

Readme file

There are 3 main file in our project:

1. 2 files jupyter notebook

*MNB.ipynb and BNB.ipynb show the training phase for Multinomial Naïve Bayes and Bernoulli Naïve Bayes and print the classification report. Two above file are both jupyter notebook

To see or run each step in those files,in command line:

`cd file` (file mean the folder that contains our project)

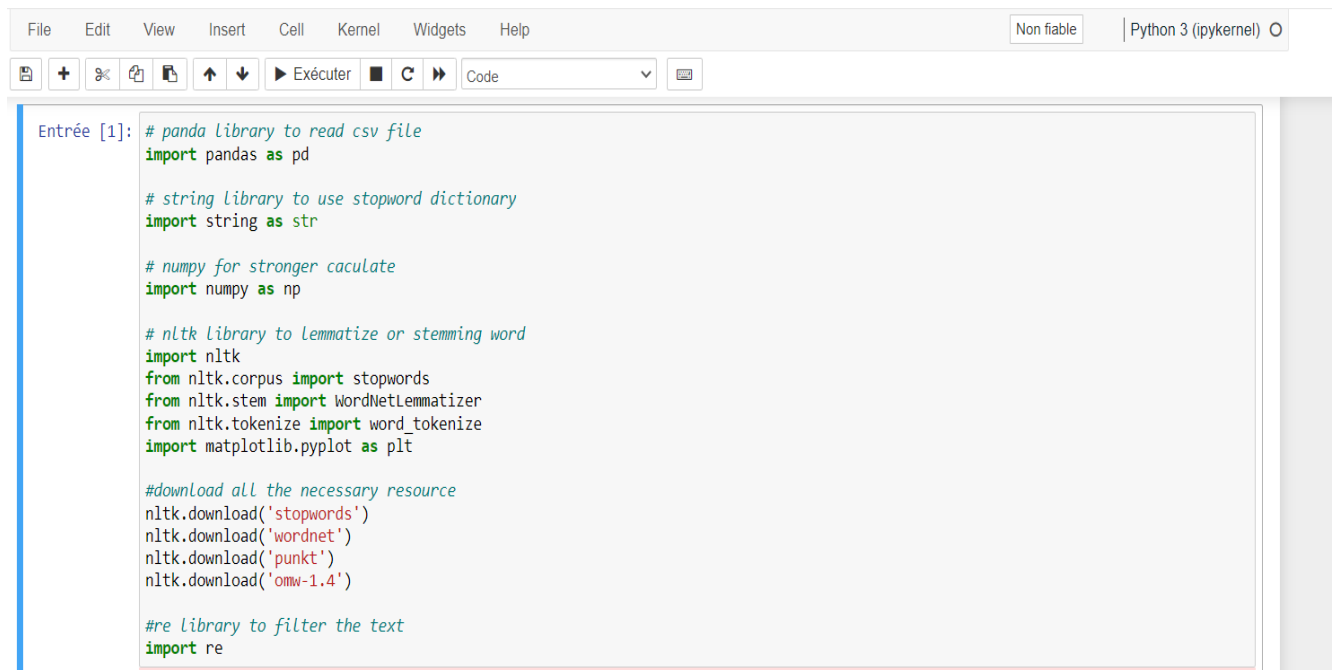
`pip install notebook` (if you haven't install notebook yet)

`jupyter notebook`

After run “jupyter notebook”, you will see the interface:



Then click BNB.ipynb or MNB.ipynb ,you will see:



The screenshot shows a Jupyter Notebook interface. At the top, there is a menu bar with 'File', 'Edit', 'View', 'Insert', 'Cell', 'Kernel', 'Widgets', and 'Help'. To the right of the menu bar, there is a status bar showing 'Non fiable' and 'Python 3 (ipykernel)'. Below the menu bar is a toolbar with icons for saving, adding, undo, redo, running, and other functions. The main area of the notebook contains a code cell labeled 'Entrée [1]:'. The code in the cell is as follows:

```
# panda library to read csv file
import pandas as pd


# string library to use stopwords dictionary
import string as str

# numpy for stronger caculate
import numpy as np

# nltk library to lemmatize or stemming word
import nltk
from nltk.corpus import stopwords
from nltk.stem import WordNetLemmatizer
from nltk.tokenize import word_tokenize
import matplotlib.pyplot as plt

#download all the necessary resource
nltk.download('stopwords')
nltk.download('wordnet')
nltk.download('punkt')
nltk.download('omw-1.4')

#re library to filter the text
import re
```

Then , just click each entrée and then click executer  to run each step.

Our project use some more python libraries ,so you have to install them first (if you haven't had yet) before launch:

pip install pandas


pip install numpy

pip install regex

pip install -U skicit-learn

pip install matplotlib

One another choice to launch jupyter notebook is on colab notebook of Google <https://colab.research.google.com/>

 Colaboratory chào mừng bạn!

TệpChỉnh sửaXemChènThời gian chạyCông cụTrợ giúp

Mục lục

Bắt đầu

Khoa học dữ liệu

Máy học


Tài nguyên khác

Ví dụ điển hình

Mục

Chào mừng bạn đến với Colab!

Nếu bạn đã quen dùng Colab, hãy xem video này để tìm hiểu về các bảng tương tác, chế độ xem lịch sử thực thi



Colab là gì?

Colab (hay còn gọi là "Colaboratory") cho phép bạn viết và thực thi Python trong trình duyệt với các lợi ích sau:

- Không yêu cầu cấu hình
- Sử dụng miễn phí GPU
- Chia sẻ dễ dàng

Click File/Tệp => Open notebook/Mở sổ tay =>Tải lên=>Chọn tệp

Ví DụGần ĐâyGoogle DriveGitHubTải Lên

Chọn tệp

Không có tệp nào được chọn

You will see:

```
+ Mã + Văn bản

# panda library to read csv file
import pandas as pd

# string library to use stopwords dictionary
import string as str

# numpy for stronger caculate
import numpy as np

# nltk library to lemmatize or stemming word
import nltk
from nltk.corpus import stopwords
from nltk.stem import WordNetLemmatizer
from nltk.tokenize import word_tokenize
import matplotlib.pyplot as plt

#download all the necessary resource
nltk.download('stopwords')
nltk.download('wordnet')
nltk.download('punkt')
nltk.download('omw-1.4')

#re library to filter the text
import re
```

Before launching ,please upload the dataset file(spam_ham_dataset.csv) by the following code:

```
from google.colab import files
uploaded=files.upload()
```

Chọn tệp Không có tệp nào được chọn Upload widget is only available when the cell has been executed in the current browser session. Please rerun this cell to enable Saving spam_ham_dataset.csv to spam_ham_dataset.csv

2. file spam_ham_app.py

We write a small app using streamlit and the code is save in spam_ham_app.py

To run this app, please install all library we mentioned in section 1 first.

Then install the streamlit by:

pip install streamlit

To run the app ,use

```
streamlit run spam_ham_app.py
```

After that, you will see the interface (the app running on port 8501 of local host) :

Email Filtering App

Enter the content of the email

Classify

Enter the content of the email and then click “Classify”. The predict result of our model will be shown below, for example:

Email Filtering App

Enter the content of the email

Please call me again

Classify

This is a ham email!