

ピープルアナリティクスによるメンタル休職の未然防止支援

技術革新に伴うコンピュータハードウェアの性能向上、ストレージの低価格化など、ディープラーニングを中心としたAIの第三次ブームの流れを受け、ビッグデータの蓄積とAIによるデータ分析のビジネス活用が進んでいる。人財分野においても、従業員データの分析（ピープルアナリティクス）の経営への活用に期待が集まっている。

日立は、人事総合ソリューション「リシテア」の事業で培った人財分野の業務知識とリシテアに蓄積された人財データに、AIの技術を組み合わせ、こうした市場の期待に応えていく。本稿では、ピープルアナリティクスによる健康経営に寄与するソリューションの商品化について紹介する。

盛井 恒男 | Morii Tsuneo

平山 純 | Hirayama Jun

1. はじめに

就業管理システムとして提供を開始した人事総合ソリューション「リシテア」は、タレントマネジメントや人事・給与管理などサービス範囲を拡大し、現在、導入企業は約1,200社、利用ユーザーは164万人に達している。近年、従業員データの分析（ピープルアナリティクス）が注目されているが、これまでは分析対象データの収集が困難なため実行に移せない企業も多かった。そこで、「リシテア」に蓄積されたデータのピープルアナリティクスへの活用を検討し、2017年2月に同ソリューションとAI（Artificial Intelligence）を組み合わせた「リシテア/AI分析」をリリースした。

2. 人事の課題

2.1

人事の声

市場の調査報告によると、ピープルアナリティクスに対する企業の期待は、生産性向上や効率向上に集中していた。リシテアユーザーを対象として実施したヒアリングでも、生産性向上に言及する企業が多く見られた。一方、人事部門における現場の対策として、メンタルヘルス疾患を原因とする休職の防止や、若手や優秀者の離職防止への活用を挙げる企業も少なくなかった。事実、メンタルヘルス疾患による休職は、復職しても作業効率が上がらない状態が続いたり、再発により休職を繰り返したりするケースが多く、生産性・効率向上の対極にある大きな問題となっている。

厚生労働省が実施した「平成28年 労働安全衛生調査(実態調査)」によると、産業別のメンタルヘルス不調による休職労働者の割合は、情報通信業が1.2%とトップであり、金融・保険業が1.0%と続いている。国内全体の平均が0.4%であるため、これらの2業種は群を抜いていることが分かる(図1参照)。内閣府の報告によれば、こうしたメンタルヘルス不調に起因する休職(以下、「メンタル休職」と記す。)によって発生する企業の負担は、従業員1人の休職当たり422万円に上るといふ。

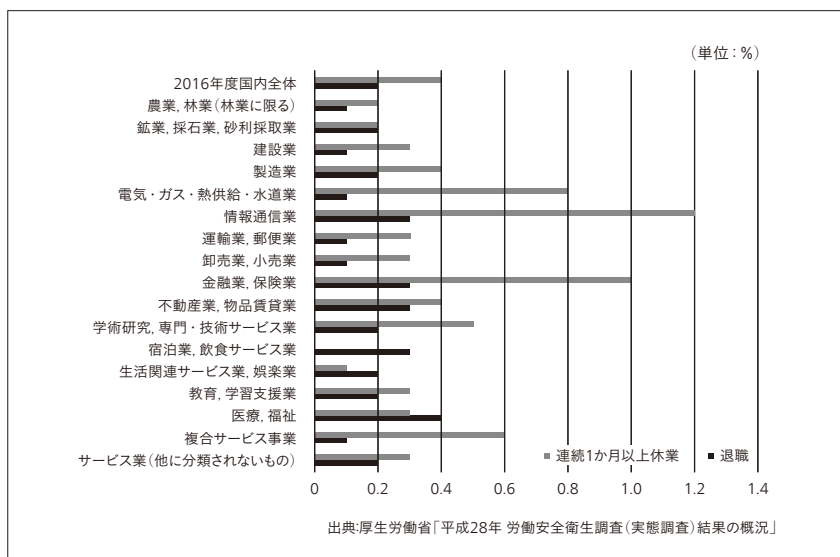
2.2

手を出せなかった未然防止

人事の現場ではメンタル休職の問題にどのように対応しているのか、株式会社日立ソリューションズの安全衛生グループの例を紹介する。日立ソリューションズのメンタル休職率はこれまでおよそ1%前後で推移してきたが、2016年度には0.77%にまで減らすことができた。最も力を入れて取り組んできたのは、復職の支援であった。一旦メンタルヘルス疾患により休職に陥った従業員は、就業規則にて定められた休職期間のうちに治療して元の職場に戻ることをめざすのだが、当人の焦りもあって十分回復しないうちに復職し、1年以内に再発するケースが多いという。2016年のメンタル休職率低下の背景には、生活のリズムをつくる復職準備期間を設けるなど、復職後の再発を減らす努力があった。さらなるメンタル休職率低下のためには未然防止の施策に着手する必要があるが、何を根拠に、どのように動けばよいか分からないのが実状であり、それこそが人事部門の長年の課題となっていた。そこで、ピープルアナリティクス活用の最初のターゲットを、メンタル休職の未然防止と設定した。

図1 | メンタルヘルス不調により
休業・退職した労働者の産業別割合

メンタルヘルス不調により休業・退職した労働者の割合を産業別に示す。



3. 人工知能による未来予測

3.1

データの収集～人事視点の想定原因とデータ

メンタル休職の未然防止策として、ピープルアナリティクスをどのように活用するか。第一の目標は、過去にメンタル休職した従業員のデータを正解データとした機械学習によって、全従業員の未来の休職可能性を予測することであった。機械学習の予測の精度は、特徴量の抽出によるところが大きい。したがって、まず、現場でメンタル休職の問題に関わってきた産業医の経験に基づき、メンタルヘルス疾患発症の要因の収集を実施した。何がメンタル休職につながったのか、どのような環境や条件がメンタルヘルス疾患を発症させたのかをブレインストーミングし、洗い出した。その結果を図2に示す。これを基に、各要因を測りうるデータの有無、入手可否について人事部門および産業医に確認・検討を依頼し、データ収集を行った。

3.2

分析アルゴリズム

収集したデータの分析のアルゴリズムには、複数の候補が挙がった。各アルゴリズムの概要と特徴を表1に示す。

予測の精度は重要であるが、この予測の目的はメンタルヘルス疾患発症の可能性が高い従業員が検出された場合、対象者をケアすることで発症に至らしめないことである。予測の精度が高くても、具体的にどの要素が対象

図2|メンタルヘルス疾患の発症要因についてのブレインストーミング

産業医を交えたブレインストーミングによって、何がメンタル休職につながったのか、どんな環境や条件がメンタルヘルス疾患を引き起こしたのかを洗い出した。

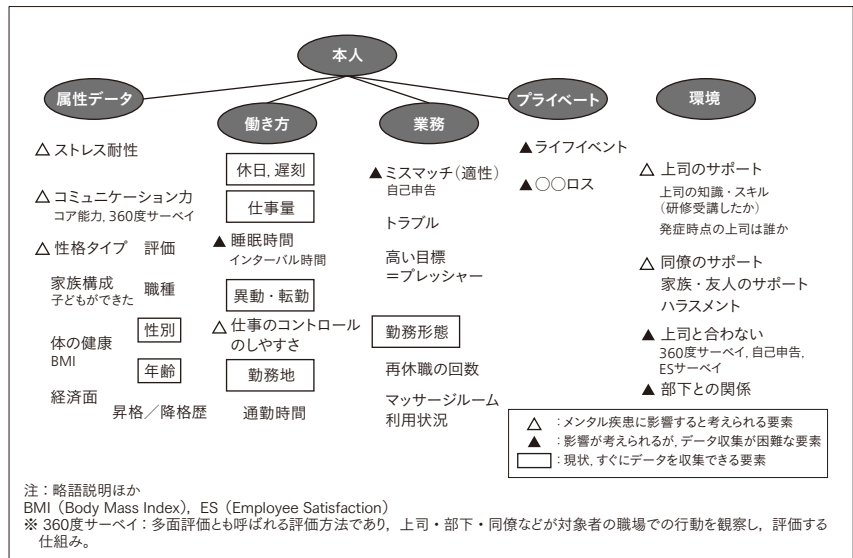


表1|機械学習アルゴリズム

収集したデータの分析に用いるアルゴリズムの候補一覧を示す。

アルゴリズム	概要	確率予測	要因把握
ロジスティック回帰	回帰によって分類する手法である。線形回帰の出力をロジット関数に入力することで、2分類問題に対応した回帰モデルをいう。統計学の分類に主に使われ、医学や社会科学でもよく使われている。	○	○
ランダムフォレスト	決定木を大量に生成し、各決定木の結果を集計して予測する手法である。各決定木は独立しており、説明変数からのサンプリングまたは学習データからのサンプリングによって、異なる特性を持つように学習する。競技系結果予測に活用されている。	○	○
勾配ブースティング	決定木を大量に生成し、各決定木の結果を集計して予測する手法である。決定木を逐次的に増やしていき、生成済みの決定木が間違えてしまうケースのラベルを更新して、新たな決定木を生成していくイメージの手法であり、最も人気がある。	○	○
サポートベクターマシン	回帰によって分類する手法である。マージン最大化を取り入れることで、少ないデータでも汎（はん）化性能が高い2分類回帰モデルを実現する。ただし学習時間は長くなる。認識性能が優れた学習モデルの1つである。	○	×
ニューラルネットワーク	脳機能に見られるいくつかの特性を計算機上のシミュレーションによって表現することをめざした数学モデルである。画像や統計など多次元量のデータで線形分離不可能な問題に対して、比較的小さい計算量で良好な解を得られることが多い。画像認識、市場における顧客データに基づく購入物の類推などに応用されている。	○	×
ディープラーニング	人間が自然に行うタスクをコンピュータに学習させる機械学習の手法の1つである。ニューラルネットワークを多層にして用いることで、データに含まれる特徴を段階的により深く学習することを可能にしたものをいう。	○	×

者のメンタルヘルス疾患発症の可能性を引き上げているのかが分からなければ、対策を立てることができないため、有益な予測とはいえない。したがって、予測のプロセスにおいて各要因の関与が完全にブラックボックスとなるものは避け、要因把握が可能なアルゴリズムの中で実際に試行し、最も精度が高かった「勾配ブースティング」を採用することにした。

4. 分析・予測

4.1

休職者の特徴～統計分析による特徴量の絞り込み

収集したデータは、欠損値の補填あるいは欠損値のあるメンバー自体を分析対象から外すことで整形し、統計

分析により休職者と非休職者間で差異の明確な変数を抽出した。その後、正規化により抽出した変数の相互の分散をそろえるなど、抽象的データ構造に変換し、次元削減・次元圧縮によりさらに変数の削減・統合を実施して特徴量を絞り込んだ。

4.2

データの増量

メンタル休職の予測に使用するデータには、従業員の労働環境や就業データが含まれるため、所属企業に依存した特徴量が抽出される。したがって予測精度の向上には、企業ごとに、当該企業内のデータに限定して分析することが有効だと考えた。しかし、機械学習の実施に際しては、データ量が問題となってくる。例えば従業員数1万人の企業の場合、メンタル休職者は年間100人程度

であり、規模が小さくなればそれだけ正解データが減ることになる。過去に遡ってメンタル休職者のデータを集めることもできるが、昨今の働き方改革の波を受けて従業員の就業状況は大きく変わりつつあり、あまりに古いデータを使うことで精度が悪化する可能性もある。

分析データには、残業時間や休暇日数などのように時系列で変動するデータと、評価結果やアンケートのように年に1~2回しか得られない点過程データがある。時系列データは分析対象期間を規定し、得られたデータの中で期間をずらして使用することで、同一人物から複数パターンのデータを切り出し、それに点過程データを組み合わせることで複数のデータセットを作ること、データ量を増やすこととした。

4.3

結果・精度

前述した手法で整備したデータを学習させて予測モデルを作成し、モデル作成に使わなかったデータを当てはめて予測を実施して、実際にメンタル休職した従業員をどこまで当てることができるかを評価した。その一例を紹介する。

この例では、時系列データは分析対象期間を6か月とし、1人の従業員のデータを1か月ずつ過去にずらすことで6人分のデータに増量して使用している。またストレスチェックの回答結果も、設問内容は伏せ、例えば「設問1の回答値が3」という単純な数字の羅列として使用した。この予測モデルに、予測対象期間のデータを与えた結果として、134人がメンタル休職の可能性が高いと判定され、その内の43人が実際にメンタル休職者であったという結果が得られた。予測モデルは、例えば「2017

年1月に休職する確率」を算出するが、予測月以降1年以内に休職した従業員が予測結果に含まれていれば的中と判定している。

予測的中率は平均すると30%程度ではあるが、実際のメンタル休職率は約1%であることを踏まえると、予測結果の上位約100人の中に30人前後の休職者を含めることができたのは評価できる結果であり、早期ケアの運用面でも十分実用に値するレベルにあると考えられる。

5. リシテア/AI分析 ～組織ストレス予測サービス

5.1

組織ストレス予測サービス

前章に示したデータ加工手法と機械学習によって、過去のメンタル休職者に近い従業員を予測するソリューションを、「リシテア/AI分析」として商品化した(図3参照)。このソリューションは、2つのサービスにより構成されている。1つは、顧客データの分析により予測モデルを作成し、予測精度を評価するとともに、メンタル休職した従業員の特徴(メンタルヘルス疾患に影響の深い要素・条件)を明確にすることで、どこに手を打てばよいかを示すコンサルティングであり、もう1つは、最新のデータから休職する可能性の高い従業員を予測モデルによって毎月予測し、早期ケアを促す運用サービスである。

このうち運用サービスについては、精度向上のため、月ごとに新たに収集した最新のデータを加えて、直近の働き方や労働環境を反映したモデルを作り直して予測する仕組みをシステム化した。これにより2か月後に休職

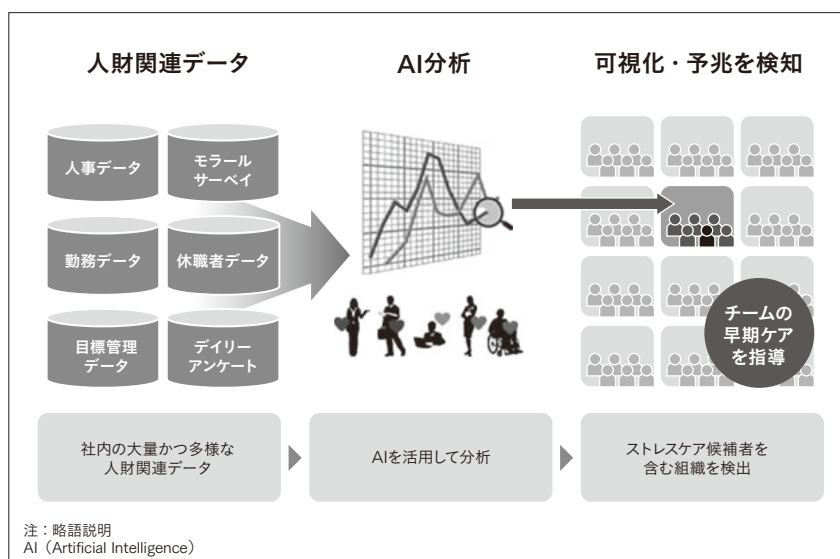


図3 | リシテア/AI分析の概要

リシテア/AI分析のソリューションイメージを示す。

する可能性のある従業員を毎月、最新のデータに基づき予測できる。

5.2

プライバシー保護への配慮

4.3節の事例ではストレスチェックの回答値を使用しているが、これは産業医からの依頼を受けて日立ソリューションズが分析・活用を代行したものであり、産業医が許可されている範囲でのデータ活用に限定されている。基本的に、分析対象データは、分析に使用することについて従業員の同意を得ていることが必要である。また同意を得たうえで分析であっても、その結果を「メンタル休職の可能性の高い人」として開示することは、該当者の処遇や評価に影響を及ぼすおそれがあるため、結果を誰に開示するかについても本人の同意を得るなど、十分な配慮が必要である。

リシテア/AI分析の運用サービスでは、顧客の状況により情報を開示する対象を制御できる仕組みを用意している。

5.3

結果の開示とアクションの促進

リシテア/AI分析は、メンタル休職の可能性の高い従業員に早期ケアを促すものである。ケアの実施者は従業員の業務に直接関わっているマネージャーが適任であり、予測の結果をマネージャーに開示できなければケアに結びつかない。日立ソリューションズでは2017年10月よりリシテア/AI分析の実運用を開始しており、現行商品は同様の運用を標準機能としている（図4参照）。

日立ソリューションズのデータ分析によるメンタル休

図4 | リシテア/AI分析運用画面例

予測の結果に応じて、「健康アラート」、「働き方ランキング」といった各種分析結果を表示する。



職者の特徴は、有給休暇取得日数、遅刻回数（始業・定刻後の出勤）、休日出勤回数などにあった。これらをメンタルヘルス疾患に影響の深い要素と捉え、個々の変数に見える化してマネージャーに「生活リズムの乱れ」として開示してアクションを促すこととした。例えば、有給休暇取得日数は休職者の方が多い傾向にあるため、組織内の取得日数の多い従業員のランキングを示し、マネージャーにはこれら従業員へのコミュニケーションの機会を増やして体調や生活環境などの確認を促すのである。

6. おわりに

メンタルヘルス疾患の発症要因はさまざまだが、実際に収集できるデータには限りがあり、前述の事例でも属性データ（ストレス耐性、性格タイプなど）、プライベートデータ（〇〇ロス、ライフイベントなど）は分析に使うことができていない。機械学習の精度向上のためには効果的な特徴量を収集することが不可欠であり、そのために例えば適性検査データ、サーベイデータ、センサーデータなどを定期的に収集・活用することで、さらなる精度向上が期待できる。一方で、これらはプライバシー保護の観点からセンシティブなデータでもあるため、メンタル休職の未然防止は、アナリティクスの進歩だけで解決できるものではないことを認識し、労使相互の理解と信頼のうえで、人事部門、現場マネージャー、産業医の連携と新たな運用の仕組みを検討していく必要がある。

参考文献

- 1) 内閣府男女共同参画局：企業が仕事と生活の調和に取り組むメリット、男女共同参画会議・仕事と生活の調和（ワーク・ライフ・バランス）に関する専門調査会（2008.4）
- 2) 厚生労働省：平成28年 労働安全衛生調査（実態調査）結果の概況、政府統計（2017.9）
- 3) PwCコンサルティング合同会社、ProFuture株式会社：ピープルアナリティクスサーベイ2017調査報告書（2017.11）

執筆者紹介



盛井 恒男

株式会社日立ソリューションズ スマートライフソリューション事業部 ライフスタイルイノベーション本部 HRテクノロジーセンタ 所属
現在、リシテア/AI分析のブラッシュアップと新たなソリューションの企画・開発に従事



平山 純

株式会社日立ソリューションズ スマートライフソリューション事業部 ライフスタイルイノベーション本部 HRテクノロジーセンタ 所属
現在、リシテア/AI分析など新たなソリューションの企画・開発に従事