🏠 ■ Modules ■ Data connection ■ Document loaders ■ Integrations ■ Jupyter Notebook

Jupyter Notebook

> Jupyter Notebook (formerly `IPython Notebook`) is a web-based interactive computational environment for creating notebook documents.

This notebook covers how to load data from a `Jupyter notebook (.html)` into a format suitable by LangChain.

```
from langchain.document_loaders import NotebookLoader
```

```
loader = NotebookLoader(
    "example_data/notebook.html",
    include_outputs=True,
    max_output_length=20,
    remove_newline=True,
)
```

`NotebookLoader.load()` loads the `.html` notebook file into a `Document` object.

Parameters:

- `include_outputs` (bool): whether to include cell outputs in the resulting document (default is False).
- `max_output_length` (int): the maximum number of characters to include from each cell output (default is 10).
- `remove_newline` (bool): whether to remove newline characters from the cell sources and outputs (default is False).
- `traceback` (bool): whether to include full traceback (default is False).

```
loader.load()
```

```
    [Document(page_content='\'markdown\' cell: \'[\'# Notebook\', \'\', \'This notebook covers how to load
data from an .html notebook into a format suitable by LangChain.\']\'\n\n \'code\' cell: \'[\'from
langchain.document_loaders import NotebookLoader\']\'\n\n \'code\' cell: \'[\'loader =
NotebookLoader("example_data/notebook.html")\']\'\n\n \'markdown\' cell: \'[\'`NotebookLoader.load()` loads
the `.html` notebook file into a `Document` object.\', \'\', \'**Parameters**:\', \'\', \'* `include_outputs`
(bool): whether to include cell outputs in the resulting document (default is False).\', \'*
`max_output_length` (int): the maximum number of characters to include from each cell output (default is
10).\', \'* `remove_newline` (bool): whether to remove newline characters from the cell sources and outputs
(default is False).\', \'* `traceback` (bool): whether to include full traceback (default is False).\']\'\n\n
\'code\' cell: \'[\'loader.load(include_outputs=True, max_output_length=20, remove_newline=True)\']\'\n\n',
metadata={'source': 'example_data/notebook.html'})]
```