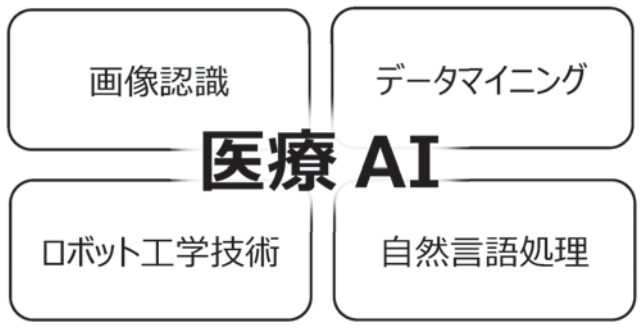
医療における情報（薬剤）の標準化を考える⑬

JA北海道厚生連網走厚生病院　佐藤 弘康

薬剤情報×AI時代における薬剤情報

はじめに

チャットGPTが社会的に話題となっており、AI（人工知能）は身近なものになりつつある。AIに単一的な定義はなく、領域や環境により様々なものがAIと呼ばれている。医療・医学においても近年、AIの利活用が進んでいるが、大域的には「画像認識」「データマイニング」「ロボット工学技術」「自然言語処理」等に分類されると考える。これらは薬剤師業務や薬学研究にも活用され始めており、今後ますます導入が進むものと思われる。



医療AIの主な分類（筆者の私見）

画像認識

現時点において最も発展している医療AIの領域と思われる。深層学習や畳み込みニューラルネットワーク等の登場により、専門医と同等あるいは上回る判定精度の画像診断支援が可能となってきており、眼底画像や内視鏡画像など、既に複数のAI搭載医療機器が国内においても承認されている。

高精度な画像認識AIを構築するためには、事前学習が重要であり、そのためには質の良い教師データが必要である。AI学習の基本課題としてよく用いられるイヌネコ分類においては、用いる学習用画像に正解ラベル（それがイヌであるかネコであるか）を付与すること（アノテーション）は容易である。

しかし、医用画像のアノテーション（例えば、その隆起が悪性であるか良性であるか）は誰にでもできるものではなく、大量な学習データの準備は容易ではない。誤ったアノテーションの画像を含むデータで学習したAIは、誤答も学習してしまう。

薬剤領域においては、薬剤画像からの医薬品特定などの用途が考えられる。薬剤画像データは用意が比較的容易であり、アノテーションの難易度も高くない。

一方で、実臨床での活用のためには販売終了品目を含めた網羅的な薬剤画像データが必要となる（イヌネコ分類AIにゾウやキリンの画像を入力しても、必ずイヌかネコかのどちらかで回答し、学習させていない回答をすることはない）

1社の製薬企業やAI開発企業では、これらの網羅的な薬剤画像データセットの用意は困難であり、公的機関あるいはそれに準じたコンソーシアム等の関与が必要と考える。

データマイニング

病歴や臨床検査値、処方データ等の医療情報から、疾病の発症や重症度、あるいは予後を予測する機械学習モデルが数多く開発され、論文公表されている。多くの報告では、複雑な条件下における予測モデルの構築において機械学習は従来の統計モデルの判定精度を上回っている。

薬剤情報領域への活用として、著者らは経口ステロイド薬の誤用量処方を判定する機械学習モデルを構築し、報告している。電子カルテや調剤支援システムには、処方監査支援機能が備わっている場合が多いが、その用量チェックのアルゴリズムは単純なものであり、幅広い臨床用量を有する薬剤への有用性は限られている。

われわれの報告では、判定精度の向上のために前回処方情報や診療科情報なども用いたが、様々な病院規模の施設でも活用可能なよう診療科のカテゴリー化も実施した。

薬剤情報へのAIの活用には診療科情報のような間接的な情報についても標準化が必要となる。近い将来、処方監査システムへAIが導入されることは、もはや自然の流れであり、どのような仕様、条件で動作しているのかを把握した上で利用することが必要である。

ロボット技術

介護支援パワードスーツや自動運転車いす、内視鏡手術支援ロボット等のロボット工学技術も医療への応用が進んでいる。薬剤師業務関連では、注射薬自動払出システムや抗癌剤調製ロボットなど調剤機器へのロボットの導入が進みつつある。

今後も技術発展、研究開発は行われ続けると考えるが、これらは緻密なパーツと複雑な制御を必要とするため高額であることが多く、商用活用の範囲は限られる可能性がある。

自然言語処理

テキスト情報の解析は古くから行われていたが、大規模言語生成モデルの登場により、非常に注目されている領域である。医療テキストデータには、診療録や各種指導記録・レポート等の大量なテキストデータが存在し、これらを解析することで有用な知識を創出しようとする領域である。

医療テキストには医薬品の情報も多く含まれるが、剤型や規格量までを含めた正式な承認医薬品名称が記述されていることは少ない。体系的分類や統計処理に用いるためには、文章中に登場する用語が何の医薬品に関する記述であるのか特定する必要があり、そのような研究も進んでいる。

一方、オープンAIが開発したチャットGPTは、3550億パラメータ数で学習した大規模言語生成モデルであり、対話形式であるため一般人にも利用しやすく、非常に話題となっている。

言語生成モデルは、他にも数多く開発されており、医療分野に特化したモデルでは、医師、看護師、薬剤師をはじめとする多くの職種の試験の合格基準に既に到達していることが報告されている。

言語生成モデルは、学習したデータから最も可能性のある文章を生成するものであり、必ずしも正解を回答するものではなく、100％の正解率とはならない。また、誤答であっても自然で違和感のない文章を回答する。そのため、教育分野や生命や健康に関与する医療分野での活用方法には熟慮が必要である。

おわりに

医療AIは間違いなく今後発展し、実臨床でも多くのものが実装されてくると思われる。そして、その進歩は近年非常に速まっている。高性能なAIの構築には、より正確にアノテーションされた巨大なデータセットが必要となり、その素材となるデータは、われわれ医療従事者の日常業務から生み出されている。質の高い処方データや指導記録の蓄積が、将来の医療の質向上に寄与するという意識が重要である。