경우 모두 Adam Optimizer을 사용한다.
4. 결과물 도출  
   학습 데이터에 대한 학습이 종료되면, 해당 모델을 사용할 수 있도록 내보낸다(Export). 완성된 모델을 저장하고 입력 데이터 이후에 발생할 사건(충돌 후 운동 상태)을 예측하는 프로그램으로 정리한다.

실험과정

1. 데이터 표본  
   본 연구의 가설에서 언급한 트랜스포머 모델을 응용한 모델은 그 사용 범위가 고전역학에 국한한다.(충돌역학에서 미시세계의 양자역학은 사용할 실효성이 없음) 허나, 인류가 알아낸 모든 물리 규칙을 적용했고, 그 결과물을 확인할 수 있는 데이터를 모으는 것은 실질적인 한계가 있다. 심지어 물리학에서 다루는 매우 일부에 해당하는 영역에 대한 데이터만 모은다고 하더라도 그 작업에 쓰이는 리소스(시간, 노동력)에 비해 오차가 매우 크다는 한계가 있다. 고로 데이터를 모으는 작업은 실제로 수행한 결과물을 모으는 것이 아닌 물리엔진으로 생성한 데이터를 사용하는 것으로 진행할 것이다. 그러나 시중에 나와있는 물리엔진은 역학적 현상을 엄밀한 물리 규칙을 이용해 해석하는 것이 목적이 아닌, 그와 비슷해 보이도록 구현하는 것이 목적이기에 이 연구의 취지와 맞지 않는다. 따라서 학습시키고자 하는 상황에 최적화된 물리 엔진의 개발이 필요하다.   
      
    제작할 상황은 다음과 같다:  
   I. 2차원(이후 3차원으로 확장) 위의 상호작용을 전제로 한다.  
   II. 2차원(이후 3차원으로 확장) 위에 두 질점이 존재한다.  
   III. 두 절점에 작용하는 힘은 마찰력과 중력 뿐이다.(우선 공기저항은 무시한다.)  
   IV. 두 질점은 서로 다른 위치에서 운동량을 가지고 출발한다.  
   V. 두 질점은 충돌할 수도 있고, 그러지 않을 수도 있다.  
   텍스트, 스크린샷, 소프트웨어, 멀티미디어 소프트웨어이(가) 표시된 사진

   자동 생성된 설명  
   위 코드는 실험 과정에서 사용한 데이터를 만들어내는 코드의 일부이다. (핵심적인 부분만 나타냄)  
     
   텍스트, 라인, 스크린샷, 도표이(가) 표시된 사진

   자동 생성된 설명  
   위 사진은 서로 다른 두 점 A, B에서 생성된 두 질점이 운동하여 한 지점에서 탄성 충돌하고, 과 으로 이동된 모습을 나타낸 것이다. 각 점 사이의 시간 간격은 0.1초 이다.   
     
   텍스트, 메뉴, 스크린샷, 번호이(가) 표시된 사진

   자동 생성된 설명  
   A와 B의 x, y좌표, A와 B의 질량을 묶어 벡터로 표현한다. 하나의 상황(event)에 대해 대략 20개의 벡터가 생성된다. 이후, 한 이벤트를 5개의 벡터들로 슬라이싱 하여(예를 들면, 1~5번 벡터, 2~6번 벡터, 3~7번 벡터와 같은 식으로 5개의 벡터들을 묶어 두는 것을 말함) 트랜스포머 모델에 입력한다.
2. 모델 구성  
   모델은 Attention Is All You Need에 등장한 트랜스포머 모델과 유사한 구조의 모델을 사용한다. 다만 일부 상수는 바꾼다.   
   손실함수는 기존의 Categorical CrossEntropy가 아닌, Mean Squared Error을 사용한다. 분류모델이 아닌 회귀모델이기 때문이다.   
   텍스트, 스크린샷, 폰트, 소프트웨어이(가) 표시된 사진

   자동 생성된 설명  
   [앞서 언급한 Attention 기법의 프로그램 구현]  
   텍스트, 스크린샷, 소프트웨어이(가) 표시된 사진

   자동 생성된 설명  
   [Encoder 레이어의 구현]  
   텍스트, 스크린샷, 소프트웨어이(가) 표시된 사진

   자동 생성된 설명  
   [Decoder 레이어의 구현]  
   텍스트, 스크린샷, 소프트웨어이(가) 표시된 사진

   자동 생성된 설명  
   [하나의 트랜스포머 모델로 엮음]  
   레이어들을 모아 하나의 트랜스포머 모델로 만드는 과정에 기존의 트랜스포머 모델과 조금의 차이가 있다. 트랜스포머 모델은 본래 언어처리를 위해 개발 되었는데, 언어모델의 데이터는 임베딩이 되지 않은 상태로 입력되는데 반해(단어의 인덱스로 구성된 벡터가 입력됨), 이 연구에서의 데이터는 이미 벡터화 된 상태로 입력된다. 고로 입력 절차에 몇 가지 변화를 주어야 한다.
3. 데이터 학습

텍스트, 스크린샷, 소프트웨어이(가) 표시된 사진

자동 생성된 설명  
학습은 위와 같은 코드로 진행한다. 대략 몇 천 번의 반복학습을 거친다. 반복학습 후 모델이 특정 수준에 수렴하면, 모델의 가중치들을 모두 저장한다.

결과 도출  
  
학습 초기 1차 실험의 결과물은 다음과 같다:  
스크린샷, 그래프, 소프트웨어, 텍스트이(가) 표시된 사진

자동 생성된 설명  
<위 – 실험군 / 아래 – 대조군>  
예측된 운동 상태의 경향과 위치는 비슷했지만, 정확하다고 할 수 없다.  
질점 이동의 모습이 직선에 가깝지 않은 모습이다.  
  
  
  
  
  
  
  
학습을 조금 더 진행한 2차 실험의 결과물은 다음과 같다:  
스크린샷, 라인, 소프트웨어, 그래프이(가) 표시된 사진

자동 생성된 설명  
이전 실험과 비슷한 결과이다. 이전 실험에 비해 조금 더 정확 해졌다고 할 수 있다.  
질점 이동의 경향이 직선에 가깝지 않은 모습이다.  
  
  
  
  
  
  
  
  
  
  
학습을 더 진행한 3차 실험의 결과는 다음과 같다:  
스크린샷, 라인, 그래프, 소프트웨어이(가) 표시된 사진

자동 생성된 설명  
기존에 비해 더 정확해졌다.  
질점 이동의 경향성이 직선에 가까워졌다고 볼 수 있다.  
  
  
  
  
  
  
  
  
  
  
  
학습을 더 진행한 4차 실험의 결과는 다음과 같다:  
스크린샷, 소프트웨어, 라인, 멀티미디어 소프트웨어이(가) 표시된 사진

자동 생성된 설명  
기대된 결과물과 거의 일치한다.   
경향성 또한 직선으로 나타난다.  
  
  
  
  
  
  
  
  
  
  
마지막으로 진행한 실험의 결과는 다음과 같다:

텍스트, 스크린샷, 도표, 그래프이(가) 표시된 사진

자동 생성된 설명결과확인을 위해 투입한 질점이동 상황이다.  
스크린샷, 텍스트, 그래프, 도표이(가) 표시된 사진

자동 생성된 설명운동상태예측 결과물-실험군스크린샷, 텍스트, 그래프, 도표이(가) 표시된 사진

자동 생성된 설명기대된 운동상태 - 대조군  
실험군과 대조군의 결과를 비교할 때, 모델이 운동상태의 양상과 거의 정확하게 일치하는 것을 알 수 있었다.

제언  
  
-높은 해상도의 결과물을 위해 회귀모델을 사용했으나, 만약 좌표계를 굉장히 작은 단위로 분리해 사용할 컴퓨팅 파워가 있다면 (그래프에 투사할 때 사용되는 최소 단위를 이용) 분류모델을 사용하는 것이 정확도를 올려줄 수 있다고 생각함.

-결과물의 정확도를 고려할 때, 현실에서 일어나는 일의 흐름을 예측하기엔 아직 부족하다고 판단됨. 허나 더 복잡한 모델과 효과적인 학습 방법을 적용하면 현실에서의 운동상태 예측도 충분히 가능해질 것이라고 기대함

-두개의 질점의 운동 예측을 넘어서, 수백만개의 질점 운동을 예측하면 고체는 물론 유체의 운동도 예측할 수 있을 것으로 기대함.

-입자의 전하와 전자의 이동까지 데이터상에 표현한다면 전자제품의 작동양상도 예측 가능할 것으로 기대함.

-이 기술을 사용해서 다양한 제품의 설계를 사전에 검증할 수 있을 것으로 기대함.

-3차원에 대한 실험을 실행하여 데이터를 확인할 예정임.