

Evaluations

Ki Hyun Kim

nlp.with.deep.learning@gmail.com

Experimental Setup

- Preprocessing
 - en: MosesTokenizer + BPE
 - Vocabulary Size: 20,525
 - ko: Mecab + BPE
 - Vocabulary Size: 29,411

- Dataset

set	lang	# lines	# tokens	# characters
train	en	1,200,000	43,700,390	367,477,362
	ko	1,200,000	39,066,127	344,881,403
validation	en	200,000	7,286,230	61,262,147
	ko	200,000	6,516,442	57,518,240
validation-1000	en	1,000	36,307	305,369
	ko	1,000	32,282	285,911
test-1000	en	1,000	35,686	298,993
	ko	1,000	31,720	280,126

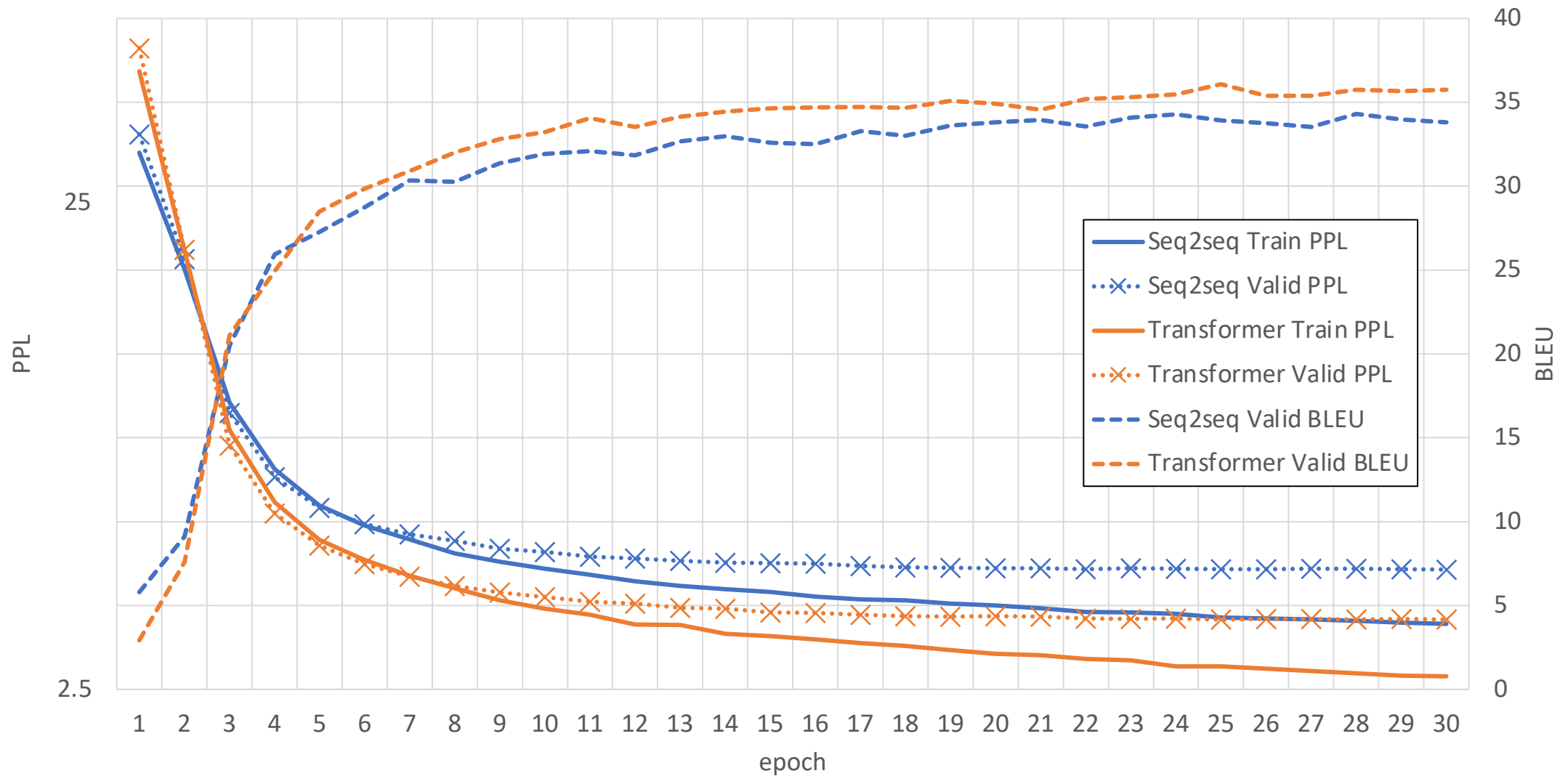
- Hyper-parameters

parameter	Seq2seq	Transformer
batch_size	320	4096
word_vec_size	512	-
hidden_size	768	768
n_layers	4	4
n_splits	-	8
n_epochs	30	30

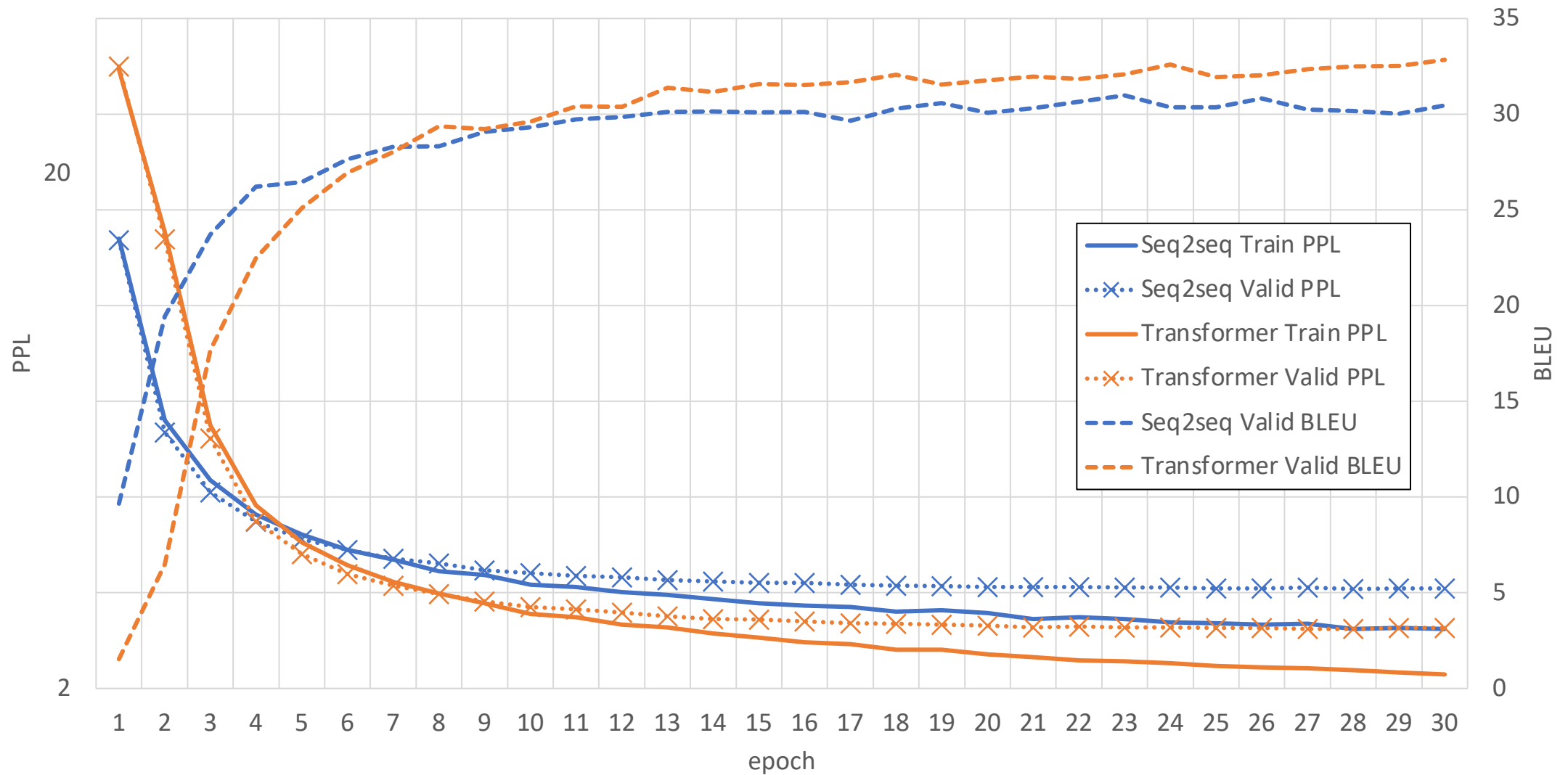
Test BLEU

	Sequence to Sequence	Transformer
ENKO	32.53	34.96
KOEN	29.67	31.84

Sequence-to-Sequence vs. Transformer (ENKO)

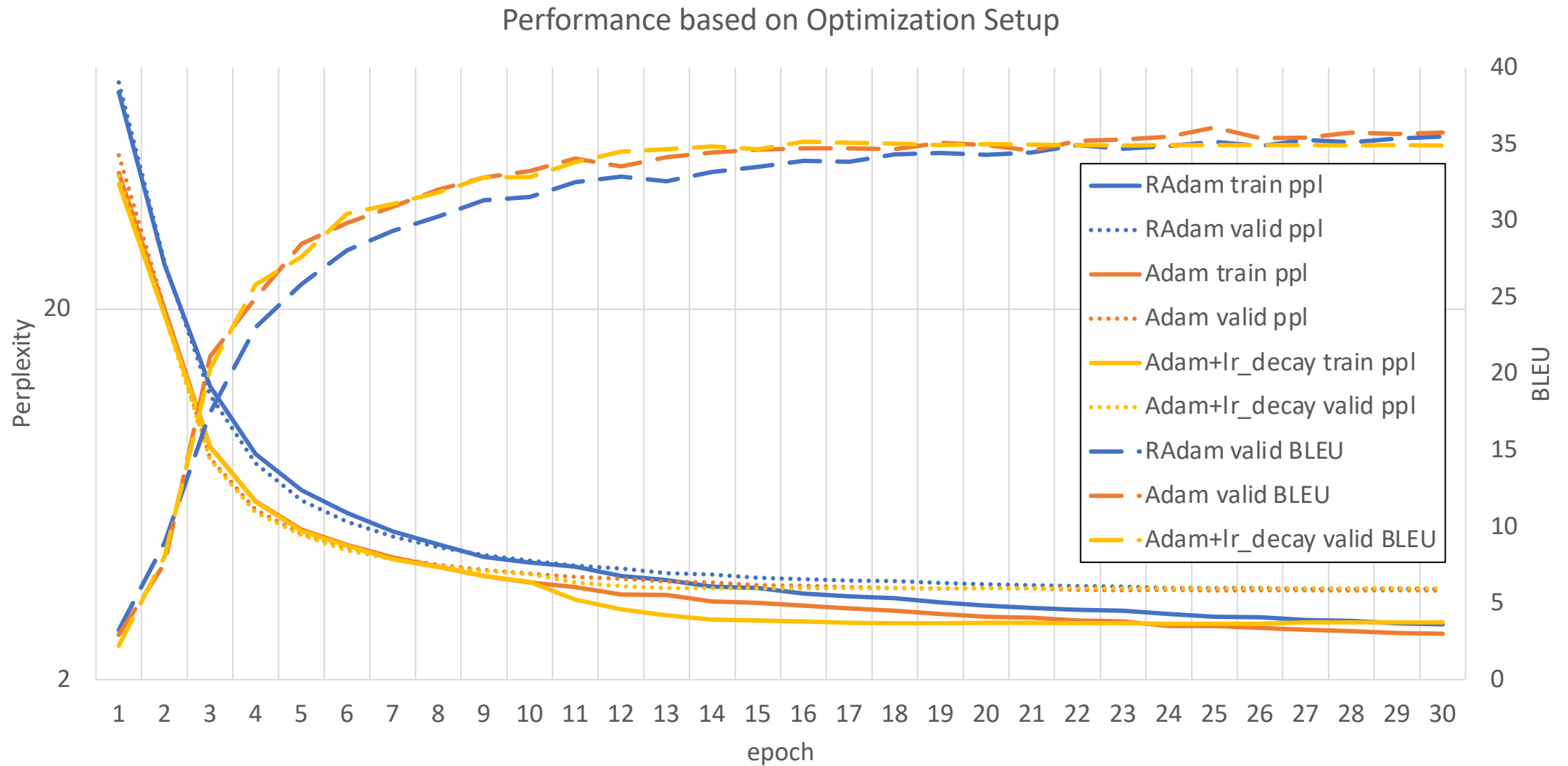


Sequence-to-Sequence vs. Transformer (KOEN)



REF	Sequence-to-Sequence	Transformer
도널드 트럼프 미국 대통령이 사우디아라비아 언론인 자말 카슈끄지 살해 사건에 대해 왕세자 무함마드 빈살만의 관여가 문제되지 않는다고 일축했다.	도널드 트럼프 미국 대통령이 사우디아라비아 언론인 자말 카슈끄지 피살 사건과 관련해 무함마드 빈 살만 왕세자의 개입이 문제가 되지 않는다고 거절했다.	도널드 트럼프 미국 대통령은 사우디 언론인 자말 카슈끄지가 살해된 것과 관련해 “모하마드 빈 살만 왕세자의 연루가 문제가 아니다”며 거절했다.
근데 그를 왜 살려줬을까요?	그런데 그는 왜 그를 구해주었나요?	그런데 그는 왜 그를 구하게 되었습니까?
택배 박스의 빈 곳에 한국 주소도 같이 붙여주시면 감사하겠습니다.	택배 박스의 빈 공간에 한국 주소를 첨부하여 주시면 감사하겠습니다.	택배 박스의 빈 공간에 대한 한국 주소를 첨부해주시면 감사하겠습니다.
이와 관련, 지난 4~5일 일본을 관통한 태풍 ‘제비’로 인해 일본 전역에 11명이 사망하고 600여명이 부상을 입은 것으로 집계됐다.	이와 관련해 지난 4~5일 일본을 휩쓸던 태풍 ‘제비’는 일본 전역에서 11명이 숨지고 일본 전역 600여명이 다쳤다고 전했다.	이와 관련, 지난 4~5일 일본을 휩쓴 태풍 ‘제비’가 11명을 사살하고 일본 전역에서 600명 이상의 부상자를 냈다는 보도가 나왔다.
수백만 가입자가 일상 환경에서 초고속 데이터 전송과 음성통화를 이상 없이 이용하도록 망 안정화와 업그레이드, 커버리지 확대를 지속하는 일은 최우선 과제다.	가입자 수백만 명이 일상 환경에 문제없이 초고속 데이터 전송과 음성 호출을 이용하는 것은 네트워크 안정화, 커버리지 고도화, 커버리지 확대 등이 최우선 과제다.	네트워크 안정화, 업그레이드, 커버리지 확대 등은 수백만 명의 가입자가 고속 데이터 전송과 음성통화를 일상 환경에서 별다른 문제 없이 이용할 수 있는 최우선 순위다.
단기적으로 시장의 변동성은 지속될 것으로 전망되지만 역사적 관점에서 볼 때 최근 시장의 변동성이 특별히 극심한 것은 아니며 장기적인 시장 펀더멘털 역시 여전히 견고해 보인다.	단기적으로 시장 변동성은 지속될 것으로 전망되지만 역사적 관점에서 보면 최근 시장의 변동성이 특별히 극단적이지 않고 장기간 시장 펀더멘털도 남아 있는 것으로 보인다.	단기적으로는 시장 변동성이 지속될 것으로 전망되지만, 역사적 관점에서 보면 최근 시장 변동성은 특별히 극단적이지 않고 장기적 시장 펀더멘털도 여전한 것으로 보인다.
기존 플라잉 프로브는 제품의 가격이 너무 고가인 데다 ICT와 같은 기존 시험 장비에 비해 시험 시간이 길어서 양산라인에 적용하기 어렵다는 단점이 있다.	기존 비행조사는 제품 가격이 너무 높고 시험시간도 ICT 등 기존 테스트 장비보다 길기 때문에 양산 라인을 적용하기가 어렵다.	기존의 비행용 탐사선은 제품 가격이 너무 비싸고 시험시간이 ICT 등 기존 시험 장비에 비해 길기 때문에 양산선에 적용하기 어렵다.
뇌졸중은 단일질환으로는 사망원인 2위에 해당하는 위험도가 높은 질환으로, 뇌졸중 의심증상이 발생하는 경우 3시간 이내에 병원으로 도착해 적절한 치료를 받는 것이 중요하다.	뇌졸중은 단일 질환으로 사망원인의 2차 흔히 사망원인이 높은 질환이며, 뇌졸중 의심환자가 발생하면 적절한 치료를 위해 3시간 이내 병원에 도착하는 것이 중요하다.	뇌졸중은 단일 질환으로 사망원인 2위에 해당하는 질환으로, 뇌졸중 의심 증상이 발생할 경우 3시간 이내에 병원에 도착하는 것이 중요하다.
문희상 의원은 조문 후 기자들과 만나 “나라가 소용돌이 한복판에 놓였을 때 국가의 큰 어른인 가르침이 그 어느 때보다 아쉬운데 돌아가시게 돼 진심으로 애도를 표한다”고 말했다.	문재인 의원은 조의 끝에 기자들과 만나 “국가가 소용돌이에 빠져 있을 때 큰 고령화를 줄 수 있는 사람이 분실이라는 점에 진심으로 애도를 표한다”고 말했다.	문희상 의원은 조문이 끝난 뒤 기자들과 만나 “국가가 소용돌이 속에 있을 때 위대한 옛 가르침을 줄 수 있는 사람을 잃었다는 것에 대해 진심으로 애도의 뜻을 표한다”고 말했다.
한 은행권 관계자는 “기존에는 자체 DSR 기준에 걸려도 심사역이 승인하면 대출이 가능했는데, 관리지표로 본격 도입되면 사실상 대출이 어려워질 것”이라고 전했다.	은행권 관계자는 “예전에는 자체 DSR 기준에 해당돼도 심사팀이 대출을 승인하면 대출을 할 수 있었는데, 경영지표로 도입되면 실질적으로 대출을 받기 어려울 것”이라고 말했다.	은행권 관계자는 “과거 심사팀이 자체 DSR 기준에 해당하더라도 대출을 승인하면 대출을 받을 수 있는데 관리지표로 도입되면 사실상 대출을 받기 어려울 것”이라고 말했다.
사회적 감사는 업무가 법규나 규정에는 부합하더라도 공평성이나 형평 등 사회적 규범에 맞지 않는다면 이를 개선하도록 권고해 조직에 변화를 유도하는 것이다.	사회적 감사는 법이나 규정을 준수하더라도 정의나 형평성 등 사회적 규범에 부합하지 않으면 업무 윤리 개선을 시키도록 유도하는 등 조직의 변화를 유도하는 것이다.	사회감사에서는 법이나 규정을 준수하더라도 정당이나 형평 등 사회규범에 맞지 않으면 업무 윤리가 향상될 수 있도록 독려해 조직의 변화를 유도한다.
마찰이 있더라도 국익 차원에서 인터넷 역사별 해소 정책을 강력하게 추진해야 한다는 목소리가 힘을 얻기 시작했다.	마찰이 일어나더라도 인터넷 역사별 정책의 해법을 구현하는 목소리가 국익을 향한 힘을 얻기 시작했다.	마찰이 발생해도 인터넷 역사별 정책의 해법을 시행하자는 목소리가 국익을 위해 힘을 얻기 시작한 것이다.
오비맥주가 주요 맥주 제품의 가격을 인상키로 하면서 하이트진로와 롯데주류 등도 조만간 맥주 가격을 올릴 가능성이 높아졌다.	오비맥주가 각종 맥주 상품 가격을 올리기로 결정한 만큼 하이트진로와 롯데주류도 가까운 시일 내에 맥주 가격을 인상할 가능성이 높다.	오비맥주가 다양한 맥주 제품 가격 인상을 결정을 내린 만큼 하이트진로와 롯데주류도 가까운 장래에 맥주 가격을 인상할 가능성이 높다.
불필요한 광고비 경쟁을 부추긴다는 지적이 나오자 오는 5월부터 광고료를 정률(주문 금액의 6.8%)로 하는 개방형 방식으로 대체하기로 했다.	불필요한 광고비 경쟁을 촉발하는 주장이 나오자마자 5월에는 광고료(주문 가격의 6.6%)를 공개적으로 대체하기로 했다.	광고비 불필요한 경쟁을 불식시키는 주장이 나오는 대로 오는 5월 광고료(발주가격의 6.8%)를 공개적으로 대체하기로 했다.
데이터를 무제한으로 쓰는 ‘헤비유저’가 데이터 이용을 독점해 중저가 요금제 가입자의 통신 품질이 낮아질 것이라는 우려도 제기한다.	무제한 데이터를 사용하는 ‘헤비 사용자’가 데이터 사용을 독점할 것이라는 우려도 제기돼 중저가 구독자 간 의사소통의 질이 낮아질 것이라는 우려도 제기된다.	무제한 데이터를 이용하는 ‘중량사용자’들이 데이터 이용을 독식하게 돼 중저가 가입자들의 통신 품질이 떨어질 것이라는 우려도 제기된다.
이 사고로 통학차량에 타고 있던 학생 7명(남3, 여4) 중 고교생 3명과 여중생 1명이 다쳐 인근 병원으로 옮겨져 치료를 받고 있다.	이 사고로 통근차량에 타고 있던 학생 7명 가운데 고등학생 3명과 여중생 1명이 부상을 입고 인근 병원으로 옮겨져 치료를 받고 있다.	통학 차량에 타고 있던 학생 7명 중 고등학생 3명과 여중생 1명이 부상을 입고 인근 병원으로 옮겨져 치료를 받고 있다.
반경 1km 내에 용현초, 신광초, 신흥여중 등 다수의 교육시설이 위치해 있고 홈플러스와 현대 유비스 병원 등이 도보 5분 거리 내에 위치해있다.	1km 이내에는 용인초등학교, 신광초등학교, 신흥여중 등 수많은 교육시설이 있으며, 걸어서 5분 거리에 홈플러스, 현대위비병원 등이 있다.	1km 이내에는 용현초등학교, 신왕초, 신흥여자중학교 등 수많은 교육시설이 있으며 도보 5분 거리에 홈플러스와 현대 유비에스 병원이 있다.

Transformer Optimization Setup (ENKO)



Transformer Optimization Setup (ENKO)

	train loss	train ppl	valid loss	valid ppl	valid BLEU	test BLEU
Pre-LN Transformer with Radam	1.04	2.82	1.26	3.51	35.49	34.61
Pre-LN Transformer with Adam	1.03	2.79	1.25	3.48	36.08	34.96
Pre-LN Transformer with Adam & LR Decay at 11-th epoch	1.05	2.87	1.26	3.53	35.15	34.66

Summary

- 압도적인 Transformer의 성능!
 - 하지만 디코딩 방법의 효율이 아쉬움 (높은 memory 사용량)
 - 따라서 경우에 따라 LSTM decoder를 사용하기도 함
 - e.g. <https://ai.googleblog.com/2020/06/recent-advances-in-google-translate.html>
- 말 많고, 탈 많던 Transformer의 최적화 문제는 일단락
 - Pre-LN Transformer에서 Adam을 쓰면 해결
 - LR decay는 실험 결과 큰 효과가 없었으나, 좀 더 튜닝해 볼 여지가 있음