

About the Beamer class in presentation making

A short story

Claudio Scheer¹ José Fernando Possebon¹

¹Pontifical Catholic University of Rio Grande do Sul - PUCRS
{claudio.scheer, jose.possebon}@edu.pucrs.br

Deep Learning, June 2020

Table of Contents

1 Proposal

2 Dataset

3 BERT

4 Results

Table of Contents

1 Proposal

2 Dataset

3 BERT

4 Results

Proposal

Table of Contents

1 Proposal

2 Dataset

3 BERT

4 Results

The Enron Email Dataset

no, i dont mind. : -)

—Original Message—

Hi how are you doing? I have a meeting from 4 to 5, do you mind waiting for me? Thanks.
John

—Original Message—

hello

The Enron Email Dataset

yuck yuck

—Original Message—

har har

The subject is: **Wine tasting.**

The Enron Email Dataset

- <https://www.kaggle.com/claudioscheer/extract-reply-emails>;
- Get only emails with less than 256 characters;

Final dataset: 40 062 input-target pairs.

Train dataset: 10 002 input-target pairs.

Input	Target
I thought this one was canceled. So I declined to get it off my calendar.	You're right.
Rex - thanks for the quick response...please note the revised letters below. Previous email had a clerical error	Who will sign theses?
Where can I get info on safety statistics for the pipes?	Ralph, please get with Rod and provide what ever he needs in the way of Safety Stats.

Table of Contents

1 Proposal

2 Dataset

3 BERT

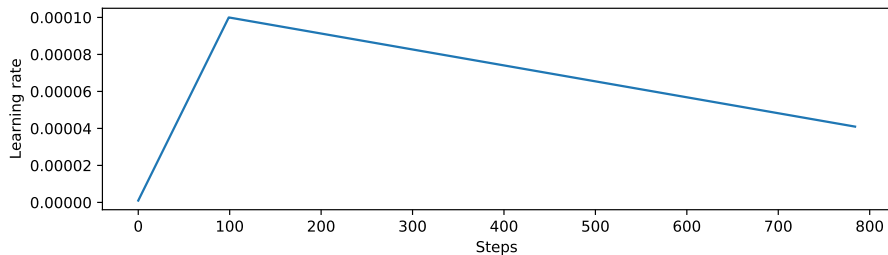
4 Results

BERT architecture

- BERT base was used;
 - 12 layers;
- Hugging Face and PyTorch was used;

Learning Rate Schedule

- Warm-up steps: 10 000;
- Learning rate: $1e-4$;



Optimizer and hyperparameters

- Adam;
 - Adam epsilon: $1e-4$;
- Batch size: 8;
- Training epochs: 32;
- Beam search hypothesis: 3;

Table of Contents

1 Proposal

2 Dataset

3 BERT

4 Results

Generated replies

Email	Reply	AI?
Shall we have the meeting tomorrow?	meeting is postponed because the 3 : 00 p. m. meeting today. meeting is canceled for today.	30%
hi	Hi how are you?	10%
you look so handsome today	why thank you, you do as well	30%