



# Select by n-gram overlap

The `NGramOverlapExampleSelector` selects and orders examples based on which examples are most similar to the input, according to an ngram overlap score. The ngram overlap score is a float between 0.0 and 1.0, inclusive.

The selector allows for a threshold score to be set. Examples with an ngram overlap score less than or equal to the threshold are excluded. The threshold is set to -1.0, by default, so will not exclude any examples, only reorder them. Setting the threshold to 0.0 will exclude examples that have no ngram overlaps with the input.

```

from langchain.prompts import PromptTemplate
from langchain.prompts.example_selector.ngram_overlap import
NGramOverlapExampleSelector
from langchain.prompts import FewShotPromptTemplate, PromptTemplate

example_prompt = PromptTemplate(
    input_variables=["input", "output"],
    template="Input: {input}\nOutput: {output}",
)

# These are a lot of examples of a pretend task of creating antonyms.
examples = [
    {"input": "happy", "output": "sad"},
    {"input": "tall", "output": "short"},
    {"input": "energetic", "output": "lethargic"},
    {"input": "sunny", "output": "gloomy"},
    {"input": "windy", "output": "calm"},
]

```

## API Reference:

- `PromptTemplate` from `langchain.prompts`
- `NGramOverlapExampleSelector` from `langchain.prompts.example_selector.ngram_overlap`
- `FewShotPromptTemplate` from `langchain.prompts`

- `PromptTemplate` from `langchain.prompts`

```
# These are examples of a fictional translation task.
examples = [
    {"input": "See Spot run.", "output": "Ver correr a Spot."},
    {"input": "My dog barks.", "output": "Mi perro ladra."},
    {"input": "Spot can run.", "output": "Spot puede correr."},
]
```

```
example_prompt = PromptTemplate(
    input_variables=["input", "output"],
    template="Input: {input}\nOutput: {output}",
)
example_selector = NGramOverlapExampleSelector(
    # These are the examples it has available to choose from.
    examples=examples,
    # This is the PromptTemplate being used to format the examples.
    example_prompt=example_prompt,
    # This is the threshold, at which selector stops.
    # It is set to -1.0 by default.
    threshold=-1.0,
    # For negative threshold:
    # Selector sorts examples by ngram overlap score, and excludes none.
    # For threshold greater than 1.0:
    # Selector excludes all examples, and returns an empty list.
    # For threshold equal to 0.0:
    # Selector sorts examples by ngram overlap score,
    # and excludes those with no ngram overlap with input.
)
dynamic_prompt = FewShotPromptTemplate(
    # We provide an ExampleSelector instead of examples.
    example_selector=example_selector,
    example_prompt=example_prompt,
    prefix="Give the Spanish translation of every input",
    suffix="Input: {sentence}\nOutput:",
    input_variables=["sentence"],
)
```

```
# An example input with large ngram overlap with "Spot can run."
# and no overlap with "My dog barks."
```

```
print(dynamic_prompt.format(sentence="Spot can run fast."))
```

Give the Spanish translation of every input

Input: Spot can run.

Output: Spot puede correr.

Input: See Spot run.

Output: Ver correr a Spot.

Input: My dog barks.

Output: Mi perro ladra.

Input: Spot can run fast.

Output:

# You can add examples to NGramOverlapExampleSelector as well.

```
new_example = {"input": "Spot plays fetch.", "output": "Spot juega a  
buscar."}
```

```
example_selector.add_example(new_example)
```

```
print(dynamic_prompt.format(sentence="Spot can run fast."))
```

Give the Spanish translation of every input

Input: Spot can run.

Output: Spot puede correr.

Input: See Spot run.

Output: Ver correr a Spot.

Input: Spot plays fetch.

Output: Spot juega a buscar.

Input: My dog barks.

Output: Mi perro ladra.

Input: Spot can run fast.

Output:

```
# You can set a threshold at which examples are excluded.
# For example, setting threshold equal to 0.0
# excludes examples with no ngram overlaps with input.
# Since "My dog barks." has no ngram overlaps with "Spot can run fast."
# it is excluded.
example_selector.threshold = 0.0
print(dynamic_prompt.format(sentence="Spot can run fast."))
```

Give the Spanish translation of every input

Input: Spot can run.

Output: Spot puede correr.

Input: See Spot run.

Output: Ver correr a Spot.

Input: Spot plays fetch.

Output: Spot juega a buscar.

Input: Spot can run fast.

Output:

```
# Setting small nonzero threshold
example_selector.threshold = 0.09
print(dynamic_prompt.format(sentence="Spot can play fetch."))
```

Give the Spanish translation of every input

Input: Spot can run.

Output: Spot puede correr.

Input: Spot plays fetch.

Output: Spot juega a buscar.

Input: Spot can play fetch.

Output:

```
# Setting threshold greater than 1.0
example_selector.threshold = 1.0 + 1e-9
print(dynamic_prompt.format(sentence="Spot can play fetch."))
```

Give the Spanish translation of every input

Input: Spot can play fetch.

Output: