

# セールスマンの活動日報に基づく商談の成否予測

千賀 裕貴<sup>1</sup> 峯 恒憲<sup>1</sup>

**概要:** 多くの企業が商談の活動内容を日報に記録している。この日報を分析し、商談の早い段階で成否予測ができれば、今後の商談活動を進める上で有益である。本稿では、この日報を分析し、商談の成否予測を行う手法を提案する。提案手法では機械学習を用いることで、商談の成否を分ける要因を抽出するとともに、進行中の商談を成功に導く要因を推定する。

## Applying Text Mining Techniques to Sales Persons' Daily Reports to Predict Business Meeting Results

SENGA HIROTAKA<sup>1</sup> MINE TSUNENORI<sup>1</sup>

### 1. はじめに

ブログや SNS などから、そこへ投稿された製品やサービスについての投稿者の意見や評価を自動的に掴む、評判分析や感情分析と呼ばれる研究が、90年代の終わりから盛んに行われている。企業でも、ナレッジマネジメントの重要性から、新たな商品開発やサービス戦略に役立つ目的で、早くから電子化文書の活用が進んでおり、営業日報やコールセンターへの問い合わせ、アンケートなど、企業内で蓄えられた大量のデータを自動的に分析し、ビジネスへ活用する試みがなされてきている [1]。たとえば顧客の要望や苦情の把握、課題の分析、営業プロセスの明示化、顧客ニーズにマッチした商品提案など、様々な用途への活用があり、その重要性は増すばかりである。

このような背景のもと、本稿では、特に、セールスマンの商談活動を記録した営業活動日報（以下、営業日報）を分析対象とする。セールスマンの商談は、1回から2回の短期の営業活動で成否が決まるものから、7回以上と長くかかるものまで様々である。営業活動の早期のうちに、成否予測が高精度でできれば、営業活動を行うセールスマンに対して意欲の提供（成立の可能性が高い場合）や、注意を促す（不成立の可能性が高い場合）ことも期待できる。同時に、営業活動の問題点の指摘とともに、活動内容の改

善を促す特徴の抽出なども期待できる。

そこで、本稿では、毎回の商談活動記録をもとに、商談の成否予測を行う手法を提案する。提案手法は、毎回の商談活動記録に機械学習手法を適用し、成否予測を行う。また、提案手法を適用することで、商談の成否を分ける要因の抽出を試み、その結果について議論する。実験の結果推定精度は F-score 0.71 が得られ、成功する商談に頻出の単語が抽出できた。

以下、2章では、関連研究について述べ、本稿で示す研究内容との違いを明確にする。3章では分析に用いたデータの概要を説明する。4章では日報のベクトル化と機械学習の手法について説明する。5章では実際に行った実験とその結果を報告する。6章では実験結果に対する考察を述べる。最後にまとめと今後の課題について述べる。

### 2. 関連研究

企業におけるナレッジマネジメントの重要性から、営業日報の分析についても早くから行われてきている。

市村ら [1] は、情報抽出技術に基づくテキストマイニング手法を日報分析に応用したシステムについて報告している。また、[2] では、知識辞書を用いた分析手法を適用した日報分析システムと、その分析用知識記述を支援する知識辞書構築支援ツールについて述べるとともに、実際の店頭管理日報分析に適用し、支援ツールの有効性を示している。

<sup>1</sup> 九州大学  
Kyushu University

杉原ら [3] は、営業支援システムに蓄積されているテキストデータから、顧客の課題の把握やニーズにマッチした商品提案などを目指し、課題記述文（「望ましくない状況や望ましいゴールといった解決・改善の対象や結果として記述される文」）の抽出を行った。

これらは、課題の抽出や、分析対象などの知識・概念の把握などを目的として行われており、本稿で述べる商談の成否予測とは趣旨も手法も異なる。

一方、割石ら [4] は、本稿が解析に利用したデータと同じセールスマンの営業日報を利用し、それに機械学習手法（サポートベクターマシン：SVM）を適用し、好成績の営業員とそうでない営業員の行動の違いを識別する要因の抽出を行ったが、本稿では、営業活動記録から、商談の成否を予測しており、分析の目的が異なる。

### 3. データ概要

今回利用したデータは IT 機器の販売やリースを行っている企業から提供して頂いたものである。データの収集期間は 2014 年 5 月から 2015 年 10 月の 18 ヶ月。営業日報の詳細を図 1 に示す。図 1 は商談ごとのデータで 1 行が一つの商談に対応する。各行には商談固有の ID、取引先の ID、商談の名称、商談の種別、登録の種別（システムで自動生成されたものが 0、営業マンが入力したものが 1）、金額、完了（予定）日、フェーズ（選択式）、確度、説明、担当者 ID が記述されている。図 2 はコメントごとのデータで 1 行が 1 つのコメントを表す。各行には商談データに対応する ID とコメントが登録された日時、登録者の ID、コメント内容が記述されている。

### 4. 分析手法

商談の成否を分ける要因を抽出する手段として、成功に終わった商談を正例、失敗に終わった商談を負例として機械学習を行う。そして正例と負例それぞれの特徴語を抽出し、その単語が出現するコメントが成否に関わる要因であると考え、コメントの分析を行う。

機械学習を行うには日報をベクトル化する必要がある。以下に日報のベクトル化と、機械学習による特徴語の抽出手法について説明する。今回機械学習には SVM と Random Forest の 2 つの手法を用い、より成績のよかった Random Forest を用いて特徴語抽出を行った。

#### 4.1 日報のベクトル化と tf-idf

1 つの商談を一つのベクトルとして表す。この手法は bag of words と呼ばれる。ベクトルの各次元はコメントに出現する一単語に対応し、その値は各単語の重みである。本稿で分析に用いた単語の品詞は名詞、動詞、形容詞、形容動詞、副詞のみである。これはこれらの品詞がテキストの特徴抽出に際して特徴を強く反映するとされているからであ

る [5]。重みの算出は tf-idf という手法を用いた。また、連続する二単語を一組としての分析も併用した。

tf-idf は tf(term frequency, 単語の出現頻度) と idf(Inverse document frequency, 逆文書頻度) を乗じて求められる。文書  $d(j)$  中の単語  $t(i)$  に関して、

$$tfidf(ij) = tf(ij) * idf(i)$$

tf(ij) は文書  $d(j)$  内の単語  $t(i)$  の出現頻度であり

$$tf(ij) = \frac{n(ij)}{\sum_k n(kj)}$$

と表される。 $n(ij)$  は単語  $t(i)$  の文書  $d(j)$  における出現回数、 $\sum_k n(kj)$  は文書  $d(j)$  に出現するすべての単語の出現回数の和である。 $idf(i)$  は総文書における単語  $t(i)$  の珍しきであり

$$idf(i) = \log\left(\frac{|D|}{|\{d : d \ni t(i)\}|}\right)$$

と表される。 $|D|$  は総文書数、 $|\{d : d \ni t(i)\}|$  は単語  $t(i)$  を含む文書数である。

ここで、同じ単語であっても正例に含まれる場合と負例に含まれる場合では重みが異なると考え、より成否の 2 値分類に適した tfidf を考案した。文書  $d(j)$  中の単語  $t(i)$  に関して

$$tfidf_x(ij) = tfidf(ij) * \frac{idf_x(i)}{idf_{x'}(i)}$$

$x$  は P または N であり、文書  $d(j)$  が正例に含まれるなら P 負例に含まれるなら N である。 $x'$  は  $x$  が P なら N、N なら P である。 $idf_x$  は正例または負例を総文書とした idf であり

$$idf_P(i) = \log\left(\frac{|P|}{|\{d : d \ni t(i), d \in P\}|}\right)$$

$$idf_N(i) = \log\left(\frac{|N|}{|\{d : d \ni t(i), d \in N\}|}\right)$$

と表す。 $|P|, |N|$  はそれぞれ正例、負例の総文書数、 $|\{d : d \ni t(i), d \in P\}|$  は正例中の  $t(i)$  を含む文書数、 $|\{d : d \ni t(i), d \in N\}|$  は負例中の  $t(i)$  を含む文書数である。

#### 4.2 Support Vector Machine

Support Vector Machine(SVM) は 2 値分類のための機械学習手法である。詳しいアルゴリズムは文献 [6] に記載されている。学習によって得られる識別関数  $f(x)$  は以下の式によって表される。

$$f(x) = \text{sign}(w^T \cdot x - h) = \begin{cases} 1 & (w^T \cdot x - h > 0) \\ -1 & (\text{otherwise}) \end{cases}$$

$x$  は入力ベクトル、 $w$  は重みベクトル、 $h$  はしきい値である。 $w$  と  $x$  の内積がしきい値を超えれば 1、超えなければ -1 を出力する。

商談ID	取引先ID	商談名	種別	登録種別	売上	完了予定日	フェーズ	確度(%)	説明	担当者ID
006100(00110000(				1	0	2015-10-23		0		3070
006100(00110000(				1	255600	2015-08-28		100		2880
006100(00110000(				1	600000	2016-03-31		0		3149
006100(00110000(				1	0	2015-12-31		10		3093
006100(00110000(				1	0	2015-07-24		0		3071
006100(00110000(				1	0	2018-03-31		0		3071
006100(00110000(				1	340000	2015-05-31		0		3111
006100(00110000(				1	0	2015-08-04		100		2962
006100(00110000(				1	1E+06	2015-10-16		20		2682
006100(00110000(				1	240000	2015-05-31		50		3066
006100(00110000(				1	240000	2015-08-31		90		3111
006100(00110000(				1	200000	2015-10-16		20		3093
006100(00110000(				1	4E+06	2015-12-18		10		3093
006100(00110000(				1	0	2015-07-31		90		2880

図 1 商談データ

商談ID	コメント日時	コメント投稿者	コメント内容
0061000000fBkG	2015-10-09 20:42:39	2530	
0061000000fBkF	2015-10-09 20:38:23	3087	
0061000000fBkE	2015-10-09 20:35:50	3087	
0061000000fBkE	2015-10-09 20:37:05	3087	
0061000000fBkE	2015-10-09 20:37:36	3087	
0061000000fBk8	2015-10-09 20:00:56	3142	
0061000000fBjh	2015-10-09 19:00:41	3016	
0061000000fBjh	2015-10-09 18:34:37	3102	
0061000000fBjh7	2015-10-09 18:30:49	3102	
0061000000fBjaz	2015-10-09 18:14:18	2739	
0061000000fBjXr	2015-10-09 18:06:51	2476	
0061000000fBjXr	2015-10-09 18:49:44	2476	
0061000000fBjRz	2015-10-09 17:55:03	3142	
0061000000fBjQl	2015-10-09 17:51:42	3066	
0061000000fBjPl	2015-10-09 17:49:42	3066	
0061000000fBjPj	2015-10-09 17:48:41	3066	

図 2 コメントデータ

### 4.3 Random Forest

Random Forest(RF)は複数の決定木を利用する機械学習手法である。決定木は木の形をした分類器であり、入力ベクトルに対してノードごとにある属性値を使って条件分岐することで予測を行う。今回用いた CART アルゴリズムでは多様性を表す Gini 係数の減少量が最大となる属性を分岐条件として選択する。Gini 係数の減少量はその属性が属性感でどれだけ特徴的であるかに相当する。学習させるベクトルの次元数を  $N$  とすると、各決定木を構築する際、ランダムに  $M (< N)$  個のベクトル要素を選ぶ。  $M$  は  $\sqrt{N}$  がよいとされている [7]。そしてそのベクトル要素から Gini 係数の減少量が最大となる要素を選び、これを分岐条件とする。同様の操作を繰り返すことで各決定木を構築する。入力ベクトルに対して、それをすべての決定木に入力し、結果の多数決を取ることで分類結果とする。ここで決定木をいくつにするかは重要である。Oshiro らによれば木の数が 128 を超えると RF の性能は大きく変化しない [4][8]。よって本稿では木の数を 128 とした。

また RF ではベクトル属性の重要度、本稿においては語の重要度を算出し、特徴語を抽出することができる [9]。ある語  $w$  の重要度は次の式で求まる。

$$importance(w) = \frac{1}{|T|} \sum_{t \in T} \sum_{l \in L_w^t} \frac{S_l}{S} * \Delta G(l)$$

$T$  は RF に含まれるすべての木の集合、 $L_w^t$  は木  $t$  において分岐条件が要素  $w$  であるノードの集合である。  $S$  は学習データの総数であり、 $S_l$  はノード  $l$  に到達した学習データの数である。  $\Delta G(l)$  はノード  $l$  における Gini 係数の減少量を表している。 RF での語重要度は正例負例を問わず正の値をとる。

### 5. 実験

1つの商談ごとに最終日を除いたすべてのコメントを一つのデータとしてベクトル化を行った。このとき出現するコメントが5件未満である語は除いた。ベクトルの次元数は単語 867, 連続する二語の組 702 をあわせて 1569, 対象の商談数 2567 であった。成功した商談と失敗した商談を分ける基準として、商談の最終的なフェーズが決定であるものを正例、敗戦・延期・消滅であるものを負例とした。単語のベクトル, 連続する二語のベクトル, それらを合わせたベクトルについてそれぞれ重みに従来 *tfidf* および改良した *tfidf* を使い, 2つの手法により機械学習を行い, 成否を推定するとともに RF により特徴語を求めた。推定精度の尺度には Accuracy と F-score を用いた。Accuracy は

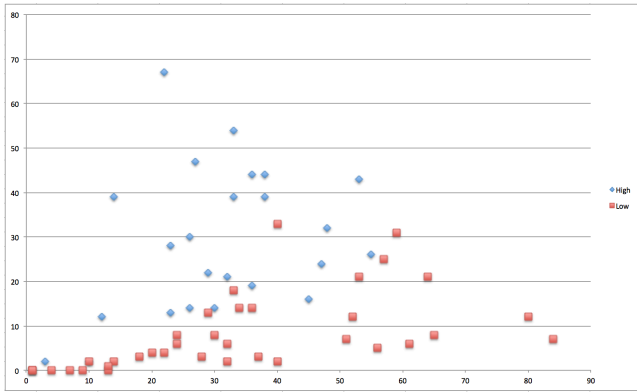


図 3 負 → 正の誤推定率

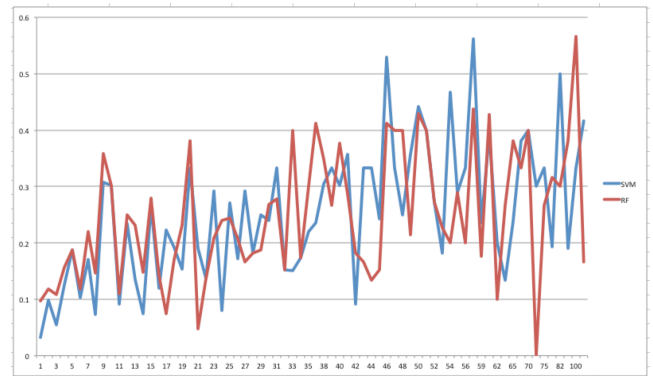


図 5 金額と誤推定率の相関

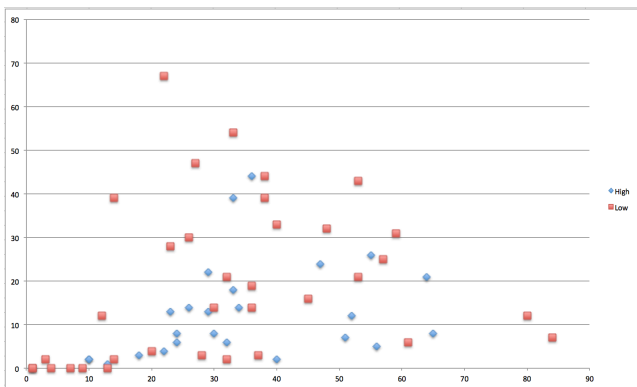


図 4 正 → 負の誤推定率

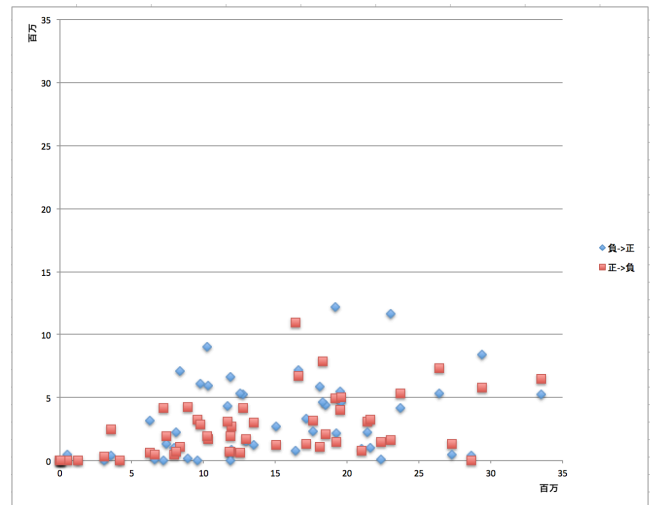


図 6 成功した商談の金額と誤推定した商談の金額

すべてのデータのうち正しく推定できたものの割合であり

$$Accuracy = \frac{|\{d : predict(d) = label(d), d \in D\}|}{|D|}$$

と表される。  $predict(d)$  は文書  $d$  に対する推定結果,  $label(d)$  は文書  $d$  の実際の分類である。 F-score は Precision(適合率) と Recall(再現率) の調和平均であり

$$F - score = \frac{2Precision \cdot Recall}{Precision + Recall}$$

Precision は正 (負) と判定したもののうち正 (負) であるものの数であり

$$Precision = \frac{1}{2} \sum_{n \in \{0,1\}} \frac{|\{d : predict(d) = n \wedge label(d) = n\}|}{|\{d : predict(d) = n\}|}$$

と表される。 Recall は正 (負) であるもののうち正 (負) と判定されたものの割合であり

$$Recall = \frac{1}{2} \sum_{n \in \{0,1\}} \frac{|\{d : predict(d) = n \wedge label(d) = n\}|}{|\{d : label(d) = n\}|}$$

と表される。 各ベクトルと手法の推定精度を表 1 に示す。改良  $tfidf$  を用いた RF が Accuracy 0.78, F-score 0.71 であり最も高い推定精度を示した。また、重みの上位 1% の語のうち、負例より正例に多く出現する、つまり  $idf_P$  より  $idf_N$  が大きい語と、逆に正例より不例に多く出現する語についても調べた。

またこうして得た結果のうち、誤推定されたものの特徴

についても調べた。図 5 図 6 は営業員一人ごとに成否の商談数を横軸に成功数, 縦軸に失敗数でプロットし, 誤推定率の高さで色分けしたものである。負 → 正の誤推定率が低い営業員は成功数が失敗数より多い、つまり成績が良い傾向にあることがわかる。一方、正 → 負の誤推定率に関してはそのような傾向がないことがわかる。

図 7 は商談の金額と誤推定率の関係である。振れ幅が大きい、金額が上がると誤推定率が上がっていることがわかる。図 8 は営業員ごとの成功した商談の金額合計と誤推定した商談の金額合計をプロットしたものである。正 → 負の誤推定に関しては成功した商談の金額に比例して金額が伸びる傾向があるが、負 → 正の誤推定には似た傾向があるもののよりばらつきが大きいことがわかる。図 7 と合わせると、金額が大きくなるほど負 → 正の誤推定率が大きくなると考えられる..

さらに、時系列的な推定の足がかりとして部分的なデータを用いたときの精度についても調べた。図 3 は商談の長さ  $N$  ごとに使うデータの日数を 1 日ずつ伸ばしていったときの F-score の推移である。初日から 2 日目にかけては概ね精度が向上しているが、それ以降は必ずしも上がり続けていない。また、2 日目時点で多くのデータが 0.6 を超え

表 1 推定精度

		従来の tf-idf			改良 tf-idf		
		単語	2 語	単語と 2 語	単語	2 語	単語と 2 語
SVM	Accuracy	0.53	0.52	0.49	0.68	0.70	0.70
	F-score	0.53	0.51	0.48	0.52	0.53	0.49
RF	Accuracy	0.71	0.70	0.72	0.76	0.76	0.78
	F-score	0.62	0.58	0.62	0.68	0.68	0.71

表 2 重み上位の単語

重み	語
16.08	(利, 予算)
14.12	納品
14.07	計上
13.62	検討
13.34	(予算, 計上)

表 3 正例に多い語

$\frac{idf_N}{idf_P}$	語
25.42	受注
13.10	注文書
5.97	(利, 金額)
4.71	(金額, 変更)
3.36	納品

表 4 負例に多い語

$\frac{idf_P}{idf_N}$	語
16.55	難しい
10.71	(日, 見積もり)
9.74	アンケート
8.18	診断
7.79	(社長, 不在)

る F-score を示していることから、商談初期のコメントのみで精度の高い推定が実現できる可能性がある。商談の件数を長さごとにグラフにしたものが図 4 である。日数が短いものほど件数は多く、長さ 1 日が最多となる。図 3 と見比べると短い商談ほど件数が多く推定精度が高いことがわかる。

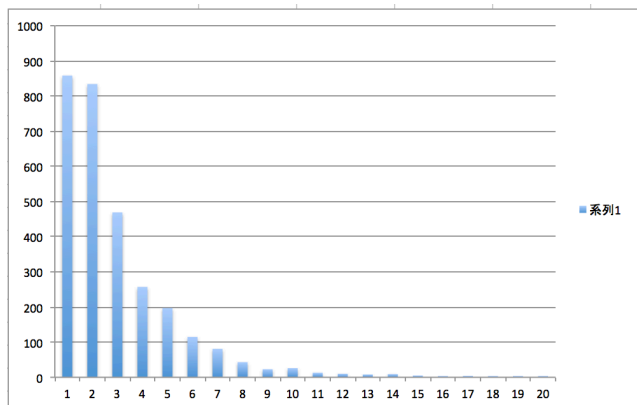


図 7 商談の長さごとの件数

のフェーズを用いることも試みた。図 9 は商談の開始時のフェーズごとに SVM と RF の F-score を示したものである。商談のフェーズに厳密な順序はないが右にあるフェーズほど遅く出現する傾向がある。見積もり書の提出フェーズでの精度が高いのはこのフェーズで始まる商談はもとより確度の高いものも多く、その傾向が文面にあらわれているためだと思われる。一方デモフェーズでの精度が低いのは、デモ自体が商談の成否の鍵となるが、その内容が文面に現れにくいと考えられる。それ以外のフェーズについては目立って大きな特徴は見られなかった。今回は開始フェーズについてしか実験していないが、商談の途中で出現するフェーズや終了前のフェーズといった観点からも分析することで新たな知見を得られる可能性がある。

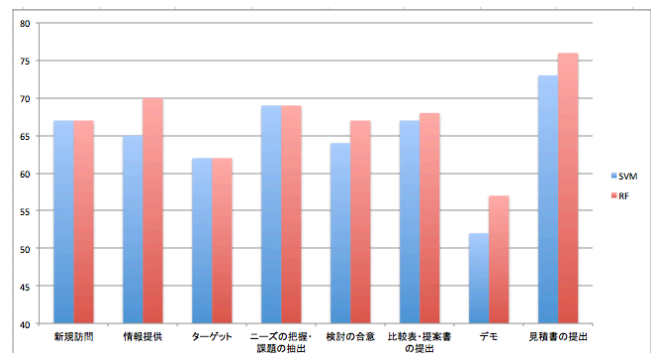


図 8 商談の開始フェーズごとの精度

商談の進捗に関する時系列以外のアプローチとして商談

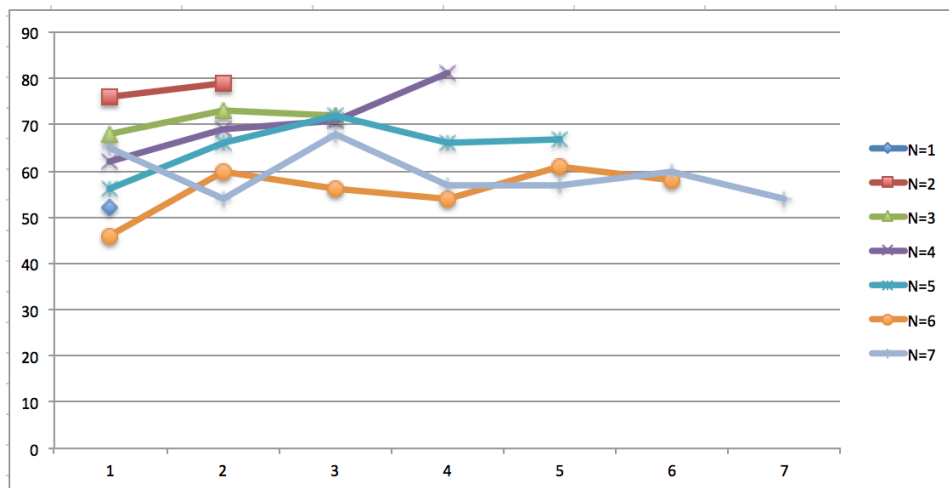


図 9 時系列的な精度変化

### 5.1 考察

重みの大きい語は「予算」「納品」などで、これらは「予算計上」「納品予定」といった文脈で現れており、商談がほぼ確定した状態でのコメントだとわかる。また、正例に多く含まれる語は重み上位の語と類似の傾向があり、正例の特徴が推定に大きく寄与していることがわかる。一方、不例に多く含まれる語について、「難しい」は「購入/工事が難しい」と直接的に成約が困難であることをしめす文脈で現れている。「アンケート」「診断」は、「セキュリティ診断アンケート」の形で現れているが、あまり有効ではないことが伺える。また(社長, 不在)の組み合わせについては現れるのが多くは短い商談の初日であり、飛び込みであることがわかる。ここからアポなしでの飛び込み営業はあまり成果を上げていないと考えられる。

### 5.2 おわりに

本稿では営業日報のコメントから商談の成否を推定すべく、成否を分ける要因の抽出を試みた。推定精度は F-score 0.7 を超えた。特徴語抽出ではアンケートや飛び込み営業が有効ではないとの推測が得られた。今後の課題としては、本稿では最終日を除くすべてのコメントを一つのベクトルとする手法と 1 日ごとに区切る手法を用いたが、異なるデータ分割の手法も検討すべきである。また、個人に注目して誤推定率の比較を行ったが、成績の良い人の失敗例と成績の悪い人の失敗例を比較することで、違った知見を得ることができる可能性がある。そして、今回は語彙にのみ注目して分析を行ったが、それぞれの語が商談のどの段階で出現するかといった情報や係り受けの情報を用いることでより高精度な分析も可能となるだろう。

### 謝辞

本研究の一部は、科研費(課題番号 26540183)の支援により行われた。また本研究の実施に当たって、理研産業株

式会社より営業活動日報データの提供を頂いた。ここに感謝の意を表します

### 参考文献

- [1] 市村由美, 鈴木 優: テキストマイニング技術と応用, 東芝レビュー, Vol. 56, No. 5, pp. 19–22 (2001).
- [2] 市村由美, 鈴木 優, 酢山明弘, 折原良平, 中山康子: 日報分析システムと分析用知識記述支援ツールの開発, 電子情報通信学会論文誌, Vol. J86-D-II, No. 2, pp. 310–323 (2003).
- [3] 杉原大悟, 大熊智子, 佐竹功次, 三浦康秀, 服部圭悟, 増市博: 営業支援システム内に蓄積されたテキストデータからの課題記述文抽出(抽出, 第 2 回テキストマイニング・シンポジウム) Extraction of Sentences Describing Problems from Sales Force Management System Texts, 電子情報通信学会技術研究報告. NLC, 言語理解とコミュニケーション 381, 電子情報通信学会 (2012).
- [4] 割石奈生, 御手洗秀一, 鈴木孝彦, 廣川佐千男: テキストマイニングによる営業日報の分析, 電子情報通信学会技術研究報告人工知能と知識処理 ai2015-32, 電子情報通信学会 (2015).
- [5] 野坂 政司 國府 久嗣: 内容推測に適したキーワード抽出のための日本語ストップワード, 日本感性工学会論文誌, Vol. 12, No. 4, pp. 511–518 (2013).
- [6] Vapnik, V. N.: *The nature of statistical learning theory*, Springer-Verlag (1995).
- [7] Trevor Hastie, Robert Tibshirani, J. F.: *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*, Springer (2009).
- [8] Baranauskas, T. M. O. S. P. A.: How Many Trees in a Random Forest?, *Machine learning Machine Learning and Data Mining in Pattern Recognition*, pp. 154–168 (2012).
- [9] Marco Sandri, P. Z.: A Bias Correction Algorithm for the Gini Variable Importance Measure in Classification Trees, *Journal of Computational and Graphical Statistics*, Vol. 17, No. 3, pp. 611–628 (2008).