

データサイエンスと機械学習で切り拓く新しい情報社会①

2019年8月20日(火) 午後

講演

「HPC X AI」

上田 修功氏 N T T株式会社

コミュニケーション科学基礎研究所

もともと機械学習をずっとやっています。現在のAIは、深層学習ブームで、ニューラルネットワークを取り込んだ機械学習です。

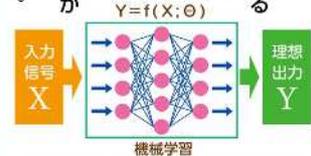
機械学習とは、演繹法で作られたデータを、帰納法的な学習によって最終的なモデルを作るというものでしたが、ニューラルネットワークを内装することにより、モデル化の必要がなくなりました。

学習には3つのパターンがあり、強化学習は、膨大な試行錯誤を繰り返す学習法で、囲碁で脚光をあびました。

教師有り学習は入力と出力のペアを膨大に与えて、未知入力に対して適切な答えを与える方法です。これはGPUや学習用のデータが出現したことにより可能となりました。

技術より時代背景がブームを作ると言えると思います。

Y=f(X;θ)



●GPUが開発される前は組み合わせ爆発により計算できませんでした。

●強化学習はとにかく膨大な試行錯誤を繰り返し学習します。

●教師無し学習はメトリックを学習します。

●教師有り学習は膨大な学習データを与えます。

●演繹方の学習と帰納法の学習。



簡単な問題はニューラルネットでできましが、複雑な問題はモデリングが不可欠です。

都市化の研究事例を紹介します。災害時には人は集団行動をするので、特定の道や広場に人が溢れかえってしまいます。なぜかという、リアルタイムの正確な情報と適切な誘導情報が無いせいで、個人の判断ができない状態だからです。

何割かの人が、赤い点線の誘導に従えば混雑を防げます。

我々が開発した近未来予測アルゴリズムを使い、集団最適誘導技術デモを行いました。まず、どの道が混むかを予測し、それを繰り返し学習しながら最適な人流の誘導を導きだします。

現代は身の回りにセンサーが張り巡らされていて、リアルタイム情報が取れます。それを解析し、ナビゲーシオンして現実世界に還元します。これを回すことで安全な生活が送れます。要素技術の開発をしています。



●競技場に向かう順路に人流を予測した誘導を与えると、混雑せずに目的地へ到着します。

●人間は集団行動をするので、人の集まる方向へ動いてしまいます。

●HPCはシミュレーションデータ駆動型科学研究で活用されると思います。



観測・実験 ↔ 理論

データ同化 ↔ モデリング

仮定・検証 ↔ 法則発見

シミュレーション

講演

「イベント・ホライズン・テレスコープによるブラックホールシャドウの撮像」

笹田 真氏 広島大学 宇宙科学センター

こちらが新聞等で有名になったブラックホールの画像です。今回、シミュレーションではなく、観測でこのようなリングが捉えられました。

なぜ、ブラックホールは「穴のあいたリング」に見えるかというと、ブラックホールには光すら引き込む強い重力があるからです。それは、中心に近づくほど強く、ある距離になると、光が外に出られなくなり、まわりを回ります。重力によって引き寄せられ、ブラックホールの周りを回ります。引き寄せられた物質同士がぶつかり合い、熱いガスとなり、光を放射します。ガスからの光がブラックホールの重力によって曲げられ、ブラックホールの「周り」から光が飛んでくるため、結果的に穴のあいたリングに見えるのです。

地球から見ると穴のように見えるのは、ブラックホールの周りの熱いガスの存在と、光がブラックホールの周りから地球へ飛んでくるためです。

中心に近いほど光が強く引張られます。この、光が外に出られなくなる境界を「事象の地平面」といいます。

ME7の中心の巨大ブラックホールです。



では、どの位の視力があればブラックホールが見えるかというと、視力が300万という望遠鏡が必要となります。そこでいくつもの電波望遠鏡を集めた、電波干涉計観測が使われました。ただ、この電波干涉計データというのは、写真そのものではなく、それをフーリエ変換したデータです。画像にするためには、逆フーリエ変換をする必要があります。このときデータ量が少ないので劣決定問題となり、これを解決しなければなりません。

これらの問題をスライスモデリングによって解決しました。ただ、与えるパラメータにより最適画像が異なるので、シミュレーション画像を元にパラメータサーチを行いました。

伝統的な手法を含む3つの独立な手法で五万通りのサーチを行い、画像の復元を行っています。

今回わかったことは、ブラックホールはほぼ真円であること、リングは非対称であること、重さが太陽の65億倍にもなります。



2次元フーリエ変換

空間周波数 U

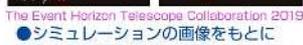
空間周波数 V

空間座標 X

空間座標 Y

●各ピクセルを条件にもとづいて評価することで、最適画像を再構成することを目指しました。

●望遠鏡を組み合わせると大きな望遠鏡を作ります。



最終的なブラックホールの画像はこちらです！

SS研 HPCフォーラム2019

データサイエンスと機械学習で切り拓く新しい情報社会②

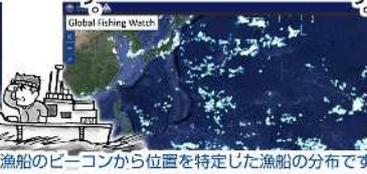
2019年8月20日(火) 午後

講演

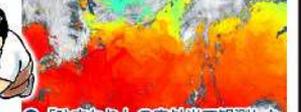
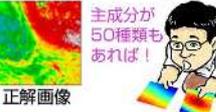
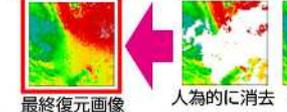
「機械学習・パターン認識の漁業への応用」

飯山 将晃氏 京都大学
学術情報メディアセンター

画像認識を水産業というフェーズで利用できないか。今、日本の水産業は減少傾向です。また、獲れる魚の量は決まっていますが、あとは「どのように効率良く獲るか」ということになります。漁船の海上での分布は毎日公表されています。この情報から見て来ることがあります。例えば潮の流れを考慮すると燃費良く航行できるとか、一隻の動きを追うと最初に獲れない場所について、あとで移動したなんてのも分かります。効率化に重要なのが燃料の節約です。よく魚の獲れる場所へまっすぐに行くための機械学習、パターン認識で出来ないかという問題です。解こうとしています。現在でも漁業者は勘と経験で漁場を決定していますが、海水温などの情報も取り入れています。



これは衛星で観測した海の表面温度です。これだけですがかなりの事が分かりますが、問題は雲のある部分の情報が出ており、その部分を学習型インヘンディング技術で補うことで情報の復元をします。まずは過去のデータで各画素での水温の組み合わせを考えました。実際、高々50個程度の「基底画像」の重み付き和で98%程度、大体の再現ができました。ただ、線形和では無理があるので、敵対的誤差を利用する方法と、過去の情報を組み込むことで、最終的にかなりの欠損を補えるようになっていきます。この海水温のデータと漁獲データを使って、漁場を推定しますが、結局は他のファクターが多過ぎて、この解析によって魚が獲れる保証はありません。ただ、傾向に関しては、傾向しているのを見つければいいので、5段階にクラス分けしたなかでも、特に魚がいそうな「豊漁クラス」を見つけておきます。ただ、漁業者にとっては、まだまだ特定範囲が広すぎるようです。



●そこにいけば魚が獲れそうという地点は見えて来ましたが…

●わざと欠損させた画像を「偽物を判定」するニューラルネットを追加することで、大体合っている状態にできます。

●「ひまわり」の赤外光で観測したデータで、雲のある部分のデータが抜けています。

●「ひまわり」の赤外光で観測したデータで、雲のある部分のデータが抜けています。

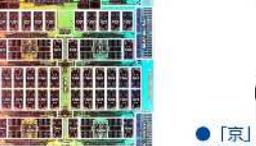
●「ひまわり」の赤外光で観測したデータで、雲のある部分のデータが抜けています。

講演

「スーパーコンピュータ「富岳」の開発」

新庄 直樹氏 富士通

先日「京」の運用が終了しまして、「京」は7年間、多くの成果をあげることができました。現在は、理研様と「富岳」の開発を進めています。今ここでは、全貌をお話する段階ではないので、いすれ今年の秋ぐらいには理研様からの発表が予定されています。「富岳」の開発目標としては、●高いアプリケーション性能、高い電力効率 ●ユーザに対する利便性 ●アプリケーションの互換性を維持するというものです。CPUは「A64FX」です。ARMv8-Aを採用した高性能かつ高効率なものです。このCPUの一番大きな特徴は高メモリバンド幅で1CPUの中でも2階層キャッシュと2階層キャッシュとキャッシュ、メモリのアウトオブオーダー制御などにより、全階層での高いスループットを実現しています。



●「京」の開発以降もFX10、FX100と開発してきました。

●「京」の開発以降もFX10、FX100と開発してきました。

●「京」の開発以降もFX10、FX100と開発してきました。

●「京」の開発以降もFX10、FX100と開発してきました。

●「京」の開発以降もFX10、FX100と開発してきました。

GPUのどなたを工夫しているかという点ですが、CPUのロード性能の改善 ●エネルギーモニター ●アナライザ ●「パワーノブ」(電力削減機能)などが挙げられます。また、「京」と「富岳」でどのくらい性能が違うかというと、1ペタと比較すると、「富岳」では「京」の設置面積で100分の1程度であるといえます。●「京」 ■計算ラックx80とディスクラックx20 ●「富岳」 ■1ラック(SSD含む)



	「京」	「富岳」
計算ノード数	7,680[-96x80]	384
I/Oノード数	480[-6x80]	
設置面積(m ²)	128[-4x32]	1.1
OS	SPARC Linux	Arm Linux



今後、実際のマシンを導入して、アプリをインストールしながら開発を進めていきます。