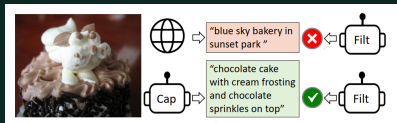


BLIP

借鉴了类似VLMo的思想，根据不同的pre-training task选择冻结不同的结构



清洗noisy web data: captioner and filter

作者：@ry664663 | 来自：知犀思维导图