Figure A1: Generated hard training pairs where positive and negative instances are highly similar. Negative instances are created by guiding the generation process with a carefully designed prompt, such as "erase 'hole' from button." This approach removes crucial features from seed samples and shifts the original semantics.