

# テキストマイニング技術を利用した お客様コールログ分析

北村 美穂子 村田 稔樹  
佐々木 美樹 奥村 晃弘

近年、データの利活用技術が急速に進み、計算処理が容易な数値や構造化データだけでなく、曖昧性を多く含んだ文章や音声データもデータ分析に利用する試みが活発になっている。

特に、お客様からのクレームや要望等が集まるコールセンターのお客様の問合せ内容記録(以降、「インバウンドのコールログ」と呼ぶ)の分析は、サービス向上や新製品の開発に直接活かすことができるため顧客志向の経営では欠かせない技術である。その一方で、分析の方法論は確立されておらず、試行錯誤的な要素が多い。

我々は、インバウンドのコールログの一般的な特徴を調査し、かつ、その分析結果を利用する現場の要望をヒアリングした。その結果から、インバウンドのコールログを分析するための4つの分析観点を策定した。さらに、4つの分析観点毎にコールログを自動分類し、観点別の時系列変化や観点間の関係性を見える化することで効果的な分析を行っている。

本稿では、まず、基本的なテキストマイニング技術について説明し、次に、我々が策定したインバウンドのコールログの分析観点について説明する。次に、本手法の適用事例として、OKIプリンターのお問合せコール分析の事例について紹介し、最後に、今後の取り組みや方向性について述べる。

## テキストマイニング技術

従来、お客様アンケートや電話の応対記録等の分析作業は、人間が読んでその傾向を把握したり、分類する必要があった。そのため、コストがかかり、かつ分析者の主観が入ることも多く、継続的な分析を行うことが難しかった。

「テキストマイニング技術」は、非定型の文章を意味ある単位(単語や句)に機械的に分解することで、単語や句を1つのデータとみなし、単語や句の出現頻度やそれらの関係性を統計的に分析する技術である。

図1に、テキストマイニング技術を模式化した図を示す。この図のようにテキストマイニング技術は、文章を単

語に分割し、単語間の係り受け関係を同定したり、意味的な曖昧性を解消したりする「言語処理技術」<sup>1)</sup>と、データマイニングと呼ばれる「統計処理技術」の2つの技術で構成されている。図1のように「統計処理技術」には様々な手法があるが、目的によって利用すべき手法は異なる。例えば、既存の分類項目に従って文書を自動分類したい場合は「カテゴライジング」手法を利用する。また、類似した文書を集めたい場合は「クラスタリング」手法を利用する。

このように、単語や句の出現の統計的な傾向をみることによって、その文書には、どのような内容が書かれているかを知ることができたり、お客様のタイプや時期による特徴を掴んだりすることができる。

テキストマイニング技術は、既にツール化され、市販されている。しかし、市販のツールを使ったとしても、要素技術の専門性が高く、かつ、分析目的に応じた手法の駆使が求められるため、ビジネス上有益な結果を得るためには、長期的な分析ノウハウの蓄積と試行錯誤が必要である場合が多い。

## インバウンドコールログを分析するための観点

我々は、インバウンドのコールログの特徴を捉えた分析観点を4つ設定した。

- ・分析観点 1. お客様の問合せの目的
- ・分析観点 2. 問合せの真の理由
- ・分析観点 3. 問合せの真の対象製品、機能、サービス
- ・分析観点 4. 対応結果

### (1) 分析観点 1. お客様の問合せの目的

インバウンドのコールログは、お客様が電話するに至った目的が必ず存在し、その目的は一般化することができる。図2は、「お客様の問合せの目的」の基本分類である。購入後のお客様は、お客様自身が困っているのか？(図2②)、それとも情報が欲しいのか？(図2③)、何らかの手続きをして欲しいのか？(図2④⑤)に大まかに分類することができる。



お客様やオペレーターの主観に基づいて入力されており基準が一定でない。このため、既にコールに付与されている値ではなく、対象となる製品やサービス名をテキストマイニング技術により決定する。

なお製品やサービス名は、商品化の際、機能や目的別に体系化されていることが多く、その体系を直接テキストマイニングで使用する。

#### (4) 分析観点 4. 対応結果

お客様の問合せに対して、コールセンターのオペレーターが最終的に、どのように対処したかを分析する。インバウンドのコールセンターでは、

- A) オペレーターの説明で解決するもの
- B) お客様になんらかの作業をお願いするもの
- C) サービス(製品)提供者側の作業が伴うものに大別される。さらに A) ~ C) は、個々の製品やサービスに応じて細分化される。

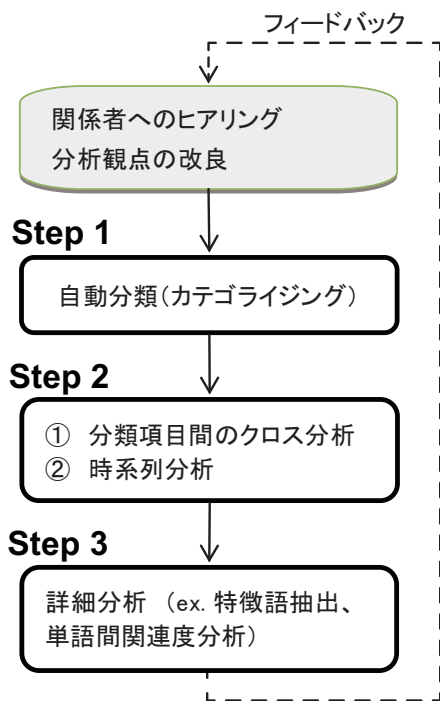


図3 コールログの分析ステップ

コールログの分析の流れを図3に示す。インバウンドのコールログは、まず上記の1.~4.の分析観点毎に分類項目を作成する。分析観点1.は、①~⑧を基に、分析観点4.はA)~C)を基に分類する。また、分析観点2.と3.は製品やサービスに依存した分類項目となる。予め設定した分類項目別に、図1の統計処理技術の「カテゴライジング」の技術を用いて自動分類する(図3のStep1)。

\*1)「COREFIDO」は株式会社沖データの登録商標です。

上記の4つの観点には、「購入前相談」では「修理」は有り得ないというように項目間に依存関係があるが、意外性のある対応結果を見つけたいため分類項目の依存関係は考慮せず独立に自動分類する。

次に、自動分類の結果から、観点別に各項目における増減の変化を時系列で確認したり、分析観点間で掛け合わせて分析する事により、どのようなお客様の困り事がどのような根本原因(真の原因)に紐づいているか、さらにはどのようなお客様の困り事に対してどのような対処を行なっているかを確認する(図3のStep2)。さらに、注目すべきお客様の困り事に対しては、その困り事に対応するコールログを抽出し、出現する単語の係り受け関係をネットワークとして可視化したり、その困り事だけに特徴的に出現する単語や表現を確認したり、コールログを直接確認したりして問題の深堀りを行なう(図3のStep3)。

#### OKI プリンターにおけるコールログ分析

OKIの複合機・プリンター向け新サービスである「COREFIDO3」\*1)は、「メンテナンスバリアフリー設計」と「クラウドメンテナンスプラットフォーム」を採用することによって、オフィスにおけるプリンターメンテナンスのコスト削減を実現した製品である<sup>2)</sup>。

プリンターのメンテナンス性を向上させるためには、お客様が利用する現場で、どのような問題が生じているのかを知り、その対策を講じることが重要である<sup>3)</sup>。我々は、その目的のために、上述のテキストマイニング技術を用いたコールログ分析を行っている。分析の進め方は図3に従う。

Step1の自動分類(カテゴライジング)は、SVM(サポートベクターマシン)という教師有り機械学習の手法を用いた。「教師有り機械学習」とは、既に人間によって分類済みのコールログから、分類の決め手となるワードの出現状況を自動的に学習し、未知のコールログに対しても自動的に分類する手法である。教師となる分類済みのコールログの作成は、最初はランダムに収集したコールログ200件を用いたが、実用に耐えうるとされる分類精度を得るために、現在では、1,000件まで増強したものをを用いて自動分類を行っている。

さらに分析の精度を向上させる試みとして、以下の3点を行っている。

##### (1) 既存テキストからの専門用語辞書の半自動作成

テキストマイニングの基本となる単位は単語である。

分析対象のコールに出現する単語を適切な単位で分割、認識することは自動分類の精度向上に役立つ。我々は、専門用語抽出ツール<sup>4)</sup>を応用して、プリンターマニュアルやコールログから、コールログに出現しやすい用語を抽出し、これらの用語は1単語として認識する。

図4に、プリンターマニュアルから専門用語抽出ツールを用いて抽出した6文字以上の専門用語上位20語を示す。抽出結果には、図4下線の「症状確認項目対処方法参照ページ」のような抽出ミスも含まれるため、目視確認し、抽出ミスは除いて辞書に登録している。

給紙ローラー
定着器ユニット
トナーカートリッジ
操作パネル
ベルトユニット
原稿ガラス
自動原稿送り装置
ネットワーク設定
メニュー項目一覧
メンテナンスユニット
イメージドラムユニット
原稿搬送ローラー
プリンタードライバー
トップカバー
用紙トレイ
メッセージ
お客様相談センター
<u>症状確認項目対処方法参照ページ</u>
原稿トレイ
スキャナー部

図4 プリンターマニュアルからの専門用語抽出結果の例

## (2) 同義表現の吸収

通常、コールログ、つまり、お客様の問合せ内容記録はコールセンターの後処理工程でオペレーターが入力することが多いが、同じ内容でもユーザーやオペレーターによって表現方法は千差万別である。例えば、図4の「自動原稿送り装置」は、ADFと表記されることが多い。また、オペレーターにより入力されたコールログの場合は、入力ミスや送り仮名の揺れも多い。このような表現の揺れを統一するために、プリンターマニュアルやコールログを用いて、同義語辞書を作成し、分析に用いている。

これらの用語集は、(1)の専門用語辞書と同様、コールログ分析用辞書として管理、増強している。

## (3) 自動分類手法の拡張

お客様の1件の問合せにおいて、お客様の問合せ目的や、話題の対象となる製品やサービスは1つとは限らな

い。特に、購入前のお客様は複数の製品やサービスに関する質問を同時にすることが多い。1件の問合せに対して、主となる内容の1つの分類項目を付与するだけでなく、複数の話題があった場合は複数の分類項目を付与する手法も用意し、分析の目的に応じて使い分けている。

## おわりに

インバウンドのコールログを分析するための方法論として、4つの分析観点(1. 問合せの目的、2. 真の理由、3. 真の対象、4. 対応結果)を策定し、この観点に基づく自動分類をベースとした分析ステップを提案した。半自動的に作成した専門用語辞書や同義辞書を利用したり、複数の自動分類手法を使い分けたりすることにより、分析精度や効果を高めている。

提案手法を用いたOKIプリンターのお客様コールログ分析は、現在継続的に行っており、その結果は、新製品の仕様策定やサービスの改善に活かされている。

本稿で紹介した「COREFIDO3」はクラウドメンテナンスプラットフォームを有している。今後は、クラウド内にあるプリンターのエラー情報とコールログ情報を組み合わせたマイニングに取り組む予定である。 ◆◆

## 参考文献

- 1) 特集「自然言語処理技術の現状と展望」、情報処理学会誌, Vol.57 No.1 (2016)
- 2) 前川, 古原, 植田, 大山, 鈴木: A3 LED カラー複合機新 MC8 シリーズ, OKI テクニカルレビュー, No.226 (2015)
- 3) 高橋, 石崎, 岩瀬, 開, 細井: OKI のプリンター / 複合機の操作性向上への取り組み, OKI テクニカルレビュー, No.222 (2013)
- 4) 下畑, 山本: 隣接文字のエントロピーに基づく定型表現の自動抽出, 沖電気研究開発 No.177 (1998)

## 筆者紹介

北村美穂子: Mihoko Kitamura. 情報・技術本部 研究開発センター センシング技術研究開発部  
 村田稔樹: Toshiki Murata. 情報・技術本部 研究開発センター センシング技術研究開発部  
 佐々木美樹: Miki Sasaki. 情報・技術本部 研究開発センター センシング技術研究開発部  
 奥村晃弘: Akihiro Okumura. 情報・技術本部 研究開発センター センシング技術研究開発部