Comparative Snippet Generation

Saurabh Jain, Yisong Miao, Min-Yen Kan
ECNLP 2022

Introduction

QA System on e-commerce reviews results one side of opinion, either positive or negative. We investigate performance of BERT to generate comparative snippet expressing both types of opinions.

Problem Statement

Given a positive and a negative opinion regarding a product, generate a single sentence comparative response.

Input: Display is awesome. Sound quality is bad.
Output: Display is awesome. However, a few users have complained that sound quality is bad.

Dataset Creation

Set of reviews
- EDU Extraction
- Segment Sentiment Classification
- Dataset Generation
- Segment Summarization
- Dataset

Dataset

- Fine-tuned base BERT model struggles to generate comparative text snippets combining a positive and a negative opinion on a product.

Fig.: Comparison of model-based generations against rule-based generations. ROUGE-L is used to measure input information preservation. ROUGE-3 and ROUGE-4 are used to measure quality of overall generations.

Examples Of Main Error Types

- Incorrect mixing: the entire set is comfortable. on the other hand, right few users have complained that right side slides down.
- Insertion of “on”: the 415’s are a great upgrade from the oem earbuds. on, it is super uncomfortable.
- Missing “Although” word: the retractil system works fine, according to a few users the pads are sort of squarish.
- Information modification: sound is pretty good. but, the movement is actually more like a saw.
- however, there are people who have complained that however, there are people who have complained that

Fig.: Pie chart shows distribution of main error types. Pie chart shows distribution of connecting strings.