Encoder-decoder approach:

인코더(encoder) : 인지 네트워크(recognition network)라고도 하며, 입력을 내부 표현으로 변환한다.

디코더(decoder) : 생성 네트워크(generative nework)라고도 하며, 내부 표현을 출력으로 변환한다.

참고할 만한 딥러닝 기법 : 오토인코더 (auto-encoder)

시퀀스: 연관된 연속의 데이터 (문장(sentences)은 문법 등의 규칙으로 연관된 일련의(sequential) 단어들 집합으로 볼 수 있고, 좋은 시퀀스의 예)

So, 시퀀스-투-시퀀스 문제란 소스 시퀀스를 결과 시퀀스로 바꾸는 문제 (예: 입력 시퀀스와 출력 시퀀스를 각각 질문과 대답으로 구성하면 챗봇으로 만들 수 있고, 입력 시퀀스와 출력 시퀀스를 각각 입력 문장과 번역 문장으로 만들면 번역기로 만들 수 있습니다. 그 외에도 내용 요약(Text Summarization), STT(Speech to Text) 등에서 쓰일 수 있습니다.) <- 대개 인코더와 디코더로 구성

참고 웹 사이트:

<시퀸스 투 시퀸스 정리>

<https://medium.com/@jongdae.lim/%EA%B8%B0%EA%B3%84-%ED%95%99%EC%8A%B5-machine-learning-%EC%9D%80-%EC%A6%90%EA%B2%81%EB%8B%A4-part-5-83b7a44b797a>

<오토인코더 정리> <https://excelsior-cjh.tistory.com/187>

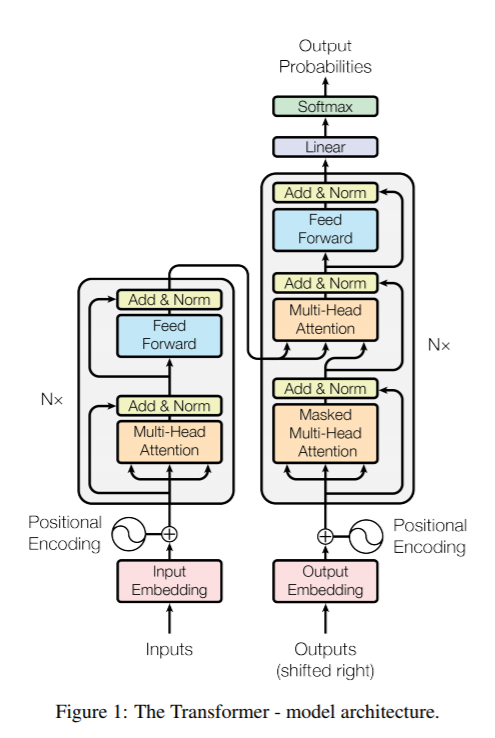
Self-Attention

Transformer: Transformer는 기계 번역 및 텍스트 요약과 같은 다양한 작업을 위해 자연어와 같은 순서가 지정된 데이터 시퀀스를 처리하도록 설계되었습니다

트랜스포머는 RNN을 사용하지 않지만 기존의 seq2seq처럼 인코더에서 입력 시퀀스를 입력받고, 디코더에서 출력 시퀀스를 출력하는 인코더-디코더 구조를 유지하고 있습니다. 다만 다른 점은 인코더와 디코더라는 단위가 N개가 존재할 수 있다는 점입니다.

이전 seq2seq 구조에서는 인코더와 디코더에서 각각 하나의 RNN이 t개의 시점(time-step)을 가지는 구조였다면 이번에는 인코더와 디코더라는 단위가 N개로 구성되는 구조입니다. 트랜스포머를 제안한 논문에서는 인코더와 디코더의 개수를 각각 6개를 사용하였습니다.

Transformer는 RNN, LSTM없이 time 시퀀스 역할을 하는 모델



Seq2Seq-Attention 모델에서의 번역 태스크의 문제는 원본 언어(Source Language), 번역된 언어(Target Language)간의 어느정도 대응 여부는 어텐션을 통해 찾을 수 있었으나, 각 자신의 언어만에 대해서는 관계를 나타낼수 없었습니다. 예를 들면 I love tiger but it is scare와 나는 호랑이를 좋아하지만 그것(호랑이)는 무섭다 사이의 관계는 어텐션을 통해 매칭이 가능했지만 it이 무엇을 나타내는지?와 같은 문제는 기존 Encoder-Decoder 기반의 어텐션 메커니즘에서는 찾을 수 없었습니다.

<참고 웹 사이트>

<https://medium.com/platfarm/%EC%96%B4%ED%85%90%EC%85%98-%EB%A9%94%EC%BB%A4%EB%8B%88%EC%A6%98%EA%B3%BC-transfomer-self-attention-842498fd3225>

Open domain: 사람과 대화하는 것처럼 다양한 주제를 탐색할 수 있는 챗봇. 이게 가능해지면 동반자(?) 때로는 치료자의 역할을 수행하기도 하고, 영화 HER에 나오는 것처럼 교감이 된다고 느낄 수도 있다. 잠재력이 크긴 하지만 구현, 그리고 평가하기가 쉽지 않다.

<참고 웹 사이트>

<http://hleecaster.com/what-are-chatbots/>

제로샷(zero-shot): 제로 샷 학습은 테스트 시간에 학습자가 학습 중에 관찰되지 않은 클래스의 샘플을 관찰하고 해당 카테고리를 예측해야 하는 머신 러닝의 문제점 설정입니다. 이 문제는 컴퓨터 비전, 자연어 처리 및 기계 인식에서 널리 연구됩니다.

GPT-2는 15억개 파라미터를 가진 대규모 번역기반 언어모델로, 800만개 텍스트의 데이터셋을 통해 학습했다. 임의의 문장을 입력하면 맥락을 고려한 종합적인 문장을 스스로 만들어낸다.

“우리는 일관된 텍스트 단락을 생성하고, 많은 언어 모델링 벤치마크에서 최첨단 성능을 달성하며, 초보적인 읽기 이해력, 기계 번역, 질문 답변, 요약 등을 모두 직무별 교육 없이 수행하는 대규모 비지도 언어 모델을 만들었다.” – openAI

<GPT-2 참고 웹사이트>

<https://openai.com/blog/better-language-models/>

Wizard-of-Oz 방식: 크라우드소싱 환경에서 두 명의 임의의 참여자가 “사용자”와 “에이전트”의 역할을 나누어 대화를 나누며 데이터를 생성하는 형태를 의미한다.

크라우드소싱(crowdsourcing) : 외부자원인 대중들의 참여를 유도하고 지원하여 비즈니스로 활용하는 것

Dialogue State Tracking (DST) : 매 대화 턴 별로 유저의 목표(goal)에 대한 가능성을 예측하는 문제

NLU (Natural Language Understanding) : 언어를 읽고 통역하는 과정. 문장의 의도 분류, 서로 다른 언어 간 번역 문장 생성, 자연어 질문에 대한 답변 추출 등이 존재

NLG (Natural Language Generation) : 듣고 이해만 하는 과정에서 더 나아가 축적되어 있는 단어들을 조합해서 직접 사용자가 이해하기 쉬운 문장으로 출력하는 것 (시스템 계산의 결과를 사람의 언어로 표현하는 기술)