

Diagrama de flujo preprocesamiento

Preprocesamiento

Limpieza de datos

Unificación de
comentarios

Regex
preprocess

Añadir etiquetas

<|BOS|>
<|EOS|>