

第6講 機械翻訳の新時代－トランスフォーマー革命と「生成A I」の驚異的進化

【学習到達目標】

- ・トランスフォーマー革命が生成A Iを誕生させたことを説明できる。
- ・トランスフォーマー革命により機械翻訳精度が向上したことを説明できる。
- ・A Iが産業、医療や農業等に应用されていることについて事例を挙げて説明できる。

1. トランスフォーマー革命

第4次A Iブームの火付け役がトランスフォーマー革命です。2017年、トランスフォーマーが、Google社で開発されました。機械翻訳システムのために発明されたものですが、今日生成A Iを開発するための基本技術となっています。

・A Iブームの歴史と予想

歴史的にみれば、人工知能というのは今まで3回、ブームを迎えています。第1次A Iブームは1956年検索推論の時代で、ゲーム等が作られました。それはビジネスにならないというので1回目の冬の時代を迎えます。

第2次A Iブームは1980年知識の時代で、エキスパートシステムや機械翻訳システムが作られました。しかし、やはりビジネスにならないというので冬の時代を迎えました。

第3次A Iブームは2013年からで、機械学習・表現学習の時代に入りました。画像認識処理にCNN、自然言語処理にRNNといったニューラルネットワークが大活躍するようになりました。

第4次A Iブームは2017年CNNやRNNを使わずに機械翻訳で最高性能を更新する革命的なトランスフォーマーが出たことによって到来しました。アテンション（Attention、注意機構）の時代と言うべきかもしれません。

この後、2030年頃から雇用の大崩壊が起こるのではないかとA Iが労働者に置き変わるのではないかとされています。

それから、シンギュラリティです。2045年にはAIが人類超えるのではないか？というようなことまで言われているわけです。

さて、ある機械がAIであるかどうか判断する1つの方法は、チューリング・テストと言われています。チューリング・テストはアラン・チューリングが最初に考案したことで、壁の向こうに人と機械があり、同じように応答するようであれば、機械は人工知能AIであると言えるということです。最初に有名になったのが、第1次AIブームのイライザです。イライザは、生成AIの元祖「チャットボット」（人工無能）です。精神分裂症の医者（AI）が、パラノイアの患者（人）と話しします。イライザが患者に話かけ、人がそれに対して答えます。今日でも「イライザ効果」として名前をとどめています。イライザ効果とは、自然な会話を行う機械を目の前にした時、あたかも本物の人間と対話しているかのような錯覚に陥ってしまう現象のことを言います。

次に登場したのは第2次AIブーム「MYCIN」というAIで、専門家システムです。MYCINの専門医の知識は、「もし～何々ならば」という「IF THEN」ルールの形で記述されていました。

第3次AIブームでは、「コンピュータが目を持った」といわれるように、CNN（畳み込みニューラルネットワーク）を使った画像認識のAIが注目を浴びました。自然言語処理では、RNN（再帰型ニューラルネットワーク）を用いたニューラル機械翻訳システム（NTM）が開発され使用されるようになりました。しかし、RNNを用いたデコーダ（生成器）には、過生成、生成不足、原構造無視などの誤動作が発生しました。この解決法として、色々な試みの中からトランスフォーマーが発明されました。

・第4次AIブーム

2023年、宿題、レポート、コンピュータプログラム、小説や論文を書いたり、翻訳する生成AI（チャットボット）であるChatGPT, BardやBing AI等が登場しています。イライザに始まり生成AIに至るチャットボットは、人間と対話（チャット）するロボットです。

・生成AIの驚異的進化

トランスフォーマーを用いたBERTやGPTが登場し生成AIの歴史が始まりました。BERTやGPTの「T」は、トランスフォーマーの頭文字に由来しています。

第4次AIブームは、Google社のトランスフォーマーの論文及び特許から始まり、図の上側のGoogle社のBERT、LaMDA、BardやGeminiに至っています。

一方、オープンAI社は、Google社を師匠にしてGPT1、GPT2、GPT3を経てGPT3.5に発展させました。生成AI「チャットGPT」を2022年公開し、2024年チャットGPTをGPT4で運用しています。

他方では、Microsoft社が、オープンAI社に巨額の出資し、チャットGPTを自社のBing AIに搭載するようになったというのが2023年前半までの状況です。

一般的に、翻訳や文書分類、質問応答といった自然言語処理における仕事の分野を「(自然言語処理)タスク」と呼びます。

生成AIの歴史を見てみると、

2017年「Transformer」：Google社開発したニューラルネットワークアーキテクチャーです。2017年にオープンソース化しました。

元祖、大規模言語モデル(LLM)です。

2018年「BERT」：Transformerによる双方向のエンコード表現で、自然言語処理タスクの検索エンジンで活躍しました。

2020年「GPT」シリーズ(Generative Pre-trained Transformer)：OpenAI社が開発しているもので、事前学習済みの文章生成型のTransformerです。2018年、2019年、2020年、2022年、2023年と大規模化しています。

2021年「LaMDA」(Language Model for Dialogue Applications)：対話アプリケーション用で自然な会話を実現する大規模言語モデルです。

2022年「ChatGPT」(チャットGPT, Chat-based Language Model for OpenAI's GPT)：対話システムや教育支援などで自然な会話を実現する大規模言語モデルです。

2023年「Bing AI」(Bing AI Chat、Copilot)：OpenAI社のGPT4を用いた、自然言語処理技術を用いた検索エンジン型のAIです。

2023年「Bard」(Google Bard,正式名はLaMDAと同じ)：対話システムや教育支援などで自然な会話を実現する大規模言語モデルです。

2024年「Gemini」(Bard改名)画像生成AIとの合体型です。

2017年以来、生成AIは驚異的な進化をとげています。トランスフォーマーの本質は機械翻訳のためのニューラルネットワークアーキテクチャーです。

第3次AIブームで一番有名なのはGoogle機械翻訳です。Google社で機械翻訳の精度をさらに向上させるために、何か方法がないかということで発明されたのが、注意機構メカニズムとエンコーダ・デコーダモデルです。

第4次AIブームを引き起こした論文の題名は「あなたに必要な全ては注意機構のみ」（出典：Ashish Vaswani et al. "Attention Is All You Need"(2017), <https://arxiv.org/1706.03762>)

トランスフォーマーは、注意機構メカニズムを介してエンコーダとデコーダを接続します。トランスフォーマーの特長は次の3つです：

- 1) CNNやRNNと異なるモデル（エンコーダ・デコーダモデル）です。
- 2) 「注意機構」メカニズムのみに基づいています。
- 3) 機械翻訳のためのニューラルネットワークアーキテクチャーです。

トランスフォーマーは、機械翻訳システムの翻訳精度を競うコンテスト「WMT 2014」で、最高得点を獲得しました。具体的には、英語からドイツ語への翻訳タスクで28.4 BLEUを、英語からフランス語への翻訳タスクで41.0 BLEUを達成しました。これは、人間の専門家でも40.0 BLEU程度なので人間を超えたと言えます。このトランスフォーマー・モデルの成果は、8台のGPUで3.5日間トレーニングして得られました。

・トランスフォーマー機械翻訳の発明

第4次AIブームのトランスフォーマー機械翻訳（TNT）の誕生です。事前学習済のビッグデータを用い、エンコーダ・デコーダ方式で機械翻訳を行います。トランスフォーマーは、第3次AIブームを牽引してきたRNNやCNNを使わずに、機械翻訳精度を当時の最高得点を獲得したことから、大きな注目を集めることとなりました。このような経緯でトランスフォーマーによって大規模言語モデルLLMと生成AIが登場するようになりました。最近では、GPT4-Turboや国産生成AIが登場し始めています。

・トランスフォーマーの学習方法

トランスフォーマーの学習方法は訓練フェーズと推論フェーズに分かれます。まず訓練フェーズでどんどん大量のデータを教え込みます。例えば、英語の「How is the weather Today」は、スペイン語では「Qué tiempo hace hoy」になると教え込んでいきます。

フランス語だったりスペイン語だったり、どんどん教え込ませ、そして次に来る語（トークン）は何になるというような穴埋めのメカニズムでどんどん覚えさせていきます。推論フェーズでは、それに対して、穴埋めしながら生成してくれるというような学習方法になっているわけです。

・アテンション

アテンションは何かというと、注意する場所を特定するということです。例えば、元画像の猫を見て、目と鼻に注目する場合、アテンションスコアをかぶせて新しい表現として取り出します。これが自己注意機構（Self-Attention：セルフ・アテンション）です。穴埋めさせるという意味でこれを「自己教師あり学習」とも言います。

・自己注意機構で文脈理解

トランスフォーマーでは、入力が入ってくると、それぞれトークンというテキストの最小単位に分割して、それをベクトル化してトークンベクトルにします。トークンベクトル毎に、アテンションスコアというものをつけます。

類似度の値が0～1の範囲になるように、ソフトマックス関数でスケーリング乗算を行い、最終的に重み付けたトークンベクトル作成し、文脈を意識したベクトルが取り出せるという仕組みです。

例えば1)「The animal didn't cross the street because it was too tired.」と2)「The animal didn't cross the to street because it was too wide.」では、1)の「it」が「animal」を2)の「it」が「street」だという文脈処理できます。

第4次AIブームを引き起こしたトランスフォーマーには、エンコーダ・デコーダ、エンコーダのみ、デコーダのみの3種類の構成が存在します。トランスフォーマーでは、エンコーダによるテキストのベクトル表現化（入力された埋め込み）に対して、キー（KEY）、クエリ（query）、バリュー（value）の3つの異なる埋め込み（ベクトル表現）を計算します。

・生成AIの教育利用

2023年7月4日永岡桂子文部科学大臣記者会見で、生成AIの学校での利用についての文部科学省の暫定的なガイドラインが公開されました。

文部科学省の暫定的なガイドラインの内容は以下の通りです。

- ①発達段階を踏まえた、事前に生成A Iの性質を十分理解させることが可能かを見極めた上で活用する。
- ②「教育活動の目的を達成する上で効果的か否かで判断すべき」との考えから、
 - 1) 活用が適切でない例、
 - 2) 活用が考えられる例の両方を示している。
- ③今後、懸念やリスクに十分対応が講じられる一部の学校で、パイロット的取組を実施し、成果・課題を検証する。というガイドラインです。

「小学校と中学校の授業での生成A Iの活用法を教えてください」と、チャットGPTとBardについてそれぞれ質問しました。

チャットGPTの回答を要約すると、「授業で様々な活用法があります。1) 質問応答のサポート、2) 文章の添削や修正、3) 創造的なライティングのサポート、4) 計算や数学のサポート、5) 外国語の学習 ツールとしての利用に適しており、教師や指導者の役割を置き換えるものではありません。」でした。

Bardの回答を要約すると、「例えば BARD は次のようなことができます
1) 生徒の質問に答える、2) 生徒に創造的なコンテンツを書くように依頼する
3) 生徒に言語を翻訳させる、4) 生徒に知識をテストする、5) 生徒が研究を行うのに役立つ、6) 生徒に個別化された指導を提供します、7) 授業を豊かに強化する強力なツールです」でした。

違いの1つは、チャットGPTは1つの回答ですが、Bardは3通りを出してきます。ちなみに、チャットGPTを使ったBing AIは、説明の根拠として3つのURLを提示しています。

2 機械翻訳精度の向上

コンピュータの機械翻訳精度は着実に向上を続けています。これはチューリングテスト・レベルの課題「そのテストで人間並みの結果を出すには、人間並みの完全な言語理解を要求される」ので、人間の能力と競い合う最後の用途のひとつです。南カリフォルニア大学のコンピュータ科学者、フランツ・ヨーゼフ・オッホは、どんな言語間でも、わずか数時間から数日で新しい翻訳システムを作りだせる手法を開発しました。必要なのはいわば「ロゼッタストーン」です。ある言語の原文テキストと、それを別の言語に翻訳したものだけなのだが、システムを作るには数百万語分の翻訳された文章が必要となります。

これは、言語学者が苦心しながら、やたらに例外がある文法規則をコード化して作っている機械翻訳システムとは、じつに対照的です。オッホの機械翻訳システムはアメリカ商務省の国立標準技術研究所が実施している翻訳システムコンテストで当時の最高得点を記録しました。

・RNNLM型デコーダによる自然言語生成技術

人間の日常的な会話で発生する自然言語生成では、常に厳密な文の構造を意識しながら単語の選択を行っているとは考えにくいといえます。話者が話そうとする情報に基づいて、前後の単語同士が流暢に接続するよう順に単語を選択しているととらえるのが妥当です。この動きを実際の文生成システムに当てはめて考えると、入力データと初期状態を受け取り、出力文の単語を逐次的に出力するオートマトン（自動機械）の一種としてとらえることができます。このオートマトンに必要となるのは、現在のシステムの状態に基づき、出力単語列の言語としての流暢さを考慮した上で、次回の出力として最も妥当な単語を決定する仕組みです。この仕組みは「言語モデル」と呼ばれます。

言語モデルの本質は「過去の入力系列から次回の出力単語の分布を求める確率モデル」です。リカレントニューラルネットワーク言語モデル「RNNLM」(Recurrent Neural Network Language Model)はその代表的なものです。

・エンコーダ・デコーダモデル

ニューラルネットワークにより構成されたエンコーダをデコーダに接続することで、エンコーダとデコーダ双方を単一のニューラルネットワークと見なして同時に学習することが可能となります。このように、中間データを介してエンコーダとデコーダが対向する形で構成されたニューラルネットワークは、「エンコーダ・デコーダモデル」と呼ばれます。

最も初期のエンコーダ・デコーダモデルとして、エンコーダにもデコーダとほぼ同様のRNNを備えたエンコーダ・デコーダ機械翻訳モデルが挙げられます。

これに注意機構を加えたモデルは、現在の機械翻訳システムの主流の構成法であり、2014年頃から著名な機械翻訳コンペティションにおいて、次々と最高精度を達成し続けています。2017年にはトランスフォーマー機械翻訳システム(TMT)が発明され今日に至っています。

Google社が2016年に発表した機械翻訳システム「GNMT」(Google Neural machine Translation)で、例えば「I have an apple」を「私はりんごを持っている」に翻訳する場合、まずエンコーダが英単語「I」「have」「an」「apple」を順に読み込み、中間データである初期状態ベクトルを生成します。デコーダは初期状態ベクトルを受け取り、自身の状態遷移にしたがって単語列「私」「は」「りんご」「を」「持つ」「て」「いる」を生成します。このモデルの場合、初期状態ベクトルが入力文に関する情報を全て蓄えていることは明らかです。

エンコーダ・デコーダ機械翻訳モデルは入出力がともに可変長の単語列であることから「Sequence to Sequenceモデル」とも、「End to Endモデル」とも称することがあります。

様々な入力に対して有効性が期待されるエンコーダ・デコーダモデルですが、従来型の自然言語生成システムには存在しなかった特有の問題が複数存在することも報告されています。主な問題は「過生成」、「生成不足」や「原構造無視」です。

過生成とは、デコーダの状態遷移が何らかの不動点に陥ってしまい、例えば「りんご」という同じ語を繰り返し、出力し続けることです。

生成不足とは、逆に、本来必要な状態遷移の一部をスキップしてしまったがために、入力として与えられた情報の一部のみが自然言語として出力されてしまうことを示します。

原構造無視とは、誤動作で原文の構造を無視して出力することをいいます。

これら問題がどんな条件で発生するのは今のところ詳しい解析がなされていません。モデルの構造を工夫することによって、間接的に回避するための手法も研究されていますが、今後の更なる研究が必要です。

3 産業におけるAI活用

・経団連のAI活用戦略

「AIの活用戦略」ということで、2019年2月19日経団連が「AI活用戦略」と題して発表したもので、経団連でまとめたAI Ready化を測るアンケートもあります。AI Readyな企業の5段階レベルの何処にいるかを自己診断できるようになっています。

・ものづくり領域における応用事例

産業（ものづくり）分野においては、さまざまな工程で自動化が進められてきました。一方、現状においても一部「人が持つ能力・感覚に頼った工程」が存在します。このような工程では、担当者ごとに習熟度の差が存在することや、担当者への負荷が高くなるという課題があります。この課題の解決を目指して「ディープラーニング等技術を活用したアプローチ」が試され始めています。

・自動車ギヤ不良品検出

製造業では、そもそも不良品が発生する頻度が少ないことから「不良品に関するデータを学習させること」自体が難しい状況にあります。次の事例では「（大量に準備することができる）良品データのみから特徴を抽出し、その特徴との差分を利用することで不良品を検出」するアプローチを採用することにより、「不良品データが少ない環境下」での不良品検知を実現しています。

【事例】2018年2月からの武蔵精密工業株式会社と株式会社A B E J A 取り組みはその一例です。自動車部品(ベベルギヤ)完成品に対する自動検査への応用となっています。

①オートエンコーダをはじめとした複数のロジック手法を組み合わせ、良品データのみから不良品を判断という仕組みです。

②人による目視検査に近い精度を実現。人に頼らざるを得なかった検査工程を自動化することにより、生産効率の大幅な向上を目指しています。

4 医療領域における応用事例

医療分野では、主に「診断支援」・「創薬」・「ゲノム解析」領域にて、ディープラーニング技術が活用されています。"ごく早期"のがんは医師でも見つけにくく、見逃してしまうおそれがあると言われていました。ディープラーニングが持つ特徴抽出能力を、このような診断支援領域に適用することにより「医師への負担緩和および見落としリスク低減」を目指す取り組みが広がっています。

・胃がんを検出する内視鏡画像診断支援

慢性胃炎の中から胃炎に類似している胃がんを拾いあげるとは、経験を積んだ医師でも難しいケースがあると言われていました。

次の事例は「内視鏡画像からの胃がん検出」領域において、（医師の診断を支援する用途としてCNNを活用している試みです。

【事例】2018年1月からの公益財団法人がん研究会有明病院と株式会社AIメディカルサービスの取り組み事例です。CNNを活用し、内視鏡画像から胃がんを検出するというものです。

- ① 6mm以上の胃がんを熟練の内視鏡医レベルに匹敵する精度で発見します。
- ② 1画像の診断にかかる時間は約0.02秒であり、人間の解析速度を遥かに凌駕しました。
- ③ 本事例発表時においてAIを活用した胃がん内視鏡診断支援システムの報告はなく、本報告は世界初の快挙です。

・大腸内視鏡検査支援

「大腸の内視鏡検査」支援事例です。内視鏡医がポリープを肉眼で検査する際、「サイズが小さいもの」や「形状が認識しにくいもの」は見逃されるケースがあるとも、言われています。次の事例では、ディープラーニングを活用した技術等を用いることにより、「内視鏡医による病変発見をサポート」するアプローチが採られています。

【事例】大腸の内視鏡検査時に撮影される画像に、深層学習を活用したAI技術等を用いることにより、大腸がんおよび前がん病変を自動検知（内視鏡医の病変の発見をサポート）します。2017年7月から、国立研究開発法人国立がん研究センターと日本電気株式会社/国立研究開発法人科学技術振興機構と国立研究開発法人日本医療研究開発機構が実施した事例ですが、（2017年7月）前がん病変としてのポリープと早期がんの発見率98%でした。動画各フレームにおける検知と“結果表示”を約33ミリ秒以内(30フレーム/秒)で行うことに成功しましたこれらの情報は、医師に対して、リアルタイムにフィードバックすることが可能です。

・広角眼底画像を用いた網膜はく離判定支援

次の事例は「診断補助」領域の「眼底画像を用いた網膜はく離判定」です。

【事例】2017年8月から社会医療法人三栄会ツカザキ病院眼科と（株）Ris tの取り組み事例です。

- ① 失明に直結する疾患である網膜剥離眼と正常眼を、網膜の広角画像のみから判別します。

②広角眼底画像を用いたA Iによる網膜剥離判定A U C 9 8%の精度でした。広角眼底画像とは、健診などで主に用いられている従来型非散瞳眼底カメラの4倍以上の画角を持ち、わずか2mmの自然瞳孔下から眼底の80%のエリアを撮影可能です。

③診断見落としを防ぐ二重チェックシステムへの応用や、眼科医療過疎地に対する遠隔診療への応用、健診による早期発見などへの応用を想定しています。

5 農業における応用事例

日本社会全体で高齢化が進む中、農業分野においても「今後、数年の内に「離農」する農家が増加」してくることが予想されています。

・収穫・仕分け支援

農業における各プロセスの中でも、「収穫・仕分け作業」は、作業者に大きな負担が発生する作業であり、「労働力不足」が課題にあげられる中、当該作業を支援する収穫ロボットの研究が進められています。

【事例】2018年5月からのパナソニック株式会社の取り組みです。

トマト収穫ロボットに、ディープラーニングを用いた画像認識技術を採用しています。重なり合って見えにくいトマトの検出等、検出率向上に寄与しています。果実検出用に、農園で撮影した画像から果実画像を2,000枚準備し、「学習済みP V A N e tモデル」をファインチューニングして、検出モデルを作成しました。果梗(かこう、果実の柄になっている部分)認識には、セマンティック・セグメンテーションアプローチを採用しています。「画素ごとに果実、主茎、果梗、その他背景の分類にラベル付けを行った1,050枚の画像」を用いて学習しています。

・ピンポイント農薬散布

害虫駆除を目的として、畑全面に農薬を散布する際に(害虫は畑に「点在」しているため)必要のない箇所にも散布してしまっていることがあります。散布する農薬を減らすことが出来れば、生産コストの抑制にもつながることから、次のような取り組みが行われています。

① 大豆畑上空においてドローンを飛行させ、畑全体を撮影します。(マルチスペクトル撮影機活用)

②大豆の害虫である「ハスモンヨトウ(蛾)の幼虫」が居そうな位置(注)を探し出し、農薬を当該箇所にピンポイントで散布します。

【事例】2017年12月からの株式会社オプティムの取り組み事例です。結果、通常栽培と比較して使用する農薬量を1/10以下に抑え、収量・品質を従前までと同等とすることに成功しました。

(注)ハスモンヨトウの幼虫は、日中、葉の裏に隠れており、上空から直接は検知できないが、「大豆の葉をハスモンヨトウの幼虫が食べるに従って、葉の色が変色する」ことを利用することで、害虫が存在することを検知しています。

課題

AIが作った小説、論文、楽曲やプログラム等の著作権はAIにあるか？について考察し、あなたの考えを800字で説明しなさい。