

スクリーピング
システム概要
まとめ

卒業研究について

蒲田 涼馬 (Ryoma Gamada)
u020010@st.pu-toyama.ac.jp

富山県立大学情報システム工学科 4 年

October 13, 2023

進捗

2/10

進捗

- ① タイトル・方針仮決定
- ② グラフの作成
- ③ 手法についての勉強
- ④ スクレイピングシステムの実装

スクレイピング
システム概要
まとめ

進捗

3/10

タイトルについて

経済情報の波及メカニズムの分析による確率的グラフィカルモデルを適用した予測(仮)

方針について

金融情報に影響を及ぼすとされている変数がそれぞれ確率のもとに動いているとし、確率的グラフィカルモデルを作成。

グラフィカルモデル上でそれぞれのデータが将来的に上がるか下がるかを表示し、最も矛盾が少なくなる形でモデルを作成。

そのモデルから未来の金融情報を分析することができるようなシステムを作成する。

為替リスク回避のための金融情報予測

波及メカニズム

4/10

波及メカニズム



図 1: 実行結果

図 2 は非循環モデルであるが、実際に作成したメカニズムでは循環している部分もあり、これについても考える必要がある。(ベイジアンネットワーク? マルコフモデル? あるいはそれ以外)

波及メカニズムと確率的グラフィカルモデル

5/10

予測手法について

スクリーピング
システム概要
まとめ

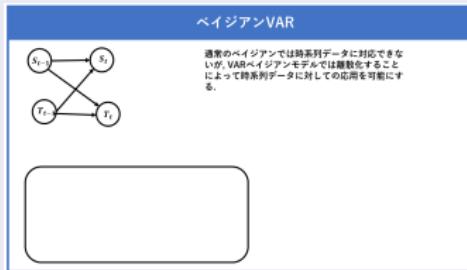


図 2: コールレート情報

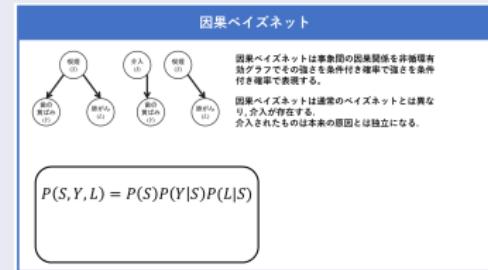


図 3: 個人支出情報

今はこれらの1つ、あるいは両方をシステムに取り入れようと考えている。

ベイジアン VARに関しては python にモジュールがある。

因果ベイズネットに関しては時系列に対応できないそうなのでもし使うならばベイジアン VARと組み合わせて使う必要がある。

3D グラフについて

6/10

3D グラフ (途中)

スクリーピング
システム概要
まとめ



図 4: モデル

今回のグラフは自分自身で座標を決定したものになっているが、最終的には自動的に作成することができるようにならう。このグラフ上で全てを表示し、分析に用いることができるようなシステムにする必要がある。

変数とデータについて

7/10

変数

- ネット上からデータを拾ってくることができる変数のみをまとめたところ、図 4 のようになった。
- 前回の研究報告時点では約 120 個の変数があったが現在は約 30 個(ほとんどが e-stat,)ほど。
 - だいたいが csv ファイルでのデータではあったがいくつかはスクリーピングが必要なものがあった。
 - また csv 形式のものでも内容がまったく違うものがほとんど。

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22
1	データコード	FMB1STRFM01STRFM01STRFM01STRFM01STRFM01STRFM01STRCOOL																				
2	系列名	東日本第一建設第一建設第二建設第三建設ヨウコウセキセイ																				
3	2023/1/1 NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	
4	2023/1/2 NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	
5	2023/1/3 NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	
6	2023/1/4 -0.016	0.001	-0.03	-0.016	0.001	-0.08																
7	2023/1/5 -0.012	0.001	-0.005	-0.012	0.001	-0.065																
8	2023/1/6 -0.009	0.001	-0.055	-0.009	0.001	-0.055																
9	2023/1/7 -0.008	0.001	-0.07	-0.009	0.001	-0.07																
10	2023/1/8 NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	
11	2023/1/9 NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	
12	2022/1/10 NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	
13	2022/1/11 -0.014	0.001	-0.087	-0.014	0.001	-0.087																
14	2022/1/12 -0.021	0.001	-0.06	-0.021	0.001	-0.06																
15	2022/1/13 -0.021	0.001	-0.07	-0.021	0.001	-0.07																
16	2022/1/14 -0.02	0.001	-0.065	-0.02	0.001	-0.065																

図 5: コールレート情報

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	
1	二人以上の会話_支店金額[円]																						
2																							
3	支店金額_施設	大分類	中分類	小分類	中間証	背号	品目分類	2018年	2019年	2020年	2021年	2022年	2023年	2024年	2025年	2026年	2027年	2028年	2029年	2030年	2031年		
4	4	1					支店金額合計	10000	10000	30000	10000	10000	10000	10000	10000	10000	10000	10000	10000	10000	10000		
5	5	2					支店世帯数	3772	7711	7700	3438	7522											
6	6	3					販賣入店数	3.02	2.99	2.98	2.98	2.97											
7	7	4					販賣生徒人	0.97	0.98	0.98	0.98	0.98											
8	8	5					販賣社員人	0.82	0.84	0.86	0.83	0.82											
9	9	6					販賣人件数	1.94	1.94	1.94	1.94	1.94											
10	10	7					支店本社員	1.35	1.34	1.32	1.31	1.34											
11	11	8					支店本社員	50.8	58.2	59.6	59.3	58.4											
12	12	9					持家率[%]	83.7	84.9	86.1	84.6	85.1											
13	13	10					家庭・居住	15.2	14.3	13.1	14.3	12.8											
14	14	11	1	1-	1-	1-	酒類販賣	334842	339025	339430	344782	352057											
15	15	12	2	1-	1-	1-	飲料	92300	92300	92300	92300	92300											
16	16	13	3	1-	1-	1-	総額	7650	7650	7650	7650	7650											
17	17	14	4	1	1	1	102 種	22851	23522	23591	24114	23212											
18	18	15	4	1	1	1	パン	36957	30394	29957	30554	32164											
19	19	16	5	1	1	1	128 食パン	9823	8964	9131	9341	9912											
20	20	17	5	1	1	1	128 沖縄パン	21484	21398	28850	22123	22225											
21	21	18	4	1	1	1	総額	17747	17668	17200	17388	17711											
22	22	19	5	1	1	1	134 生ごみ	3495	3418	3368	3313	3094											

図 6: 個人支出情報

スクレイピング

8/10

スクレイピング

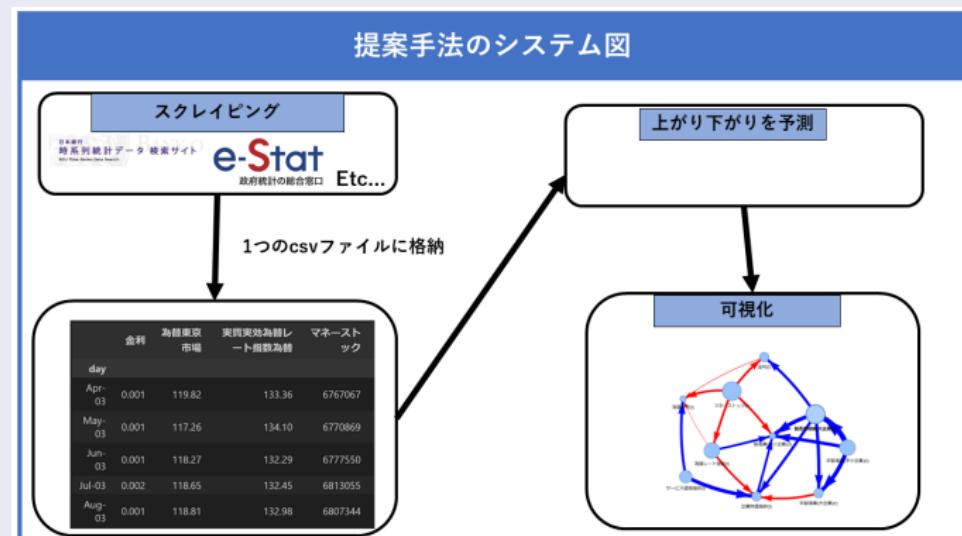
日本銀行時系列サイトから自動でデータを持ってくるシステムを作成した。

データは csv でダウンロードしたものになるが、これらが一つの csv にまとまっているければ、扱うのが少し面倒くさいため、1つの csv にまとめる必要がある。

スクリーピング
システム概要
まとめ

システム

グラフの作成、因果性の分析をした。



まとめ

10/10

進捗まとめ

LiNGAM を用いて因果性を求め、グラフの作成をした。手法についての勉強をした。

今後やること

使う手法を決めてそれについての実装をする。