Adversarial Training for Aspect-Based Sentiment Analysis with BERT

Akbar Karimi Leonardo Rossi Andrea Prati

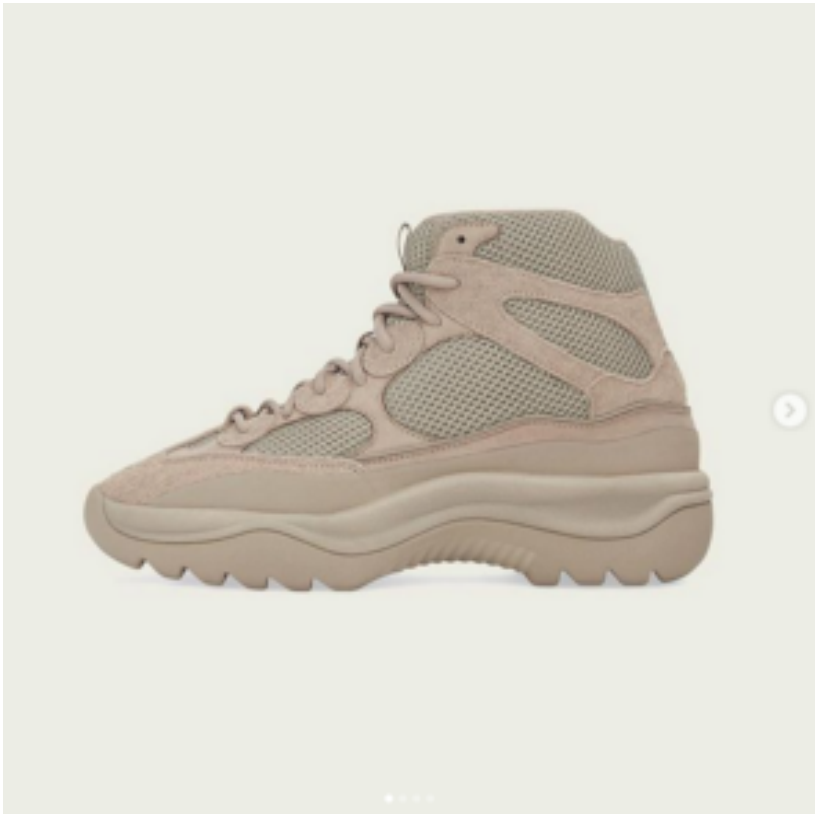
University of Parma
Dep. of Engineering and Architecture
IMP Lab

Video Presentation for ICPR 2020

Source Code: https://github.com/IMPLabUniPr/BERT-for-ABSA

Problem 1: Aspect Sentiment Classification (ASC)

How do consumers feel about a product (service) and its aspects?



They look good and comfy

Army boots

My grandma wears those

Problem 2: Aspect Extraction (AE)

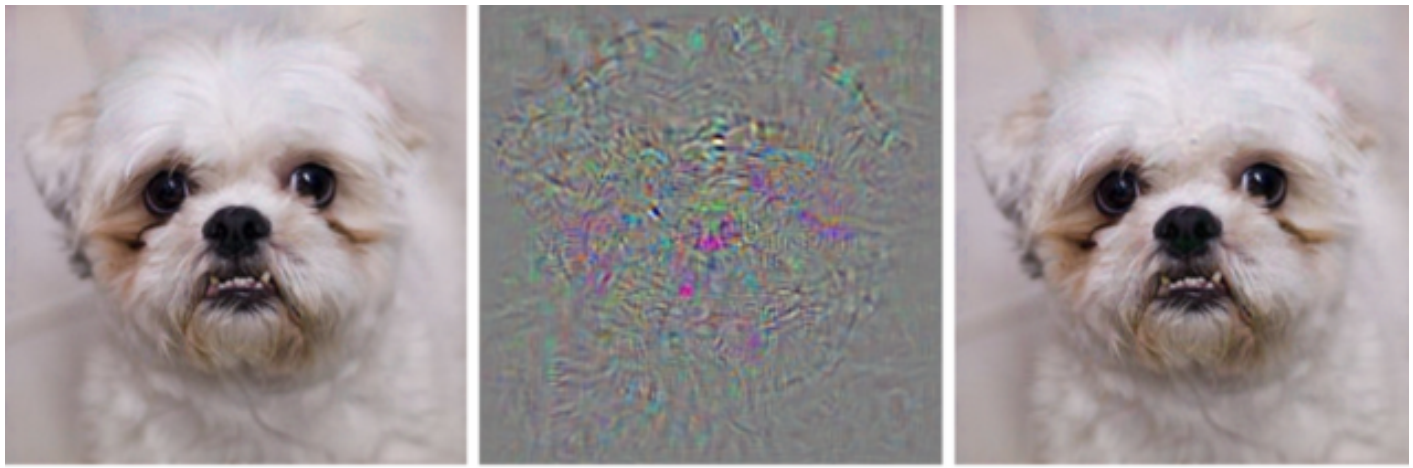
What aspects of a product (service) do they discuss?

patterns
Love the contrasting patterns

colour
Yes this colour is amazing

Adversarial Examples for Image

- Sometimes classifiers act strangely



Adversarial Examples for Text

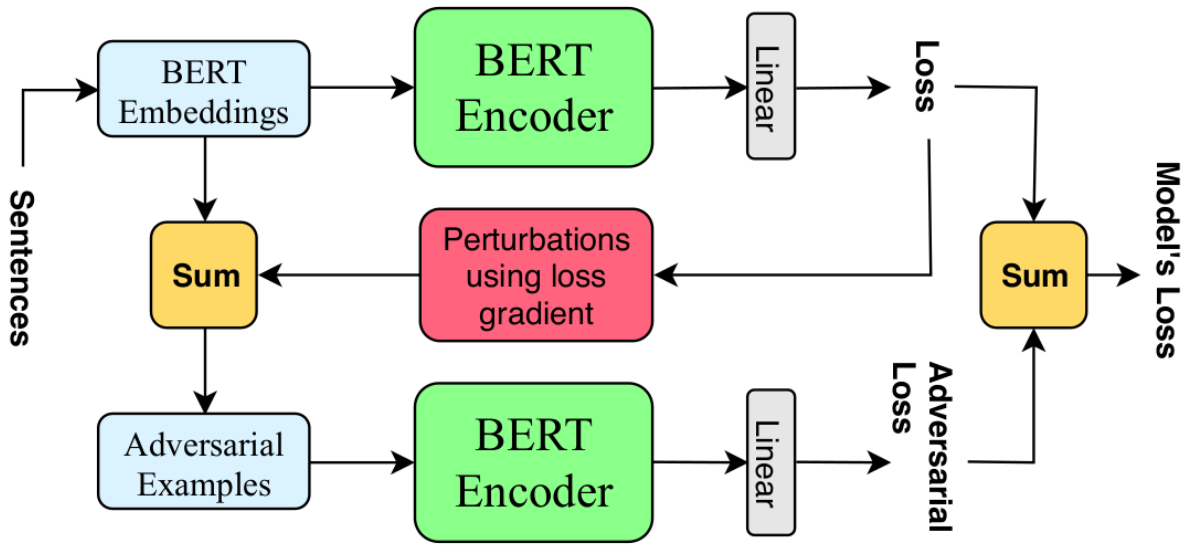
- Perturbations on input word embeddings:

$$x = x - \epsilon \frac{g}{\|g\|} \tag{1}$$

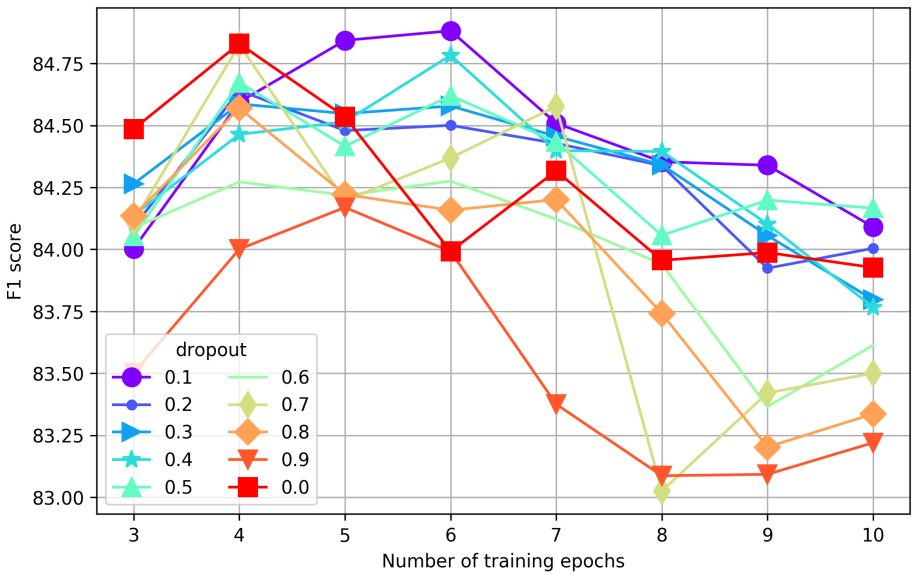
x: input embedding
ϵ: size of perturbation
g: gradient of loss w.r.t. x

BERT Adversarial Training (BAT) Model

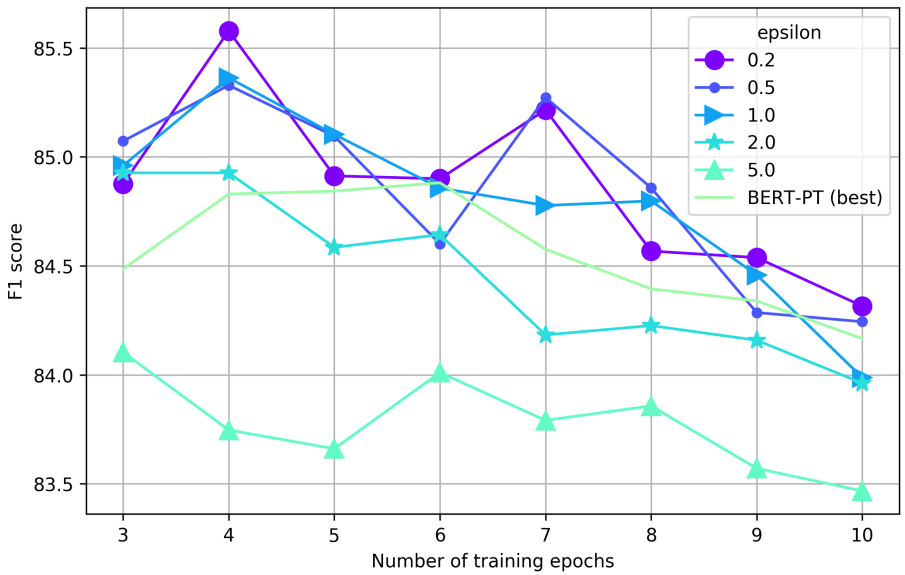
- Input embeddings + adversarial examples



Experiments: Aspect Extraction (AE)



(a) BERT-PT (Laptop)



(b) BAT (Laptop)

Results

- BAT improves post-trained BERT

Domain	Laptop	Restaurant
Methods	F1	F1
BERT-base (2018)	79.28	74.1
BERT-PT (2019)	84.26	77.97
BERT-PT (best)	**84.88**	**80.69**
BAT (Ours)	**85.57**	**81.50**

Table 1: Aspect Term Extraction

Domain	Laptop	Restaurant
Methods	Acc	Acc
BERT-base (2018)	75.29	81.54
BERT-PT (2019)	78.08	84.95
BERT-PT (best)	**78.89**	**85.92**
BAT (Ours)	**79.35**	**86.03**

Table 2: Aspect Sentiment Classification

ICPR 2020
10 | 15 Jan 2021
Virtual Conference, Milan, Italy
Poster: 2417

Akbar Karimi

University of Parma
Dep. of Engineering and Architecture
IMP Lab