Comparing predictions v/s actual using zero-shot model on VALIDATION SET