DAX25-03【書籍】よくわかるディープラーニングの仕組み【谷田部卓】

1. 過学習と対策
   1. 過度に教師データに依存した（汎化できていない）状態
   2. 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
2. ニューラルネットワークの歴史
   1. ディープラーニング
      1. 2011年音声認識コンテスト
      2. 2012年画像認識コンテスト
3. ディープラーニングのビジネス（実用化）
   1. 2種類
      1. CNN（畳み込みニューラルネットワーク）系
      2. RNN（LSTM）系
   2. CNNが得意とする
      1. パターン認識系
         1. 精度で既に人間を凌駕
      2. 自然言語処理系
      3. 操作系
   3. 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
   4. 文章の自動要約や意味抽出まで進んでいくと、 事務職の大半はディープラーニングで置き換えが可能と考えられる
   5. 自動運転は数年で実用化が始まり、 普及も急速に進んでいくはず
4. ディープラーニングにおける日本の役割
   1. 技術を自社だけで囲い込もうとする意識が強すぎるため、 研究成果を公開したり、 アイデアを募るような研究ができない体質
   2. 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
      1. http://gigazine.net/news/20160929-partnership-ai-facebook-amazon-google-ibm-microsoft/

ドキュメントを参照: [20160929-partnership-ai-facebook-amazon-google-ibm-microsoft](http://gigazine.net/news/20160929-partnership-ai-facebook-amazon-google-ibm-microsoft/)

* 1. エンジニアを囲い込んだりせず、 基礎研究の段階から幅広い英知を求めていくべき

1. AIの学習方法と幼児教育のアナロジー
   1. 偏りのないあらゆるパターンで、 かつ大量の訓練データを与えることにより改善できる
   2. ディープラーニングは、 その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
   3. 過学習が生じる前の幼児の段階で、 出来る限り様々な体験をさせるべき
   4. 実体験が無理なら、 色々な絵本を読み聞かせることで、 たとえ仮想でも体験の幅が格段に広がり、 過学習を生じさせない柔軟な頭脳になるかもしれない
   5. 視覚や聴覚から入力されてくる空間的・時系列的な情報をパターン化・抽出化することで大幅に情報圧縮し、 有限の脳に大量の情報を詰め込めることができるようになった。
   6. 「言語」は、 その抽象化された情報、 すなわち記録を外部化・表出したもの。 さらに「文字」の発明が、 記憶を外部に長期間記録することを可能とし、 より大量の情報を人類が共有できるようになった
   7. 人間は、 言語機能はまだDNAに組み込まれていないか、 言語機能をハードウェアとしてDNAに取り込むことにより、 柔軟に変更可能なソフトウェアとして、 後から学習できるように機能分離されているのかもしれない。
   8. 「水を得た魚のように」などとアナロジーを好んで使う。 「ディテールをそぎ落として抽象化した言葉」。 パターンのほうがイメージしやすく、 情報量が少ないほうがマッチングが容易「
   9. 大人になると抽象的概念を使って複雑な事象も考えることができるようになる。
   10. 知識や経験を積むことで、 ニューラルネットワークの階層が深くなっていくかのようです。 人間も過学習におちいらないように、 すなわち「偏見」をもたないように、 幅広く見識を積むべきなのでしょう