



HACARUS

ファルコバイオシステムズ様ご提案

課題: カルテ文章チェック

カルテ文書を解析し、記載ミスの可能性がある箇所を抽出し指摘する

- 文書の主題(例: 胃癌)を検出
- 主題から外れる用語を検出

判 断	s/o Adenocarcinoma
推定病変	腺癌の疑

細胞所見は正しく腺癌の話題か？

細胞所見

核腫大、核形不整を呈する異型細胞を散在性から小集塊にて認めます。
核は偏在傾向にあり腺癌細胞を疑います。
精査を希望します。

トピックモデル Latent Dirichlet Allocation

トピックモデルは一つの文書が複数のトピックからなることを仮定する言語モデルであり、文書クラスタリングによく使われている。

今回はトピックモデルの 1つである LDA を使って文書の単語のトピックを推論し、異なるトピックの単語が文書に入っている場合に異常なものとして検出する。

トピックモデルを使って主題や異常単語を検出

胃癌の中には、胃の壁を硬く厚くさせながら広がっていくタイプがあり、これをスキルス胃癌といいます。早期のスキルス胃癌は内視鏡検査で見つけることが難しいことから、症状があらわれて見つかったときには進行していることが多く、治りにくい癌です。

主題: 胃癌
異常な単語: なし

胃癌の中には、**子宮**の壁を硬く厚くさせながら広がっていくタイプがあり、これをスキルス胃癌といいます。早期のスキルス胃癌は内視鏡検査で見つけることが難しいことから、症状があらわれて見つかったときには進行していることが多く、治りにくい癌です。

主題: 胃癌
異常な単語: 子宮

[Demo](#)

課題: ポジティブ or ネガティブな文書判断 (概要)

病理組織
検査結果

①②③ : Stomach / Biopsy / Group 1 / Non-neoplastic lesion without or with atypia / Fundic gland polyp / No malignancy

所 見

①②③ : いずれも同様の所見で、胃底腺の過形成によるポリープです。上皮の異型性はほとんど認めません。悪性所見は認めません。H. pyloriは明かではありません。

検査結果と所見とで内容に矛盾がないか判定したい

- 感情分析にかけて文書ごとにポジティブ度を算出
 - 検査結果と所見とでポジティブ度を比較
- 矛盾を発見した場合は根拠となる表現を提示
 - 検査結果 → No malignancy
 - 所見 → 悪性所見を認める

感情分析について

1. 表現とスコアの辞書を利用し、文章に含まれるポジネガ表現の平均スコアを算出
 - 例
 - +0.9 異型性はほとんど認めません
 - +1.0 悪性所見は認めません
 - → +0.95 (平均スコア)
 - 単純で分かりやすいが、日本語用の辞書は一般向けで単語を基準とする([単語感情極性対応表](#)、[日本語評価極性辞書](#))。精度向上には単語の組み合わせに注目した前処理と辞書が必要になる([Du et al., 2019](#))。英語の辞書も必要。
2. ニューラルネットワークを用い、ポジティブかネガティブか確率的に分類
 - fastTextなど文脈を考慮できるモデルが提案されている ([商品の☆数を推定する例](#))
 - 教師データがポジティブかネガティブかラベル付けしておく必要がある

血液検査別人検知

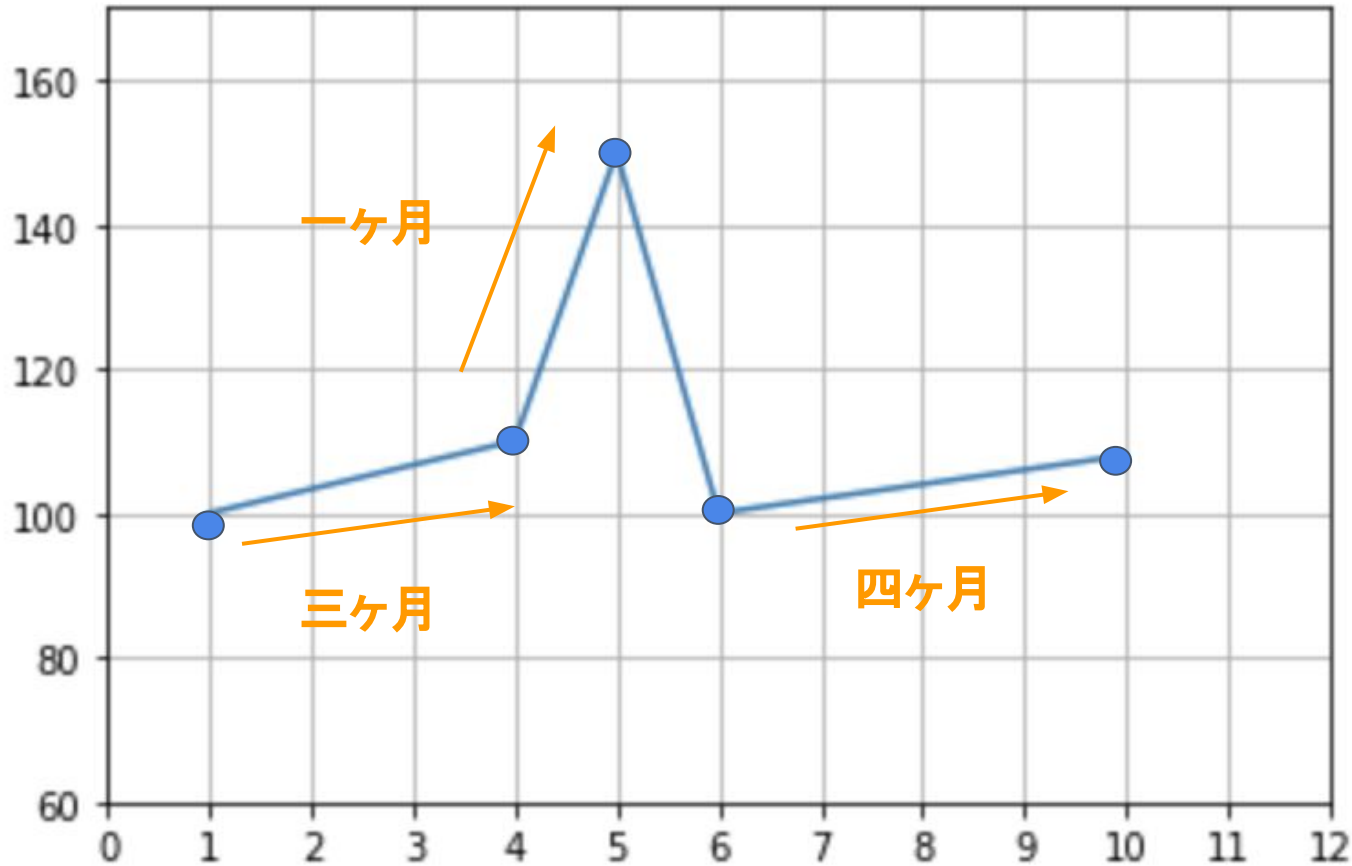
受付日	今回値	前回値(1)	前回値(2)	前回値(3)	前回値(4)	前回値(5)													
検体番号	20/02/03	19/10/30	19/10/02	19/01/05	17/12/16	17/01/10													
性別/生年月日	男性 S60/08/13	男性 S50/09/13	男性 S50/09/13	男性 S60/08/13	男性 S60/08/13	男性 S60/08/13													
カルテN°	3907	5394	5394	3907	3907	3907													
材料コメント	99	99	99	99	99	99													
コード	略称	結果値	付加	再	結果値	付加	再	結果値	付加	再	結果値	付加	再	結果値	付加	再	結果値	付加	再
0003	T-BIL	0.6			0.5			0.6			0.6			0.5			0.7		
0037	T-P	6.8			6.7			7.5			7.0			7.0			7.0		
0038	7/7'ミン	5.0			4.5			4.6			4.9			5.0			5.0		
0039	A/Gt	2.8			2.0			1.6			2.3			2.5			2.5		
0067	AST	21			30			28			19			17			16		
0071	ALT	20			40			45			19			16			14		
0073	ALP	195			202			199			199			202			184		
0075	LD	186			149			180			167			156			158		
0079	G-GT	34			31			42			24			22			20		
0143	CK(CPK)	112			170			102			89			70			63		
0261	T-CHO	175			185			204			207			173			179		
0265	TG	63			97			118			66			97			60		
0351	ナトリウム	137			140			137			144			142			139		
0353	カルシウム	104			103			102			103			103			102		
0352	カリウム	10.5 EE	\$		4.6			5.3			5.2			4.2			9.2 EE	\$	
0425	BUN	10.6			15.4			16.4			13.5			10.9			10.7		
0426	クレアチニン	0.73			0.93			0.88					0.71				0.73		
0522	BS	97						96					103				84		
5588	BS BL	21.4																	
0807	CRP Q	0.05L			0.05L			0.05L			0.05L			0.05L			0.05L		

別人？

情報一覧を表示中です

別人？

血液検査別人検知



定期的な検査ではないことや、
データ数が少ない点から考えて
機械学習の代わりに統計方法で
判定することが妥当

血液検査別人検知

今回値		前回値(1)		前回値(2)		前回値(3)		前回値(4)		前回値(5)	
今回値		前回値(1)		前回値(2)		前回値(3)		前回値(4)		前回値(5)	
受付日	20/02/03	19/10/30		19/10/02		19/01/05		17/12/16		17/01/10	
検体番号											
性別/生年月日	男性 S60/08/13	男性 S50/09/13		男性 S50/09/13		男性 S60/08/13		男性 S60/08/13		男性 S60/08/13	
カルテNo	3907	5394		5394		3907		3907		3907	
材料コメント	99	99		99		99		99		99	
コード	略称	結果値	付加再	結果値	付加再	結果値	付加再	結果値	付加再	結果値	付加再
0003	T-BIL	0.6		0.5		0.6		0.6		0.5	
0037	T-P	6.8		6.7		7.5		7.0		7.0	
0038	アルブミン	5.0		4.5		4.6		4.9		5.0	
0039	A/G	2.8		2.0		1.6		2.3		2.5	
0067	AST	21		30		28		19		17	
0071	ALT	20		40		45		19		18	
0073	ALP	195		202		199		199		202	
0075	LD	186		149		180		167		184	
0079	G-GT	34		31		42		156		158	
0143	CK(CPK)	112		170		102		24		20	
0261	T-CHO	175		185		204		70		63	
0265	TG	63		97		118		173		179	
0351	ナトリウム	137		140		137		97		60	
0353	カルシウム	104		103		103		142		139	
0352	カリウム	10.5 EE	\$	4.6		5.3		103		102	
0425	BUN	10.6		15.4		12.1		4.2		9.2 EE	\$

統計
処理

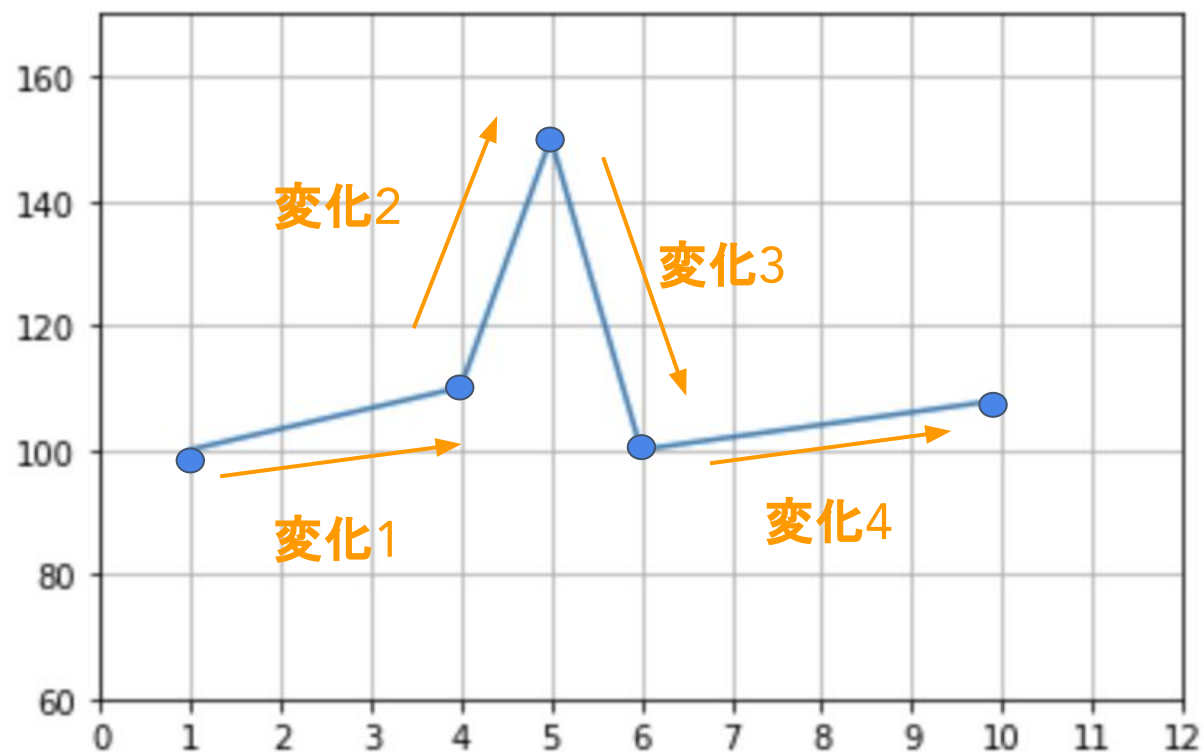
数値の正常変化を測定項目ごとにモデル化。必要に応じて相関関係にある他の項目やデモグラフィック情報等も組合せてモデル化する。

例：
30代、糖尿病歴あり男性
の血糖値変化をモデル化

正常データ

血液検査別人検知

別人データが入った場合



別人データの場合、時系列で見ると値の急激な変化が短時間に発生すると考えられる。
この急激な変化のうち正常な変化と考えられる範囲から外れるものを別人として検出する。