

Prediction of patient journey trajectory using Transformer-based object tracking system

Noriko F. HIROI Kanagawa Institute of Technology

特願2025-106097

Keywords: AI, 健康情報, 未来予測, 時系列データ, ブラウザアプリ

技術概要

Transformerベースの多変量時系列データによる軌跡予測システム応用の一つです。健康診断の血液検査データを学習させたAIを軸とした健康アドバイス提供ツールを開発しました。新しい血液検査情報を入力すると、現時点より未来の検査結果がどのような変化を辿るかを予測します。学習に用いたデータ群と比較して、予測される深刻な疾患を事前に回避するための健康アドバイスを提供します。入力データはcsvファイルの形式でアップロードでき、操作は全てブラウザで完了します。

想定される活用事例

生活習慣改善のアドバイスを求める人、未病状態の維持に関心のある人を対象に、病院等で予測データと健康アドバイスを提供し、個人の健康意識改革を促し未病維持を実現します。未病人口の増加に伴う医療費削減に関心のある行政機関による病院等と連携した活用を想定しています。血液検査と本アプリケーションの利用のみで将来の糖尿病患者数を減らせた場合、10%減で約2000億円規模の医療費削減が見込まれます。

技術的背景

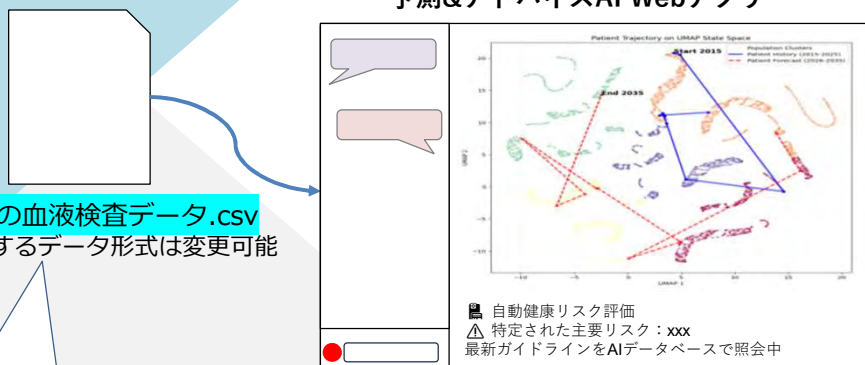
Transformerは自然言語を扱う生成AIとして設計、開発されたが、公開後様々な応用が行われている。当研究室の高橋による船舶軌跡予測、座間による細胞軌跡予測もそれに応用の一つにあたる。大規模データに基づく優れた文脈予測能から、Transformerの応用でどのようなことができ、どのようなことができないかを調べ、どのような新しいアルゴリズムが必要かなどについて、世界的な関心が高まっている。

今回の展示物

本研究では、Transformerを応用したObject trackingシステムによる軌跡予測を応用し、言語以外の時系列ベクトルとして、血液検査などで得られる健康情報を学習させ、新規の健康情報に対し、未来予測と健康状態カテゴリ推定、およびそれに基づく健康アドバイス生成を行うシステムの構築を目的とする。

血液検査データから将来かかりやすい疾患をAI予測

予測&アドバイスAI Webアプリ



トレーニングデータを変更することで、
予測に使うデータ・予測目的の変更可能

チャット形式で質問可能

使用データ例

[illegible]

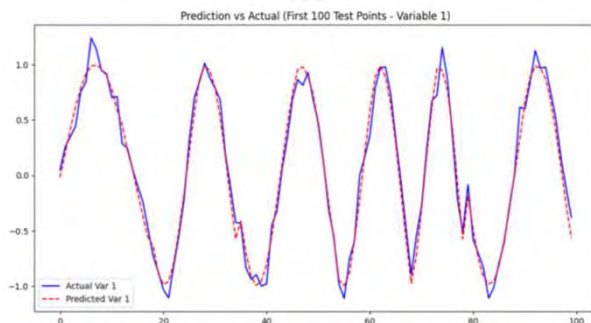
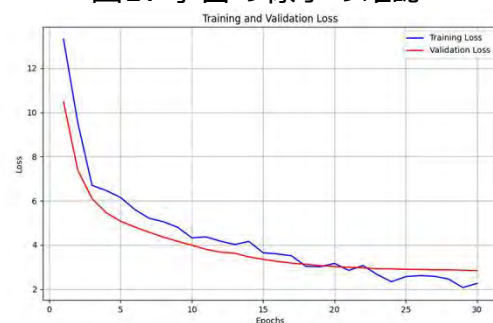
データについて

1,000人分、
21年間の情報
から構成される
37,800ポイント
×4変数の
データを学習

学習過程の確認

上: test dataと学習後モデルによる予測の差分変化の様子
下: test dataと学習後モデルによる予測の具体的な軌跡

図1: 学習の様子の確認



方法概要

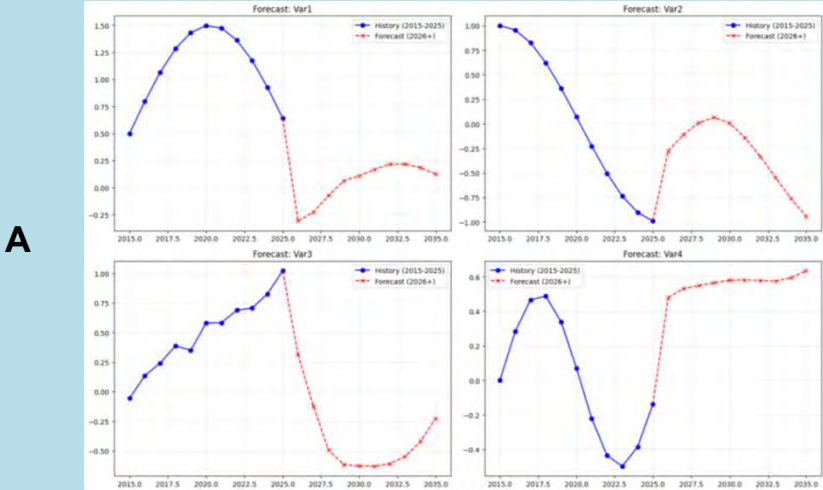
1. 時系列データ解析用Transformerの準備
当研究室高橋によって開発された時系列データ解析用Transformerを、当研究室内のLinux Machine内に移植した。
これまで解析してきた船舶軌跡データ等で、予定通りの動作をすることを確認した。またテストケースとして公表されている細胞移動情報を含むCellTrackerの細胞軌跡情報をcsvの形で取得、学習データ、および予測の精度テストに使用した。**図1**は学習過程のlossの変化(左)および予測軌跡(赤線)と実際のデータのオーバーレイ(下)である。概ね予定の学習が進められることを確認した。

2. 健康情報データの準備
健康情報として日本人の死因のトップ3(がん、脳血管疾患、心血管疾患)に関する指標、およびそれらの疾患を共通して憎悪する要素として糖尿病の指標について、実データをもとに1,000人分の仮装データを作成し、学習用データとした。学習用データは37,800ポイント分を用意し、上述のテストケースと同様にTransformerの学習を行った。

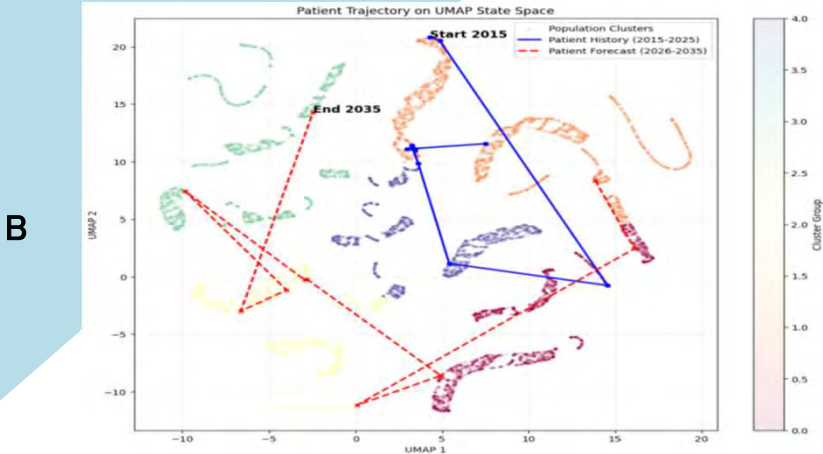
3. 学習用データのグループ分け
学習に用いた仮想患者をUMAPによりデータドリブンにグループ分けした。これを、予測データから助言生成する際に利用した。

4. 学習後Transformerを利用した将来予測の実行および軌跡の描画
学習後Transformerを利用し、学習に用いていない1人分の新規のデータ11年間分を作成し、その後の10年分の予測を行った。

出力例 – 健康情報データの移り変わり –



A.新規に11年分のデータを元に10年分の予測を実行



B.予測された結果がどの患者グループに属するかtrajectoryを推定

出力例 – 健康アドバイス –

予測データを元に健康アドバイスを提示する機能をLLMを介して実現。LLMに対してはRAGを用いる方法のほか、ローカルでの利用を想定して、low-grade ModelへのFine-tuningによる精度確保を行なったSLMを準備可能です。

 健康リスク評価

=====

△ 特定された主要リスク：がん（悪性新生物）（Var1）
（レベル：平均より0.27標準偏差上） ...
最新ガイドラインをAIデータベースで照会中 ...

高リスクがん患者向けの、食事・運動・生活習慣改善に焦点を当てた優先順位付け済み実践的推奨事項トップ10を、ご要望の形式で提示します。

がん（悪性新生物）予防のためのトップ10推奨事項

- 1.健康的な体重の達成と維持：肥満は複数のがんの重大なリスク因子です。食事と運動による持続可能な減量に注力してください。BMIを健康的な範囲内に保つことを目標とします。
- 2.植物性中心の食事を優先：果物、野菜、全粒穀物を重視。赤身肉と加工肉の摂取を減らす。多様な栄養摂取のため、食品の多様性を優先。
- 3.定期的な身体活動：週に少なくとも中強度150分、または高強度75分の有酸素運動に加え、週2回以上の筋力トレーニングを目標とする。継続的な運動が重要。
- 4.すべてのタバコ使用を中止する：喫煙はがん予防可能な最大の原因です。喫煙や無煙タバコの使用を直ちにやめるための支援とリソースを求めましょう。受動喫煙を避けてください。

特願2024-132109
「大規模言語モデルによるインタラクティブ薬剤情報提供マシン」
正確な知識とインタラクティブなコミュニケーション機能を備えた”専門家”モデル構築技術で出願。現在の主流モデルよりパラメータ数の少ないモデルをベースに必要な専門知識を事後学習させることで正確性を実現。2024年度大学見本市に薬剤情報モデルの例で出展しました。

特願2025-106097
出願者
高橋昂也・座間快人・柴田陽平・遠藤空・広井賀子
情報処理装置、情報処理システム、方法およびプログラム

共同研究開発を希望しています。
お問合せ先：研究推進機構
E-MAIL: liaison@kait.jp
TEL：046-291-3109

