特集号
招待論文

機械学習のビジネス適用事例紹介
―電話オペレータ支援と保険支払査定の事例から―

壁谷 佳典[†1]　坪井 祐太[†1]　吉田 一星[†1]　豊島 浩文[†1]　岡原 勇郎[†1]

[†1] 日本アイ・ビー・エム（株）

機械学習および自然言語処理の技術革新は目覚ましく，企業が蓄積する大量データへのそれらの技術の適用が注目されている．しかしながら，実業務への詳細な適用事例があまり公開されていないために，具体的な課題やメリットを把握できず，適用を躊躇することがあると考えられる．そこで本稿では，2つの具体的な応用例を取り上げ，自然言語を介する業務に機械学習技術を適用する際の考慮点や実際の効果について説明する．

1. はじめに

これまでのITシステムは，明確な要件定義のもとで，基幹業務を確実かつ堅牢に遂行するために構築されてきた．しかし，人でなければ判断が難しいと考えられてきた意思決定の支援に，ITシステムが必要とされるようになってきている．こういった新たな応用に対しては，従来のシステム構築手法では実用的な精度や性能を達成できないことも多かった．しかし，最近の人工知能や機械学習の手法により，実用的なシステムの構築と運用が可能になってきている．機械学習は，大量の訓練データからデータに内在する特徴を学習し，新たに発生したデータに対して学習結果を元にその新しいデータに特有の特徴を検出する技術である．計算機の処理能力が向上し，また大規模データが利用可能になるにつれ，機械学習は画像認識・音声認識・自然言語処理など各種の要素技術での活用にとどまらず，多くの分野へ本格的に応用されている．

本稿では，機械学習が実業務にもたらす効果と機械学習の適用における考慮点を，2つの応用例を取り上げて具体的に示す．

1つは，企業のコールセンターにおける電話オペレータ支援である．電話オペレータは商品やサービスの利用者からの問合せに回答するため，いち早くデータベースから問合せ関連情報を探す必要があり，機械学習はその探索精度向上に用いられる．

もう1つは，保険業界での応用として保険支払査定の自動化である．査定の自動化処理の入力データは，自然文で記述された診断書であるため自然言語処理・機械学習技術の理想的な応用例の1つで，保険の査定業務の中で自然と発生するデータが訓練データとして使用できる．本稿では，第2章で機械学習の概要を述べ，第3章で機械学習を業務に適用するときの考慮点を述べる．そして第4章，第5章では先述した具体的な事例について述べる．訓練データ収集の仕組みづくりから行った事例として，電話オペレータ支援を第4章で述べ，訓練データがすでに十分多く存在して，機械学習の適用が比較的ストレートに行える事例として保険の支払査定の自動化を第5章で述べる．最後に第6章で本稿をまとめる．

2. 機械学習概要

機械学習は，既知のデータ（訓練データ）を用いてデータの特徴を学習したあと，新しいデータが入力されたときに学習の結果を用いて何らかの予測や判定を行う．機械学習は，訓練データの性質によって，教師なし学習と教師あり学習の2つに大別される．教師なし学習は，訓練データの特徴を何らかの形で抽出して出力するが，期待される出力の正解は存在しないので機械学習の出力の意味づけを人間が行う必要がある．これに対し，教師あり学習は，訓練データとして「入力に対して期待される出力が付与されているデータ」を使用し，入力に対して期待される出力を行うためのルールを学習する．教師あり学習を使うことで，ビジネスルールの記述が困難な業務のシステムによる自動化が可能である．また，そもそもビジネスルールの記述が不要なためビジネスルールが変化した場合にもルールの修正作業が発生しないので，システムのメンテナビリティの向上を図ること

ができる．本稿では，教師あり学習の技術を使った業務改善について議論する．

教師あり学習を業務に適用するためには，当然のことながら訓練データが必要となる．しかしながら，実際には訓練データとなるデータが存在しないケースも存在する．たとえば，機械学習を用いてユーザの嗜好に合った商品推薦を行うシステムを新規に開発する場合，訓練データとして利用可能なユーザの商品選択に関する行動ログデータは，まったく存在しない可能性がある．したがって，訓練データとして利用可能なデータを効果的に収集する仕組みや，その仕組みを業務プロセスの中で実施することに対する顧客の理解が，機械学習の業務への適用において重要になると筆者らは考えている．

3. 機械学習の業務適用の考慮点

3.1 機械学習と業務の親和性

本節では，機械学習の利用に適した業務の要件を整理する．業務によっては，機械学習を使わずとも通常のルールベースのシステムで十分もしくはそのほうが効果的・効率的であるケースも多い．本当に機械学習を採用すべきかどうかを判断するため，業務が以下の要件のどれかに該当するかどうかを判断材料とされたい．

1. ルールが複雑すぎる業務

業務オペレーションが複雑すぎて，人がルールを表現できないような業務が該当する．ルールを表現できなければプログラミングすることは非常に困難である．このような業務には機械学習の適用を検討すべきである．ルールは表現可能だが，例外処理が多いためにルールの人手によるメンテナンスが現実的でない，という場合も該当する．

2. ルールの変更が頻繁に発生する業務

ルールが絶えず変更されるような業務が該当する．ルールベースで実装した場合，ルールのメンテナンスにかかるコストが膨大になるだけでなく，頻繁に変更されるルール間の整合性を保つことが困難になる．その企業の状態（提供しているサービス／プロダクトの変化）や，社会的な情勢の変化といった外部要因にビジネスロジックが影響を受けるものが該当する．

3. 長年の経験と高度なスキルが求められる業務

企業にとって専門領域での高度なスキルを持った人材を育成するためには，長い期間と多大なコストが必要である．対象となる業務が高度専門家に依存しており，なおかつマニュアルやルールが可視化されていない場合に機械学習の導入効果は大きい．

3.2 機械学習以外の技術との共存

機械学習を業務に適用する場合に，実現する機能を機械学習のみで実装することが本当に最善であるかを考慮する必要がある．昨今の機械学習に対する期待の大きさから，ややもすると何でも機械学習に処理させようと考えてしまいがちである．筆者らの経験上，ルールベースのシステムと機械学習とをうまく組み合わせることが，現実的に良い解を得る上で重要である．以下，組合せによって得られるメリットを述べる．

1. ルールベースのシステムは，訓練データが存在しないケースにもうまく対応できる．第1章で述べたように，新規サービスの開始時などに訓練データが存在しない場合がしばしばある．機械学習を用いたシステムは訓練データが存在しないと機能しないため，いわゆるコールドスタート問題が生じる．この問題の解決策として，新規サービスを機械学習とルールベースあるいはその他の手法の2つの組合せで実現させて，サービス開始直後は機械学習を用いずにサービスを運用する方法がある．サービスの継続に伴い，訓練データとして利用可能なシステムログデータが蓄積されてきたら，機械学習の学習を行ってサービスの精度をさらに向上させる．事例1では，機械学習と情報検索の2つを組み合わせてこのような運用を実施している．
2. ルールベースのシステムは，業務担当者の知見を簡単にシステムに反映できるというメリットがある．一般に，機械学習は訓練データに依存しており決定的な判断を保証することは難しいが[1]，法規制などの理由で決定性を保証したい場面はあり得る．その場合，確実に例外処理を実行させるために，ルールベースのほうが確実かつ低コストで実現できる．後で述べる事例2では，機械学習とルールベースのシステムを併用している．

4. 事例 1　電話オペレータ支援

本章では，1つ目の事例として，機械学習技術を使った電話オペレータ支援について示す．

4.1 コールセンター業務内容と典型的な課題

本節では，電話オペレータ支援におけるビジネス課題や技術的な課題について整理する．

4.1.1　電話応対時間の短縮化

企業のコールセンターの典型的な課題の1つが，電話応対時間の短縮化である．コールセンターは，エンドユーザからのサービスや商品に関する電話問合せ応対業務を日々行っている．問合せ内容は，新規サービスの申し込みや解約，企業が運営しているWebサイトの操作方法，支店の場所の確認など多岐にわたる．

電話問合せのオペレーションの難易度は，問合せ内容に依存する．問合せ内容が簡単でかつオペレータに一定の熟練度があれば，資料を参照することなく直ちに回答できる．これに対し，比較的難易度の高い問合せや，正確性を問われる問合せに対しては，Webサイトを検索したり，紙の業務マニュアルを確認しながら応対を行う．このような電話応対中の調査・確認の作業負担が回答の遅延につながり，エンドユーザの満足度低下を引き起こす可能性がある．

4.1.2　自然文を扱う難しさ

電話オペレータを支援するような，高いユーザビリティが必要となるシステムは，自然文による入力を処理できることが必須となる．ここでいう自然文とは普段人間が話す話し言葉を指す．しかし，システムに自然文を処理させる際に以下の2つの課題がある．

1. 自然文の表現の多様性への対応

自然文では1つのことを言及するのにさまざまな表現方法がある．たとえば，あるサービスの値段を尋ねる質問として「このサービスの値段を教えてください」「このサービスはいくらですか？」「このサービスは何円ですか？」といったようにさまざまな表現が存在する．システムはこのような多様性に対処することが求められるが，一般にこれらの質問が本質的に同じことを尋ねていることを機械的に解決することは難しい．この課題をLexical Gapと呼ぶ．

2. コンテキストを理解した上での自然文中の重要な表現（特徴）の特定

たとえば「定期預金を始めたい」というクエリ（問合せ）を表す自然文に対して，候補となる回答用のFAQ（Frequently Asked Questions）文書が2つあり，それぞれ「定期預金の開設方法を教えて」と「投資信託を始めてみたい」というタイトルであるとする．クエリに対して，前者は「定期預金」という単語（特徴）がマッチし，後者は「始める」という特徴がマッチする．どちらもマッチしている特徴の数は同じ1だが，人間の場合はこのクエリが持つコンテキストにおいては「定期預金」というキーワードが重要であることが容易に判断できる．そのため前者のFAQ文書がより適切であることが判断できる．しかし，システムにこの正しい判断を行わせることは容易ではない．

4.2　ソリューションとシステム概要

本節では，前節で述べた課題への解決方法の1つである電話オペレータ支援システムについて述べる．最初に電話オペレータ支援システムの仕組みの概要を述べ，次に課題への具体的な対応方法を述べる．

4.2.1　システム概要

電話オペレータ支援システムは，電話オペレータに必要な情報を，タイムリーに提供できるシステムとして設計されている．以下，図1で挙げる電話オペレータ支援システムの処理の流れを述べる．電話オペレータ支援システムへの入力は，エンドユーザと電話オペレータの通話である．通話は音声データとして電話オペレータ支援システムに送信される．その音声データは，支援システム内の音声認識エンジンによってテキ

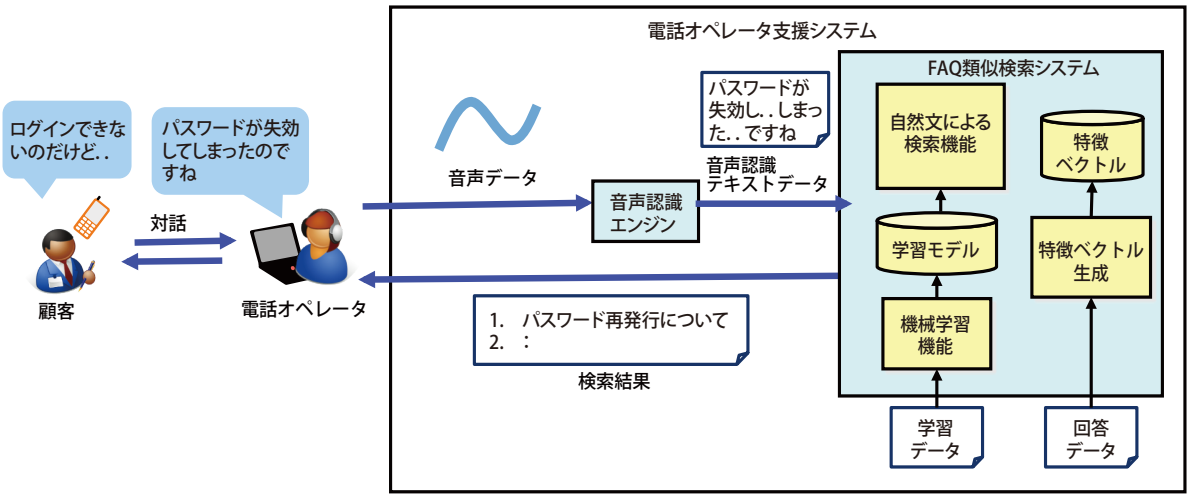

図1　電話オペレータ支援システム

ストデータに変換される．変換されたテキストデータは，自然文をクエリとして受理する検索エンジンに入力され，検索エンジンは最も通話と関連があると判断された回答文書を，電話オペレータにランキング付きで返す．電話オペレータは，現在通話で応対中の問合せ内容に合致した回答文書へのリンクをシステムのユーザインタフェース上でクリックして，必要な情報を入手する．図2では対話と検索処理のタイミングのイメージを表している．電話オペレータとエンドユーザの通話が行われている間，絶えず検索処理は実行される．

4.2.2　各課題への取り組み

顧客と電話オペレータの通話を直接システムに送信するアーキテクチャとなっている理由は，4.1.1節で述べた電話オペレータの作業短縮のためである．この仕組みにより，電話オペレータはキーワードなど通話から得られる情報を画面から能動的に入力する必要がなく，通話が進むにつれて回答文書を自動的に取得することが可能となる．音声認識エンジンとして，IBM Watson APIが提供する機能のひとつであるWatson STT（Speech To Text）[2]を採用している．

もう1つの課題である自然文の扱いに対し，自然文をクエリとして文書検索を行うFAQ類似検索システムを採用している．FAQ類似検索システムは日本アイ・ビー・エム（株）東京基礎研究所で開発された，機械学習による検索精度向上の機能を備えた自然文類似検索システムである．FAQ類似検索システムには，カテゴリ学習・ランキング学習・クリックログ学習という3種類の異なる機械学習のコンポーネントが実装されている．これ以降は，それらの類似検索に果たす役割についてそれぞれ説明する．

1．カテゴリ学習機能

カテゴリ学習は，4.1.2節で述べたLexical Gapの課題を克服するために用意された自然文の分類器である．これを，入力されるクエリおよび回答文書の分類に適用して，クエリと回答データ間の意味的な近さを判別する．たとえば，クエリと回答文書が同じカテゴリに分類されたときに，回答文書に対して検索スコアに加点するようにする．この結果，たとえば「値段を教えてください」と「いくらですか」のように文字列上の一致がないクエリと回答文書の組に対しても，適切な検索スコア計算ができる可能性がある．

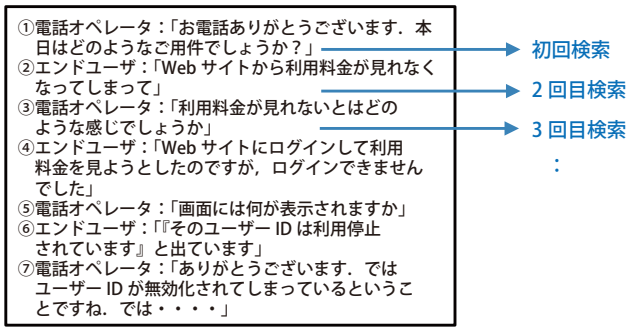

図2　架空の通話と検索処理

カテゴリ学習を実施するためには，クエリを表す自然文に対してカテゴリ（フラグ）が付与された訓練データが必要になる．電話オペレータの通話記録が，しばしばこの目的に利用可能である．コールセンターの典型的な運用では，電話オペレータが業務の一環として通話記録（メモ）を残す．その際，オペレータが通話内容をあらかじめ決められたカテゴリに仕分けして記録していることが一般的である．この場合は，通話記録をそのままカテゴリ学習の学習データとして利用することができる．

2．ランキング学習機能

ランキング学習は，4.1.2節で述べた重要な特徴を特定するための機械学習コンポーネントである．ランキング学習は，訓練データとしてクエリとそのクエリに対して適合する文書とのペアを入力として，クエリと文書との適合ペアを正例，適合しないペア（ランダムにクエリと文書を選ぶことで生成可能）を負例とみなし，ランキング算出の際に用いる各特徴の重みを分類問題に帰着させて学習する．学習の結果得られた重みを用いて回答文書のスコアリングを行うことにより，文書ランキングの精度を向上させることができる．

3．クリックログ学習機能

クリックログ学習も，ランキング学習と同様に特徴の重みを適正化する機械学習コンポーネントである．ランキング学習とは異なり，クリックログ学習は，クエリに含まれる特徴と回答文書そのものとの関連性の強さを学習する．たとえばクエリに「手数料」という単語（特徴）が含まれていて，検索結果の中から「手数料はいくらになりますか」という回答文書が選ばれてクリックされた場合，クリックログ学習はこのキーワードと回答文書との関連性が強いことを学習し，この回答文書に対してのみ「手数料」の単語の重みを大きくする．その結果，次のクエリにクリックログで学習した特徴が含まれていた場合，その特徴に関してクリックログに記録された回答文書が優先的にユーザに返されるようになる．

4.3　データがない，それが問題だ

4.2節で述べた機械学習コンポーネントを利用して類似検索の精度を向上させるためには，学習データ，つまり問合せへの応対に有益であった回答文書付きの通話データ（以後，正解付き通話データと呼ぶ）が必要になる．本節では，正解付き通話データを業務プロセスの中で自然に収集できるようにするための考慮点を2つ述べる．

1．ユーザビリティの高いインタフェース

正解付き通話データを集めるためには，電話オペレータにクリックされた個々の回答文書が，問合せ応対の役に立ったかどうかをフィードバックしてもらう必要がある．クリックの有無だけを用いて有益性を判断するモデルも研究されている[3],[4]が，実際のフィードバック以上に確実なものはない．電話オペレータが快適にフィードバックの入力を行えるようにするため，なるべく作業負荷を与えない画面設計が望ましい．電話オペレータ支援システムでは，電話オペレータの通話完了後に通話中でクリックされた回答文書の一覧を画面に提示し，実際に有用だった回答文書にチェックしてもらうインタフェースを用意した．その結果，電話オペレータは通話後に落ち着いてフィードバックを入力することができる．学習データとは直接関係ないが，自由コメント欄も用意し，電話オペレータの生の意見を汲み取れる設計となっている．

2．電話オペレータの協力体制の整備

システムに対する期待の高さから，最初から賢いシステムが提供されると誤解されることも少なくない．その結果，学習データが不十分もしくはまったくない状態での検索性能と期待とが乖離し，電話オペレータのフィードバックが得られず，学習データが蓄積されないという負のスパイラルに陥る可能性がある．したがって，機械学習は学習データがあって初めて有効に機能することを理解してもらい，積極的なフィードバックの入力を促すべきである．企業内のシステムはインターネット上の不特定多数向けサービスに比べて，トランザクション数が極端に少ないにもかかわらず要求される精度が高く，かつ短い期間で開発・運用に漕ぎ着けることが求められる．そのため，業務の一環としてフィードバックを入力する運用体制を確立してもらう工夫をすることが望ましい．

4.4　効果と考察

この節では，電話オペレータからのフィードバックによって作成された学習データを用いた学習の実システムの精度改善効果について述べる．

4.4.1　評価対象および学習データの作成方法

検証対象の回答データは顧客サイトのFAQを用いた．学習データは電話オペレータに運用してもらい，フィードバックを得ることで生成された正解付き通話データを使用した．収集されたデータを取得期間の違いによって2つに分け，その一部を評価データ，残りを学習データとした．学習による効果を確認するために，すべての学習コンポーネント（ランキング学習／カテゴリ学習／クリックログ学習）を利用したケース，学習コンポーネントを2つずつ利用したケース（ランキング学習／カテゴリ学習のみ，ランキング学習／クリック学習のみ，クリック学習／カテゴリ学習のみ），学習コンポーネントをまったく利用しないケース，のそれぞれに対して検索精度を評価した．

4.4.2　評価結果

評価した結果を図3に記す．図3には，適用した学習の種類およびそのときの回答文書の再現率（Recall）を記している．再現率の@Nは，全評価データに対して検索ランキングの順位がN位以内に入った割合を示す．たとえば評価データが10件あり，そのうち正解の順位が1位のものが4件あった場合は，@1の値は40%となる．

結果から，いずれの学習も再現率の向上に貢献している．最も貢献度が高いのがランキング学習であり，すべての@Nに対して10ポイントから20ポイントに近い改善効果が確認された．

4.5　今後の課題

実業務で生じるデータを用いて，機械学習による精度改善を実現することができた．それにより，オペレータ支援システムが実業務の支援に有用であることが確認でき，本番環境で利用されるに至った．本節では，

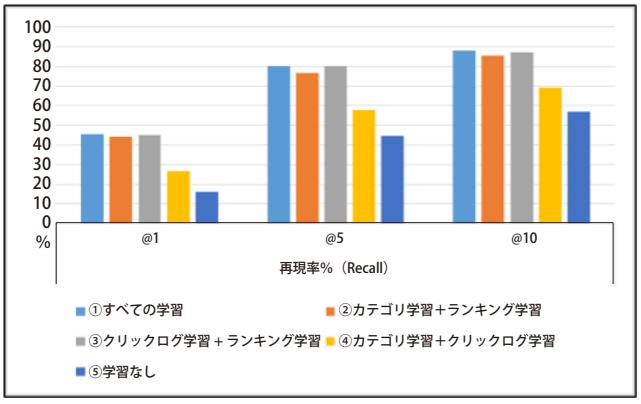

図3　学習の利用と再現率の関係

現在未解決の課題を整理する．

4.5.1　電話オペレータへレスポンスのタイミング

電話オペレータへの検索結果の表示タイミングを適切に決定することは，通話データがリアルタイムに流れてくる状況では自明ではない．業務上，電話オペレータは通話のなるべく早い段階で正しい回答文書が提示されることを期待するのは自然であろう．ここでは図2の例を用いて説明する．図での会話における問合せの核心は⑥⑦の発話にある．したがって，検索システムが適切な回答文書を発見できるタイミングが⑥⑦以降である可能性が高い．しかし，これより早いタイミングで適切な回答文書を返すことができれば，ユーザの待ち時間短縮の観点で望ましい．たとえば④もしくは②の時点でシステムが適切な回答データを返すことが望ましく，有効な施策を検討中である．

4.5.2　フィードバックの取得

学習データは，電話オペレータが回答文書をクリックし，フィードバックを返すことで生成される．この方式の課題は，フィードバックを得るには正解となる回答文書が電話オペレータに視認されることが前提となっていることである．たとえば，画面上に表示できる回答データが10件の場合，クエリに対する真の正解回答文書が11位以降の順位だった場合は何らフィードバックを得ることができない．この問題を緩和するため，ランキングにある種の“揺らぎ”を入れるなどして低いランキングの回答文書にも上位にくる可能性を持たせる仕組みが考えられる．たとえばContextual Banditアルゴリズム[5]に見られるように各特徴に対する重みを確率的に与えることで，ランキングを確率的に変化させるなどといった方法が考えられる．

5.　事例2　保険支払査定の自動化

本章では保険会社における保険支払査定の自動化における機械学習の応用について解説する．

5.1　支払査定業務内容と典型的な課題

支払査定とは，保険会社に対して被保険者からの保険金請求に対して支払内容を決定するための審査のことであり，保険会社における基幹業務の1つである．保険請求に対して適切に保険金を過不足なく支払うことは非常に重要である．補償内容以上に保険金を支払ってしまうことは保険会社の利益を圧迫する一方で，補償された内容に対して十分な支払を行わないことは被保険者にとって不信感を招く．過去には保険金の不適切な不払いが社会問題化したこともある[6]．

支払査定では，病院から提供される診断書などを参照し補償内容や支払内容を決定する必要があり，幅広い医療知識が必要とされる．また，日本の保険の補償内容は選択可能な多数の特約オプションなど多岐にわたるため，支払査定を行うには保険商品の約款にも精通している必要がある．さらに，業務の性格上査定の正確さがきわめて重要であるため，多くの保険会社は1件の査定に対して複数人が2重3重のチェックを行っており，人的負担が大きい．一方で，支払査定を早く行うことで保険金支払までにかかる時間を短縮することができるため，顧客サービス向上の観点では迅速さも同時に求められる業務である．

5.2　機械学習の適用方法

先に述べたように，支払査定は複数人で多重に行われていることが多い．最初の査定をシステムが行い，システムの査定結果を人がチェックする業務フローに変更することにより，人的負担の軽減，査定の迅速化，正確性の向上が実現できた（図4）．

保険会社には，過去に数十万から数千万件の大規模な保険金請求のデータとそのデータに対する人が行った査定の結果が存在している．このデータを利用すると，保険金請求情報を入力とし査定結果を教師信号として機械学習を適用することが可能である．機械学習アルゴリズムの入力となる特徴量として，保険金請求情報内の「性別･年齢･入院期間などの定型項目」と「傷病名や経過欄などに書かれた自然文に自然言語処理を施して得られる単語や単語間の関係（係り受け）」[7]を用いた．予測対象は支払を決定づける傷病コードや手術コードなどのカテゴリであり，構造化文書へのマルチラベル分類☆1問題として扱った．

また，2重に行っていた査定業務の片方を自動化することで，後続の人手による査定の結果を新しい訓練

☆1　1つ以上のラベルを同時に付与する問題設定．

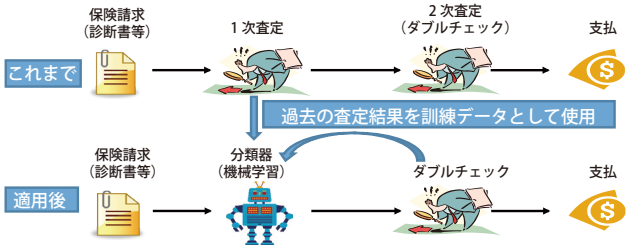

図4　保険査定業務への機械学習適用前後の比較

データとして使用できるため，継続的にシステムを改善し続けることが可能となる．支払査定は，保険会社において通常業務として行われているものであるから，通常多大なの労力を要する訓練データ作成の業務を新たに定義することなく大量の訓練データを継続的に得ることができる．この点において，支払査定は機械学習に非常に適した業務適用形態であると考えられる．

100万件規模の訓練データを用いることで，適合率・再現率とも実用レベルの学習結果を得ることができた．高い性能が実現できた理由として，大規模の訓練データが利用可能であることに加えて，支払査定は知識のある専門家の間では，主観に依存せずそれほど判断に違いが出にくい内容であり，正解データとしての質が良いため学習しやすかったことも考えられる．

また，定期的な更新を効率的に行うために，学習アルゴリズムにオンライン学習アルゴリズムを用いて，訓練データの増加分だけを用いて学習するシステムを構築した．

一方，実業務に適用する際には機械学習結果のコントロールが困難である点が課題となった．保険の支払査定においては，法令遵守や監督省庁からの指導対応のために特定のケースにおいて決められた判断をする必要がある．しかし，機械学習結果は訓練データに依存しているため，訓練データの変更に伴って意図しない予測をするようになることはあり得る．また，誤認識を想定できたとしても，これまでうまくいっていたデータに対しての結果に副作用なしに誤認識だけを修正することは容易ではない．訓練データや特徴量を調整することで特定のデータに対する誤認識を修正できたとしても，少し異なるデータがきた場合や，訓練データがさらに追加されたときにも常に同じ挙動を保証できるものではない．

この課題に対応するために，ルールベースと機械学習ベースの2つの仕組みを統合した構成で運用することとした（図5）．この構成の重要な点は，ルールが適合した事例に対してはルールを優先することである．ルールを優先することで，人手による修正が必ずシステムに反映されることが保証できる．そして，ルールがカバーできない事例を機械学習した分類器で対応する構成になっている．訓練データが多い場合には平均的には機械学習だけを使った方が良い分類精度を期待できるが，前述したように業務的に決定的な分類結果が得られることが非常に重要であるため，ルールによるコントロールを優先した構成になっている．なお，ルールだけを用いたシステムよりも提案する統合構成の方が分類性能が高くなることが確認されている．

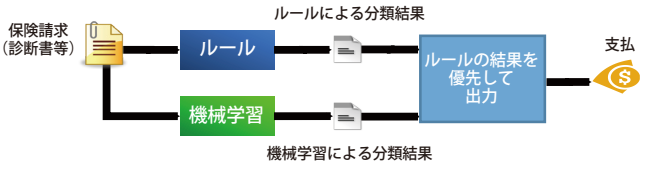

図5　機械学習よりルールを優先した分類システム

5.3　支払査定事例まとめ

本節では，保険の支払査定における機械学習の適用事例を紹介した．ここで紹介した仕組みは，人手で書類を読み何らかの判断をしているほかの業務にも適用可能であるが，保険の支払査定システムのように機械学習を効果的に利用するための性質として，1）業務の中で自然に訓練データが作られること，2）判断には専門知識が必要だが専門家の間では判断が一貫していること，および3）複数人で判断を確認する業務であること，があげられる．1）の性質により，大規模訓練データが追加のコストなしで得られる．2）の性質により，人的負荷の高い作業を高い分類性能のシステムで置き換えることができ費用対効果が高い．3）の性質により，複数人での判断の一部をシステムに置き換えることで人的負荷を下げられるとともに，人手で判断したデータを使って継続的にシステムを改善するための訓練データを得ることができる．

また，機械学習の応用では機械学習の出力に人が介入する仕組みの重要性も再度強調したい．保険の支払査定システムでは，ルールを優先した構成とすることで，業務要件を満たすことが可能になった．出力結果の調整のために人が介入できないことが機械学習を適用するための障害となっているケースがあり，ルールを優先した構成は最適ではないが妥当な解の1つであると考える．

6.　おわりに

本稿では，2つの事例を用いて機械学習技術の実業務適用における考慮点をまとめた．ここで述べた事例は機械学習技術を利用していることを除けば，従来のITを使った業務プロセス改善プロジェクトと同じである．その観点においては従来同様ユーザとの協力・理解がプロジェクト成否に大きく関係する．従来のプロジェクトと異なる点は，機械学習の精度がプロジェクト成

否の重要なファクタとなっているため，その精度をプロジェクトの初期段階で素早く検証すること，あるいは検証に特化したPOC（Proof Of Concept）プロジェクトを別途実施すること，の重要性が高いことである．昨今のクラウドサービスの成熟により，プロジェクトのスモールスタートが容易になってきていることも機械学習の活用の幅を広げるのに役立ってきていると考えている．本稿が今後の機械学習を使った業務改善の一助になれば幸いである．

参考文献

1）Sculley, D., Holt, G., Golovin, D., Davydov, E., Phillips, T., Ebner, D., Chaudhary, V. and Young, M.: Machine Learning: The High Interest Credit Card of Technical Debt, SE4ML: Software Engineering for Machine Learning, NIPS 2014 Workshop (2014).

2）Speech to Text Watson Developer Cloud，http://www.ibm.com/smarterplanet/us/en/ibmwatson/developercloud/speech-to-text.html（2016 年 5 月現在）

3）Craswell, N., Zoeter, O., Taylor, M. and Ramsey, B: An Experimental Comparison of Click Position-bias Models, In WSDM '08: Proceedings of the First ACM International Conference on Web Search and Data Mining, pp.87-94 (2008).

4）Fan, G., Chao, L. and Wang, Yi-Min.: Efficient Multiple-click Models in Web Search, In Proceedings of the 2nd ACM International Conference on Web Search and Data Mining, pp.124-131, 2009b (2009).

5）Chapelle, O. and Li, L: An Empirical Evaluation of Thompson Sampling, NIPS (2011).

6）鳳佳世子：保険金不払い問題の概要と課題，No. 572，調査と情報 (2007).

7）坪井祐太：機械学習による自然言語処理－言葉を扱う技術とビッグデータの接点，No. 83，ProVISION (2014).

壁谷 佳典（非会員）yokabeya@jp.ibm.com
2003 年東京大学大学院理学系研究科修士課程修了．同年日本アイ・ビー・エム（株）入社．

坪井 祐太（正会員）yutat@jp.ibm.com
2002 年日本アイ・ビー・エム（株）入社．2009 年奈良先端科学技術大学院大学博士後期課程修了．言語処理学会会員．博士（工学）．

吉田 一星（正会員）issei@jp.ibm.com
2001 年東京大学大学院数理科学研究科修士課程修了．同年日本アイ・ビー・エム（株）入社．

豊島 浩文（非会員）toyos@jp.ibm.com
1985 年日本アイ・ビー・エム（株）入社．プロジェクトマネジメント学会会員

岡原 勇郎（非会員）okahara@jp.ibm.com
1985 年日本アイ・ビー・エム（株）入社．

採録決定：2016 年 6 月 23 日
編集担当：上條浩一（日本アイ・ビー・エム（株））