

Choose token  $\hat{v}$  to replace

$$\hat{v} = \operatorname{argmax}_{v \in \mathcal{V}_{\text{cand}}} \Delta \log \Pr(\mathbf{y} | f_{\text{sub}}(\mathbf{X}', v); \theta)$$

$$\mathbf{X}' = p(\mathbf{X})$$

$\mathbf{X}$

a gorgeous, witty, seductive movie.

$\langle T \rangle$

$\langle T \rangle$

$\langle T \rangle$

$\langle \text{mask} \rangle$  .