

Multi Language Support for Virtual Assistants

Prise en charge multilingue pour les assistants virtuels

对虚拟助手的多语言支持

پشتیبانی چند زبانه برای دستیاران مجازی

Supporto multilingue per assistenti virtuali

仮想アシスタントの多言語サポート

Soporte multilinguaje para asistentes virtuales

वर्चुअल असिस्टेंट के लिए मल्टी लैंग्वेज सपोर्ट

가상 어시스턴트를위한 다국어 지원

Suporte em vários idiomas para assistentes virtuais

April 2020

Overview

Overview

Goals:

- Extending the current capabilities of Almond to other languages in a cost and time efficient manner
- Avoiding template development for each new language



Overview

Goals:

- Extending the current capabilities of Almond to other languages in a cost and time efficient manner
- Avoiding template development for each new language



Solution:

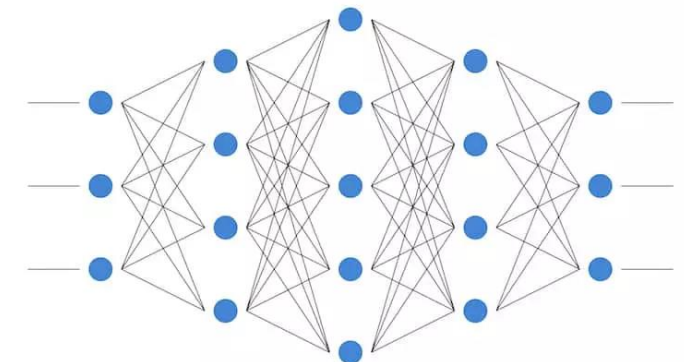
Data collection strategy:

- Using neural machine translation models to produce translated sentences
- Improving translation quality using domain-dependent rules



Training strategies:

- Joint and sequential training
- Enforcing low variance on encoded outputs on same sentences from different languages



Data Collection method

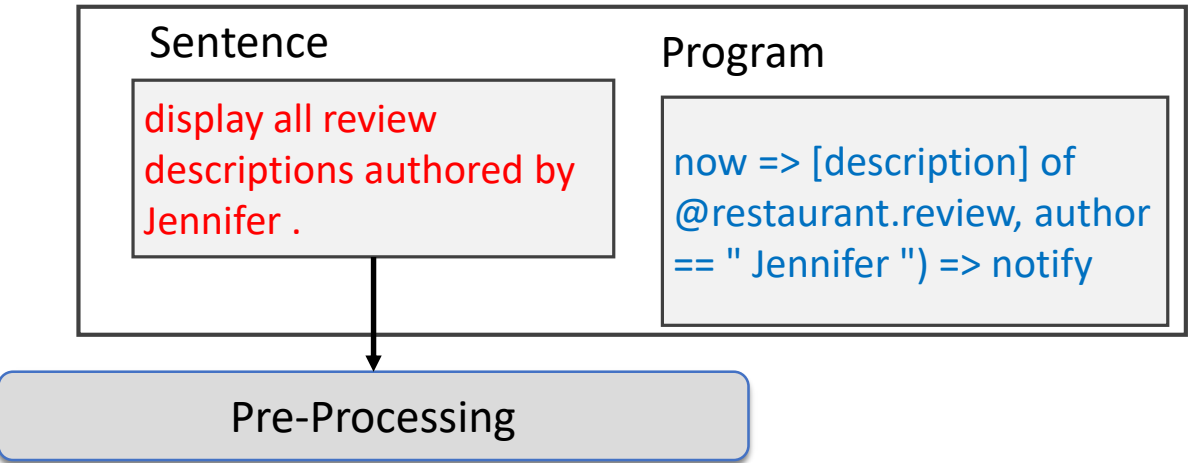
Data Collection method

English Dataset

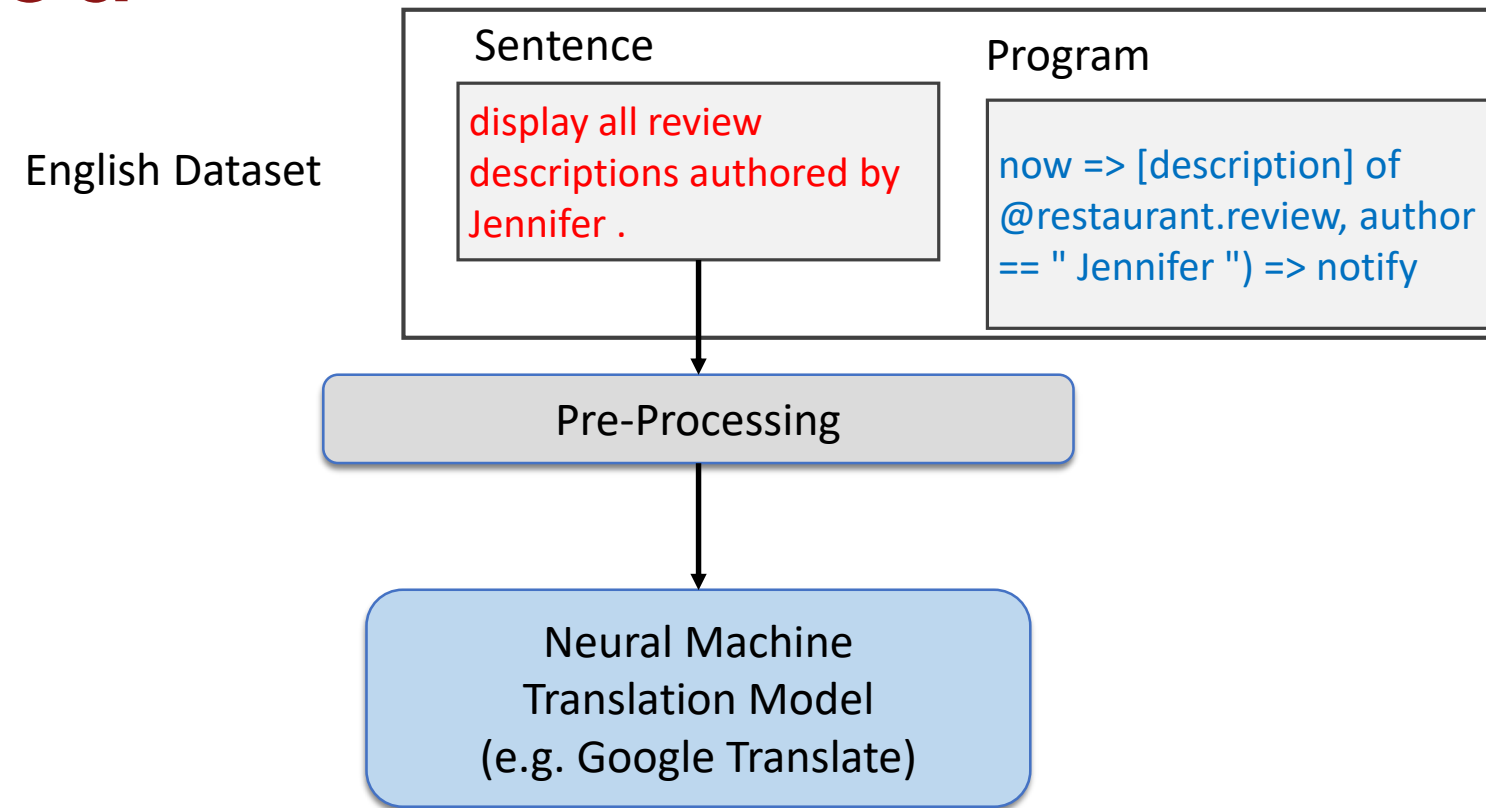
Sentence	Program
<code>display all review descriptions authored by Jennifer .</code>	<code>now => [description] of @restaurant.review, author == " Jennifer ") => notify</code>

Data Collection method

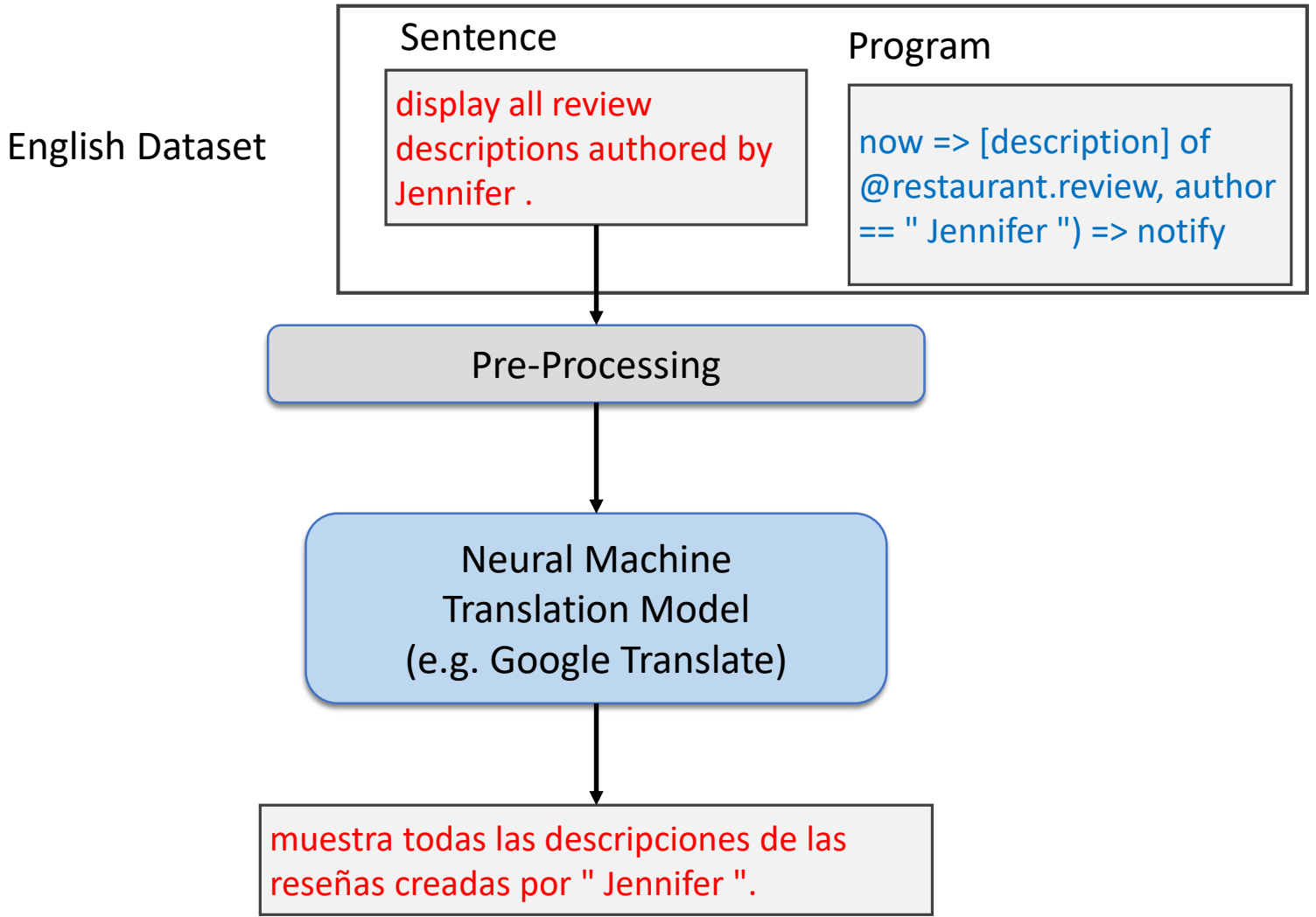
English Dataset



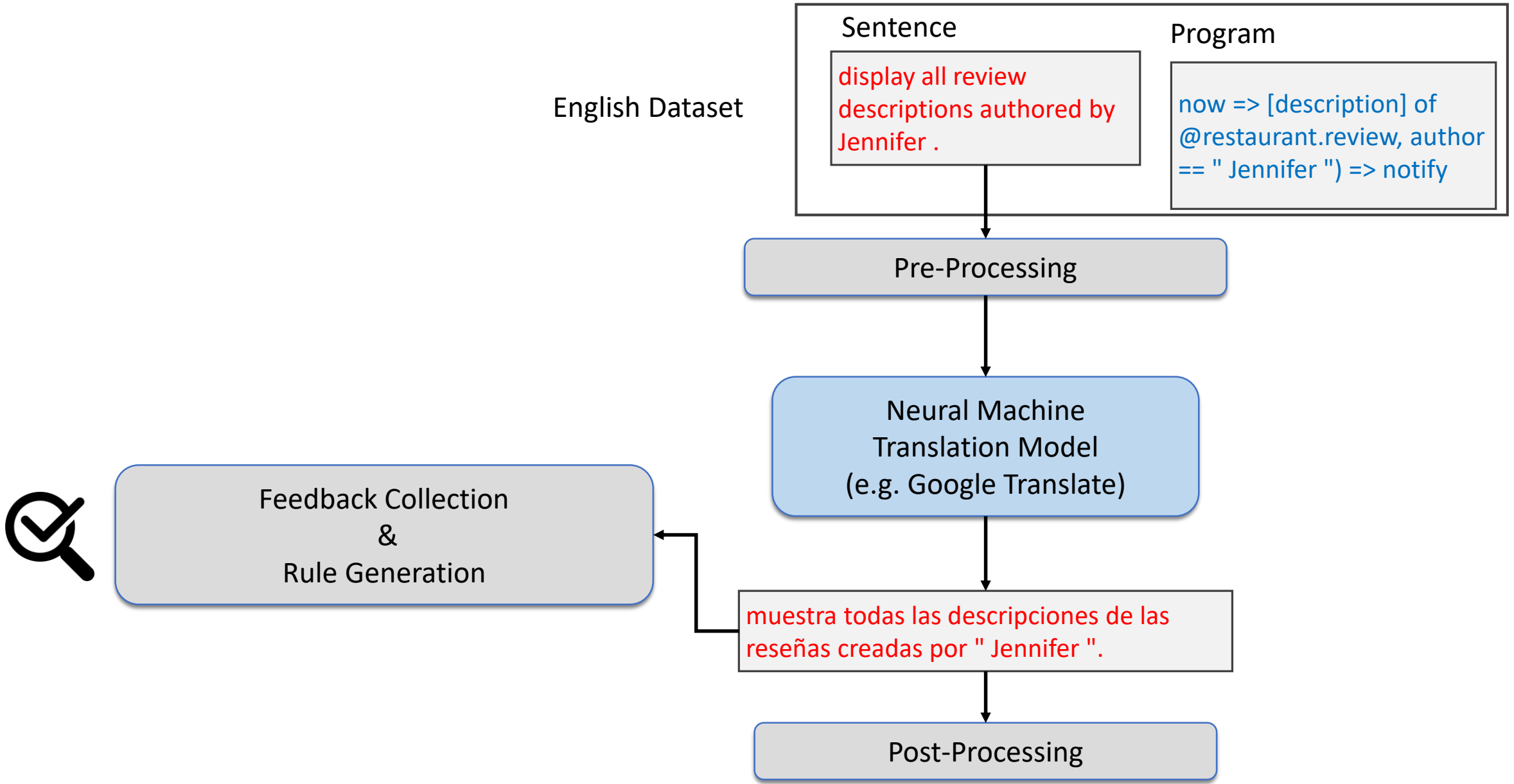
Data Collection method



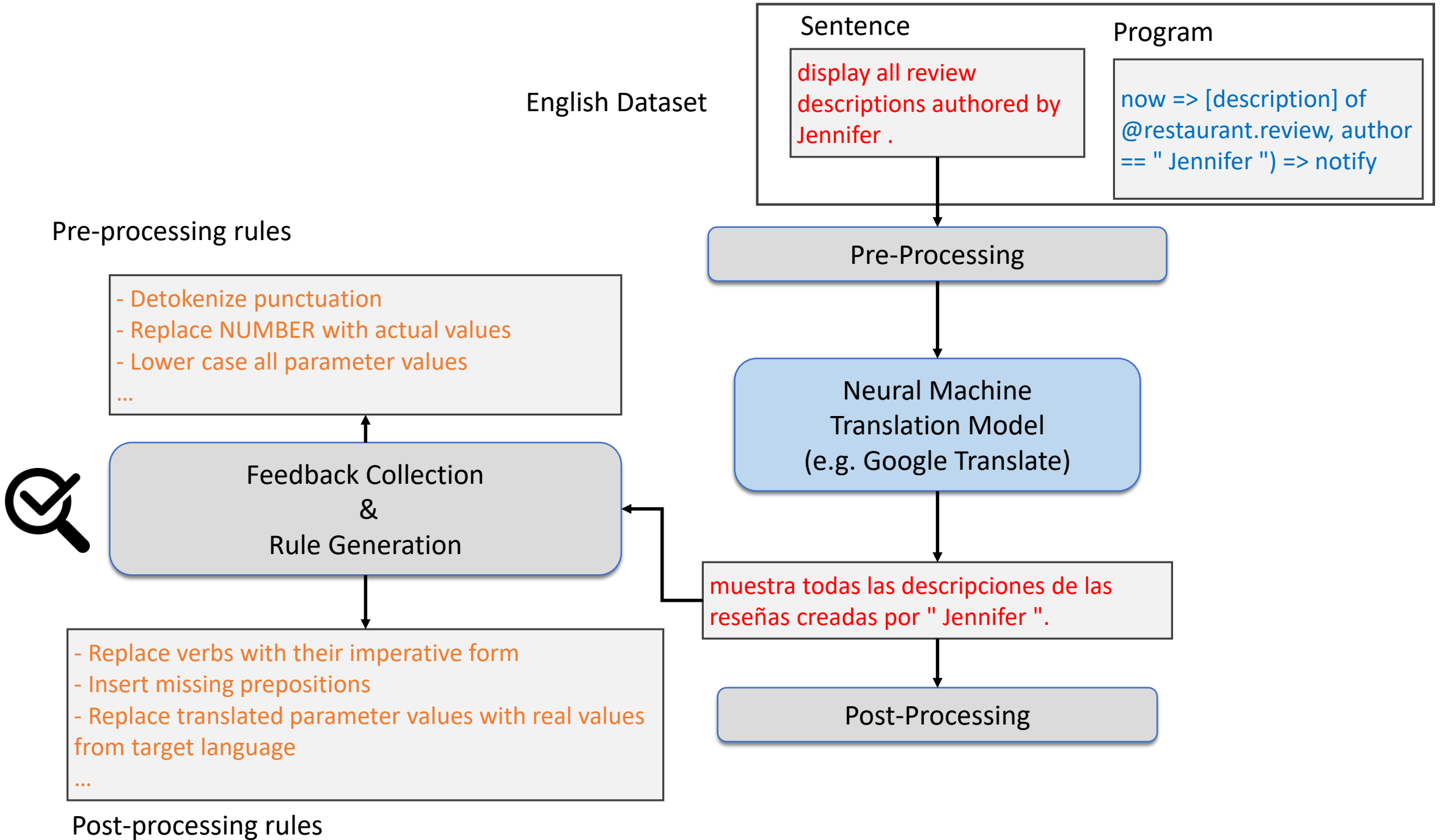
Data Collection method



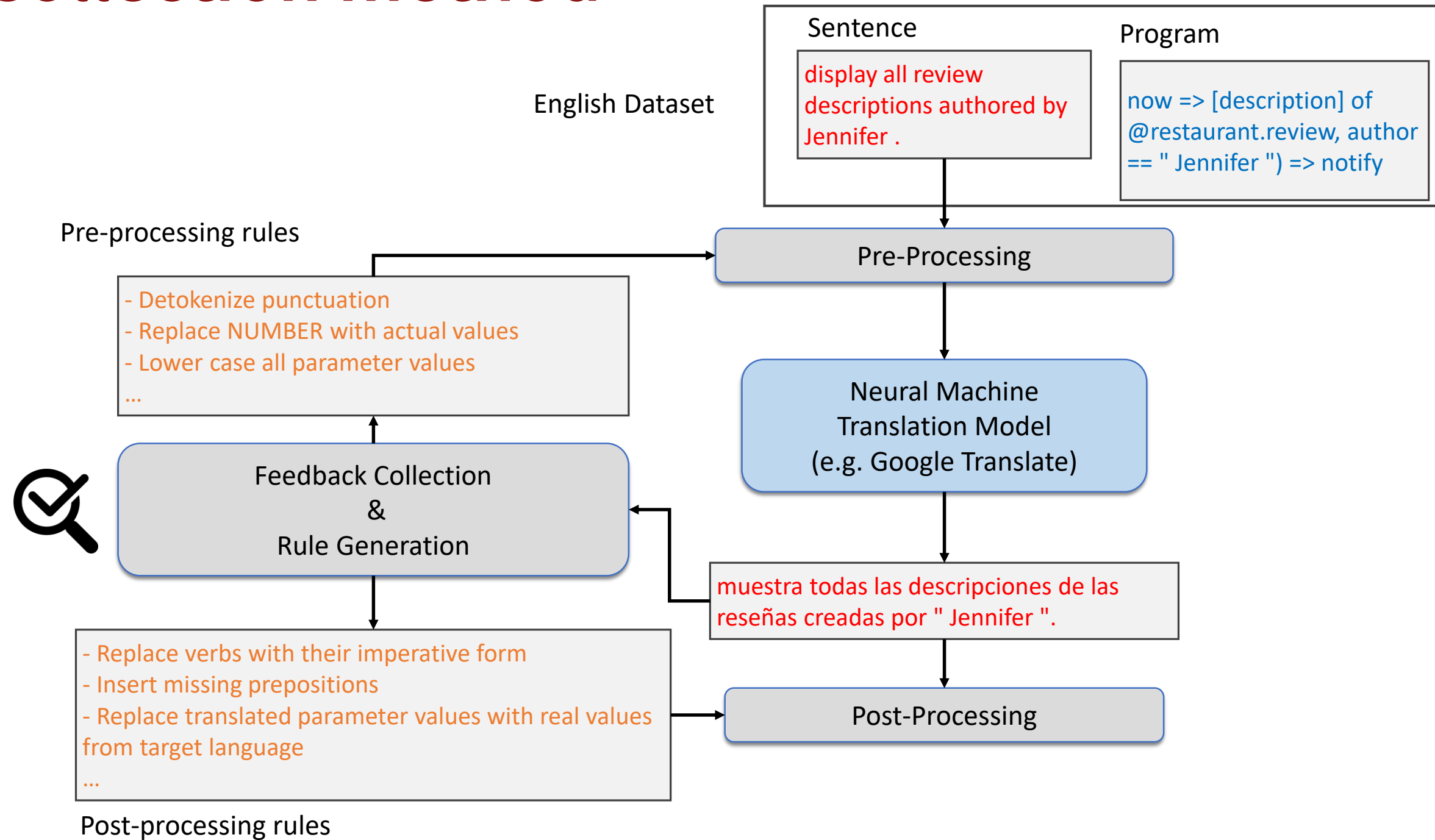
Data Collection method



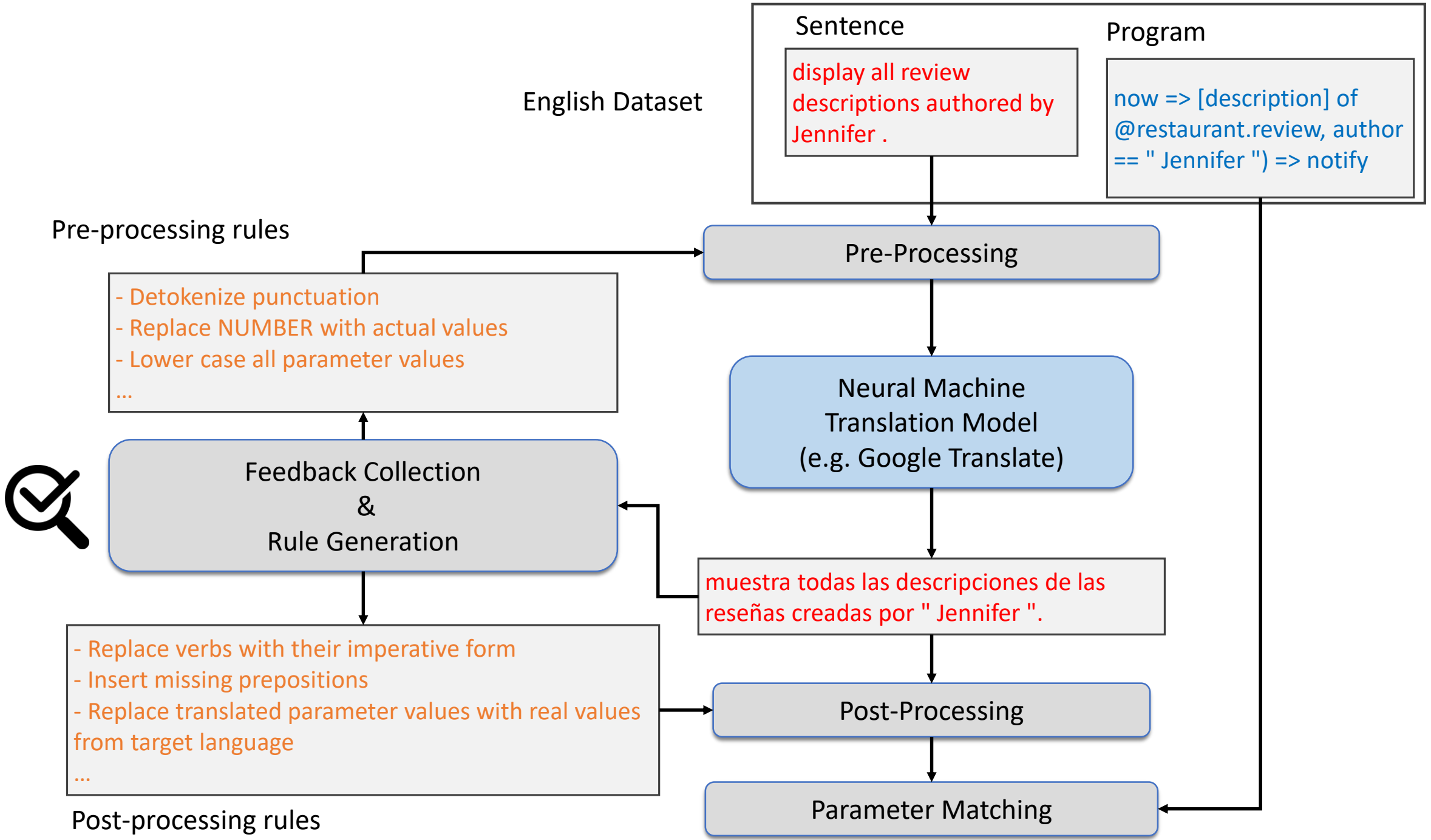
Data Collection method



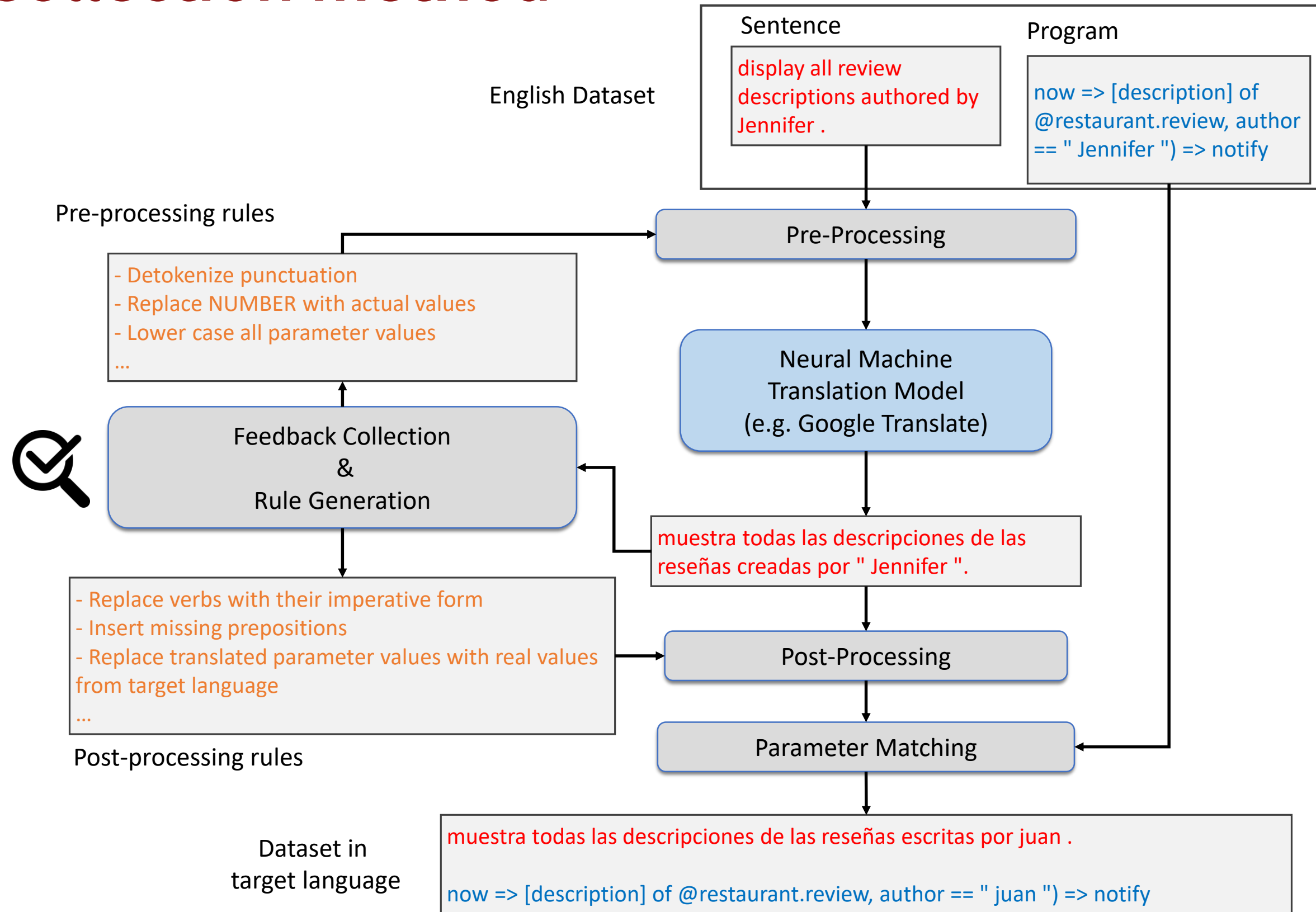
Data Collection method



Data Collection method



Data Collection method



Naive Training

display all review descriptions authored by Jennifer .

muestra todas las descripciones de las reseñas creadas por Jennifer .

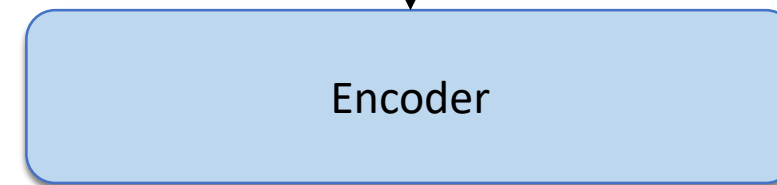
显示Jennifer撰写的所有评论描述。

Naive Training

display all review descriptions authored by Jennifer .

muestra todas las descripciones de las reseñas creadas por Jennifer .

显示Jennifer撰写的所有评论描述。

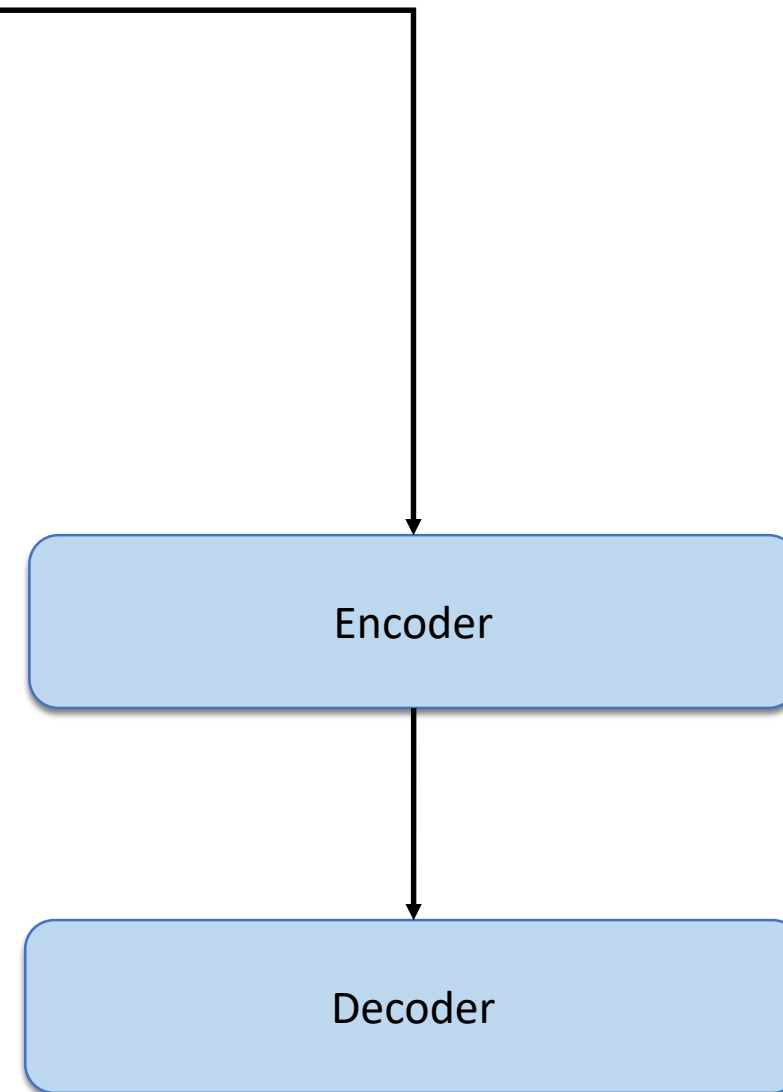


Naive Training

display all review descriptions authored by Jennifer .

muestra todas las descripciones de las reseñas creadas por Jennifer .

显示Jennifer撰写的所有评论描述。

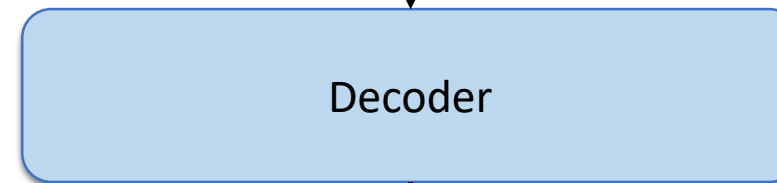
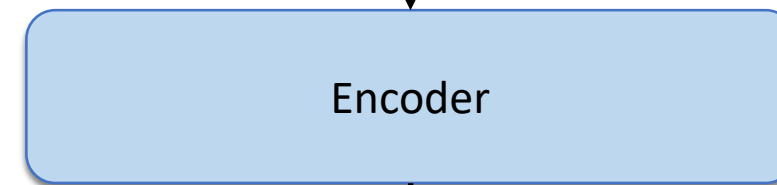


Naive Training

display all review descriptions authored by Jennifer .

muestra todas las descripciones de las reseñas creadas por Jennifer .

显示Jennifer撰写的所有评论描述。



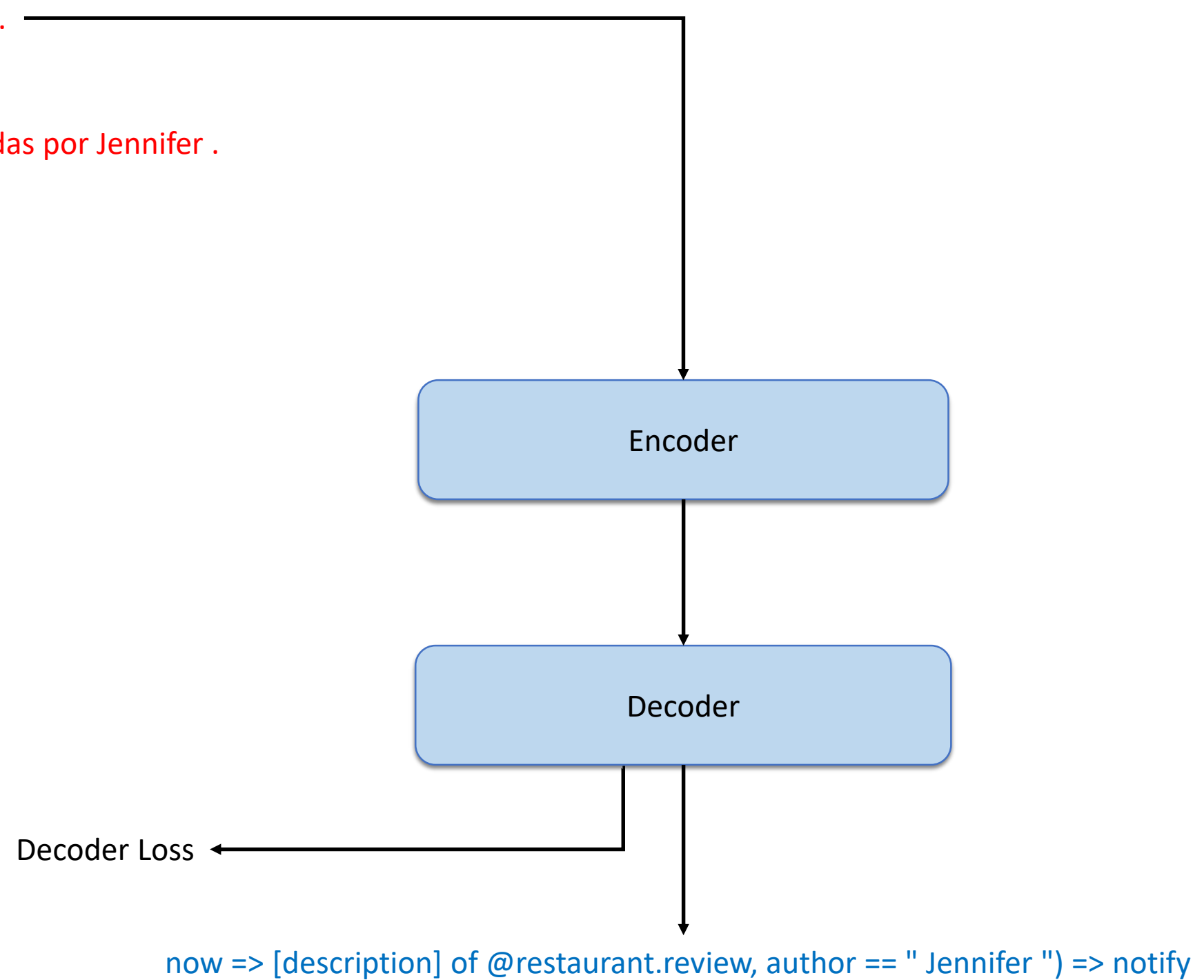
now => [description] of @restaurant.review, author == " Jennifer " => notify

Naive Training

display all review descriptions authored by Jennifer .

muestra todas las descripciones de las reseñas creadas por Jennifer .

显示Jennifer撰写的所有评论描述。

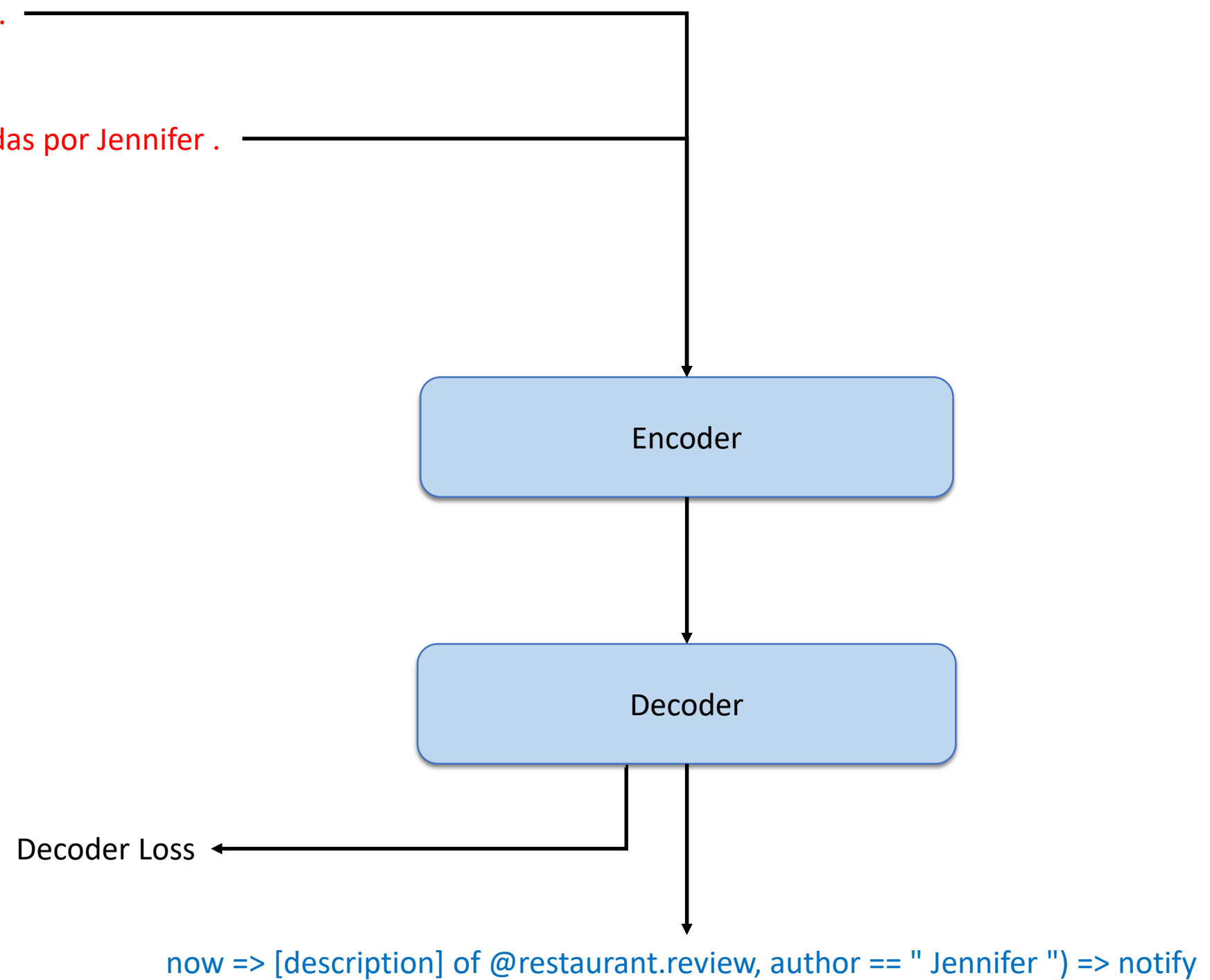


Naive Training

display all review descriptions authored by Jennifer .

muestra todas las descripciones de las reseñas creadas por Jennifer .

显示Jennifer撰写的所有评论描述。

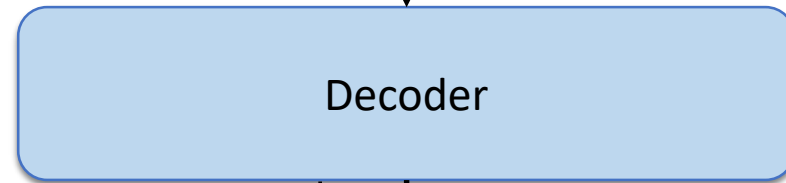
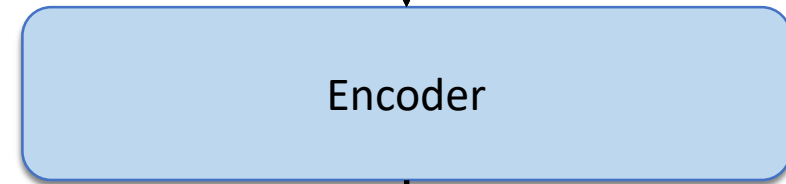
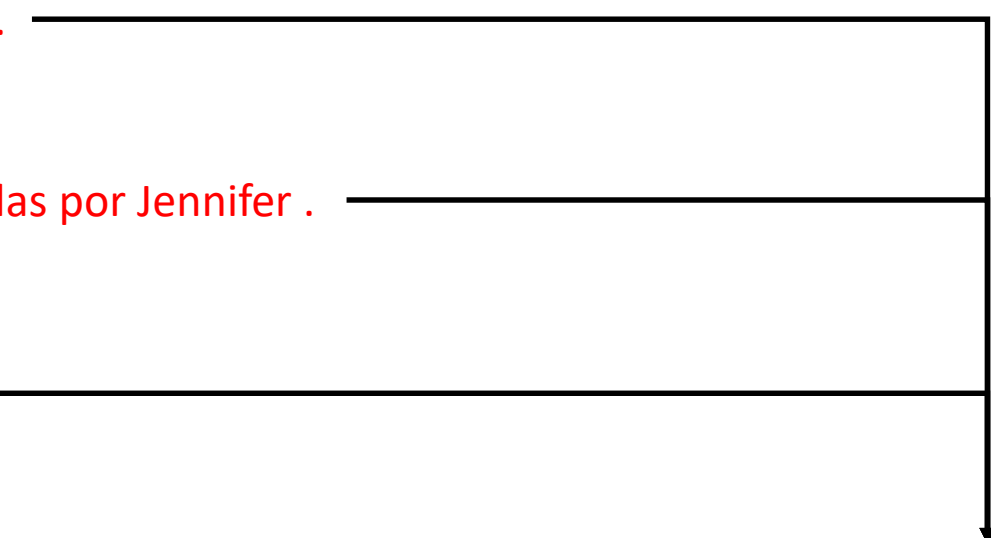


Naive Training

display all review descriptions authored by Jennifer .

muestra todas las descripciones de las reseñas creadas por Jennifer .

显示Jennifer撰写的所有评论描述。



Decoder Loss

now => [description] of @restaurant.review, author == " Jennifer ") => notify

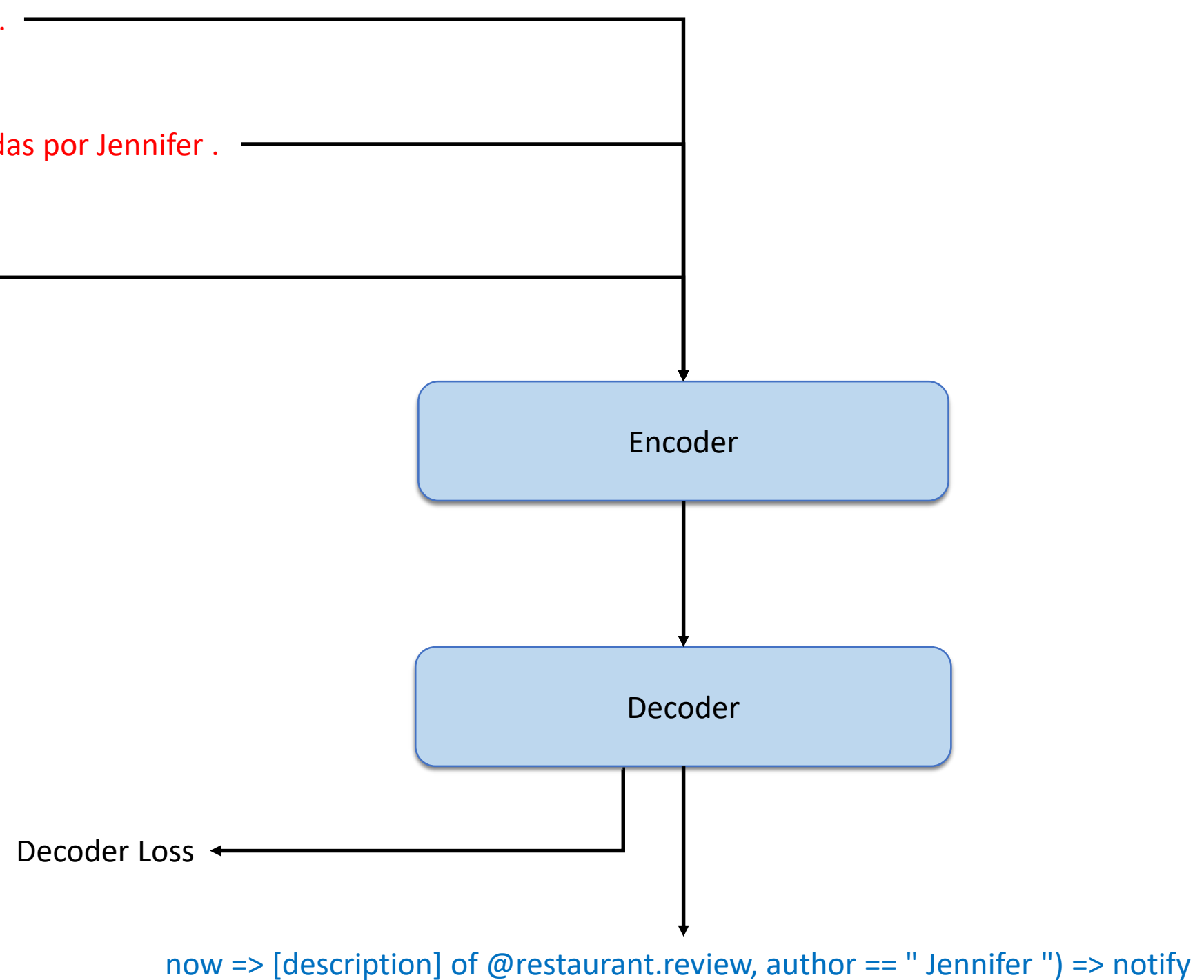
Naive Training

display all review descriptions authored by Jennifer .

muestra todas las descripciones de las reseñas creadas por Jennifer .

显示Jennifer撰写的所有评论描述。

We are not using the “knowledge” that these sentences are semantically equivalent



Training with sentence batching

display all review descriptions authored by Jennifer .

muestra todas las descripciones de las reseñas creadas por Jennifer .

显示Jennifer撰写的所有评论描述。

Training with sentence batching

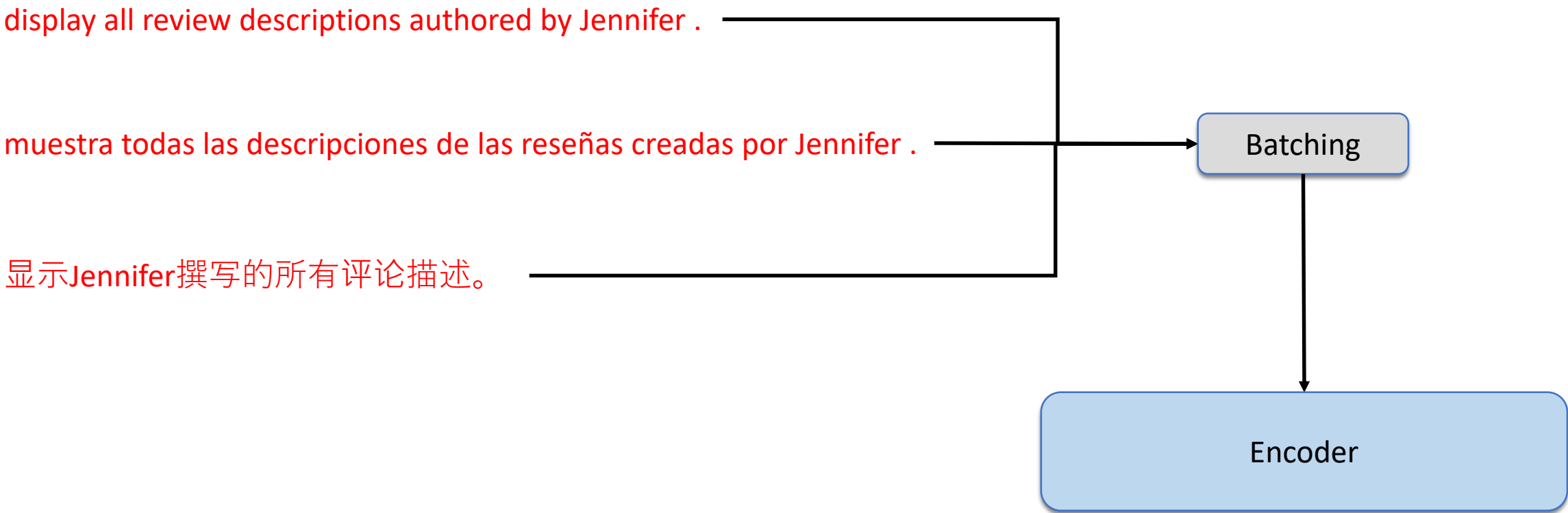
display all review descriptions authored by Jennifer .

muestra todas las descripciones de las reseñas creadas por Jennifer .

显示Jennifer撰写的所有评论描述。



Training with sentence batching

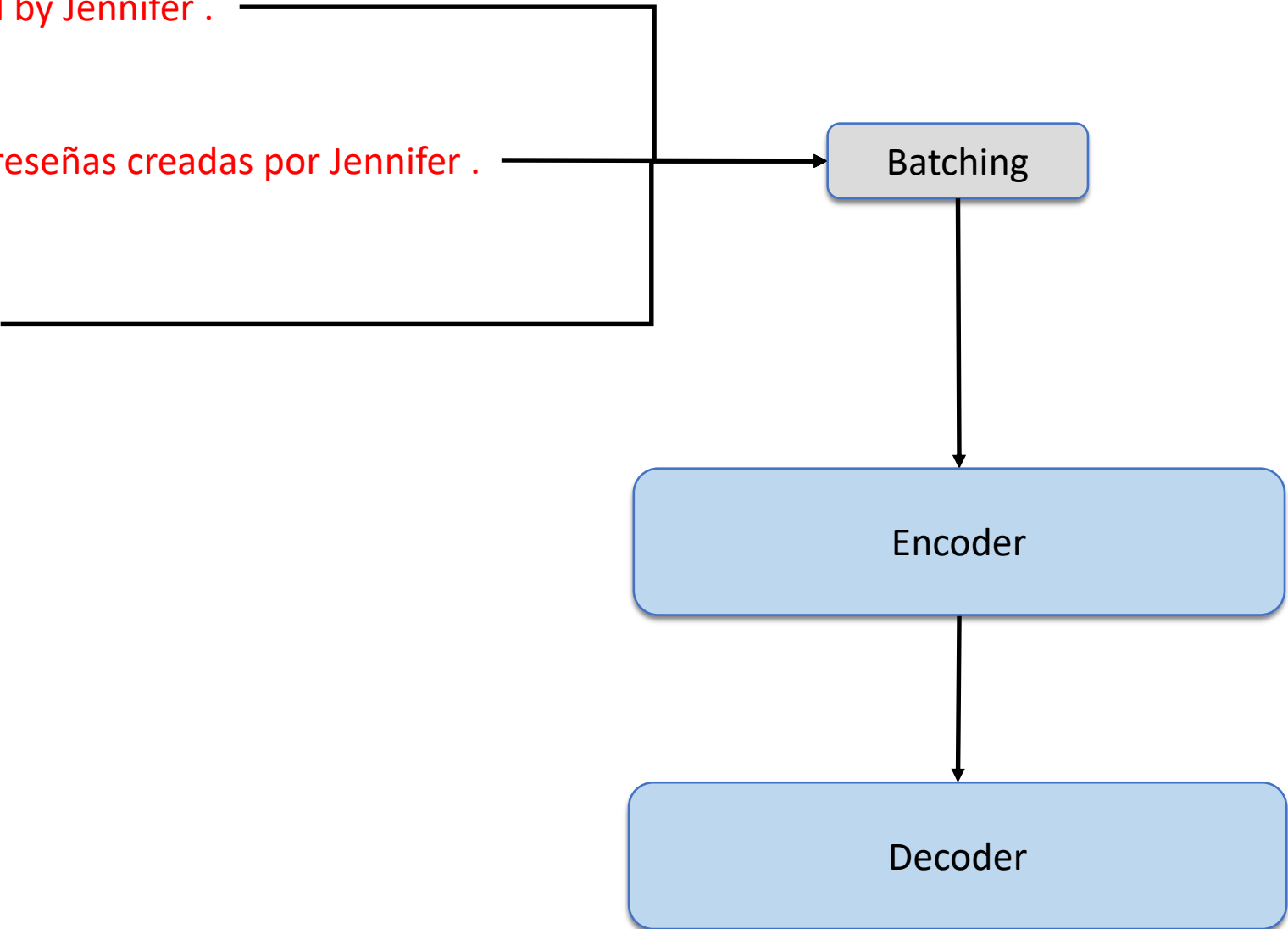


Training with sentence batching

display all review descriptions authored by Jennifer .

muestra todas las descripciones de las reseñas creadas por Jennifer .

显示Jennifer撰写的所有评论描述。

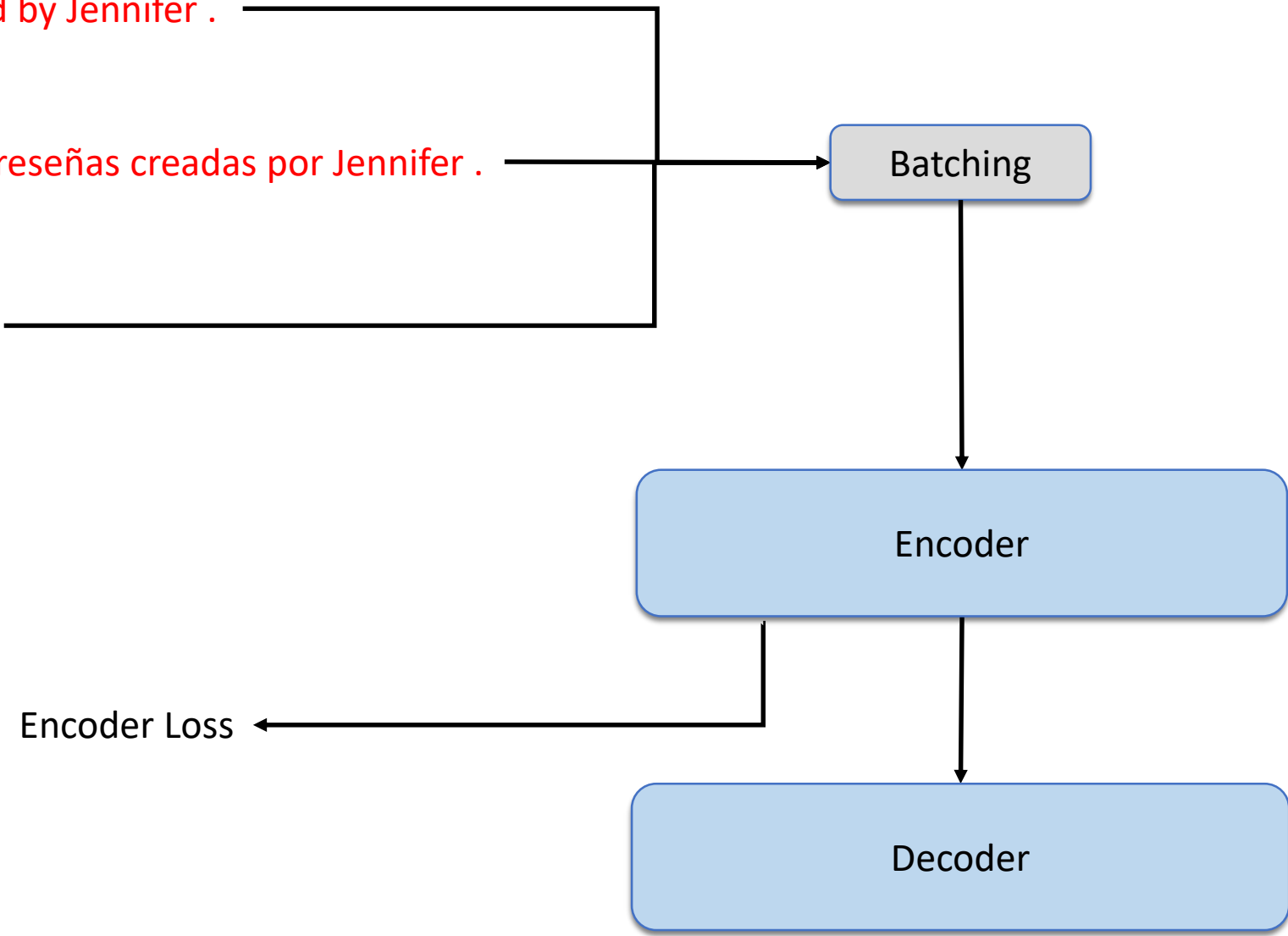


Training with sentence batching

display all review descriptions authored by Jennifer .

muestra todas las descripciones de las reseñas creadas por Jennifer .

显示Jennifer撰写的所有评论描述。

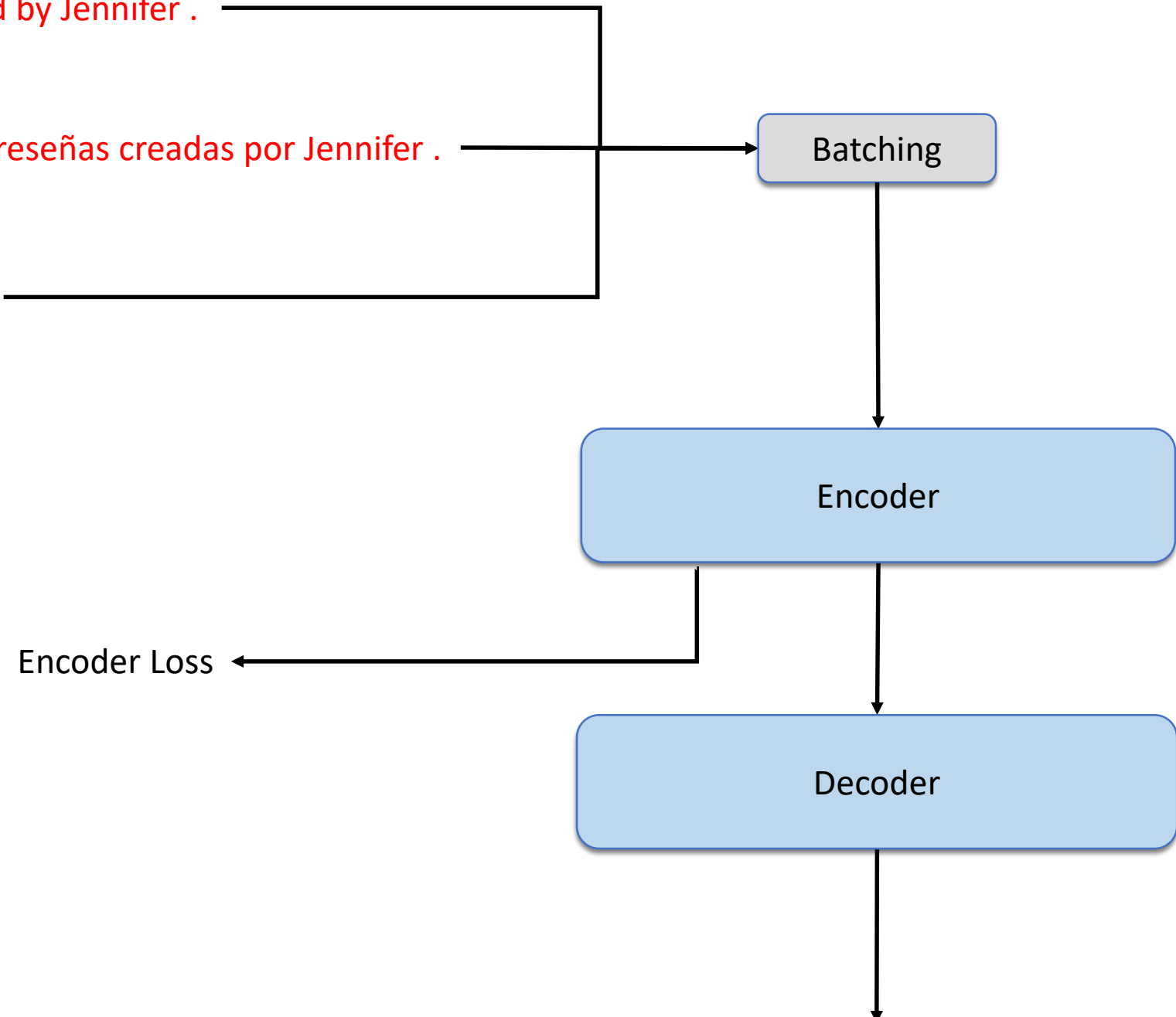


Training with sentence batching

display all review descriptions authored by Jennifer .

muestra todas las descripciones de las reseñas creadas por Jennifer .

显示Jennifer撰写的所有评论描述。



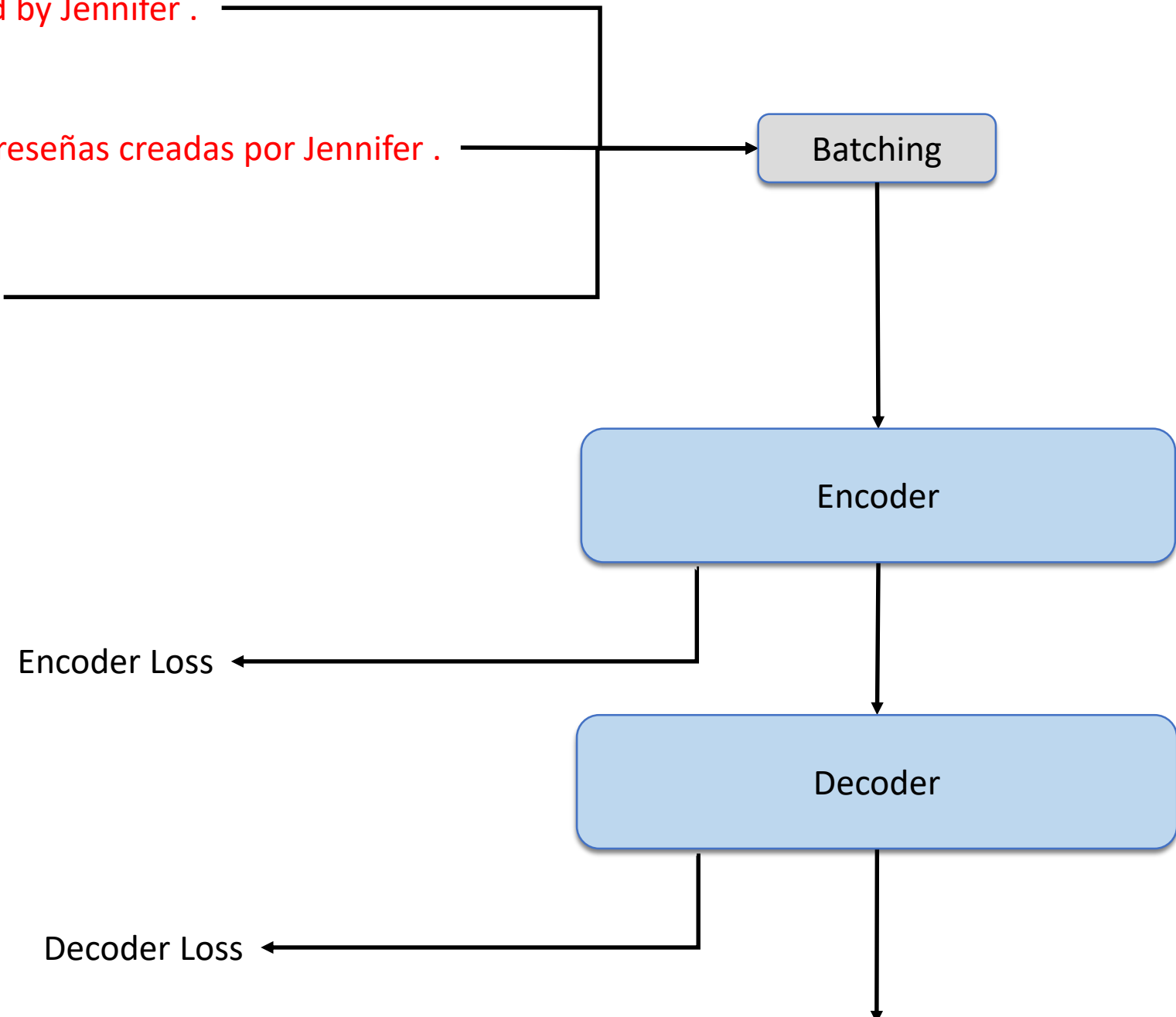
now => [description] of @restaurant.review, author == " Jennifer ") => notify

Training with sentence batching

display all review descriptions authored by Jennifer .

muestra todas las descripciones de las reseñas creadas por Jennifer .

显示Jennifer撰写的所有评论描述。



now => [description] of @restaurant.review, author == " Jennifer ") => notify

Training with sentence batching

display all review descriptions authored by Jennifer .

muestra todas las descripciones de las reseñas creadas por Jennifer .

显示Jennifer撰写的所有评论描述。

We now use both losses to guide the training

Encoder Loss

Decoder Loss

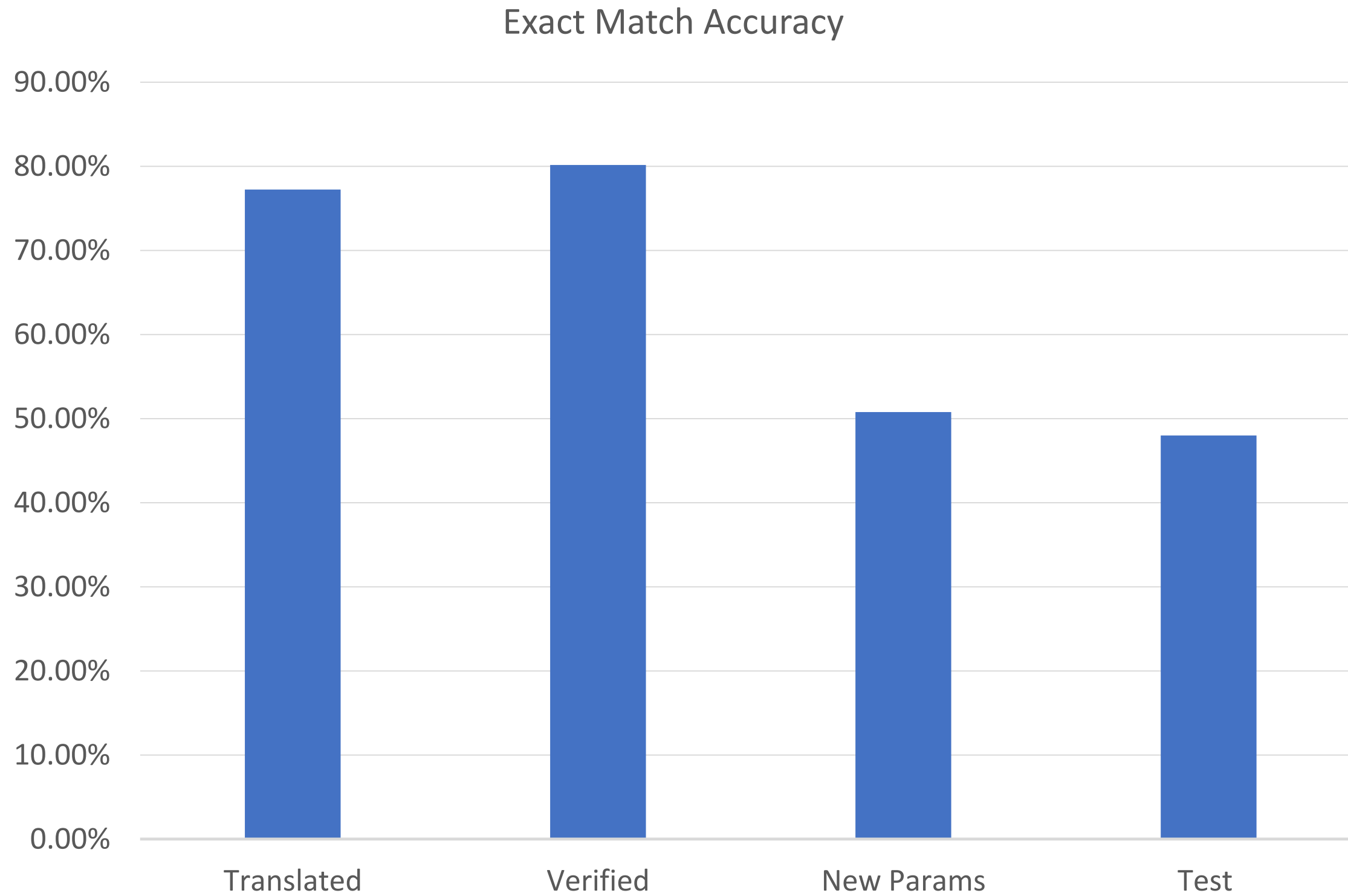
now => [description] of @restaurant.review, author == " Jennifer ") => notify

Batching

Encoder

Decoder

Experiment results (Farsi)



Challenges

Challenges

- Google translate is not perfect
- Identifying Language specific traits (single/ plural, missing prepositions, ...)
- Closing the gap between evaluation accuracy and test (real data) accuracy
- Automating and improving collection of natural parameter values for each language
- ...

Challenges

- Google translate is not perfect
- Identifying Language specific traits (single/ plural, missing prepositions, ...)
- Closing the gap between evaluation accuracy and test (real data) accuracy
- Automating and improving collection of natural parameter values for each language
- ...

Bonus:

- Started code is available free of charge!
- 18/6 project technical support
- Optional happy hours to celebrate our results
- Will be featured as a contributor in our EMNLP paper