

Donation Lookup Bot

Chatbot hỗ trợ tra cứu thông tin ủng hộ đồng bào bị ảnh hưởng bởi bão lũ - qua tài khoản Vietcombank
0011001932418 từ 1/9-10/9/2024


- 1. Phân tích dữ liệu**
- 2. Các chức năng chính của bot**
- 3. Build Chatbot với Rasa Pro**

1. Phân tích dữ liệu

1. Phân tích dữ liệu

Dữ liệu gốc gồm khoảng 12k trang PDF về việc sao kê tài khoản (xem tại [đây](#))

Phần dữ liệu chúng ta cần quan tâm như hình bên gồm có **Ngày giao dịch/ Số chứng thực, Số tiền ghi có và Nội dung chi tiết.**



SAO KÊ TÀI KHOẢN
ACCOUNT STATEMENT

Tên tài khoản/Account name: MAT TRAN TO QUOC VN - BAN CUU TRO TW

Ngày thực hiện/ Date: 11/09/2024

Địa chỉ/ Address: 46 TRANG THI, HANOI

Chi nhánh thực hiện/ Branch: SỞ GIAO DỊCH

Số tài khoản/Account number: 0011001932418

CIF: 0002040837

Loại tiền/Currency: VND

Từ/ From: 01/09/2024 Đến/ To: 10/09/2024

Ngày GD/ TNX Date Số CT/ Doc No	Số tiền ghi nợ/ Debit	Số tiền ghi có/ Credit	Số dư/ Balance	Nội dung chi tiết/ Transactions in detail
01/09/2024 5213.45946		50.000		292976.010924.013647.xin cam on
01/09/2024 5090.85797		20.000		VCB.CTDK.31/03/2024.ADIDA PHAT. CT tu 0481000755821 toi 0011001932418 MAT TRAN TO QUOC VN - BAN CUU TRO TW
01/09/2024 5241.83107		29.000		MBVCB.6916176124.CAO VIET TUAN chuyen tien.CT tu 0101001222009 CAO VIET TUAN toi 0011001932418 MAT TRAN TO QUOC VN - BAN CUU TRO TW
01/09/2024 5218.87149		3.000		272986.010924.101858.DO DUC LOI chuyen

1. Phân tích dữ liệu

Dữ liệu sau khi được xử lí và chia sẻ tại nhóm [AIO Question / Answer and Competition](#)

	Date	Amount	Content	CT Code
0	01/09/2024	50000	NaN	5213.45946
1	01/09/2024	20000	VCB.CTDK.31/03/2024.ADIDA PHAT. CT tu 04810007...	5090.85797
2	01/09/2024	29000	MBVCB.6916176124.CAO VIET TUAN chuyen tien.CT ...	5241.83107
3	01/09/2024	3000	272986.010924.101858.DO DUC LOI chuyen tien	5218.87149
4	01/09/2024	3000	020097040509011046122024JDC5013867.96713 .1046...	5388.96713

Gồm các cột:

- Date:** Ngày giao dịch
- Amount:** Số tiền giao dịch
- Content:** Nội dung chuyển khoản
- CT Code:** Mã số chứng thực

```
RangeIndex: 200364 entries, 0 to 200363
Data columns (total 4 columns):
#   Column      Non-Null Count  Dtype
---  ---
0   Date        200364 non-null object
1   Amount      200364 non-null int64
2   Content     194545 non-null object
3   CT Code     200364 non-null object
dtypes: int64(1), object(3)
```

1. Phân tích dữ liệu

Date	Amount	Content
01/09/2024	50000	
01/09/2024	20000	VCB.CTDK.31/03/2024.ADIDA PH
01/09/2024	29000	MBVCB.6916176124.CAO VIET T
01/09/2024	3000	272986.010924.101858.DO DUC L
01/09/2024	3000	272986.010924.101858.DO DUC L

Ngày GD/ TNX Date Số CT/ Doc No	Số tiền ghi nợ/ Debit	Số tiền ghi có/ Credit	Số dư/ Balance	Nội dung chi tiết/ Transactions in detail
01/09/2024 5213.45946		50.000		292976.010924.013647.xin cam on
01/09/2024 5090.85797		20.000		VCB.CTDK.31/03/2024.ADIDA PHAT. CT tu 0481000755821 toi 0011001932418 MAT TRAN TO QUOC VN - BAN CUU TRO TW
01/09/2024 5241.83107		29.000		MBVCB.6916176124.CAO VIET TUAN chuyen tien.CT tu 0101001222009 CAO VIET TUAN toi 0011001932418 MAT TRAN TO QUOC VN - BAN CUU TRO TW
01/09/2024 5218.87149		3.000		272986.010924.101858.DO DUC LOI chuyen tien
01/09/2024 5388.96713		3.000		020097040509011046122024JDC5013867.96713 .104607.Vietcombank:0011001932418:DANG THI THANH MHS533583
01/09/2024 5212.91607		3.000		529474.010924.112032.BUI VAN DAI chuyen tien
01/09/2024 5220.28596		3.000		577301.010924.113215.DO DUC LOI MHS 4242
01/09/2024 5388.33462		3.000		020097040509011137472024Y7ZE058213.33462 .113747.Vietcombank:0011001932418:NGUYE N THI VAN chuyen tien
01/09/2024 5216.68569		3.000		761664.010924.121847.NGUYEN VAN TUAN chuyen tien
01/09/2024 5387.37142		3.000		0200970405090112202020246WV1059470.3714 2.122015.Vietcombank:0011001932418:NGUYE N QUANG THO 010166
01/09/2024 5161.94345		10.000		PARTNER.DIRECT_DEBITS_VCB.ZLP.ZP6R 34H953F5.20240901.Nguyet chuyen tien qua Zalopay
01/09/2024 5219.94264		6.000		266303.010924.122804.NGUYEN THI MAO Chuyen tien

Content ở cột đầu bị thiếu

1. Phân tích dữ liệu

Ngày GD/ TNX Date Số CT/ Doc No	Số tiền ghi nợ/ Debit	Số tiền ghi có/ Credit	Số dư/ Balance	Nội dung chi tiết/ Transactions in detail
01/09/2024 5213.45946		50.000		292976.010924.013647.xin cam on

03/09/2024	10000	MBVCB.6934364247.gia dinh Dung Thuy Giang chuc to quoc khoe manh, binh an, hanh phuc, cam on CT tu 1018144916 HOANG THANH THUY toi 0011001932418 MAT TRAN TO QUOC VN - BAN CUU TRO TW
03/09/2024	100000	0200970415090316001320248Tkn135640.57821. 160013.NGUYEN NHU DUNG chuyen gop
03/09/2024	12000	633739.030924.184314.NGUYEN VAN LINH Chuyen tien
04/09/2024	20000	718788.040924.001536.LE THI MY HANH chuyen tien
04/09/2024	80000	MBVCB.6938069039.kho khan CT tu 1013662745 toi 0011001932418 Uy Ban Trung uong Mat tran To quoc Viet Nam Postal address: Telex : (0805) 411504 VCB - VT 198 TRAN QUANG KHAI AVENUE Swift : BFTV
04/09/2024	5000000	028520.040924.081441.NGUYEN VAN HUNG ET21212050967727

Ở cột cuối, thông tin trích ra bị thừa

01/09/2024	5219.1607	3.000	529474.010924.112032.BUI VAN DAI chuyen tien
01/09/2024	5219.28596	3.000	577301.010924.113215.DO DUC LOI MHS 4242
01/09/2024	5318.33462	3.000	020097040509011137472024Y7ZE058213.33462 .113747.Vietcombank:0011001932418:NGUYE N THI VAN chuyen tien
01/09/2024	5116.68569	3.000	761664.010924.121847.NGUYEN VAN TUAN chuyen tien
01/09/2024	5187.37142	3.000	0200970405090112202020246WV1059470.3714 2.122015.Vietcombank:0011001932418:NGUYE N QUANG THO 010166
01/09/2024	161.94345	10.000	PARTNER.DIRECT_DEBITS_VCB.ZLP.ZP6R 34H953F5.20240901.Nguyet chuyen tien qua Zalopay
01/09/2024	5219.94264	6.000	266303.010924.122804.NGUYEN THI MAO chuyen tien

Postal address:
198 TRAN QUANG KHAI AVENUE
HANOI - S.R. VIETNAM

Telex : (0805) 411 504 VCB - VT
Swift : BFTV VNVX
Website: www.vietcombank.com.vn
Contact center: 1900.54.54.13

1. Phân tích dữ liệu

Cách tự chuyển file pdf thành csv tham khảo từ J2Team ở [đây](#) thì gặp lỗi sau

tiktok.com - To exit full screen, press Esc

14/09/2024 6:16 AM Python Source File 5 KB

CSVFileView - D:\CODE\saoke\output.csv

File Edit View Options Help

date	transaction_code	amount	transaction_detail
'01/09/2024'	5213.45946	50000	47.xin cam on
'01/09/2024'	5090.85797	20000	4.ADIDA PHAT. CT tu
'01/09/2024'	5241.83107	29000	011001932418 MAT
'01/09/2024'	5218.87149	3000	BAN CUU TRO TW
'01/09/2024'	5388.96713	3000	CAO VIET TUAN
'01/09/2024'	5212.91607	3000	1001222009 CAO VIET
'01/09/2024'	5220.28596	3000	2418 MAT TRAN TO
'01/09/2024'	5388.33462	3000	TRO TW
'01/09/2024'	5216.68569	3000	58.DO DUC LOI chuyen
'01/09/2024'	5387.37142	3000	
'01/09/2024'	5161.94345	10000	6122024JDC5013867.96713
'01/09/2024'	5219.94264	6000	0011001932418:DANG
'01/09/2024'	5215.97152	3000	04.NGUYEN THI MAO
'01/09/2024'	5219.24714	10000	
'02/09/2024'	5212.22965	10000	39.chuc mung ngay 29
'02/09/2024'	5245.21394	10000	chu nhia Viet Nam

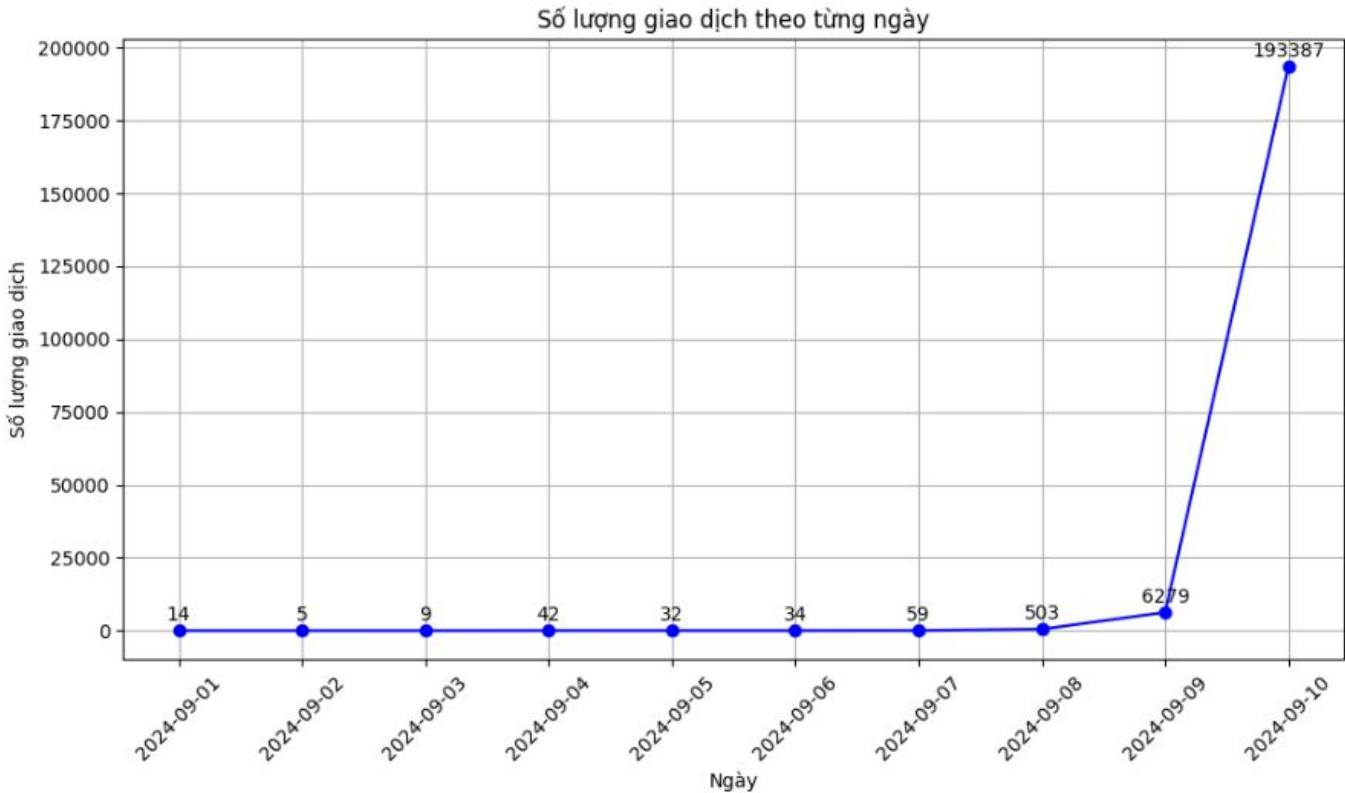
Ngày GD/ TNX Date Số CT/Doc No	Số tiền ghi nợ/ Debit	Số tiền ghi có/ Credit	Số dư/ Balance	Nội dung chi tiết/ Transactions in detail
01/09/2024 5213.45946		50.000		292976.010924.013647.xin cam on
01/09/2024 5090.85797		20.000		VCB.CTDK.31/03/2024.ADIDA PHAT. CT tu 0481000755821 toi 001 1001932418 MAT
				TRAN TO QUOC VN - BAN CUU TRO TW
01/09/2024 5241.83107		29.000		MBVCB.6916176124.CAO VIET TUAN chuyen tien.CT tu 0101001222009 CAO VIET TUAN toi 0011001932418 MAT TRAN TO QUOC VN - BAN CUU TRO TW
01/09/2024 5218.87149		3.000		272986.010924.101858.DO DUC LOI chuyen tien
01/09/2024 5388.96713		3.000		020097040509011046122024JDC5013867.96713 .104607.Vietcombank:0011001932418:DANG THI THANH MHS533583
01/09/2024 5212.91607		3.000		529474.010924.112032.BUI VAN DAI chuyen tien
01/09/2024 5220.28596		3.000		577301.010924.113215.DO DUC LOI MHS 4242
01/09/2024 5388.33462		3.000		020097040509011137472024Y7ZE058213.33462 .113747.Vietcombank:0011001932418:NGUYE N THI VAN chuyen tien
01/09/2024 5216.68569		3.000		761664.010924.121847.NGUYEN VAN TUAN chuyen tien
01/09/2024 5387.37142		3.000		02009704050901122020246WV1059470.3714 2.122015.Vietcombank:0011001932418:NGUYE N QUANG THO 010166
01/09/2024 5161.94345		10.000		PARTNER.DIRECT_DEBITS_VCB.ZLP.ZP6R 34H953F5.20240901.Nguyet chuyen tien qua Zalopay
01/09/2024 5219.94264		6.000		266303.010924.122804.NGUYEN THI MAO Chuyen tien

1. Phân tích dữ liệu

Data từ thành viên trong nhóm AIO mặc dù bị thiếu và thừa ở một số dòng nhưng không bị lệch nhiều như bên nhóm J2TEAM nên sẽ sử dụng data từ nhóm AIO để xử lý.

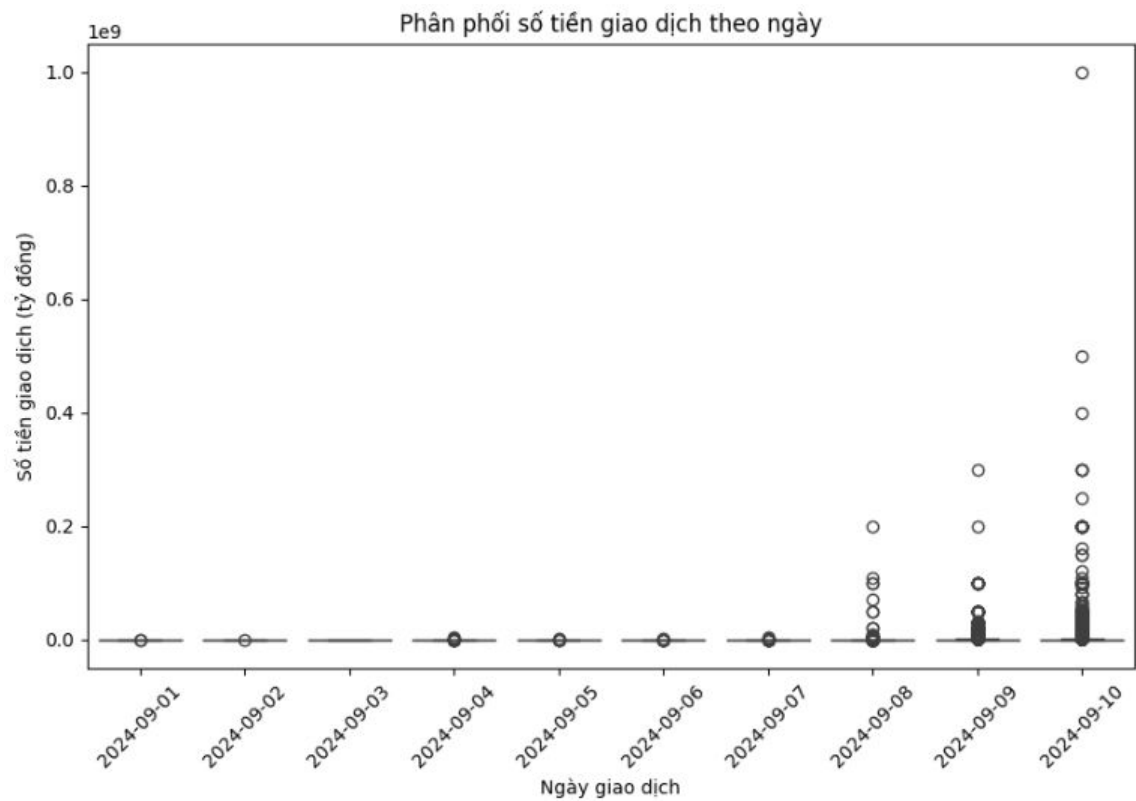
1. Phân tích dữ liệu

Số lượng giao dịch trong từng ngày từ 01/09/2024 đến 11/09/2024



1. Phân tích dữ liệu

Số lượng giao dịch trong từng ngày từ 01/09/2024 đến 11/09/2024



1. Phân tích dữ liệu

Phân tích nội dung chuyển khoản

- **Từ khóa nổi bật nhất:** "ung", "ho", "dong", "bao": Đây là từ lớn nhất, chiếm vị trí trung tâm, liên quan đến việc ủng hộ đồng bào bị ảnh hưởng bởi bão lũ
- **Nhóm từ liên quan đến ngân hàng và tài chính:** "CT" (có thể là chuyển tiền, chứng thực), "chuyen tien", "Swift", "Vietcombank", "BFTV", "VCB", "Telex", "TRO" (có thể là trợ lý) "Quoc te" (quốc tế), "toi" (tối), "tu" (từ), "dong bao" (đồng bào)
- **Tên địa danh và khu vực:** "HANOI", "mien Bac" (miền Bắc), "AVENUE", "KHAI AVENUE"
- **Tên người hoặc phần của tên:** "NGUYEN", "THI", "TW" (có thể là Trung ương)
- **Các từ khác:** "Yagi": Tên của cơn bão gây thiệt hại "VN": Viết tắt của Việt Nam



Word clouds được vẽ từ file dữ liệu .csv

1. Phân tích dữ liệu

Các n-gram theo thứ tự giảm dần

n_gram	count
(ung, ho)	54225
(dong, bao)	52474
(CUU, TRO)	42819
(TO, QUOC)	42357
(TRAN, TO)	42336
(MAT, TRAN)	42326
(BAN, CUU)	42308
(TRO, TW)	42303
(QUOC, VN)	42286
(VN, -)	41948
(-, BAN)	41942
(toi, 0011001932418)	41107
(chuyen, tien)	41084
(0011001932418, MAT)	40745
(ho, dong)	38897
(khac, phuc)	23522
(bao, lu)	22168
(lu, lut)	15854
(tien, ung)	13999
(bao, so)	13773

n_gram	count
(TRAN, TO, QUOC)	42327
(MAT, TRAN, TO)	42325
(BAN, CUU, TRO)	42307
(CUU, TRO, TW)	42303
(TO, QUOC, VN)	42285
(-, BAN, CUU)	41942
(QUOC, VN, -)	41938
(VN, -, BAN)	41937
(toi, 0011001932418, MAT)	40745
(0011001932418, MAT, TRAN)	40745
(ho, dong, bao)	38020
(ung, ho, dong)	21676
(tien, ung, ho)	13731
(ho, khac, phuc)	13479
(bao, so, 3)	12042
(TRAN, QUANG, KHAI)	12028
(HANOI, -, S.R)	12027
(VNVX, HANOI, -)	12027
(BFTV, VNVX, HANOI)	12027
(-, BFTV, VNVX)	12027

n_gram	count
(MAT, TRAN, TO, QUOC)	42321
(BAN, CUU, TRO, TW)	42294
(TRAN, TO, QUOC, VN)	42285
(-, BAN, CUU, TRO)	41942
(TO, QUOC, VN, -)	41938
(VN, -, BAN, CUU)	41937
(QUOC, VN, -, BAN)	41937
(toi, 0011001932418, MAT, TRAN)	40745
(0011001932418, MAT, TRAN, TO)	40745
(ung, ho, dong, bao)	21159
(AVENUE, Swift, :, BFTV)	12027
(KHAI, AVENUE, Swift, :)	12027
(QUANG, KHAI, AVENUE, Swift)	12027
(www.vietcombank.com.vn, Contact, center, :)	12027
(TRAN, QUANG, KHAI, AVENUE)	12027
(198, TRAN, QUANG, KHAI)	12027
(Swift, :, BFTV, VNVX)	12027
(BFTV, VNVX, HANOI, -)	12027
(VNVX, HANOI, -, S.R)	12027
(HANOI, -, S.R, .)	12027

n_gram	count
(MAT, TRAN, TO, QUOC, VN)	42282
(-, BAN, CUU, TRO, TW)	41939
(TRAN, TO, QUOC, VN, -)	41938
(QUOC, VN, -, BAN, CUU)	41937
(TO, QUOC, VN, -, BAN)	41937
(VN, -, BAN, CUU, TRO)	41937
(toi, 0011001932418, MAT, TRAN, TO)	40745
(0011001932418, MAT, TRAN, TO, QUOC)	40745
(VNVX, HANOI, -, S.R, .)	12027
(www.vietcombank.com.vn, Contact, center, :, 1...	12027
(:, www.vietcombank.com.vn, Contact, center, :)	12027
(VIETNAM, Website, :, www.vietcombank.com.vn, ...	12027
(-, VIETNAM, Website, :, www.vietcombank.com.vn)	12027
(S.R, ., VIETNAM, Website, :)	12027
(-, S.R, ., VIETNAM, Website)	12027
(HANOI, -, S.R, ., VIETNAM)	12027
(Website, :, www.vietcombank.com.vn, Contact, ...	12027
(BFTV, VNVX, HANOI, -, S.R)	12027
(QUANG, KHAI, AVENUE, Swift, :)	12027
(:, BFTV, VNVX, HANOI, -)	12027

1. Phân tích dữ liệu

Bi-gram với các giá trị count=1, >=100, >=1000

n_gram	count	n_gram	count		
(483633.100924.141748.IBFT, HOANG)	1	(vi, dong)	103	(3, yagi)	1447
(Yagi, 483633.100924.141748.IBFT)	1	(VU, NGOC)	103	(con, mien)	1446
(tien, 574167.100924.141743.PHAN)	1	(NGOC, DUNG)	103	(MIEN, BAC)	1433
(263175.100924.141749.DONG, GOP)	1	(NGUYEN, MAI)	103	(THI, MINH)	1432
(3, YAGI-100924-14:17:47)	1	(lu, sau)	103	(TRANG, toi)	1426
(YAGI-100924-14:17:47, 263175)	1	(an, nan)	103	(se, cung)	1420
(263175, 428175.100924.141748.Ung)	1	(NGOC, NHU)	103	(gay, ra)	1403
(428175.100924.141748.Ung, ho)	1	(chong, thien)	103	(quyen, gop)	1373
(lu, 428187.100924.141748.NGO)	1	(va, cac)	103	(Bac.CT, tu)	1370
(428187.100924.141748.NGO, MINH)	1	(THANH, TU)	103	(QUA, BAO)	1354
(574167.100924.141743.PHAN, THI)	1	(KIM, PHUONG)	102	(PHAM, THI)	1324
(Thanh-WindWin, Fanpage)	1	(tinh, anh)	102	(PHUONG, chuyen)	1315
(Yagi, 265326.100924.141748.Ung)	1	(., MONG)	102	(thanh, pho)	1292
(265326.100924.141748.Ung, ho)	1	(THU, NGAN)	102	(THIET, HAI)	1282
(lu, 020097041509101417482024Vkeo213214.15914)	1	(NGUYEN, THAO)	102	(THI, HUYEN)	1251
(020097041509101417482024Vkeo213214.15914, .14...	1	(BAC, VUOT)	102	(NGOC, ANH)	1244
(.141749.NCT-WAYV, WINWIN)	1	(deu, duoc)	102	(huong, boi)	1238
(WINWIN, Dong)	1	(VAN, DUC)	102	(se, voi)	1235
(Tu, Thanh-WindWin)	1	(My, ung)	102	(tinh, thanh)	1234
(807439.100924.223851.cua, it)	1	(KHANH, ung)	102	(THAO, chuyen)	1233
				(voi, dong)	1233

- Bi-grams với count = 1 phản ánh những giao dịch cụ thể, không lặp lại và chứa mã số giao dịch chi tiết.
- Bi-grams với count ≥ 100 liên quan đến tên cá nhân, tổ chức quen thuộc và các cụm từ phổ biến trong cứu trợ.
- Bi-grams với count ≥ 1000 tập trung vào việc **quyên góp**, **chuyển khoản**, và phản ánh thiệt hại do thiên tai, cho thấy các giao dịch cứu trợ có tính hệ thống và thường xuyên.

1. Phân tích dữ liệu

Tri-gram với các giá trị count=1, >=100, >=1000

n_gram	count	n_gram	count	n_gram	count
(-, Nguoi, ham)	1	(THI, NGOC, LAN)	103	(dong, bao, thiet)	1535
(Yagi.CT, tu, 9815999909)	1	(bao, lu, Mien)	103	(bao, so, 3.CT)	1527
(tu, 9815999909, NGUYEN)	1	(bao, ,, lu)	103	(hai, do, con)	1516
(9815999909, NGUYEN, BICH)	1	(binh, an, nan)	103	(LINH, chuyen, tien)	1506
(NGUYEN, BICH, DUNG)	1	(Bac, bi, lu)	103	(UNG, HO, KHAC)	1497
(TRO, TW, 020097041509101419052024ot8I215801.1...	1	(long, nho, den)	103	(phuc, sau, bao)	1493
(.CT, tu, 1039655070)	1	(HOA, chuyen, tien.CT)	103	(PHUC, HAU, QUA)	1488
(TW, 020097041509101419052024ot8I215801.19468....	1	(HO, BAO, LU)	103	(KHAC, PHUC, HAU)	1484
(020097041509101419052024ot8I215801.19468.1, 4...	1	(Hoa, ung, ho)	103	(dong, bao, vuot)	1477
(41905.Thai, Hoang, ung)	1	(THI, LAN, HUONG)	103	(ung, ho, mien)	1468
(lu, lut, MBVCB.6992768292.ung)	1	(chuyen, tien, chung)	102	(cuu, tro, dong)	1448
(lut, MBVCB.6992768292.ung, ho)	1	(THU, HANG, toi)	102	(so, 3, yagi)	1436
(MBVCB.6992768292.ung, ho, thiet)	1	(MINH, ANH, toi)	102	(TRANG, toi, 0011001932418)	1426
(3.CT, tu, 0381000441708)	1	(Cau, mong, moi)	102	(som, vuot, qua)	1420
(tu, 0381000441708, NGUYEN)	1	(THI, NHUNG, toi)	102	(phuc, bao, yagi)	1376
(0381000441708, NGUYEN, THI)	1	(long, nho, gui)	102	(ba, con, mien)	1361
(TRO, TW, MBVCB.6992768391.PHAN)	1	(DIEP, chuyen, tien)	102	(do, bao, yagi)	1335
(TW, MBVCB.6992768391.PHAN, THI)	1	(HO, LU, LUT)	102	(qua, sau, bao)	1314
(MBVCB.6992768391.PHAN, THI, MY)	1	(tinh, anh, huong)	102	(qua, bao, lu)	1310
(807439.100924.223851.cua, it, long)	1	(CHUYEN, TIEN, UNG)	102	(chuyen, tien, cuu)	1308

- Tri-grams với count = 1 phản ánh các giao dịch riêng lẻ, thường bao gồm mã giao dịch cụ thể và tên cá nhân.
- Tri-grams với count ≥ 100 thể hiện sự tham gia mạnh mẽ của các cá nhân, tổ chức trong các hoạt động từ thiện liên quan đến lũ lụt.
- Tri-grams với count ≥ 1000 cho thấy sự phổ biến của các cụm từ liên quan đến quyên góp, cứu trợ và thiệt hại do thiên tai, với nhiều giao dịch liên quan đến các vùng bị ảnh hưởng bởi bão lụt.

2. Các chức năng chính của bot

2. Các chức năng chính của bot

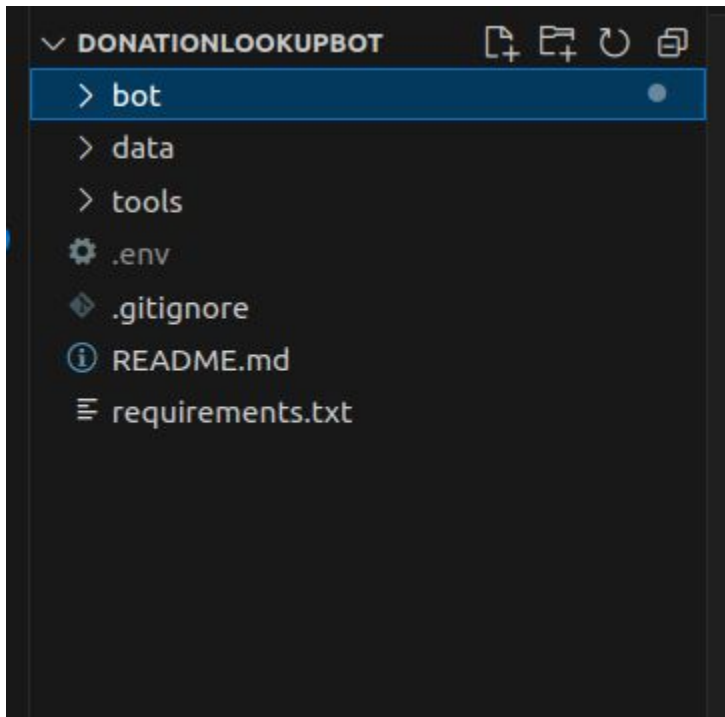
- Truy vấn thông tin giao dịch
- Tính tổng số tiền ủng hộ trong khoảng thời gian nhất định.
- Vẽ biểu đồ thể hiện số lượng giao dịch trong khoảng thời gian nhất định.



Lưuồng hoạt động cơ bản

3. Build Chatbot với Rasa Pro

3. Build Chatbot với Rasa Pro



Tạo trước folder DonationLookupBot với cấu trúc như sau:

- Folder "bot" sẽ chứa các file liên quan đến chatbot như flows, config, ...
- Folder "data" chứa các file liên quan đến data gốc cũng như data đã được xử lý
- Folder "tools" chứa các file dùng để xử lý, chuyển đổi data.

3. Build Chatbot với Rasa Pro

Dựa trên hướng dẫn tại [đây](#), tạo môi trường ảo với python=3.9.19 và cài đặt Rasa Pro version 3.9.2

```
(dlubot) namdv@namdv-Latitude-5490:~/VSCode/DonationLookupBot$ rasa --version
Rasa Version      :      3.9.2
Minimum Compatible Version: 3.7.0
Rasa SDK Version  :      3.9.0
Python Version    :      3.9.19
Operating System  :      Linux-6.8.0-40-generic-x86_64-with-glibc2.35
Python Path       :      /home/namdv/anaconda3/envs/dlubot/bin/python3
```

3. Build Chatbot với Rasa Pro

▼ DONATIONLOOKUPBOT

> bot

▼ data

```
(dlubot) namdv@namdv-Latitude-5490:~/VSCode/DonationLookupBot$ rasa init --template tutorial
```

Welcome to Rasa! 🤖

By installing and using this software, you agree to be bound by the terms and conditions of the Developer Terms available at <https://rasa.com/developer-terms>. Please review the Developer Terms carefully before proceeding.

To get started quickly, an initial project will be created.

If you need some help, check out the documentation at <https://rasa.com/docs/rasa-pro>.

Now let's start! 👉

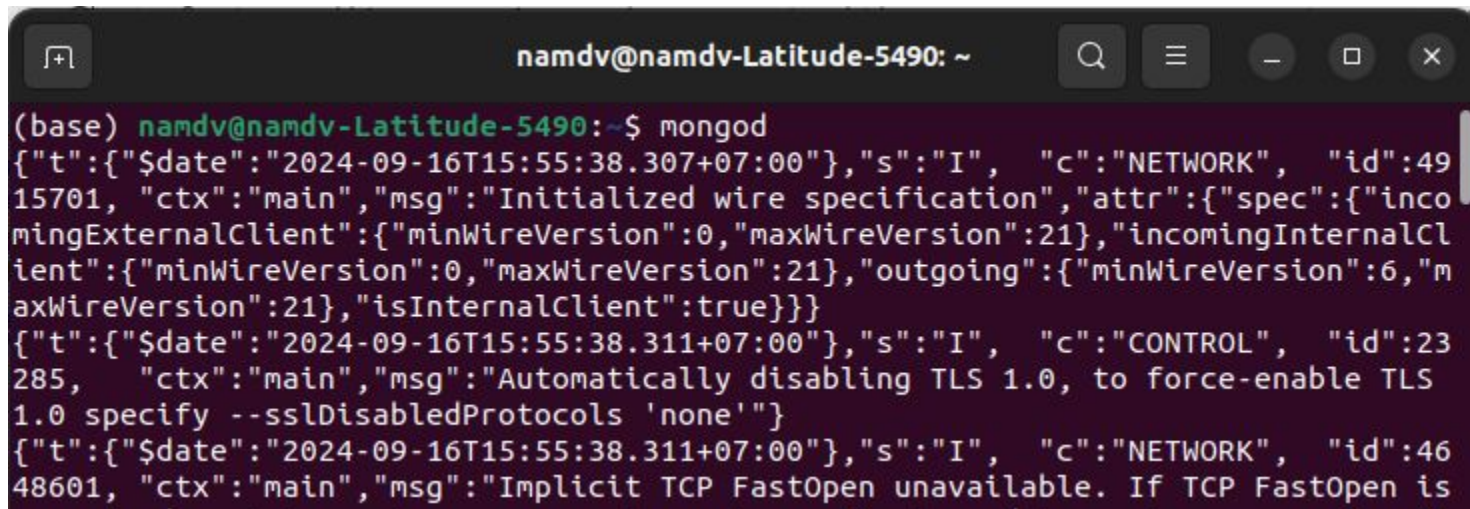
? Please enter a path where the project will be created [default: current directory] bot

Tạo template vào folder "bot"

3. Build Chatbot với Rasa Pro

Tải mongodb the hướng dẫn [sau](#).

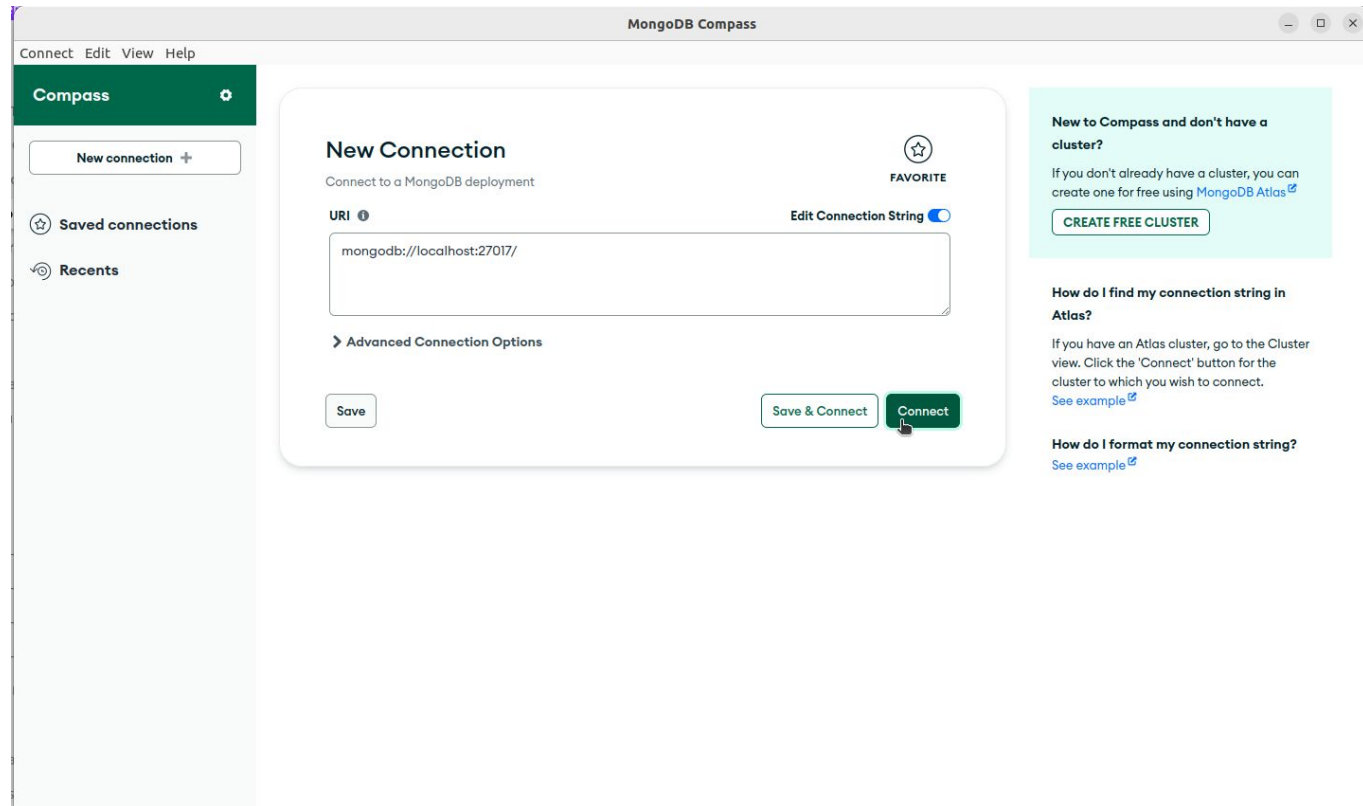
Sau khi cài đặt hoàn thành mở terminal và chạy lệnh như hình

A terminal window with a dark background and light-colored text. The window title is 'namdv@namdv-Latitude-5490: ~'. The prompt is '(base) nambv@namdv-Latitude-5490:~\$'. The command 'mongod' has been executed, and the output shows three log messages. The first message is about initializing the wire specification. The second message is about automatically disabling TLS 1.0. The third message is about implicit TCP FastOpen being unavailable. The terminal window has standard macOS window controls (red, yellow, green buttons) and a search icon in the top right corner.

```
(base) nambv@namdv-Latitude-5490:~$ mongod
{"t":{"$date":"2024-09-16T15:55:38.307+07:00"},"s":"I", "c":"NETWORK", "id":4915701, "ctx":"main","msg":"Initialized wire specification","attr":{"spec":{"incomingExternalClient":{"minWireVersion":0,"maxWireVersion":21},"incomingInternalClient":{"minWireVersion":0,"maxWireVersion":21},"outgoing":{"minWireVersion":6,"maxWireVersion":21},"isInternalClient":true}}}}
{"t":{"$date":"2024-09-16T15:55:38.311+07:00"},"s":"I", "c":"CONTROL", "id":23285, "ctx":"main","msg":"Automatically disabling TLS 1.0, to force-enable TLS 1.0 specify --sslDisabledProtocols 'none'"}
{"t":{"$date":"2024-09-16T15:55:38.311+07:00"},"s":"I", "c":"NETWORK", "id":4648601, "ctx":"main","msg":"Implicit TCP FastOpen unavailable. If TCP FastOpen is
```

3. Build Chatbot với Rasa Pro

Mở MongoDB Compass và nhấn Connect



3. Build Chatbot với Rasa Pro

Tạo database mới



Create Database

Database Name

dlubot_db

Collection Name

transactions

☐ Time-Series

Time-series collections efficiently store sequences of measurements over a period of time. [Learn More](#)

> Additional preferences

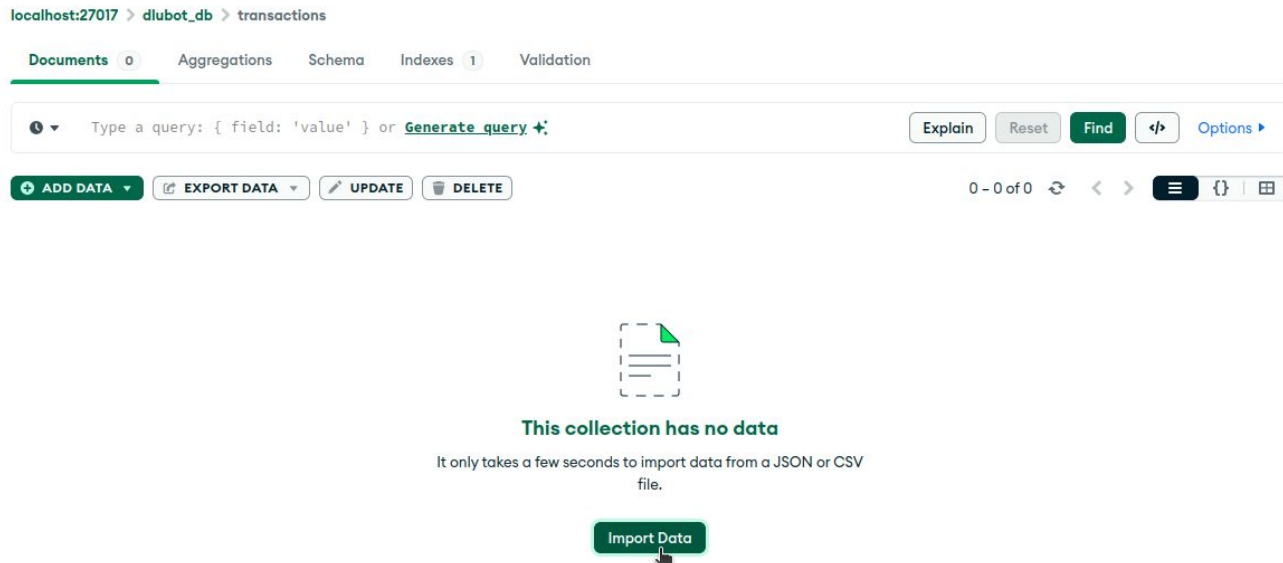
(e.g. Custom collation, Capped, Clustered collections)

Cancel

Create Database

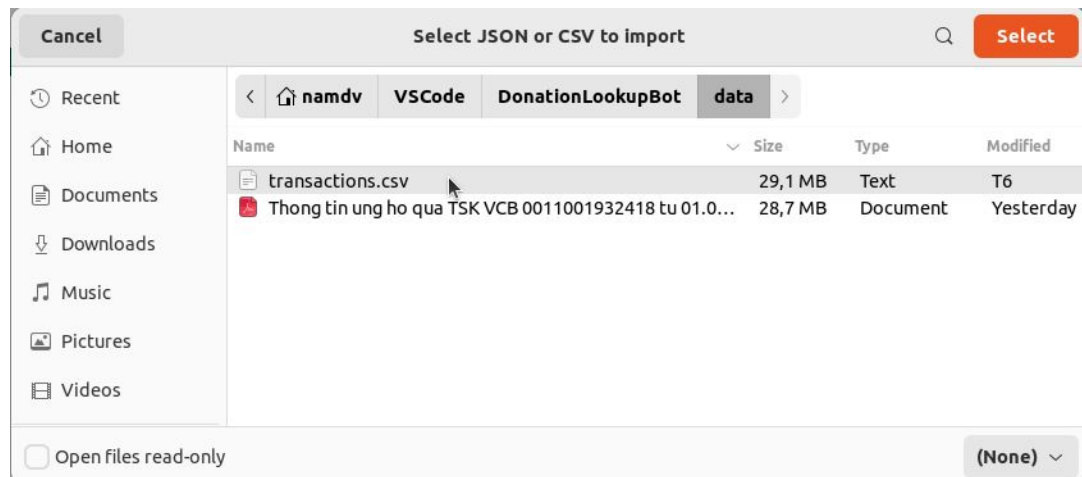
3. Build Chatbot với Rasa Pro

Chọn Import Data



3. Build Chatbot với Rasa Pro

Chọn file transactions.csv



3. Build Chatbot với Rasa Pro

Chọn Import

Import

To collection dlubot_db.transactions

Import file: transactions.csv

Options

Select delimiter Comma

☒ Ignore empty strings

☐ Stop on errors

Specify Fields and Types

Learn more about data types

	<input checked="" type="checkbox"/> Date	<input checked="" type="checkbox"/> Amount	<input checked="" type="checkbox"/> Content	<input checked="" type="checkbox"/> CT Code
	String	Int32	String	Mixed
1	01/09/2024	50000	empty string	5213.45946
2	01/09/2024	20000	VCB.CTDK.31/03/2024.ADIDA PHAT. CT t...	5090.85797
3	01/09/2024	29000	MBVCB.6916176124.CAO VIET TUAN chuye...	5241.83107
4	01/09/2024	3000	272986.010924.101858.DO DUC LOI chuye...	5218.87149
5	01/09/2024	3000	020097040509011046122024.JDC5013867...	5388.96713

Cancel

Import

3. Build Chatbot với Rasa Pro

Một số documents trong database

localhost:27017 > dlubot_db > transactions

Documents 200.4K Aggregations Schema Indexes 1 Validation

Type a query: { field: 'value' } or [Generate query](#)

Explain Reset Find Options

ADD DATA EXPORT DATA UPDATE DELETE

1 - 20 of 200364

_id: ObjectId('66e7f615e262e494424a1ff6')

Date : "01/09/2024"

Amount : 50000

CT Code : 5213.45946

_id: ObjectId('66e7f615e262e494424a1ff7')

Date : "01/09/2024"

Amount : 20000

Content : "VCB.CTDK.31/03/2024.ADIDA PHAT. CT tu 0481000755821 toi 0011001932418 ..."

CT Code : 5090.85797

_id: ObjectId('66e7f615e262e494424a1ff8')

Date : "01/09/2024"

Amount : 29000

Content : "MBVCB.6916176124.CAO VIET TUAN chuyen tien.CT tu 0101001222009 CAO VIE..."

CT Code : 5241.83107

_id: ObjectId('66e7f615e262e494424a1ff9')

Date : "01/09/2024"

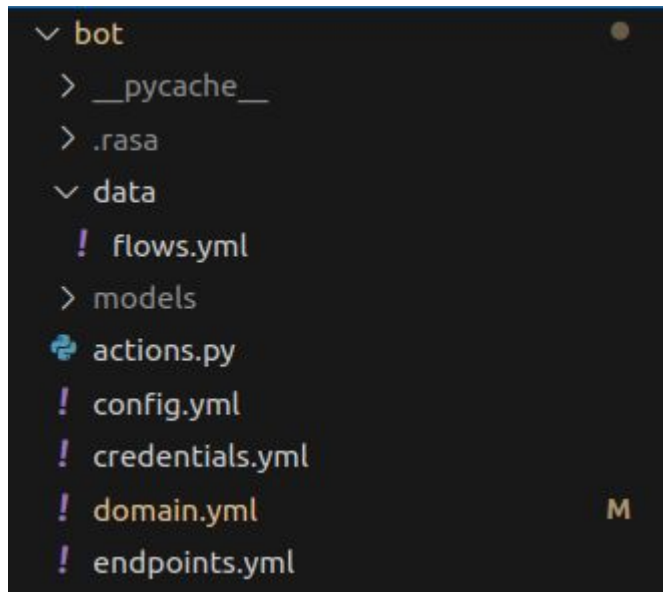
Amount : 3000

Content : "272986.010924.101858.DO DUC LOI chuyen tien"

CT Code : 5218.87149

3. Build Chatbot với Rasa Pro

Trong folder bot gồm các file và folder sau



Hiện tại chỉ cần quan tâm các file

flows.yml: chứa các luồng hội thoại bot sẽ thực hiện

actions.py: các hành động mà bot sẽ thực hiện trong từng flows

config.yml: Tập cấu hình xác định các thành phần và chính sách mà mô hình sẽ sử dụng để đưa ra dự đoán dựa trên đầu vào của người dùng.

domain.yml: xác định toàn bộ bối cảnh mà chatbot sẽ hoạt động.

Chi tiết các file xem tại source code sau <https://github.com/dauvannam1804/DonationLookupBot>

3. Build Chatbot với Rasa Pro

Sau khi đã cài đặt các flows, action tương ứng, di chuyển đến thư mục bot và train chatbot

```
(dlubot) namdv@namdv-Latitude-5490:~/VSCode/DonationLookupBot$ cd bot  
(dlubot) namdv@namdv-Latitude-5490:~/VSCode/DonationLookupBot/bot$ rasa train
```

3. Build Chatbot với Rasa Pro

Nếu chưa khởi kết nối với database, mở một terminal mới và chạy lệnh sau

```
(base) namdv@namdv-Latitude-5490:~/VSCode/DonationLookupBot$ mongod
{"t":{"$date":"2024-09-17T08:03:00.085+07:00"},"s":"I", "c":"CONTROL", "id":23285, "ctx":"main","msg":"Automatically disabling TLS 1.0, to force-enable TLS 1.0 specify --sslDisabledProtocols 'none'"}
{"t":{"$date":"2024-09-17T08:03:00.085+07:00"},"s":"I", "c":"NETWORK", "id":4915701, "ctx":"main","msg":"Initialized wire specification","attr":{"spec":{"incomingExternalClient":{"minWireVer
```

mongod

python3 bot

3. Build Chatbot với Rasa Pro

Sau đó chạy action server

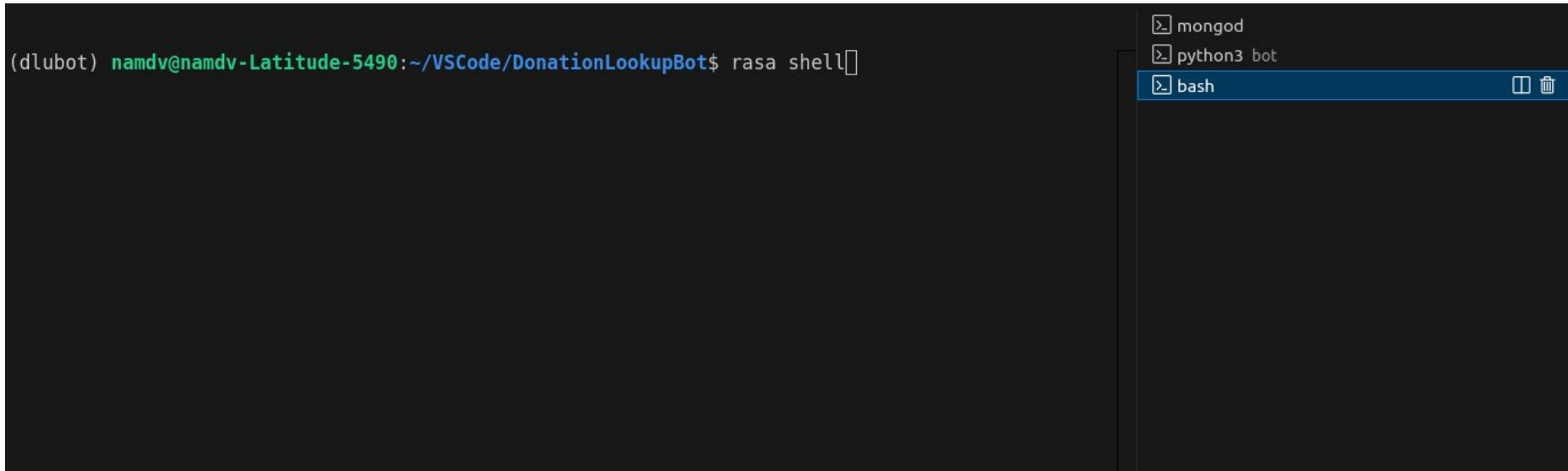
```
(dlubot) namdv@namdv-Latitude-5490:~/VSCode/DonationLookupBot/bot$ rasa run actions
2024-09-17 08:03:08 INFO      rasa.tracing.config - No endpoint for tracing type available in en
dpoints.yml, tracing will not be configured.
2024-09-17 08:03:09 INFO      rasa_sdk.endpoint - Starting action endpoint server...
2024-09-17 08:03:10 INFO      rasa_sdk.executor - Registered function for 'action_search_transac
tion'.
2024-09-17 08:03:10 INFO      rasa_sdk.executor - Registered function for 'action_calculate_tota
l_donations_within_date_range'.
2024-09-17 08:03:10 INFO      rasa_sdk.executor - Registered function for 'action_draw_transacti
on_chart_within_date_range'.
2024-09-17 08:03:10 INFO      rasa_sdk.endpoint - Starting plugins...
2024-09-17 08:03:10 INFO      rasa_sdk.endpoint - Action endpoint is up and running on http://0.
0.0.0:5055
2024-09-17 08:03:10 INFO      rasa_sdk.tracing.config - No endpoint for tracing type available i
n endpoints.yml, tracing will not be configured.
2024-09-17 08:03:10 INFO      sanic.server - Starting worker [73254]
```

mongod

python3 bot

3. Build Chatbot với Rasa Pro

Mở thêm một terminal mới để bắt đầu chạy chatbot



The image shows a VS Code terminal window with a dark theme. The terminal prompt is `(dlubot) namdv@namdv-Latitude-5490:~/VSCode/DonationLookupBot$`. The command `rasa shell` has been entered, and the cursor is at the end of the line. To the right of the terminal, there is a sidebar showing a list of open terminals: `mongod`, `python3 bot`, and `bash`. The `bash` terminal is currently selected and highlighted in blue. There are also icons for opening a new terminal and deleting a terminal in the sidebar.

```
(dlubot) namdv@namdv-Latitude-5490:~/VSCode/DonationLookupBot$ rasa shell
```

3. Build Chatbot với Rasa Pro

Chat với chatbot

```
Bot loaded. Type a message and press enter (use '/stop' to exit):
Your input -> Tìm kiếm thông tin giao dịch
2024-09-17 08:05:05 INFO      openai      - message='OpenAI API response' path=https://api.openai.com/v1/embeddings processing_ms=
21 request_id=req_4f975aca30b47920b6b5b3963cc5c0b5 response_code=200
2024-09-17 08:05:06 INFO      openai      - message='OpenAI API response' path=https://api.openai.com/v1/chat/completions processi
ng_ms=182 request_id=req_63b7c70ccecb3d2a444e9faab3880b186 response_code=200
Nhập thông tin cần tìm kiếm
Your input -> CAO VIET TUAN
2024-09-17 08:05:16 INFO      openai      - message='OpenAI API response' path=https://api.openai.com/v1/embeddings processing_ms=
41 request_id=req_18ba9b5a978c878552ed031a3a3a0046 response_code=200
2024-09-17 08:05:17 INFO      openai      - message='OpenAI API response' path=https://api.openai.com/v1/chat/completions processi
ng_ms=542 request_id=req_7bc75c87c9d169a6e5053f61500638bb response_code=200
# Kết quả search với 'CAO VIET TUAN' #
DATE: 01/09/2024
AMOUNT: 29000
CONTENT: MBVCB.6916176124.CAO VIET TUAN chuyen tien.CT tu 0101001222009 CAO VIET TUAN toi 0011001932418 MAT TRAN TO QUOC VN -
BAN CUU TRO TW
CT CODE: 5241.83107
What else can I help you with?
Your input -> Tính tổng số tiền giao dịch từ ngày 1/9/2024 đến ngày 4/9/2024
2024-09-17 08:05:25 INFO      openai      - message='OpenAI API response' path=https://api.openai.com/v1/embeddings processing_ms=
38 request_id=req_2586a4c13ab5f09284ca5a1fbe50481c response_code=200
2024-09-17 08:05:26 INFO      openai      - message='OpenAI API response' path=https://api.openai.com/v1/chat/completions processi
ng_ms=667 request_id=req_c4b4c647944ab6571960e20c0522bf50 response_code=200
Tổng số tiền ủng hộ trong khoảng thời gian 01/09/2024-04/09/2024 là 9513000 vnd
What else can I help you with?
Your input -> Tính tổng số tiền giao dịch trong ngày 9/9/2024
2024-09-17 08:05:34 INFO      openai      - message='OpenAI API response' path=https://api.openai.com/v1/embeddings processing_ms=
21 request_id=req_f6a15fc4cdb73b7be20d526a280fd71c response_code=200
2024-09-17 08:05:35 INFO      openai      - message='OpenAI API response' path=https://api.openai.com/v1/chat/completions processi
ng_ms=409 request_id=req_b83129e40d0e91903ef9b8e958447d10 response_code=200
Tổng số tiền ủng hộ trong ngày 09/09/2024 là 7868744495 vnd
What else can I help you with?
Your input -> Về biểu đồ thể hiện số lượng giao dịch từ 1/9/2024 đến 10/9/2024
2024-09-17 08:05:40 INFO      openai      - message='OpenAI API response' path=https://api.openai.com/v1/embeddings processing_ms=
34 request_id=req_7884ad4255ed5e42f7944988d3950c7e response_code=200
2024-09-17 08:05:41 INFO      openai      - message='OpenAI API response' path=https://api.openai.com/v1/chat/completions processi
ng_ms=337 request_id=req_e2cfc73602a70f2773142324924247f0 response_code=200
Biểu đồ mà bạn cần đã được lưu tại /home/namdv/VSCode/DonationLookupBot/transaction_charts/chart_17092024_080541.png
What else can I help you with?
```

DONATIONLOOKUPBOT

bot

data

tools

transaction_charts

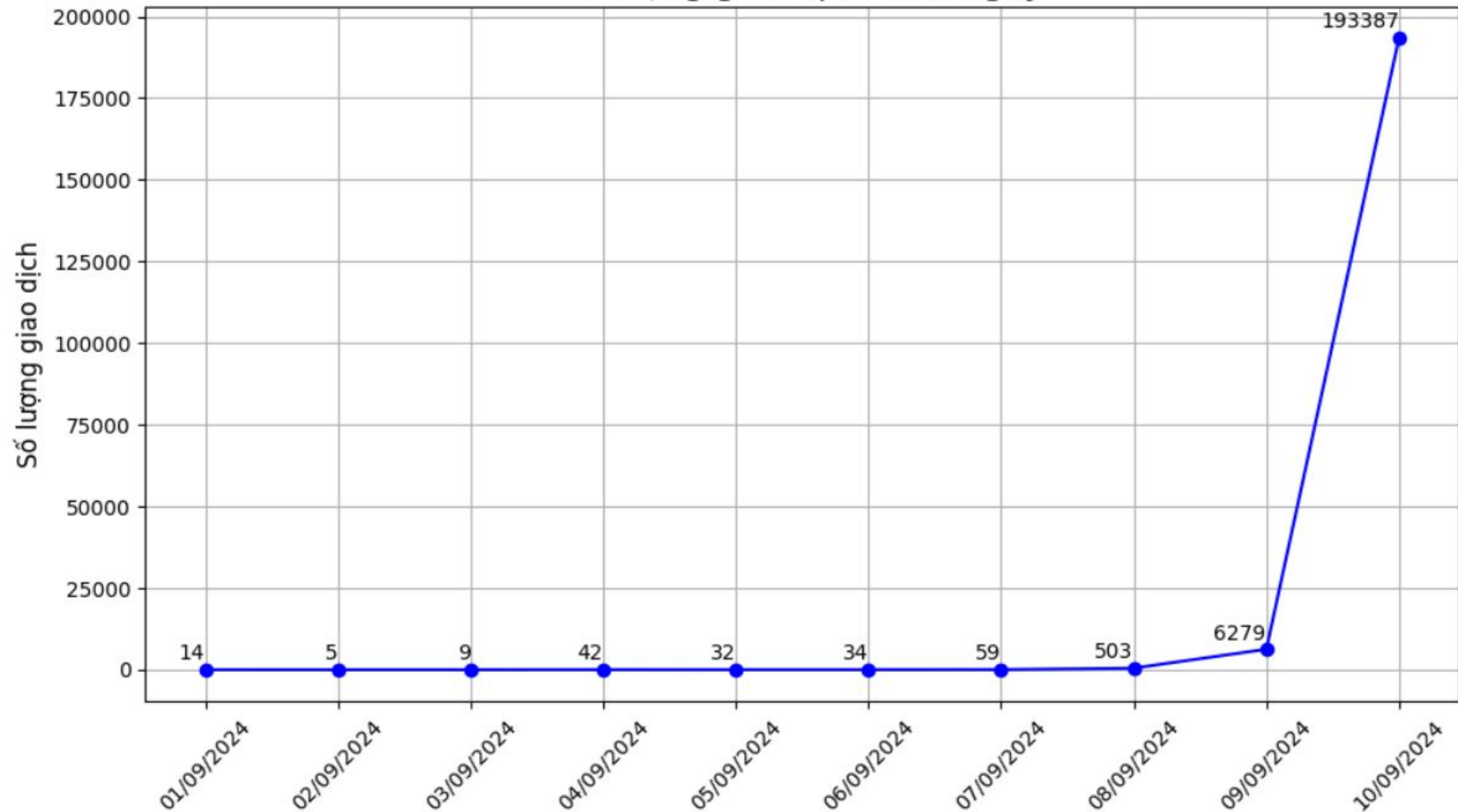
chart_17092024_080326.png

.env

.gitignore

README.md

Số lượng giao dịch theo ngày



Tham khảo

<https://www.facebook.com/share/p/GMxWhnZ1tgXKdNzq/>

https://www.tiktok.com/@juno_okyo/video/7414532506705300737?is_from_webapp=1&sender_device=pc&web_id=7414513215453398549

<https://www.facebook.com/share/v/ruQizoxuEWu3Pe1B/?mibextid=K35XfP>

<https://fpt.ai/vi/bai-viet/xay-dung-chatbot-tu-kinh-nghiem-thuc-te-voi-fptai-conversation/>

<https://vinbigdata.com/chatbot/kich-ban-chatbot-hieu-qua-cho-doanh-nghiep.html>

<https://fchat.vn/blog/19-kich-ban-chatbot-mau-chuyen-nghiep-nhat-hien-nay.html>

https://viblo.asia/p/function-calling-loi-giai-cho-he-thong-rag-linh-hoat-va-hieu-qua-y37Ld1k0Vov#_2-van-de-thuong-gap-khi-xay-dung-he-thong-rag-1