Ze2nb: Zeppelin to jupyter notebook API

Wenqiang Feng and Ryan Blue

June 21, 2019

CONTENTS

Welcome to our **ze2nb: Zeppelin to jupyter notebook API**! This library is based on Ryan Blue's Jupyter/Zeppelin conversion: https://github.com/rdblue/jupyter-zeppelin. The PDF version can be downloaded from HERE.

You can install the `ze2nb` from [PyPI](https://pypi.org/project/ze2nb):

```
pip install ze2nb
```

CHAPTER

ONE

PREFACE

Chinese proverb

Good tools are prerequisite to the successful execution of a job. – old Chinese proverb

1.1 About

1.1.1 About this API

This document is the `API` book for our ze2nb: Zeppelin to jupyter notebook [zeppelin2nb] API. The PDF version can be downloaded from HERE. This library is based on Ryan Blue's Jupyter/Zeppelin conversion: [jupyter-zeppelin]. **You may download and distribute it. Please be aware, however, that the note contains typos as well as inaccurate or incorrect description.**

The `API` assumes that the reader has a preliminary knowledge of `python` programing and `Linux`. And this document is generated automatically by using sphinx.

1.1.2 About the author

- **Wenqiang Feng**
  - Data Scientist and PhD in Mathematics
  - University of Tennessee at Knoxville
  - Webpage: http://web.utk.edu/~wfeng1/
  - Email: von198@gmail.com
- **Ryan Blue**
  - github: https://github.com/rdblue

- **Biography**

  Wenqiang Feng is Data Scientist within DST's Applied Analytics Group. Dr. Feng's responsibilities include providing DST clients with access to cutting-edge skills and technologies, including Big Data analytic solutions, advanced analytic and data enhancement techniques and modeling.

  Dr. Feng has deep analytic expertise in data mining, analytic systems, machine learning algorithms, business intelligence, and applying Big Data tools to strategically solve industry problems in a cross-functional business. Before joining DST, Dr. Feng was an IMA Data Science Fellow at The Institute for Mathematics and its Applications (IMA) at the University of Minnesota. While there, he helped startup companies make marketing decisions based on deep predictive analytics.

  Dr. Feng graduated from University of Tennessee, Knoxville, with Ph.D. in Computational Mathematics and Master's degree in Statistics. He also holds Master's degree in Computational Mathematics from Missouri University of Science and Technology (MST) and Master's degree in Applied Mathematics from the University of Science and Technology of China (USTC).

- **Declaration**

  The work of Wenqiang Feng was supported by the IMA, while working at IMA. However, any opinion, finding, and conclusions or recommendations expressed in this material are those of the author and do not necessarily reflect the views of the IMA, UTK and DST.

1.2 Acknowledgement

At here, Wenqiang Feng would like to thank **Weiyu Wang** at Missouri University of Science and Technology and **Jiangtao (Lotto) Xie** at Purdue University for the unit testing and valuable disscussion.

1.3 Feedback and suggestions

Your comments and suggestions are highly appreciated. I am more than happy to receive corrections, suggestions or feedbacks through email (Wenqiang Feng: von198@gmail.com) for improvements.

CHAPTER

TWO

HOW TO INSTALL

2.1 Install with `pip`

You can install the `ze2nb` from [PyPI](https://pypi.org/project/ze2nb):

```
pip install ze2nb
```

2.2 Install from Repo

2.2.1 Clone the Repository

```
git clone https://github.com/runawayhorse001/ze2nb.git
```

2.2.2 Install

```
cd zeppelin2nb
pip install -r requirements.txt
python setup.py install
```

2.2.3 Uninstall

```
pip uninstall ze2nb
```

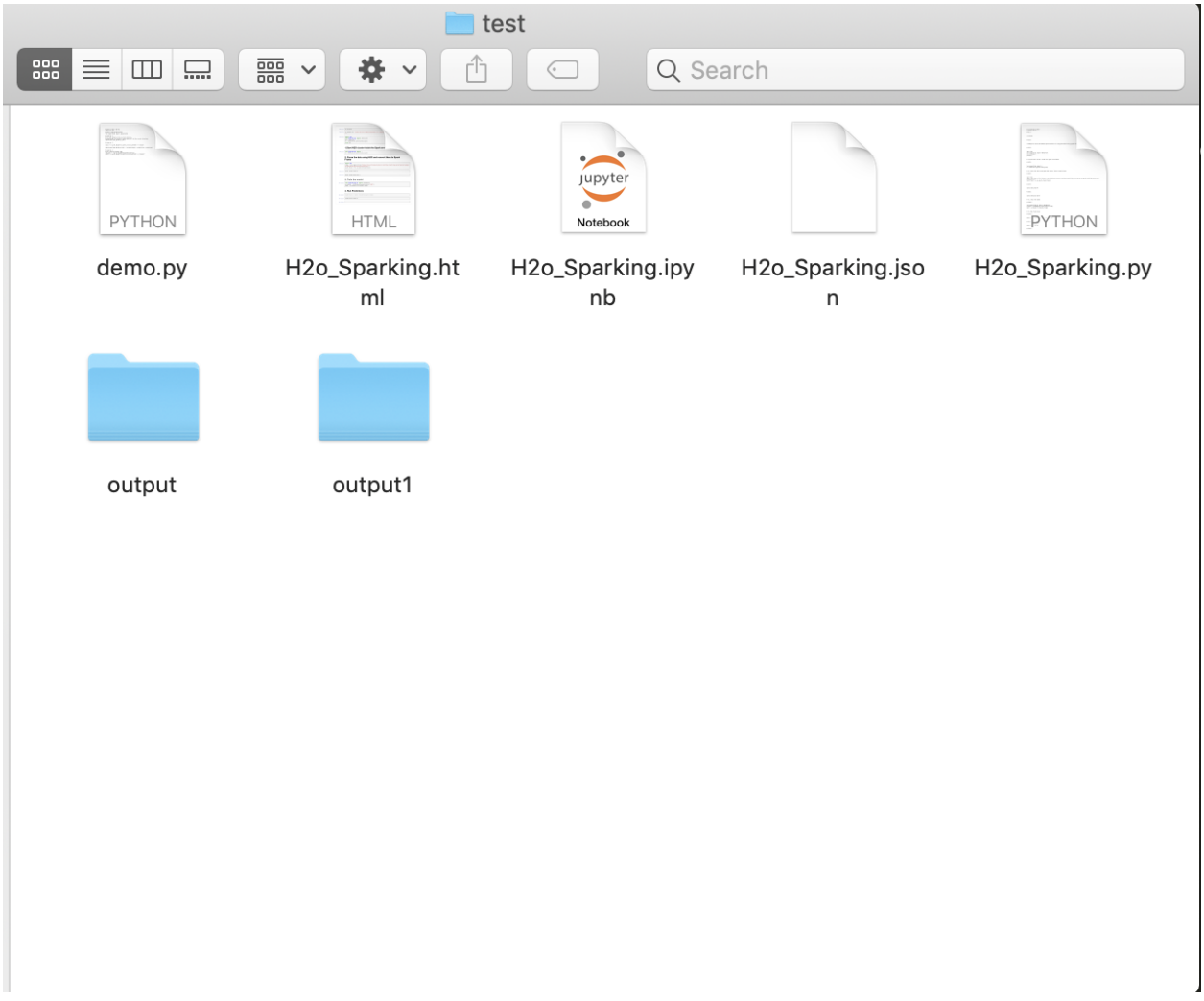
2.2.4 Test

```
cd zeppelin2nb/test
python demo.py
```

test.py

```
# import python library
import os, sys

# import zeppelin2nb module
from ze2nb import ze2nb


# scenario 1
# file and output at the current directory
# output path, the default output path will be the current directory
ze2nb('H2o_Sparking.json')

# scenario 2
output = os.path.abspath(os.path.join(sys.path[0])) +'/output'
ze2nb('H2o_Sparking.json', out_path=output, to_html=True, to_py=True)

# scenario 3
# with load and output path
load_path = '/Users/******/Documents/MyJson/'
output = os.path.abspath(os.path.join(sys.path[0])) +'/output1'
ze2nb('H2o_GBM.json', load_path=load_path, out_path=output, to_
↪html=True, to_py=True)
```

Results:


test
Search
PYTHON
demo.py
HTML
H2o_Sparking.html
jupyter
Notebook
H2o_Sparking.ipynb
H2o_Sparking.json
PYTHON
H2o_Sparking.py
output
output1

CHAPTER

THREE

ZEPPELIN TO JUPYTER NOTEBOOK

3.1 Basic Functions

3.1.1 mkdir

ze2nb.ze2nb.**mkdir**(*path*)

Make a new directory

Parameters **path** – the directory path

Author Wenqiang Feng

Email von198@gmail.com

3.1.2 file_load

ze2nb.ze2nb.**file_load**(*file_name*)

load zeppelin .json file

Parameters **file_name** – the input .json file name

Returns encoded text content from .json file

Author Wenqiang Feng

Email von198@gmail.com

3.1.3 table_cell_to_html

ze2nb.ze2nb.**table_cell_to_html**(*cell*)

Formats a cell from a Zeppelin TABLE as HTML.

Parameters **cell** – cell from Zeppelin

Returns zeppelin TABLE as HTML

Author Ryan Blue

Github https://github.com/rdblue

3.1.4 table_to_html

ze2nb.ze2nb.**table_to_html**(*tsv*)

Formats the tab-separated content of a Zeppelin TABLE as HTML.

Parameters **cell** – cell from Zeppelin

Returns zeppelin TABLE as HTML

Author Ryan Blue

Github https://github.com/rdblue

3.1.5 convert_parsed

ze2nb.ze2nb.**convert_parsed**(*zeppelin_note*)

Converts a Zeppelin note from parsed JSON to a Jupyter Notebook.

Parameters **zeppelin_note** – encoded JSON file

Return notebook_name the JSON notebook name

Return notebook_name the parsed Jupyter notebook content

Author Ryan Blue and Wenqiang Feng

Github https://github.com/rdblue

Email von198@gmail.com

3.1.6 write_notebook

ze2nb.ze2nb.**write_notebook**(*notebook_name*, *notebook*, *out_path=None*)

Writes parsed JSON notebook to a Jupyter notebook .ipynb file from the notebook name.

Parameters

- **notebook_name** – JSON notebook name and output Jupyter notebook name
- **notebook** – parsed JSON notebook contents
- **out_path** – Jupyter notebook output path, the default output path is current directory.

Author Wenqiang Feng and Ryan Blue

Email von198@gmail.com

Github https://github.com/rdblue

3.2 ze2nb

ze2nb.ze2nb.**ze2nb**(*file_name*, *load_path=None*, *out_path=None*, *to_nb=True*, *to_html=True*, *to_py=True*)

Parameters

- **file_name** – the input JSON file name
- **load_path** – the load path for the input JSON file
- **out_path** – the output path for the converted files
- **to_nb** – the flag for keeping .ipynb
- **to_html** – the flag for converting to .html
- **to_py** – the flag for converting to .py

Author Wenqiang Feng and Ryan Blue

Email von198@gmail.com

Github https://github.com/rdblue

CHAPTER

FOUR

CONVERTING DEMOS

The following demos are designed to show how to use `zepplin2nb` to convert the `.json` to `.ipynb`, `.py` and `.html`.

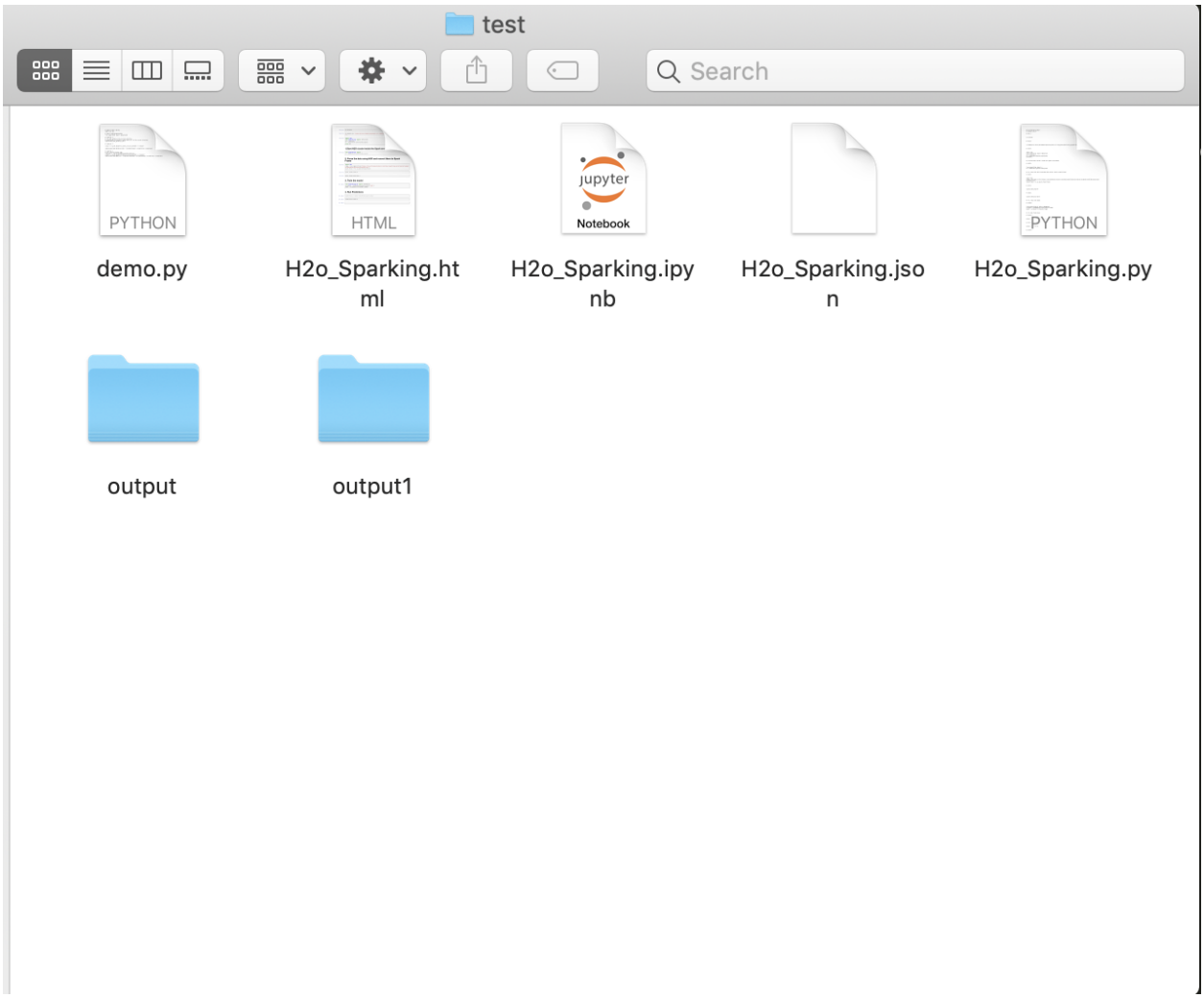
4.1 Converting in one function

For example:

```
# import python library
import os, sys

# import zeppelin2nb module
from ze2nb import ze2nb


# scenario 1
# file and output at the current directory
# output path, the default output path will be the current directory
ze2nb('H2o_Sparking.json')

# scenario 2
output = os.path.abspath(os.path.join(sys.path[0])) +'/output'
ze2nb('H2o_Sparking.json', out_path=output, to_html=True, to_py=True)

# scenario 3
# with load and output path
load_path = '/Users/******/Documents/MyJson/'
output = os.path.abspath(os.path.join(sys.path[0])) +'/output1'
ze2nb('H2o_GBM.json', load_path=load_path, out_path=output, to_
↪html=True, to_py=True)
```

4.2 Converted results

Result:

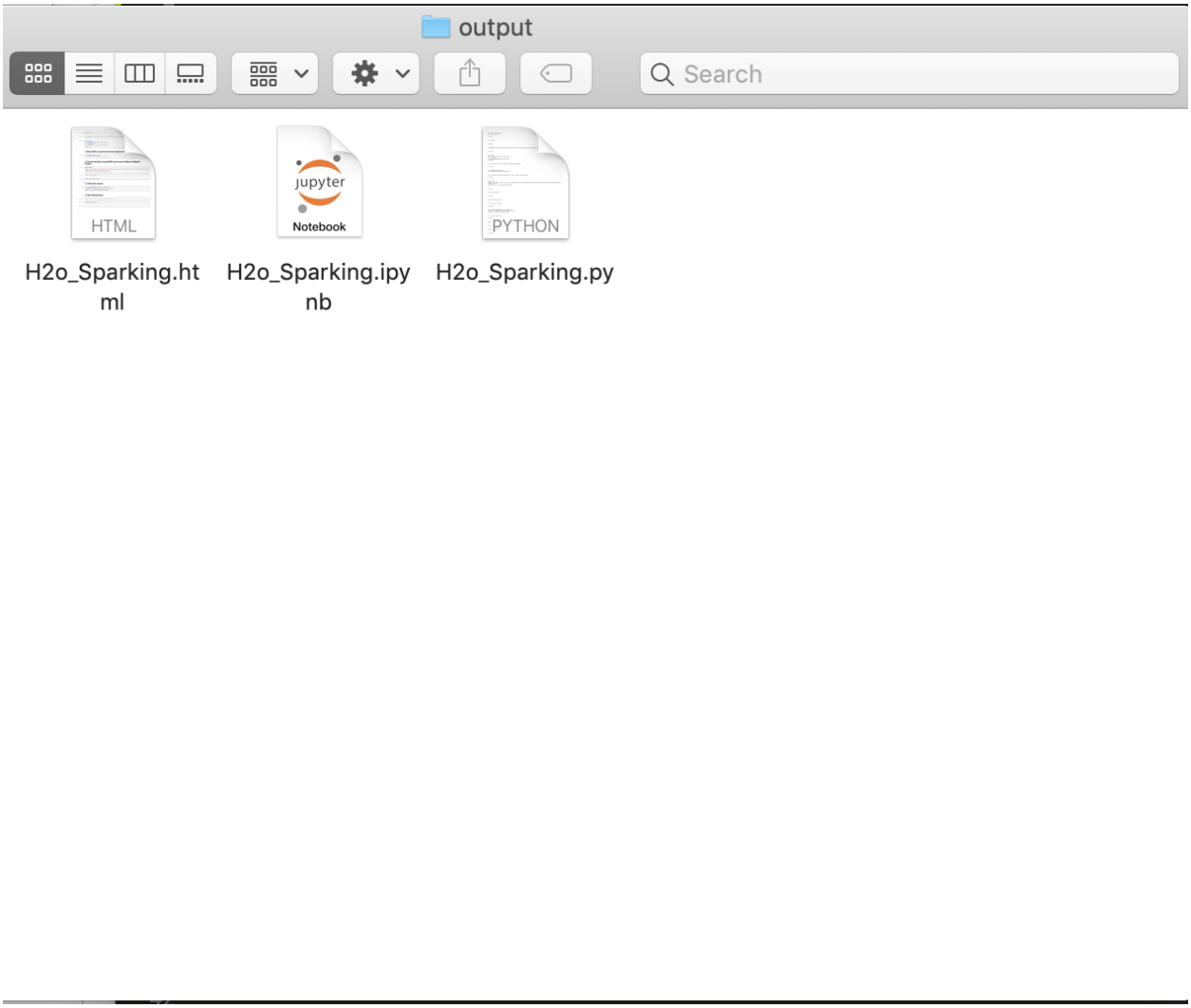


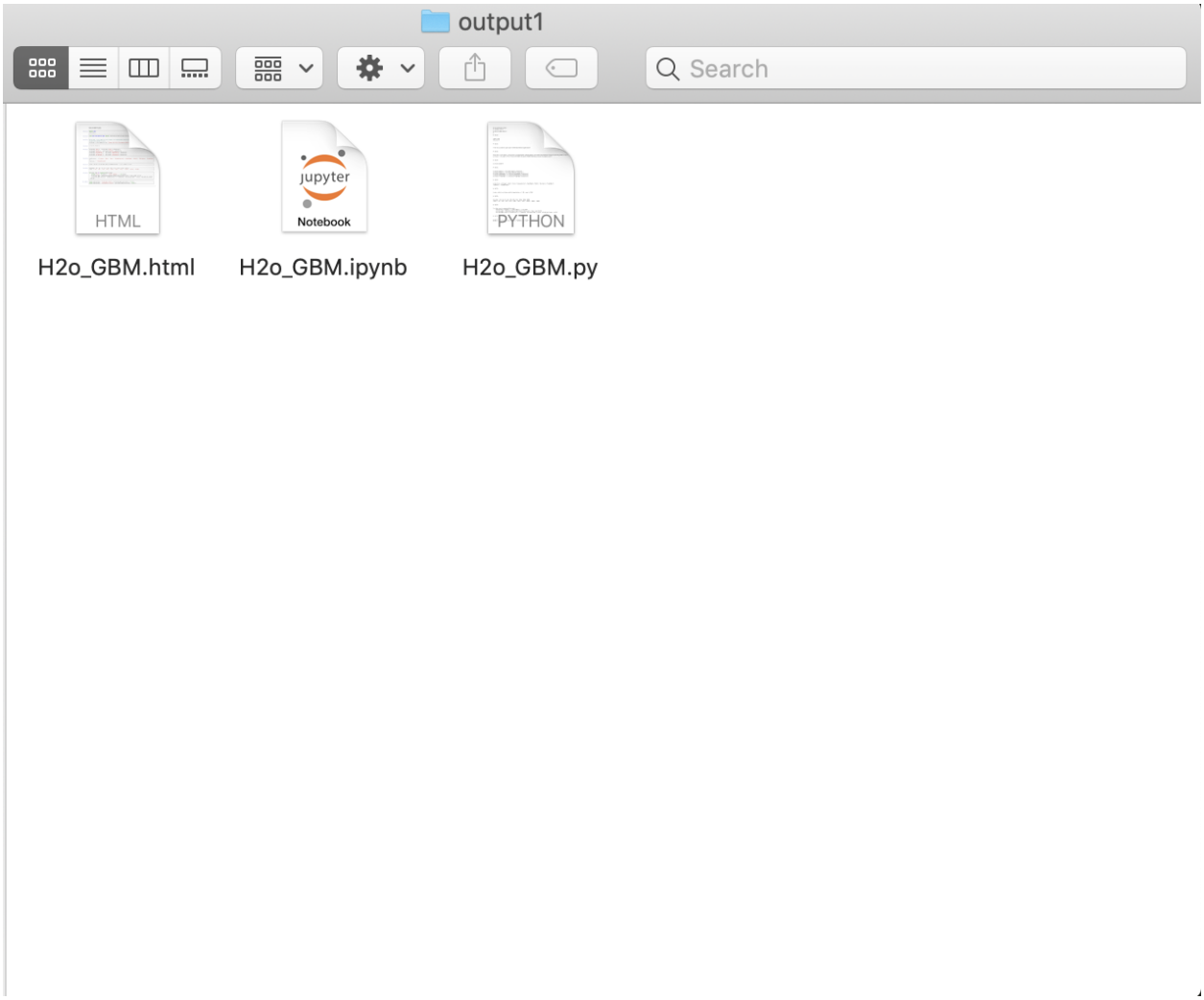
Results in output:

Results in output1:

```
       .,,.
     ,;;*;;;;,
    .-'``;-');;.
   /'   .-.   /*;;
 .'     \d     \;;                .;;;,
/  o      `      \;      ,__.     ,;*;;;*;,
\__, _.__,'     \_.-') __)--.;;;;;*;;;;,
 `""`;;;\        /-')_) __)   `\' ';;;;;;
    ;*;;;         -') `)_)   |\ |  ;;;;*;
```

(continues on next page)


output
Search
HTML
jupyter
Notebook
PYTHON
H2o_Sparking.html
H2o_Sparking.ipynb
H2o_Sparking.py


output1
Search
HTML
jupyter
Notebook
PYTHON
H2o_GBM.html
H2o_GBM.ipynb
H2o_GBM.py

(continued from previous page)

```
   ;;;;|         `---`     O | | ;;*;;;
   *;*;\|                    O  / ;;;;;*
  ;;;;;/|     .-------\        / ;*;;;;;
 ;;;*;/ \     |          '.    (`. ;;;*;;;
 ;;;;;'. ;    |            )    \ | ;;;;;;
 ,;*;;;;\/    |.          /    /` | ';;;*;
  ;;;;;;/     |/         /    /__/    ';;;
  '*wf*/      |         /     |       ;*;
       `"""`          `"""`      ;'
```

CHAPTER

FIVE

MAIN REFERENCE

BIBLIOGRAPHY

[zeppelin2nb] Wenqiang Feng and Ryan Blue. Zeppelin notebook to jupyter notebook Library API, 2019.

[jupyter-zeppelin] Ryan Blue. Jupyter/Zeppelin conversion, 2017.