

一般口演

一般口演26

機械学習・HISデータ解析

2017年11月23日(木) 09:15 ~ 10:45 F会場 (10F 会議室1004-1005)

[4-F-1-OP26-5] 機械学習を用いた糖尿病外来患者の血糖管理不良予測

倉沢 央^{1,2}, 林 勝義^{1,2}, 藤野 昭典², 芳賀 恒之², 脇 嘉代¹, 野口 貴史^{1,3}, 大江 和彦¹ (1.東京大学, 2.日本電信電話株式会社, 3.国立成育医療研究センター)

【背景】糖尿病治療では、血糖管理不良が長期間続くと腎症等の合併症リスクが高まる。血糖管理不良を早期に予測できれば、治療を強化すべき患者を抽出でき、合併症発症の抑制に繋がると期待できる。筆者らは過去に受診中断リスク予測技術を報告した。本研究では通院治療を継続中の患者を対象に血糖管理不良の予測に発展させた。【目的】糖尿病患者の血糖管理不良を予測する。【方法】血糖管理不良の定義を「受診日を起点に、将来64週のHbA1cの最悪値が合併症抑制の血糖管理目標値7%を下回らず、かつ、過去64週の最悪値を上回る」とした。電子カルテから抽出した受診日を起点に過去64週のHbA1c値を入力し、血糖管理の不良／良好のいずれかを出力するモデルを深層学習で構築した。予測精度は10分割交差検定で評価した。対象は①2004年以降、東京大学医学部附属病院の糖尿病・代謝内科の外来に来院し、②電子カルテに糖尿病の病名コード（ICD10 E10からE14）が確定病名として付与されて転帰していない、を全て満たす患者とした。対象期間は2006年11月27日から2016年1月29日とした。正解ラベルは6,039人の患者に平均30.9件（標準偏差23.1）付与され、血糖管理不良は75,174件、良好は111,469件だった。検査は不等間隔で実施されるため、検査値は欠損を考慮したオートエンコーダで経時変化を表す低次元特徴量に変換することで予測精度向上が可能かを併せて検討した。【結果】検査値を変換せずに用いるとAUC値が0.72、F値が0.57のモデルが構築された。低次元特徴量へ変換すると予測精度が向上し、AUC値が0.78、F値が0.64だった。【考察】血糖管理不良を予測し、その精度を検証したのは本研究が初である。また、低次元特徴量への変換が予測精度の改善に寄与することを確認した。今後は特徴量を増やして精度向上を図る。

機械学習を用いた糖尿病外来患者の血糖管理不良予測

倉沢 央^{*1*2}、林 勝義^{*1*2}、藤野 昭典^{*2}、芳賀 恒之^{*3}、脇嘉代^{*1}、野口 貴史^{*1*4}、大江 和彦^{*1}

*1 東京大学、*2 日本電信電話株式会社、*3 NTT アドバンステクノロジー株式会社、*4 国立成育医療研究センター

Machine learning-based Prediction of poor glycemic control in patients with diabetes

Hisashi Kurasawa^{*1*2}, Katsuyoshi Hayashi^{*1*2}, Akinori Fujino^{*2}, Tsuneyuki Haga^{*3},
Kayo Waki^{*1}, Takashi Noguchi^{*1*4}, Kazuhiko Ohe^{*1}

*1 The University of Tokyo, *2 Nippon Telegraph and Telephone Corporation

*3 NTT Advanced Technology Corporation, *4 National Center for Child Health and Development

[Background] Poor glycemic control in patients with diabetes results in the progression of diabetes-related complications and reduced quality of life. [Objective] Predict poor glycemic control in patients with diabetes. [Methods] Poor glycemic control was defined as the maximum HbA1c level for 64 weeks from a hospital visit being higher than both 7% and the maximum HbA1c level for the past 64 weeks. A deep learning algorithm was used to build both a binary classification model for the prediction of poor glycemic control and a feature extraction model from the past HbA1c levels with unequal intervals. The 10-fold cross validation was used to evaluate prediction performance. All experiments were performed using electronic medical records from the University of Tokyo Hospital that included those of 7,180 patients, who had hospital visits after January 1, 2004, and who had diagnostic codes indicative of diabetes. 86,299 records of 4,635 patients were poor glycemic control and 154,912 records of 6,905 patients were good glycemic control. [Results] The ROC AUC and F-measure of the prediction without the feature extraction model were 0.72 and 0.52, respectively. These values increase to 0.80 and 0.61 by using the feature extraction model, respectively. [Conclusions] Our findings may provide clinical information concerning the development of clinical decision support systems for the initial screening of poor glycemic control.

Keywords: Diabetes, Poor glycemic control, Missing value, Deep learning

1. はじめに

糖尿病は、インスリンの作用不足によって細胞に糖を取り込めなくなり、血糖値が慢性的に高くなってしま病態である¹⁾。糖尿病の患者数は近年増加傾向にあり、2014年における糖尿病の患者数は316万人に達している²⁾。糖尿病は重症化すると、腎症等の合併症を引き起こし、患者自身のQOL (Quality of Life) の低下や医療費の増大につながる。

糖尿病重症化予防は「治療継続者の割合の増加」と「血糖コントロール指標におけるコントロール不良者の割合の減少」が具体的なアクションとして位置づけられている³⁾。このうち後者は、HbA1cが8.4%以上の血糖管理不良者を平成22年度の1.2%から平成34年度までに1.0%へ減らす数値目標が設定されている。これに関連して、日本糖尿病学会は第56回日本糖尿病学会年次学術集会にて「熊本宣言2013」を発表し、合併症予防のための管理目標値としてHbA1c7%未満を設定し、予防に尽力している⁴⁾。

良好な血糖管理を達成し、その状態を維持することを目標に、血糖管理不良の患者には早期から治療を強化する¹⁾。具体的には、食事・運動指導、経口血糖降下薬療法、インスリン療法、経口血糖降下薬とインスリンの併用療法、強化インスリン療法へと治療を強化していく。しかし、血糖管理の急激な是正あるいは厳格すぎる血糖管理は、重篤な低血糖や最小血管症の増悪といったリスクを伴うため⁵⁾、患者の病態に応じて治療強化を慎重に進める必要がある。

これまでも血糖管理不良と関連する因子を明らかにすることが試みられてきた。先行研究では年齢や糖尿病治療期間、治療内容、人種や民族、糖尿病の家族歴などが因子として挙げられている⁶⁻¹²⁾。しかしながら、血糖管理不良のリスクの高い患者を絞り込めたととしても、将来の血糖管理状態を正しく予測することは難しく、治療強化を進める判断への参考にはならなかった。

患者の血糖管理の予測については、血糖値の短期的な変動、特に低血糖の予測が取り組まれてきた¹³⁻¹⁶⁾。先行研究では、血糖自己測定(SMBG)や持続血糖測定(CGM)で得られた、ほぼ一定間隔の血糖値をもとに数時間から1日以内に低血糖になるか否かを予測し、その精度を確認している。しかし、糖尿病患者の将来の血糖管理不良の予測にこれらの手法を適用することは難しい。なぜなら、一般に、糖尿病患者のHbA1c検査は外来受診と同一日に実施されるため、不等間隔で計測された値を入力として受け付けるモデル設計が必要であるからだ。さらに、HbA1c値は病態以外に季節の影響も受けることが知られているため¹⁷⁻²²⁾、長期的な傾向の考慮が望ましい。

そこで本研究は、糖尿病患者の将来の血糖管理不良を高い精度で予測することを目的とする。過去のHbA1c値から血糖管理不良を予測するモデルに深層学習²³⁾を用いた。欠損を考慮したオートエンコーダ²⁴⁾を用いて不等間隔のHbA1c値から経時変化を表す低次元特徴量を抽出することで、予測精度向上を図った。

2. 血糖管理不良の予測モデル

2.1 血糖管理不良の定義

血糖管理不良の定義は「受診日を起点に、将来64週のHbA1cの最悪値が合併症抑制の血糖管理目標値7%を上回り、かつ、過去64週の最悪値を上回る」とした。血糖管理目標値は熊本宣言2013の値を採用した。この目標値はHbA1cの上限値を設定しているため、対象期間中に検査されたHbA1cの最悪値を判定に用いた。また、HbA1c値は季節の影響を受けることが知られており¹⁷⁻²¹⁾、寒い季節ほど高く、暖かい季節に低くなる傾向がある。それゆえ、過去と将来の最悪値は1年よりも若干長い64週の期間の値から算出することにした。図1に血糖管理不良と判定されるHbA1cの例を示す。

この例では、将来 64 週の HbA1c 最悪値が 7% を上回り、かつ、過去 64 週の最悪値を上回るため、血糖管理不良と判定される。

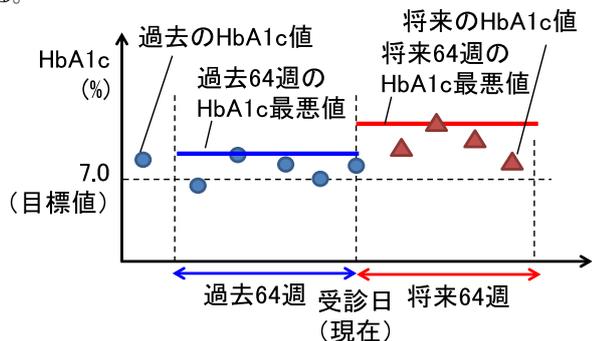


図 1. 血糖管理不良と判定される HbA1c の例

2.2 予測モデルの構成

予測モデルは、図 2 に示すように、2 種類のオートエンコーダと二値分類器で構成した。2 種類のオートエンコーダは、外来受診日までの HbA1c 値から HbA1c 経時変化を表す特徴ベクトル x を抽出する。二値分類器は、 x から血糖管理不良を 0、血糖管理良好を 1 と定義されたクラス y を分類する。それぞれ多層ニューラルネットワーク (NN) で設計した。各多層 NN の設計は 3 章と 4 章にて説明する。

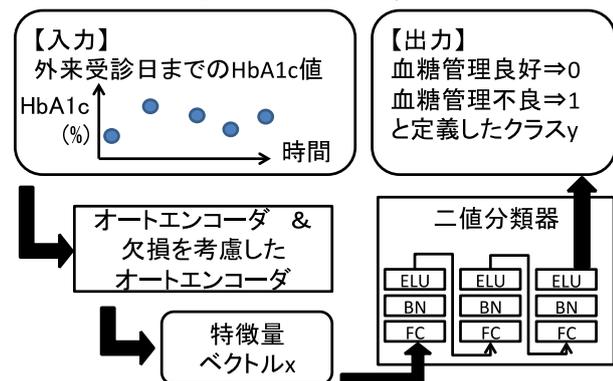


図 2. 予測モデルの設計

2.3 予測モデルの評価法

10 分割交差検定を用いて、Accuracy が最も高くなるときの ROC AUC 値、F 値、Accuracy を算出し、モデルの予測精度を評価した。ROC AUC 値は血糖管理良好と不良に分離する性能を表す尺度、F 値は血糖管理不良患者のみを抽出する性能を表す尺度、Accuracy は正解率である。いずれの尺度も 1 に近いほど性能が高いことを表す。なお、交差検定では患者単位でデータセットを 10 分割し、すべての患者が訓練データとテストデータのどちらか一方にのみ含まれるようにして、未知の患者に対する血糖管理予測性能を評価した。

モデルの学習には Adam²⁵⁾ を勾配法とし、パラメータは既存研究を参考に $\alpha = 0.001$ 、 $\beta_1 = 0.9$ 、 $\beta_2 = 0.999$ 、 $\varepsilon = 10^{-8}$ 、epoch = 1,000 を設定した。モデルの実装および学習に Chainer²⁶⁾ を用いた。実験評価には NVIDIA Tesla P100 を搭載した NVIDIA DGX-1 を用いた。

3. HbA1c 経時変化を表す特徴ベクトルの抽出

3.1 不等間隔データの課題

HbA1c は過去 1、2 ヶ月の血糖の平均的な状態を調べるために測られる。一般に、糖尿病患者の HbA1c 検査は外来受

診と同日に実施される。通院間隔は病態や患者や医師の都合により変化するため、HbA1c 検査実施間隔もそれに応じて変化し、間隔は一定でない。

これまででも不等間隔の計測データから特徴量を抽出することが試みられてきた。最も単純なアプローチは、間隔を無視し、最大値や最小値、分散値といった統計値に計測値を抽象化する手法である。しかし、これら統計値は時間情報を失うため、計測値の経時変化を十分に特徴量として反映することはできない。そこで、計測間隔が一定のデータとして量子化・整形し、計測されていない時刻の値を推定したうえで、特徴を抽出するアプローチが一般に用いられている。例えば、スプライン補間等で欠損値を補間し、Symbolic Aggregate Approximation (SAX)²⁷⁾ のような符号化技術や FFT のような周波数解析といった特徴抽出手法が適用される。しかし、正確に欠損補間できるとは限らないため、欠損補間に特徴抽出して得た特徴量は、計測値の特性を正確に反映しているとは限らないという問題があった。この問題は、特徴量の質の低下に繋がる。

そこで、本研究では欠損を考慮したオートエンコーダを新たに考案し、欠損補間と特徴抽出を同時に行い、特徴量の質の向上を図った。

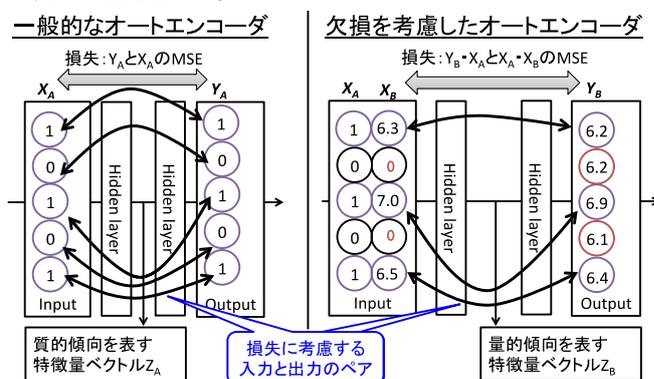


図 3. 2 種類のオートエンコーダ

3.2 欠損を考慮したオートエンコーダの提案

オートエンコーダは、深層学習の一種で、次元圧縮に用いられる手法である。この手法は、入力データから特徴量に変換するエンコーダと、特徴量から入力データと同じ次元の出力データに変換するデコーダの 2 つで構成される。特徴量の次元数は、エンコーダの出力層とデコーダの入力層のユニット数によって設計される。学習では、出力データと入力データの距離を損失にして、出力データが入力データに近づくように収束させる。特徴量の次元数を入力よりも小さい値に設定することで、エンコーダで入力データの次元圧縮が行える。

筆者らは、オートエンコーダの構造を参考に、不等間隔データから欠損補間と特徴抽出を同時に行う、「欠損を考慮したオートエンコーダ」を提案する。本手法は、入力データと出力データの差が最小になるようにパラメータを学習するオートエンコーダにおいて、式 1 に示すように、入力データで欠損している要素以外で計算される入出力ベクトルの差 $L(w)$ が最小になるようにパラメータ w を学習した。f はオートエンコーダで定義される関数を表す。

$$L(w) = X_{An} \cdot (X_{Bn} - Y_n(w))$$

$$Y_n(w) = f(X_{An}, X_{Bn}; w) \quad (\text{式 1})$$

図 3 に 2 種類のオートエンコーダの概要図を示す。学習が収束すると、入力データの欠損値が補間されたデータが出力データとして得られる。同時に、欠損に影響を受けない低次元の特徴量に変換するエンコーダが得られる。

欠損を考慮したオートエンコーダ構築の前処理として、不等間隔の計測データを計測間隔が一定のデータに量子化する。そして、計測値があれば1、欠損していれば0の値で構成された計測の質的傾向を表す特徴量ベクトル X_A と、欠損値を仮の値(例えば、0)で埋めた計測値の量的傾向を表す特徴量ベクトル X_B に整形する。欠損を考慮したオートエンコーダは、入力は X_A と X_B を連結した特徴量ベクトル、出力は X_B と同じ次元で構成される Y_B 、損失は $Y \cdot X_A$ と $Y \cdot X_B$ のユークリッド距離を用いる。

3.3 2種類のオートエンコーダの設計

筆者らは、HbA1c 検査有無を表す質的特徴量 Z_A と、HbA1c 経時変化を表す量的特徴量 Z_B の2つを抽出した。 Z_A は、 X_A を入力とし、通常のオートエンコーダを用いて抽出した。一方、 Z_B は、 X_A と X_B を連結した特徴量ベクトルを入力とし、欠損を考慮したオートエンコーダを用いて抽出した。そして、 Z_A と Z_B を連結したものを HbA1c 経時変化を表す特徴量 x として、予測モデルに用いた。

各オートエンコーダの層は、入力層、中間層(エンコーダの中間層、エンコーダの出力層、デコーダの入力層、デコーダの中間層)、出力層で構成される。全層を全結合(FC)層にし、中間層は Batch Normalization (BN)²⁸ で正規化した後に活性化関数 Exponential Linear Unit (ELU)²⁹ を適用した。中間層の数は3を基準値として0から4の間で、中間層のユニット数は256を基準値として64から512までの間で、それぞれ変化させた。量子化は1週間単位とした。予測モデルで参照する期間は、外来受診日を起点とした過去64週を基準値として64から192週までの間で変化させた。過学習回避のため、入力層に近い中間層2つに Dropout³⁰ を50%の確率で適用した。エンコーダの出力層(デコーダの入力層)のユニット数は8に設定した。つまり、HbA1c 経時変化を表す特徴量は16次元で得られる。

3.4 比較手法

評価実験では、Statistics、Spline+SAX、Raw の3つの特徴量ベクトルを比較手法として用いた。

Statistics は、計測頻度、最大値、最小値、平均値、中央値、標準偏差の6種類の統計値で構成された、6次元の特徴量である。これらは時間情報を全く含んでいない。

Spline+SAX は、スプライン補間で内挿した後に、時間軸8分割、値を表す5符号と欠損を表す1符号のSAXで変換した、8次元の特徴量である。外挿は内挿に比べて精度が低いことが知られているため、時間軸8分割のうち外挿が必要な領域については、欠損を表す符号に変換した。

Raw は X_A と X_B を連結したベクトルであり、予測モデルで参照する期間の週数の2倍の次元の特徴量である。例えば、過去64週を参照する場合は128次元となる。

さらに、Spline+SAX で用いたスプライン補間と欠損を考慮したオートエンコーダとで、欠損補間の性能を比較した。テストデータの各計測値を欠損させ、2手法で推定した値と実際の計測値との根平均二乗誤差を10分割交差検定で算出した。

4. 二値分類器の設計

二値分類器の多層 NN 入力層のユニットは x の次元数とし、出力層にはソフトマックス関数を適用し、交差エントロピーで損失を求めた。全層を FC 層にし、中間層は BN で正規化した後に活性化関数 ELU を適用した。中間層の数は3を基準

値として0から4の間で、中間層のユニット数は256を基準値として64から512までの間で、それぞれ変化させた。

5. 評価実験と結果

5.1 対象

①2004年以降、東京大学医学部附属病院の糖尿病・代謝内科の外来に来院し、②電子カルテに糖尿病の病名コード(ICD10 E10からE14)が確定病名として付与されて転帰していない、患者7,180名の電子カルテデータを医学部倫理委員会の承認を得て使用した(承認番号10705)。対象期間は2006年11月27日から2016年1月29日とし、受診日を起点に過去と将来64週に4度以上HbA1c検査を実施したケースを対象レコードとした。

レコード総数は241,211件(患者7,180人に平均33.6件(標準偏差26.6件))、このうち血糖管理不良は86,299件(患者4,635人に平均18.6件(標準偏差14.5件))、血糖管理良好は154,912件(患者6,905人に平均22.4件(標準偏差17.3件))であった。HbA1c 値を1週間単位で量子化すると、平均5.9週(標準偏差2.6週)間隔で検査が実施され、受診日を起点に過去64週のうち平均53.6週(83.7%)の検査値が欠損していた。

表1. 予測モデルの性能評価結果

	ROC AUC 値	F 値	Accuracy
Statistics	0.47	0.00	0.64
Spline+SAX	0.73	0.49	0.69
Raw	0.72	0.52	0.68
提案手法	0.80	0.61	0.73

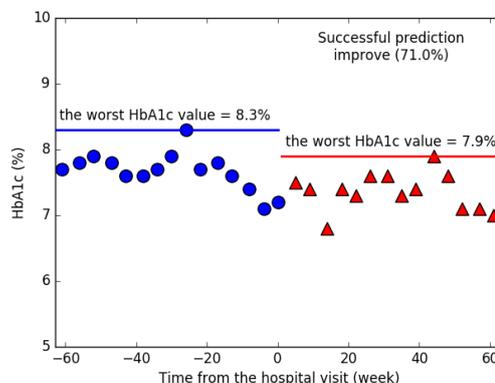


図4. 予測の成功例-1

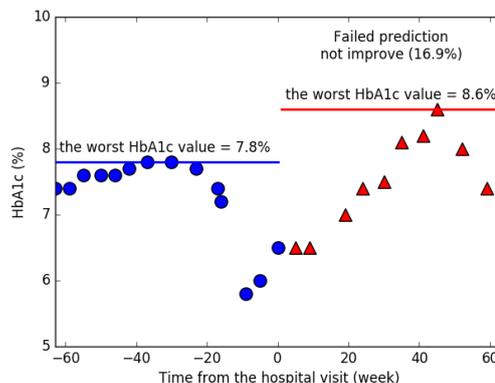


図5. 予測の失敗例-1

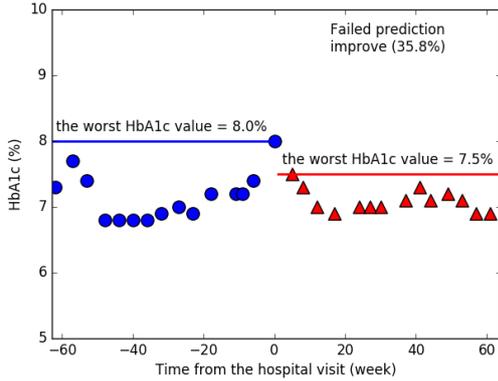


図 6. 予測の失敗例-2

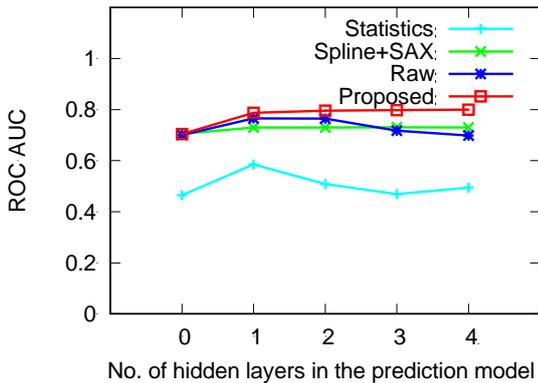


図 7. 予測モデルの中間層の数の予測への影響

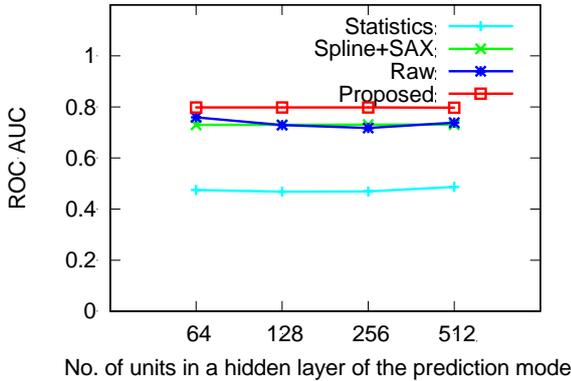


図 8. 予測モデルの中間層のユニット数の予測への影響

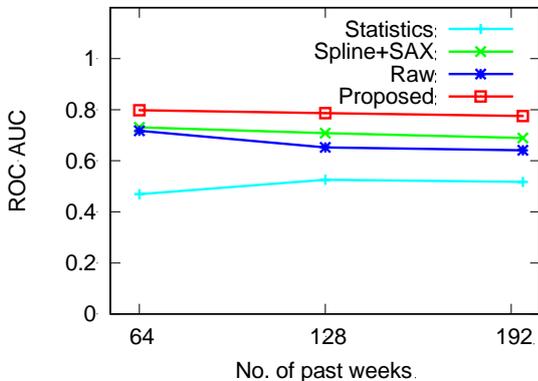


図 9. 予測モデルの参照期間の予測への影響

5.2 予測モデルの性能評価

予測モデルを Accuracy 最大化で学習して得た評価結果を表 1 に示す。提案手法の ROC AUC 値は 0.80、F 値は 0.61、Accuracy は 0.73 であり、最も優れた性能を示した。比較手法

の中では Statistics の予測が最も悪く、F 値は 0 であった。Spline+SAX と Raw はほぼ同じ性能だった。提案手法が Raw よりも高い性能を示したことから、低次元特徴量への変換が予測精度の改善に寄与することを確認できた。

予測が成功した 1 例と失敗した 2 例を図 4 から図 6 に示す。図 4 は 71.0% の確率で血糖管理良好と予測し、成功であった。図 5 は 16.9% の確率で血糖管理不良 (83.1% の確率で血糖管理良好) と予測し、失敗であった。この例では受診日直前 20 週まで HbA1c 7% 未満の目標を達成していたため、血糖管理不良となる予測が難しかったと思われる。図 6 は 35.8% の確率で血糖管理良好 (64.2% の確率で血糖管理不良) と予測し、失敗であった。この例では受診日までの HbA1c 値の上昇傾向が続くと予測し、失敗したと思われる。

予測モデルの中間層の数とユニット数を変化させたときの ROC AUC 値を図 7 と 8 に示す。中間層の数に関しては、提案手法は 1 以上では予測性能に大きな改善は見られなかった。これに対して、Statistics と Raw は中間層の数が 2 以上で予測性能が悪化した。中間層のユニット数に関しては、Raw が 128 以上で若干の予測性能悪化が見られたが、他の手法への影響は小さかった。

予測モデルで参照する期間を変化させたときの ROC AUC 値を図 9 に示す。Statistics を除くいずれの手法も参照期間を伸ばしても予測性能に大きな改善は見られなかった。

5.3 欠損を考慮したオートエンコーダの性能評価

エンコーダの中間層の数とユニット数を変化させたときの ROC AUC 値を図 10 と 11 に示す。提案手法は中間層の数が 2 以上、ユニット数が 64 以上では予測性能に大きな改善はなかった。中間層の数とユニット数を変化させたときの特徴量を、デコーダで変換し、経時変化の推定結果として図示した一例が図 12 と図 13 である。青丸は過去 64 週の HbA1c の計測値を表して、線は推定された経時変化を表している。予測には寄与していないが、中間層の数とユニット数を増やしたほうが経時変化がよりフィットすることが確認できる。

欠損を考慮したオートエンコーダとスプライン補間とで欠損補完の性能を平均二乗誤差で比較した結果を表 2 に示す。内挿はスプライン補間の方が優れ、外挿および内挿と外挿を合わせた平均では欠損を考慮したオートエンコーダが優れていた。

表 2. 欠損補完の平均二乗誤差

	内挿	外挿	平均
欠損を考慮したオートエンコーダ	0.57	1.6	0.86
スプライン補間	0.41	16	7.14

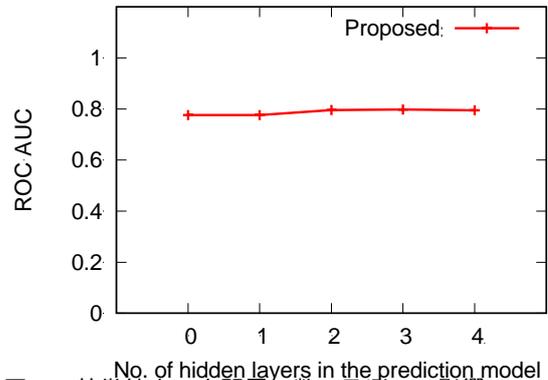


図 10. 特徴抽出の中間層の数の予測への影響

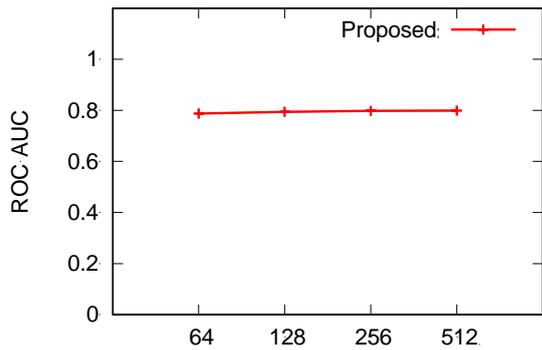


図 11. 特徴抽出の中間層のユニット数の予測への影響

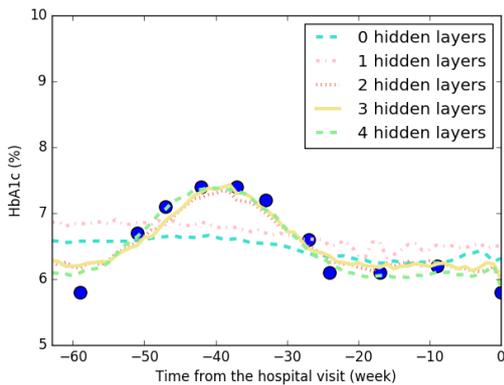


図 12. 特徴抽出の中間層の数の特徴抽出例

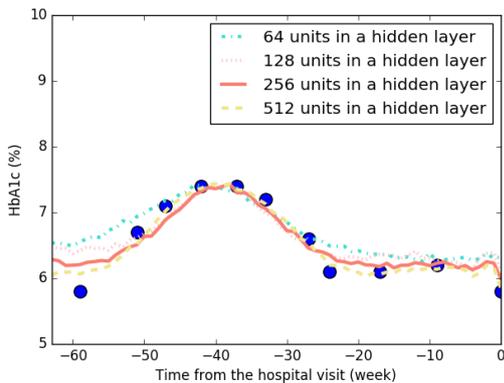


図 13. 特徴抽出の中間層のユニット数の特徴抽出例

6. 考察

糖尿病患者の血糖管理不良を予測し、その精度を検証したのは本研究が初である。また、低次元特徴量への変換が予測精度の改善に寄与することを確認した。

6.1 モデルの有用性

性能評価に用いた HbA1c 値のデータは、糖尿病治療中の患者のものであるため、モデルの予測対象とした血糖管理不良は、担当医が患者の血糖管理不良を予測できなかったか、もしくは予測できたが治療強化しても血糖管理の改善が難しかったケースである。モデルの性能は ROC AUC 値が 0.80 で程々であったが、担当医が見逃しやすい血糖管理不良の兆候や、治療をさらに強化したほうが良いとされる病態を検出するのに有用であると期待される。

しかしながら、血糖管理不良を未然に防ぐには血糖管理不良を精度高く予測するだけでは不十分であり、患者に効果的

な治療強化法の選定も必要である。今回のモデルでは、HbA1c 値のみのデータを用いたが、今後は期間中に実施した治療強化法を特徴量に加えるなどして、治療強化した場合の血糖管理予測ができるようにモデルを改良する方針である。

また、筆者らのテレメディシン領域の研究結果からは、患者に生活習慣に関するフィードバックを行うと、特に食事内容に関連した自己管理が促進されることが明らかになっており³¹⁾³²⁾、これに類似して患者に血糖管理不良の予測結果を提示することが、食生活の改善や運動習慣の増加などの行動変容に効果があるかも検証する予定である。

6.2 モデルの汎用性

本研究は一施設の電子カルテデータを使って評価したが、モデルの汎用性を調べるためには、病院の機能別区分や地域が異なる他の病院のデータでも予測精度を評価する必要がある。さらに、HbA1c 以外の検査での予測性能の違いについても今後検証したい。

6.3 モデルの設計

予測モデルの設計については、中間層の数やユニット数を変化させても提案手法の性能は大きく変化しなかったことから、現在の設計から大きな変更は必要ないものと思われる。一般に、中間層の数やユニット数が増えると、モデルの表現力が高まる一方で、過学習の危険が高まる。Statistics と Raw は、過学習したことで予測性能が悪化したと推察される。一方、提案手法は欠損を考慮したオートエンコーダを用いて経時変化を表す低次元特徴量に変換していたため、過学習を回避できたと思察される。

予測モデルで参照する期間についても将来 64 週の血糖管理を予測するには過去 64 週の経時変化でほぼ十分であると推察される。Statistics のみ性能が改善したのは、参照範囲拡大でサンプル数が増えたことで統計値の信頼区間幅が狭まり、特徴量の質が向上したためと思われる。

エンコーダの設計は、中間層の数を 4 まで深くしても予測精度は大きな改善は見られなかった。今後はさらに層を深くして予測精度向上が見込めるか検証を深めたい。層を深くする場合は勾配消失といった新たな問題も懸念されるため、ResNet³³⁾のように層の入力を参照した残差関数を導入するなどの工夫が必要と考えている。

6.4 糖尿病重症化予防に向けて

糖尿病重症化予防は、血糖管理不良者の減少だけでなく、治療継続者の割合の増加も同様に課題である。これについては、筆者らは過去に受診中断リスク予測技術を報告している³⁴⁾³⁵⁾。電子カルテデータを用いて、予約不履行(受診が途絶えるきっかけとなり得る予約外来の不受診)を AUC = 0.958、F 値 = 0.704 で予測し、受診中断リスク順位(将来の受診中断日までの日数の長さによる患者の順位付け)を正解率 = 0.706 で予測した。受診中断を避けるために積極的に支援すべき患者の絞りこみや、支援を開始すべき時期の見極めに役立てることが期待できる。

今回のモデルと併用して治療支援に役立てることで、「治療継続者の割合の増加」と「血糖コントロール指標におけるコントロール不良者の割合の減少」の2つの糖尿病重症化予防の目標達成に寄与できると考えている。

7. おわりに

機械学習を用いた初の試みとして、電子カルテデータから

抽出した糖尿病患者の受診歴と HbA1c 値をもとに、将来の血糖管理不良を予測するモデルを設計した。東京大学医学部附属病院に通院する糖尿病患者の電子カルテデータを用いて、10 分割交差検定でモデルを評価し、ROC AUC 値が 0.80 の予測性能を示すことができた。このモデルは、患者への治療強化時期の絞り込みに役立つと期待された。今後はさらなる予測性能改善に向けて、モデルを改善する予定である。

利益相反(COI)

本研究は東京大学 COI 拠点「自分で守る健康社会」における日本電信電話株式会社との共同研究により実施された。共著者の脇嘉代は東京大学と NTT ドコモとの社会連携講座に所属している。

参考文献

- 1) 糖尿病診療ガイドライン 2016. 日本糖尿病学会, 2016.
- 2) 平成 26 年国民健康・栄養調査. 厚生労働省, 2014.
- 3) 健康日本 21(第二次). 厚生労働省, 2012.
- 4) 熊本宣言 2013. 日本糖尿病学会, 2013.
[http://www.jds.or.jp/common/fckeditor/editor/filemanager/connectors/php/transfer.php?file=/uid000025_6B756D616D6F746F323031332E706466 (cited 2017-Aug-9)].
- 5) Gerstein HC, Miller ME, Byington RP, et al. Effects of intensive glucose lowering in type 2 diabetes. *N Engl J Med.* 2008 Jun 12;358(24):2545-59.
- 6) Khatrab M, Khader YS, Al-Khawaldeh A, and Ajlouni K. Factors associated with poor glycemic control among patients with type 2 diabetes. *J Diabetes Complications.* 2010 Mar-Apr;24(2):84-9.
- 7) Crowley MJ, Holleman R, Klamerus ML, Bosworth HB, Edelman D, and Heisler M. Factors associated with persistent poorly controlled diabetes mellitus: clues to improving management in patients with resistant poor control. *Chronic Illn.* 2014 Dec;10(4):291-302.
- 8) Wu M, Wen J, Qin Y, et al. Familial History of Diabetes is Associated with Poor Glycaemic Control in Type 2 Diabetics: A Cross-sectional Study. *Sci Rep.* 2017 May 3;7(1):1432.
- 9) Harris MI, Eastman RC, Cowie CC, Flegal KM, and Eberhardt MS. Racial and ethnic differences in glycemic control of adults with type 2 diabetes. *Diabetes Care.* 1999 Mar;22(3):403-8.
- 10) Goudswaard ANI, Stolk RP, Zuithoff P, and Rutten GE. Patient characteristics do not predict poor glycaemic control in type 2 diabetes patients treated in primary care. *Eur J Epidemiol.* 2004;19(6):541-5.
- 11) Sanal TS, Nair NS, and Adhikari P. Factors associated with poor control of type 2 diabetes mellitus: a systematic review and meta-analysis. *Journal of Diabetology.* 2011;3 (1):1-10.
- 12) Juarez DT, Sentell T, Tokumaru S, Goo R, Davis JW, and Mau MM. Factors associated with poor glycemic control or wide glycemic variability among diabetes patients in Hawaii, 2006-2009. *Prev Chronic Dis.* 2012;9:120065.
- 13) Sudharsan B, Peeples M, and Shomali M. Hypoglycemia Prediction Using Machine Learning Models for Patients With Type 2 Diabetes. *J Diabetes Sci Technol.* 2015 Jan; 9(1): 86-90.
- 14) Cox DJI, Gonder-Frederick L, Ritterband L, Clarke W, and Kovatchev BP. Prediction of severe hypoglycemia. *Diabetes Care.* 2007 Jun;30(6):1370-3.
- 15) Plis K, Bunescu RC, Marling C, Shubrook J, and Schwartz F. A Machine Learning Approach to Predicting Blood Glucose Levels for Diabetes Management. *Proceedings of AAAI Workshop: Modern Artificial Intelligence for Health Analytics.* 2014;35-39.
- 16) Dassau E, Cameron F, Lee H, et al. Real-Time hypoglycemia prediction suite using continuous glucose monitoring: a safety net for the artificial pancreas. *Diabetes Care.* 2010 Jun;33(6):1249-54.
- 17) Gikas A, Sotiropoulos A, Pastromas V, Papazafropoulou A, Apostolou O, and Pappas S. Seasonal variation in fasting glucose and HbA1c in patients with type 2 diabetes. *Prim Care Diabetes.* 2009 May;3(2):111-4.
- 18) Mianowska B, Fendler W, Szadkowska A, et al. HbA1c levels in schoolchildren with type 1 diabetes are seasonally variable and dependent on weather conditions. *Diabetologia.* 2011 Apr;54(4):749-56.
- 19) Sohmiya M, Kanazawa I, and Kato Y. Seasonal changes in body composition and blood HbA1c levels without weight change in male patients with type 2 diabetes treated with insulin. *Diabetes Care.* 2004 May;27(5):1238-9.
- 20) Ishii H, Suzuki H, Baba T, Nakamura K, and Watanabe T. Seasonal variation of glycemic control in type 2 diabetic patients. *Diabetes Care.* 2001 Aug;24(8):1503.
- 21) Higgins T, Saw S, Sikaris K, et al. Seasonal variation in hemoglobin A1c: is it the same in both hemispheres? *J Diabetes Sci Technol.* 2009 Jul 1;3(4):668-71.
- 22) Sakura H, Tanaka Y, and Iwamoto Y. Seasonal fluctuations of glycosylated hemoglobin levels in Japanese diabetic patients. *Diabetes Res Clin Pract.* 2010 Apr;88(1):65-70.
- 23) LeCun Y, Bengio Y, and Hinton GE. Deep Learning. *Nature* 2015; 521:436-44.
- 24) Hinton GE and Salakhutdinov RR. Reducing the dimensionality of data with neural networks. *Science.* 2006 Jul28;313(5786):504-7.
- 25) Kingma DP and Jimmy LB. Adam: A method for stochastic optimization, arXiv:1412.6980, 2014.
- 26) Chainer. [<https://chainer.org/> (cited 2017-Aug-9)].
- 27) Lin J, Keogh E, Lonardi S, and Chiu B. A Symbolic Representation of Time Series, with Implications for Streaming Algorithms, *Proceedings of the 8th ACM SIGMOD Workshop on Research Issues in Data Mining and Knowledge Discovery* 2003;2-11.
- 28) Ioffe S and Szegedy C. Batch normalization: Accelerating deep network training by reducing internal covariate shift. arXiv:1502.03167, 2015.
- 29) Clevert D, Unterthiner T, and Hochreiter S. Fast and Accurate Deep Network Learning by Exponential Linear Units (ELUs). arXiv:1511.07289, 2016.
- 30) Srivastava N, Hinton G, Krizhevsky A, Sutskever I, and Salakhutdinov R. Dropout. A simple way to prevent neural networks from overfitting. *The Journal of Machine Learning Research* 2014:1929-1958.
- 31) Waki K, Fujita H, Uchimura Y, et al. DialBetics: A Novel Smartphone-based Self-management Support System for Type 2 Diabetes Patients, *J Diabetes Sci Technol* 2015;8(2), 209-15.
- 32) Waki K, Aizawa K, Kato S, et al. DialBetics With a Multimedia Food Recording Tool, FoodLog: Smartphone-Based Self-Management for Type 2 Diabetes, *J Diabetes Sci Technol* 2015; 9(3), 534-40.
- 33) He K, Zhang X, Ren S, and Sun J. Deep Residual Learning for Image Recognition. arXiv:1512.03385, 2015.
- 34) Kurasawa H, Hayashi K, Fujino A, et al. Machine-learning-based prediction of a missed scheduled clinical appointment by patients with diabetes. *J Diabetes Sci Technol.* 2016 May 3;10(3):730-6.
- 35) 倉沢 央, 林 勝義, 藤野 昭典, ら. 機械学習を用いた糖尿病外来患者の受診中断リスク予測. 第 59 回日本糖尿病学会年次学術集会, 2016.