Inference Parameters

Inference parameters for vLLM APIs.

(sampling-params)=

Sampling Parameters

```{eval-rst}
.. autoclass:: vllm.SamplingParams
    :members:
```

(pooling-params)=

Pooling Parameters

```{eval-rst}
.. autoclass:: vllm.PoolingParams
    :members:
```