Error per sample on the TEST set using zero-shot model

