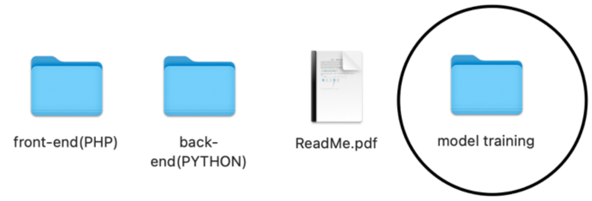
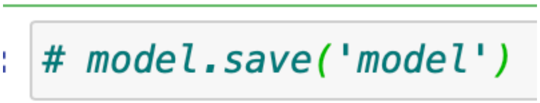
Part 1 Model training

In this project, 2 models, Neural network and SVC, are in model training folder.

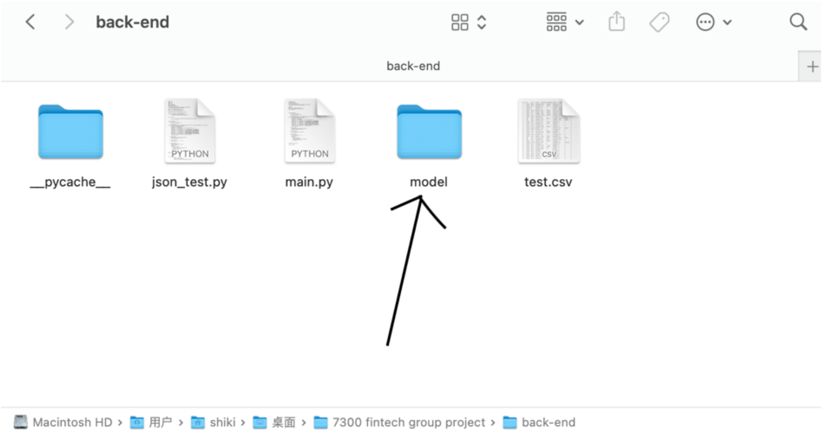


Two model use data.csv for training, the neural network’s performance is better so select it as the final AML detection model. The model is saved in /back-end/model folder, so the system doesn’t need to run the training’s process.

If you want to run the code, you should use anaconda jupyter notebook to start the code and import the required libraries. The last code in file “neural network.ipynb” is not necessary. It is used to save the model to “model” folder, which is already saved in back-end folder.



```
# model.save('model')
```

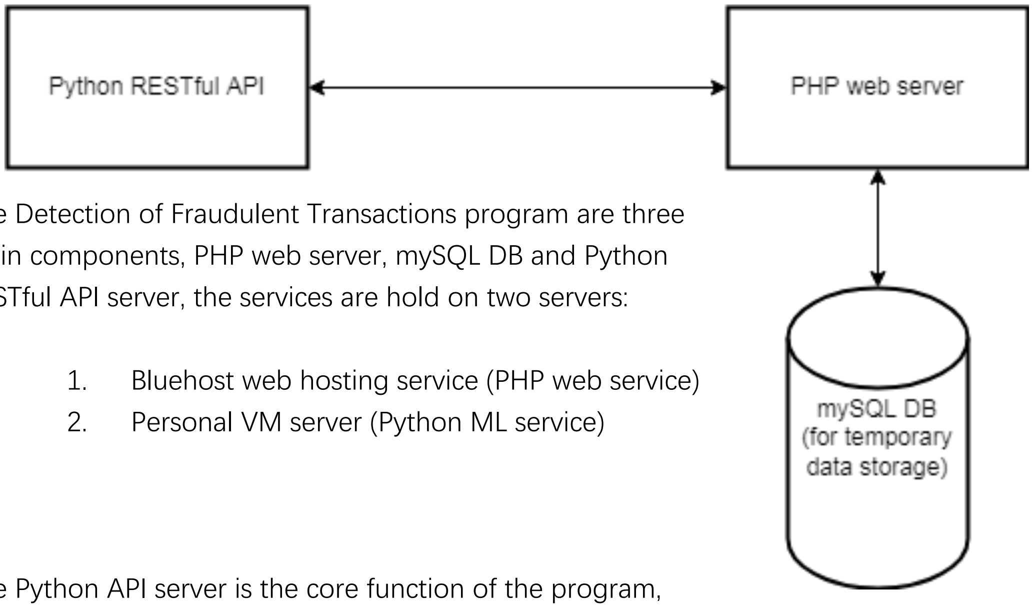

back-end
back-end
__pycache__
json_test.py
main.py
model
test.csv
PYTHON
PYTHON
CSV
Macintosh HD › 用户 › shiki › 桌面 › 7300 fintech group project › back-end

Part 2 Web UI

Program URL: https://wtj.iah.mybluehost.me/fintech/index.php

Username: fintech
Password: Fintech0403

Program structure:



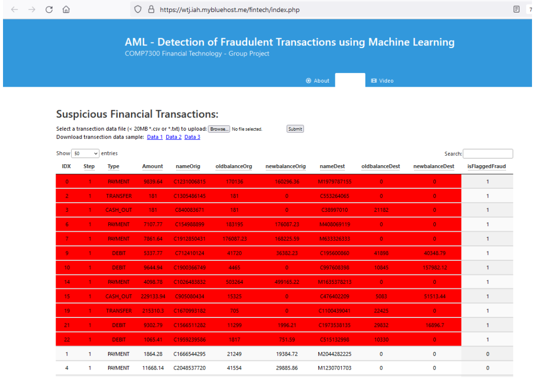
The Detection of Fraudulent Transactions program are three main components, PHP web server, mySQL DB and Python RESTful API server, the services are hold on two servers:

1. Bluehost web hosting service (PHP web service)
2. Personal VM server (Python ML service)

The Python API server is the core function of the program,
It has used machine learning to training large number of transactions data to find out which is the fake transaction.

Program flow:

Users upload a text file (transactions data) from the UI interface (PHP web server), the PHP web server will store the text file data to mySQL DB, in the same time it will sent the text file to Python API server and make a HTTP POST request, the Python API will process the text file to find which transactions are fake, then, it will return a result jSON back to PHP web server, the PHP server will flag false records on the MySQL DB with the result of jSON and finally display the result to the UI:



AML - Detection of Fraudulent Transactions using Machine Learning
COMP7300 Financial Technology - Group Project

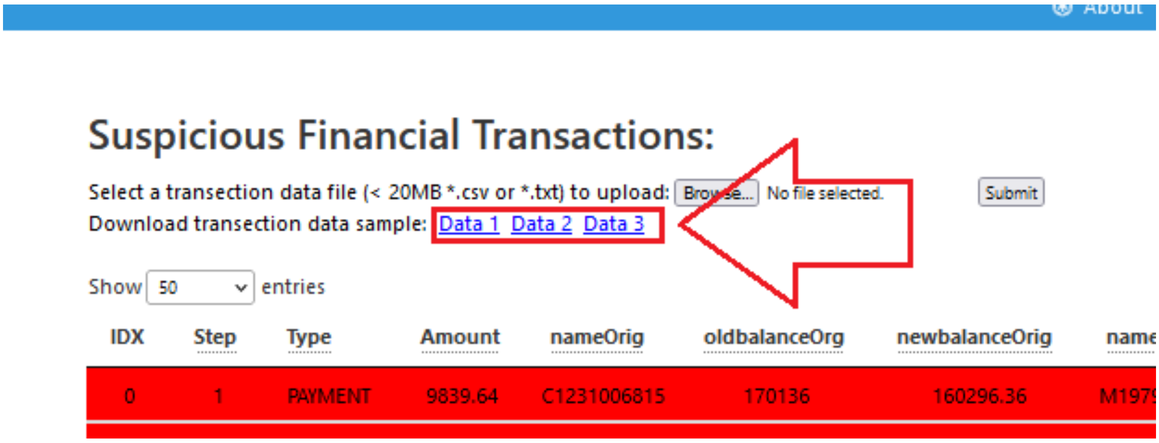
Suspicious Financial Transactions:

Select a transection data file (< 20MB *.csv or *.txt) to upload: Browse... No file selected. Submit
Download transection data sample: Data 1 Data 2 Data 3

Show 50 entries Search:

IDX	Step	Type	Amount	nameOrig	oldbalanceOrg	newbalanceOrig	nameDest	oldbalanceDest	newbalanceDest	isFlaggedFraud
0	1	PAYMENT	9839.64	C1231006815	170136	160296.36	M1979787155	0	0	1
2	1	TRANSFER	181	C1305486145	181	0	C553264065	0	0	1
3	1	CASH_OUT	181	C840083671	181	0	C38997010	21182	0	1
6	1	PAYMENT	7107.77	C154988899	183195	176087.23	M408069119	0	0	1
7	1	PAYMENT	7861.64	C1912850431	176087.23	168225.59	M633326333	0	0	1
9	1	DEBIT	5337.77	C712410124	41720	36382.23	C195600860	41898	40348.79	1
10	1	DEBIT	9644.94	C1900366749	4465	0	C997608398	10845	157982.12	1
14	1	PAYMENT	4098.78	C1026483832	503264	499165.22	M1635378213	0	0	1
15	1	CASH_OUT	229133.94	C905080434	15325	0	C476402209	5083	51513.44	1
19	1	TRANSFER	215310.3	C1670993182	705	0	C1100439041	22425	0	1
21	1	DEBIT	9302.79	C1566511282	11299	1996.21	C1973538135	29832	16896.7	1
22	1	DEBIT	1065.41	C1959239586	1817	751.59	C515132998	10330	0	1
1	1	PAYMENT	1864.28	C1666544295	21249	19384.72	M2044282225	0	0	0
4	1	PAYMENT	11668.14	C2048537720	41554	29885.86	M1230701703	0	0	0

The red color records are the transaction record that may fake, every time a text file is uploaded, the old records will be deleted, to facilitate the testing procedure, we provide three transactions data (Data 1, Data 2 and Data 3) for download:



Suspicious Financial Transactions:

Select a transection data file (< 20MB *.csv or *.txt) to upload:

Download transection data sample: Data 1 Data 2 Data 3

Show 50 entries

IDX	Step	Type	Amount	nameOrig	oldbalanceOrg	newbalanceOrig	name
0	1	PAYMENT	9839.64	C1231006815	170136	160296.36	M197

The columns of the text file:
idx,step,type,amount,nameOrig,oldbalanceOrg,newbalanceOrig,nameDest,oldbalanceDest,newbalanceDest,isFlaggedFraud

The transaction data must following the above format, the column description will display when we mouse over the every columns of the program. Other basic function are column sorting, keyword real time searching.