

▢ Bib20-04人工知能に関する基礎知識（まとめ）

■ 【2018年2月20日】

■ 要約

▢ 人工知能とは

- 人間の脳のように物事を学習したり、意志決定したりする能力を持つコンピュータのプログラム

▢ 脳の神経細胞

- 構成単位であるニューロン
- 脳は複数のニューロンにまたがって分散的に情報を保持し、また、新しい情報が入ってきたときは脳全体で学習し、全体で保管しているのではないかとされている

▢ 人工知能の歴史

▢ 1940～60年代、第1次AIブーム

- 1943年脳のニューロンの工学的模倣として、形式ニューロンが提案された
- 1956年、初めてAIという言葉が用いられた。
- 単純な問題は処理できても、多くの要因が絡む課題を解くことができず、ブームは終焉

▢ 1980年代、第2次AIブーム

- 従来型の記号処理を発展させて、推論機能を持たせようとした論理プログラミングで始まった
- 論理を表現する推論マシンの開発を柱として
- 今まで、学習のための情報が少なく、また、コンピュータの処理能力が低かったため、広くAIが普及できなかった
- ニューロン系の仕組みが再び注目され始めた。ニューロンの中間層（隠れ層）を1層追加した「3層ニューラルネット」が誕生
- ニューロコンピュータ

▢ 多様な推論を行うエキスパートシステム


- ▢ 問題解決を行う専門家の知識を場合分け規則などの形でコンピュータに実装
 - 人間が機械に理解できる推論の情報を記述する必要があった

▢ ナレッジエンジニアリング



- ルール（規則）ベースの「人工知能」に人の脳内にある知識を移植する役割とされた
- 小規模：専門家の知識をプログラミングしたもの（エキスパートシステム）で、人間がコンパクトに自分の常識を一断面をコーディングした概念体系。第2次AIブーム

▢ AIが実用レベルに達したのは

- 今のAIは人間の能力を補完できる部分が多くなった

- 同じ基本機能を用いて対象を取り換えて、正解データを作り、
だけで、新たなプログラミングなしで、精度向上をさせることができる
- 2011年音声認識コンテスト
- 2012年画像認識コンテスト
- 2016年9月末、Facebook, Amazon, Google, IBM, Microsoftの5社が、AIに関して歴史的な提携を発表
- <http://gigazine.net/news/20160929-partnership-ai-facebook-amazon-google-ibm-microsoft/> 
- 基礎となる学問分野・情報科学の進展
 - 世の中にあふれている情報を全て把握し、これを図書館学的立場から収集、蓄積、利用するということになれば、下記のような情報に関係する様々な学問を身につける必要がある。【長尾先生】
- A：自然言語分野
 - ・形態素解析、重要語の抽出、シソーラス・オントロジーの作成、かな漢字変換、固有名詞・未知語の認識、辞書学
 - ・クラスタリング、分類学
 - ・構文解析、類似意味の文・文章の同定、否定文の解析と対立する肯定文の同定
 - ・人とシステムとの対話、文生成
 - ・テキストマイニング、文章分析、自動要約、機械翻訳、質問応答システム、対話システム
- B：音声・音楽分野
 - ・音声分析、音楽分析、アクセント、イントネーション抽出、ポーズ区間の検出、メロディの抽出
 - ・音声認識、音楽から楽譜へ
 - ・音声合成、音楽生成
- C：画像・映像分野
 - ・パターン認識理論、特徴抽出理論、文字認識、画像ディジタル化・圧縮技術、スペクトル分析
 - ・画像処理・変換技術、歪補正技術、画像認識・理解、画像検索技術、電子透かし技術
 - ・映像処理、映像圧縮、映像認識技術、部分映像検索技術
 - ・マンガ、アニメ生成技術、ゲーム
 - ・コンピュータグラフィックス
 - ・マルチメディア技術、ヒューマン・インターフェイス
- D：コンピュータ・ソフトウェア、情報通信
 - ・プログラミング、ソフトウェア、システム、OS
 - ・データベースの理論と実際
 - ・ソフトウェア工学、デバック技術、ゲームソフト

- •ネットワーク・ソフト、ウェブ技術、検索エンジン技術
- •暗号理論、セキュリティ技術
- E：知識工学、人工知能
 - •知識表現、セマンティック・ウェブ
 - •辞書学、百科辞典学
 - •編集工学
 - •推論技術
 - •エキスパートシステム、問題解決、学習
 - •著作権、知的所有権、クリエイティブ・コモンズ
- F：図書館学、図書館情報学
- ビッグデータが利用可能に
 - 大量のデータが発生、流通し、手軽に使えるようになった
 - インターネット上での玉石混合の情報に加え、信頼性の高い情報がオープンデータとして利用可能になってきた
 - インターネット上に分散するデータ群を、あたかも一つのデータベースに見せる技術も発達
- ビッグデータを活用するためにはAIが必要。
 - データ・マイニング
 - パターンやデータ間の相互関係を発見する
 - データ・アナリティクス
 - 対象とする情報についての何らかの結論をいくつか引き出す目的で生データを検査・分析する科学的手法
 - 推論を行い、意思決定を行う
- コンピュータの高性能化
 - CPUのパワーが前回の人工知能ブーム時の何千倍、何万倍になった
 - 単純作業の反復、総当たりチェックする能力が指数関数的に向上
 - スパコンを利用しなくても、クラウドサービスレベルのコンピュータで大量の情報の試行錯誤的ぶんせきが可能になった
- 専門家の知識だけでは過学習になる可能性があるが、ビッグデータの分析から得られた実績値による知識を自動生成
- それにより、ディープラーニングが可能になった
- 2020～2030年代
 - 弱いAIにより、既存の仕事の一部がなくなる。AIが幅広く普及
- 2045年
 - シンギュラリティ
- 従来のシステム開発

- モデル化、アルゴリズム化が可能で大量の情報を正確に処理する
を利用した情報処理システム
 - 人がモデルを想定して最適なアルゴリズムを検討し、プログラミング
 - ディープラーニングは、モデル化作業、プログラミングが不要
- 人工知能の振る舞いの段階によるカテゴリー分け
- 総務省「インテリジェント化が加速するICTの未来像に関する研究会 報告書2015」
P.13 
- <従来の人工知能>
- カテゴリー1：単なる制御（言われた通りにやる）
- 【アルバイト】言われたことだけをこなすレベル
 - -温度が上がるとスイッチを入れる。下がるとスイッチを切る。
 - -洗濯物の重さで洗い時間を調整。
- カテゴリー2：対応のパターンが非常に多い
- 【一般職員】たくさんのルールを理解し判断するレベル
 - （探索や知識を使って、言われた通りにやる）
 - -探索や推論。将棋や囲碁で、決められたルールにしたがって、手を探す。
 - -知識。例えば、与えられた知識ベースを使い、検査の結果から診断内容や処方する薬を出力する
- <以下が第3次人工知能ブームの主たる対象>（※著者注）
- カテゴリー3：対応のパターンを自動的に学習（重みを学習する）
- 【課長クラス】決められたチェック項目に従って業務の改善をしていくレベル
 - -機械学習
 - -駒がこういう場所にあるときは、こう打てばよいということを学習。
 - -この病気とこの病気はこういう相関があるということを学習。
- カテゴリー4：対応のパターンの学習に使う「特徴量自体」も学習（変数も学習する）
- 【マネージャクラス】チェック項目まで自分で発見するレベル
 -  -（特徴）表現学習。ディープラーニングはこの一種
 - -駒の位置だけでなく、複数の駒の関係性をみたほうがいい。
 - -こういった一連の症状が、患者の血糖異常を表し、複数の病気の原因になっているようだ。
- かつて人工知能と呼ばれていたものが実用化されひとつの分野を構成すると、人工知能と呼ばれなくなる「AI効果」【人工知能の名付け親であるジョン・マッカーシー】
- 人工知能の分類
- 「強いAI」、「弱いAI」
- 強いAI

- 人間の脳と同じような振る舞い、原理の知能を作る
- 全くの未知の事態に遭遇して、創造的に問題解決できるという知識の定義をクリアする本物のAI。現場で自分で知識を獲得しながら、人類初めての対処法を考え出すことができる
- 弱いAI
 - 人間の能力を補佐・拡大する仕組みを作る
- 「大規模知識・データ」、「小規模知識・データ」
 - 小規模：専門家の知識をプログラミングしたもの（エキスパートシステム）で、人間がコンパクトに自分の常識を一断面をコーディングした概念体系。第2次AIブーム
 - 大規模：ビッグデータを投入することで知識を得る。第3次AIブーム
- 「専門的」、「汎用的」
 - メタ知識：知識獲得・知識創造のための知識
 - 汎用のAI：メタ知識をもって、未知の事態にも、ある程度対応できるAI,汎用の学習能力をもったAI
 - 専用のAI：例えば、草花の名前であれば何百種・何千種から判定できるような専用の画像認識エンジン
- 従来までの人工知能
 - コンピュータへの命令としてプログラムを書くときに人の手ですべてをやるには大変な部分を自動で機械に学習してもらおう
 - 教師データありの学習
- 教師データは、専門家が持つ形式知である知識、暗黙知であるスキル
 - 過学習と対策
 - 過度に教師データに依存した（汎化できていない）状態
 - 原因として、教師データが足りなくてデータに偏りがあるため
- 人が教師データを使ってラベル付けしてきた
 - 大変なスキルと工数が必要だった
- 機械学習とは
 - 様々なデータから予測モデルを作成し、覚えていないものを予測する学習。教師ありと教師なしがある。
- 過学習
 - 例えば、教科書に書かれている例題をいくら正確に説くことができたとしても、ちょっとひねったテストの問題では歯が立たないといった勉強の仕方では意味がない
- 実用化レベルに達した人工知能（人工知能のブレークスルー）
 - AIの学習方法は、幼児教育のアナロジー
- 人間に近い学習過程を辿れるようになった人工知能

- ▣ 豊富な知識の習得
 - 人間は教師に教わって学習することもある。自分で気づいて学ぶこともある。これは人工知能も同じ
- ▣ 迅速で高度な思考
 - 豊富な思考回路（アルゴリズム）を持ち、例外的な処理を含めて、迅速に処理する思考能力
- ▣ ニューラルネットワーク
 - 脳の構成単位であるニューロンを簡易的に模倣して工学的に再現する試み
 - 人間の知能は、問題の解き方をモジュール化して、問題に応じてモジュールを再活用し、また複数のモジュールを組み合わせ活用することができる
- ▣ トップダウンアプローチ
 - 脳の構造そのものを生理学的に解明し、その構造を工学的に再現しようという自然科学寄りのアプローチ。「強いAI」
- ▣ ボトムアップアプローチ
 - 超高速計算や大量のメモリーを駆使して、本来機械が得意だった能力をもっと生かやすく、使いやすくするために、音声認識、自然言語処理、画像認識などの人間的なインターフェース、流儀を機械に学ばせる、実用的なアプローチ。人間の能力をますます強力にして、人間の問題解決を支援する「弱いAI」
- ▣ ニューラルネットワーク
 - 大切なのは情報と情報の関連性。重要な情報は強く、重要でない情報は細く。
- ▣ 多層に重なるニューラルネットワーク
 - 情報に対して「関連性」や「重要性」を踏まえて、情報を扱っていく
 - 入力層、出力層、その中間で実質的な思考を行う部分が「隠れ層」
- ▣ AIが自ら辞書を整備していく
 - 初めは人が
 - その後はAIにより、「局所解」
 - 人がアルゴリズムを選択し、試行錯誤（実行、評価の繰り返し）により最適解を導く
 - 強化学習
- ▣ ディープラーニング
 - ▣ ディープラーニングでの学習を簡単に言うと
 - 生データとその正解ラベルや別の生データの対応関係をトレーニング。人間があらかじめ与えた正解を出せずに失敗したとき時には正解に至る確率を上げるべく、各層間の結合線上の重みを調整するやり方での学習⇒トレーニング
 - ディープラーニングは、原理を生物の脳に求めており、ニューロンとシナプスを模したものがその原型
 - ラベルが付いていないデータを用いて、

- ディープラーニングは、「適切な特徴抽出能力をもつ教師なしワークを多層にして構築したもの」
- 特徴抽出能力とは、
 - 特徴抽出能力とは、何が関連して何が関連していないかを理解できる能力
- 画期的なこと
 - 非常に画期的だったことは、ニューラルネットワークの情報同士のつながりがある意味「自動設定」できるようにしたこと
 - 人間が教えなくても自ら特徴を抽出して理解することができる
 - 従来は特徴量抽出手法を、場合に応じて使い分けを人の手で意図的に行う必要があった。しかし、ディープラーニングでは、特徴量抽出の部分自体も学習による自動で獲得することができるようになった
- 課題
 - 学習させるデータの課題
 - ただ大量であるだけでなく、質が求められる
 - 実際には、学習において教師ありのラベル付きデータが必要となるため、データに対してラベル付を人の手で行うことになる。データ量が増えるほど学習精度はあがる一方、ラベル付する量や時間的コストも大きくなる
 - 根本的なディープラーニングの手法の見直し
 - 手法自体の発展が目覚ましいわけではなく、コンピュータ計算処理能力がやっとな理論に追いついた状況
 - 実際の脳のニューロンの構造的には、横に広く浅いことが知られているが、現在のディープラーニングは縦に層を深くするほど精度が改善しているのが現状
 - 今後はデータを見て、ディープラーニング自体が自からのネットワーク構造を最適化して決定していくような仕組みも必要となる
 - 「No Free Lunch定理」どんな問題やどんなデータに対しても最高の精度を出せる万能なアルゴリズムは存在しないという定理
- エージェント指向
 - API連携
 - 2005年以降10年ほどで、APIの活用が当たり前となった
 - 全てのアプリケーションを自作する必要がない
 - 企業は今、自社の様々な情報やサービスへのアクセスを、APIを通じてオープン化しようとしている
 - 5W1Hメタデータを活用したデータ連携
 - 商用ベースでAPIを提供し、APIが相互に機能提供し、通信することで協同で問題解決を行い、業務フローを回し、売り上げをシェアできるようになってきた

- エージェントが互いに適切な相手を見つけたり、「会話」するお作法（プロトコル）を決め、標準化する必要がある
- ディープラーニングの未来
 - まだ、「自分自身が世界の中で存在していることを意識し、物事に接して感情と理性で考え、判断し、発言し、行動する」ような、「強いAI」はまだ存在していない。
「物事を忘れたり、嘘をついたり」といった人間の「弱み」も真似できていない
 - 情報の関連性を理解して特徴を見出すというのは非常に人間らしい「理解」の方法
 - 次第に人間が知らない特徴に人工知能が気づくようになり、人間の代わりに名前を付けるようになるかも知れない
 - ディープラーニングを超える人工知能の技術が生まれてくる可能性もありますし、画像や文章を認識できるようになるだけでも、画像や文書の分別などを仕事とする人はいらなくなる。簡単なタスクを人工知能がこなせるようになるだけでも、人間の仕事が減ることは間違いない
- 機械が人間を越える段階（シンギュラリティ）
- 人工知能の実用化のための技術要素
 - 画像認識
 - 画像の認識、分析と自動分類
 - 音声認識
 -
 - 自然言語処理
 - 複雑に絡み合った多種多様な曖昧さの認識
 - 文字列から単語を切り出す（分かち書き）
 - 発音の曖昧さ
 - 品詞の曖昧さ
 - 修飾語句のい掛かり先の曖昧さ
 - 単語やフレーズの意味の曖昧さ
 - 文脈解釈の曖昧さ
 - 複数のエージェントの協調
 - 自動要約
 - 情報の構造化
 - 再編成して見える化
 - 曖昧な指示の具体化
 - 学習データ
 - 構造化された学習データ
 - XML

- ▢ RDF、トリプル
 - SPARQLでアクセス
- ▢ 構造化されていない学習データ
 -
- ▢ 人工知能を活用した事例
 - ▢ IBMワトソン
 - ▢ 専用AIの集合体
 - 各専門分野の知識を、その専門にある程度あわせた構造で持つ、また、専門ごとに追加的に知識を獲得する方法（アルゴリズム）も微妙に違う
 - 処理方式の主要な部分はAIではない
 - ▢ アプリケーションシステム
 - 機械翻訳
 - 自動運転
 - ▢ ルンバ
 - 部屋の形状や家具の配置の地図を作成し、無駄の少ない移動方法を考え、かつ、同じ場所を通過せずに効率よく掃除する（小規模知識・データ）
 - ▢ 自由回答テキストの自動解析、集計
 - 人はチェックと高度な分析に専念
 - ▢ 定性分析
 - ▢ 類似検索
 - 長い全テキスト中の全単語に精妙な重みづけを自動で行って、記事中のほぼ全単語で検索、互いの類似性を判定するような仕組み
 - 対話履歴から学習してパーソナライズする機能や自動要約機能も容易に付加
 - 自動要約
 - 定性的な評価を定量化
 - ▢ 知識創出システム
 - ▢ 人とAIの協調による創造的な課題解決
 - AIの助けを借りて新サービスの構想を練り開発する
 - 人とシステムの得意な作業が連携
- ▢ ビジネス分野
 - ▢ 自動運転自動車
 - 自動車がセンサー機器、電気自動車（EV）、AI機能でモジュール化
 - ADAS（先進的運転支援システム）
 - ▢ フィンテック（Fintech）

■ 金融 + 技術

▢ X-tech

- AI技術をコアとして、その業界では新参者である企業が、今までにない価値や仕組みを提供する動向
- X-techの各分野で培われたテクノロジーが、別の分野に適用され、新しい付加価値を生み、従来ボトルネックとなっていた課題を解決していく

▢ X-techを支えるオープンデータ

- 機械可読形式、API、フリー

■ MarcheTech（マーケティング×Tech）

■ Media-Tech

▢ 人工知能に活用に適した業務

▢ 新サービス、新製品の要件

- ニーズの存在
- AIによるコストダウン、スピードアップ、精度・性能の向上
- ニーズとその具体的実現手法を結びつける技術、サービス提供の仕組みを、利益の出るコストで実現するアイデアの存在

▢ AIを直接、形にした新サービス、新規事業

- ディープラーニングによる画像認識のサービス化

- 強いAIが実現する前であっても、次のようなことはAIを活用してビジネスモデルが変わる

▢ ビジネスプロセスにおける業務とシステムの役割分担の見直し

- AIの実用段階に達した今、いままで人が担ってきた部分の作業も、精密化すると一部がAIを活用できる、したほうが効率がいいことが現れる

▢ IoTの活用

- IoTを活用することにより、人をはるかに超えるポイントでの情報をM2Mにより収集し、学習用データとして利用できる

▢ IoTの要素

- 例えば、RFID
- モノ、センサー、プロセッサ、通信機能、データ処理

▢ 丸暗記型の業務

- 大量お情報から関連しそうな情報を抽出し、信頼性が高い情報を提示する

▢ 事実に基づいた調査レポート

- 私見、創造がなく、引用文献に記述された内容にのみに基づく論文は、AIでも作成可能

▢ ルーティンワーク、マニュアル化された仕事

- 職員⇒非正規職員、外部委託⇒アルバイト⇒人工知能

- ▢ QA集に基づいたレファレンスサービス
 - ヘルプデスク業務は、数年でディープラーニングもよる自動応答に置き換わっていく
- ▢ 知識データベースの構築
 - 正解データを用意して、自動で特徴の抽出、認識を行い知識データベース化
- ▢ 一般論
 - ▢ 様々な分野融合する際の柱
 - 第1：「AI for Human Life」サービス業との融合
 - 第2：「AI for Science」基礎科学研究との融合
 - 第3：「AI for Manufacturing」製造業との融合
 - ▢ 既存サービスの改善と効率化
 - 「監視、見守り、点検」あるいは、生データの分類・整理に相当するタスクがどこにあるかを見つけること
 - 監視
 - 見守り
 - 点検
 - 分類・整理
 - ▢ 新たなサービスの創造
 - ハッカソン、アイデアソンで、様々なアイデアが。。。。
 - ▢ また、マッシュアップで完成度の高いアプリを作る場合も。
 - Ruby on RailsやPythonのDjangoなどのアプリケーション開発フレームの上で
- ▢ 課題
 - 著作権
- ▢ 人工知能の活用方法
 - ▢ オープンソース機械学習パッケージ
 -
 - ▢ 学習データ
 - JSTシソーラスmap
 - DBPedia
- ▢ 人工知能で変わる社会、人工知能の脅威
 - ▢ 人工知能がより賢い人工知能を生み出し続ける「シンギュラリティ（技術的特異点）」が起こるかも知れない
 - （2045年問題）質的にも人間の理解や発想の能力を超えて、超・知性として発達し始める特異点

- 今まででは考えられなかったような知的労働の世界までロボット
とになる

▢ これから必要な人材

▢ データサイエンティスト

- 技術
- ビジネス
- 分析
- コミュニケーション能力
- 主流のプログラミング言語でプロトタイプを開発でき、数学、統計学、確率論、コンピュータサイエンスの確かな素養のある人材
- かつ、ビジネス課題への感覚が鋭く顧客に共感できること

▢ 人間だからできる仕事は

- 与えられた仕事をこなすだけではなく、相手が必要としているモノを理解し、要求以上のモノを提供することで人工知能との差別化が図れる

▢ 創造的な仕事

- 新しいものを生み出せる仕事をするか、人間らしい温かみのある仕事ができなければ仕事を失うことになる
- 事実のみを知識とする創造性のない論文はAIで自動作成される

▢ 人間の感性に訴える芸術、芸能関係の仕事

- 人間の感性や感覚に訴えるモノが多く、人間の特性や文化を正しく理解していなければこの仕事はできない

■ 最終判断をする仕事

■ 人と人のコミュニケーション能力の必要なサービス業

▢ 進化と発展を生み出す仕事

▢ AIを活用するためシステム技術者

- 課題解決に有用な学習データの選別とアルゴリズムの選択。AIによる試行錯誤の結果の評価。ある程度精度が上がるとAIの出力を確認する作業になっていく

▢ 信頼性の高い教師データ（辞書類）の作成

- 初めは人が作成し、ある程度精度が上がるとAIの出力を確認する作業になっていく

▢ 教育に関わる仕事

- 「なぜできないのか」を想像して教えることは人間でも難しい。まして人工知能だとさらに難しい
- 「授業は人工知能」で「質問は教員」のようにタスクを分ける塾や学校が現れる可能性は大いにある

▢ 図書館情報システムは

▢ 文献情報検索から事実検索へ

- レファレンスサービスにおいて、「どの本に書いてある」では [Expand - Collapse](#)
いい情報は、どの本のどこにどのように書いてあるので、質問の解答は何々である」

☐ 参考文献

- 絵でわかる人工知能 明日使いたくなるキーワード68【2017年 三宅陽一郎】
- ビジネスマンのためのビッグデータ解析：知の集合体・最先端人工知能の活用
- 人工知能「超入門」ディープラーニングの可能性と脅威（Impress QuickBooks）
- よくわかるディープラーニングの仕組み【谷田部卓】
- ビジネスで使う機械学習【谷田部卓】
- 非エンジニア、文系、ビジネスマンのための人工知能入門：数式が苦手なあなたにおすすめ
- 人工知能（AI）活用時代に必要とされる能力とは？ビジネスで差がつく「データサイエンス力」
- グーグルに学ぶディープラーニング（日経ビッグデータ）
- 2020年を見据えたグローバル企業のIT戦略 IoT編
- 平成28年度情報通信白書【総務省】
- 人間の仕事を奪う「AI」の過去・現在・未来 | 蘊蓄の筆筒 100章 | 経済ニュースの新基準【2017年04月15日東洋経済オンライン】
- IT人材白書2017【2017年4月25日IPA】
- 第四次産業革命を視野に入れた知財システムの在り方について【2017年4月19日METI】
- 科学技術イノベーション総合戦略2017（案）【2017年4月21日】
- 新産業ビジョン【2017年5月METI】
- 人工知能技術戦略（案）
- 知的財産推進計画2017（2017年5月16日内閣官房知的財産戦略本部）
- 「デジタルトランスフォーメーション」【2016年9月ベйкаレント・コンサエルトディング】