

因果推論

NRIデジタル株式会社
データサイエンティスト

2020年08月08日

NRI

Share the Next Values!



ガイドンス

- 以下の講義内容で進めていきます。
 - 主には基礎の考え方を中心に、プログラミングで実演していきます。
- 質疑応答は、チャットなどでいつでも受け付けます。
 - 1コマ目：10時から11時半まで
 - ・ 統計分析の復習
 - 2コマ目：12時半から14時まで
 - ・ 相関と因果の考え方
 - ・ 効果量の測定方法
 - ・ DIDの考え方
 - 3コマ目：14時半から16時まで
 - ・ バイアスの除去方法
 - ・ 傾向スコアマッチング
 - ・ IPV法
 - ・ 効果量の測定方法
 - 4コマ目：16時半から18時まで
 - ・ 検証実験の設計
 - ・ 検出力解析
 - ・ 効果検証データの解析
 - ・ 検証結果の妥当性

自己紹介

■ 田村光太郎 博士（理学）、技術経営修士（専門職）

- 株式会社野村総合研究所
- （出向）NRIデジタル株式会社 データサイエンティスト

■ 2016.4～2018.3

- 東京工業大学科学技術創成研究院
 - ・ ビッグデータ数理科学研究ユニット特任助教

■ 2018.4～

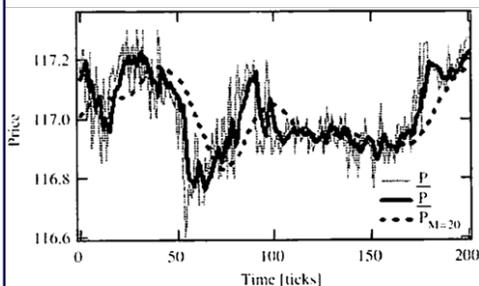
- 株式会社野村総合研究所
- NRIデジタル株式会社（出向）
- ビジネスデザインユニット データサイエンティスト
- 金融ITイノベーション事業本部（兼務）



時系列解析とは異なる世界

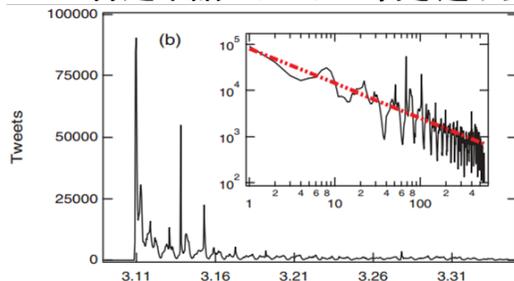
- 複雑ネットワークという計算社会学の一分野。
 - 経済物理学、複雑系物理学などとも深いつながりがある。
- モノとモノとのつながりをベースに、システムを評価する分野。

為替価格



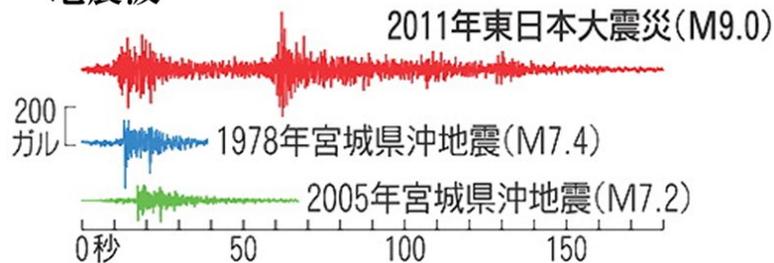
M.Takayasu et. al. Physica A(2006)

特定単語のブログの書き込み数

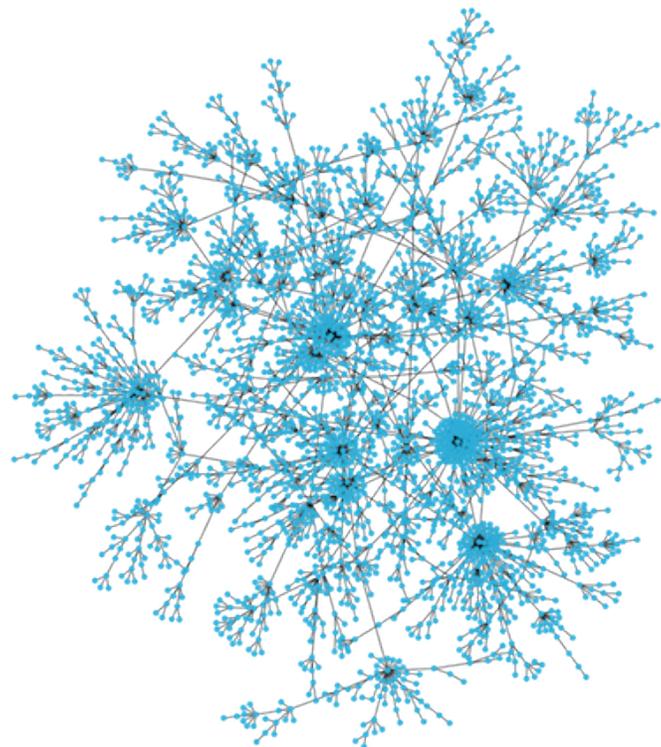


Y.Sano et. al. PRE (2013)

地震波



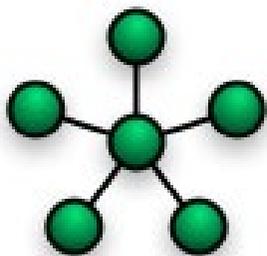
ネットワーク：つながりの解析



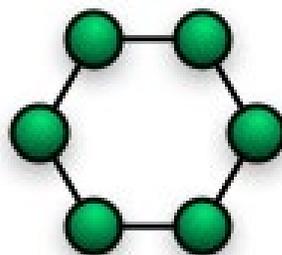
つながりに注目するという視点

- 人、モノ、企業は、つながり方でリスクや役割は異なる。

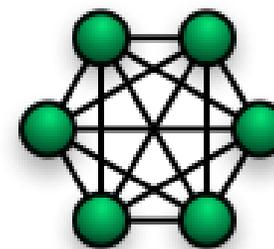
例：同じ6個のモノをつなげたときの形状を考えよう



Star



Ring



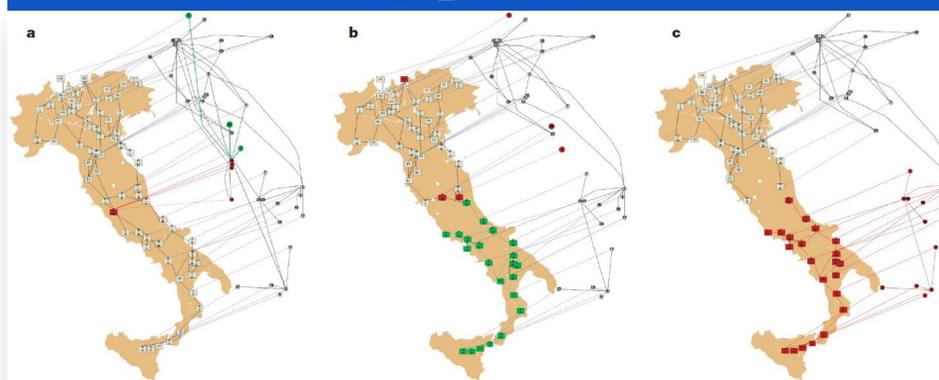
Fully Connected

鉄道路線のネットワーク



「地下鉄路線図」東京都交通局

イタリア送電網ネットワーク

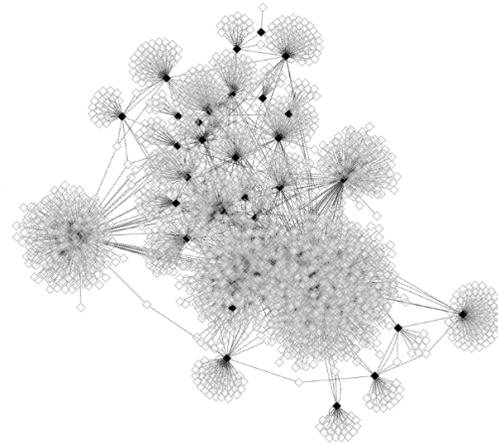
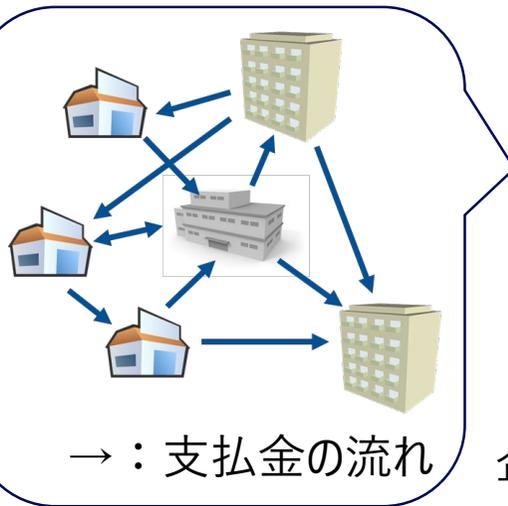


Sergey V. Buldyrev et al., "Catastrophic cascade of failures in interdependent networks", Nature 464, (2010)

時系列解析とは異なる世界

- 東工大在籍時には、帝国データバンクとの共同研究に主に携わる。
 - 企業間のお金の流れをネットワーク上の流れとして、モデル化して日本企業の経済システムを評価
 - 「Diffusion-localization transition caused by nonlinear transport on complex networks」
 - <https://www.nature.com/articles/s41598-018-23675-x/>

100万社の企業間取引データの解析（株式会社帝国データバンクとの共同研究）

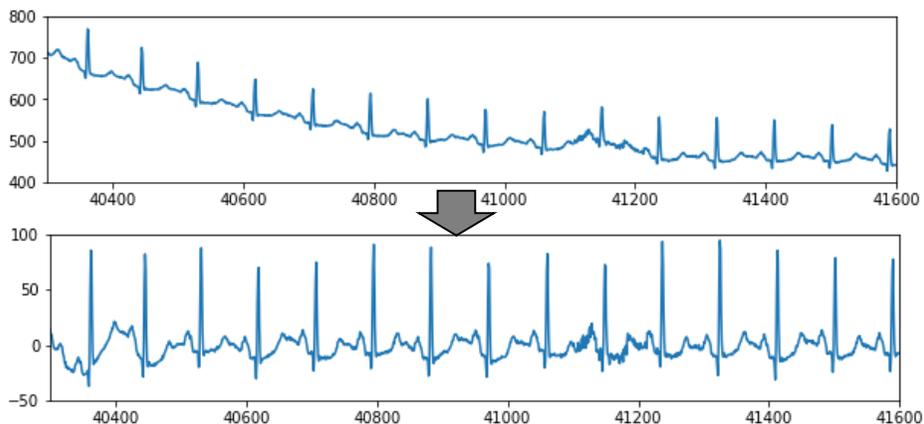


企業間取引ネットワークの一部

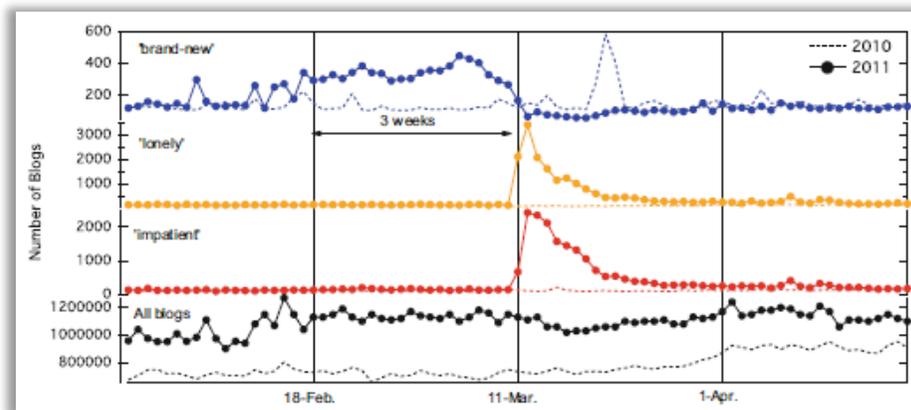
- 企業間の取引額推定
- 企業の売上推定モデル

- ネットワーク上の拡散
- ネットワーク上の流域

などなど、さまざまなデータ形式・データ解析に関わった

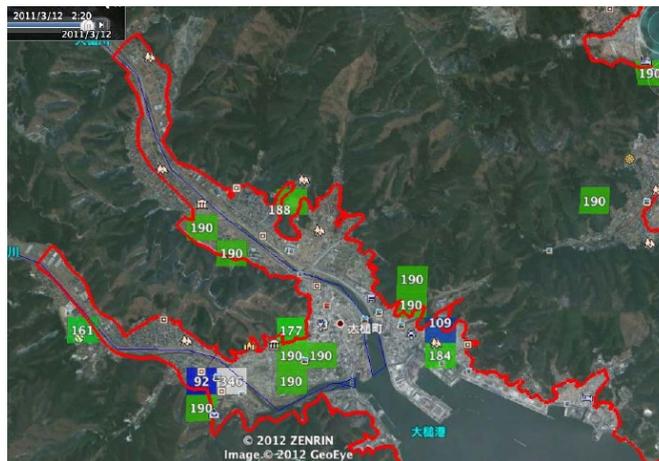


ホルター心電図データの不整脈検知



Y. Sano et al., ICSMS+EC2014: Chapter 26, (2015)
SNSデータへのPOMS適用による感情分析

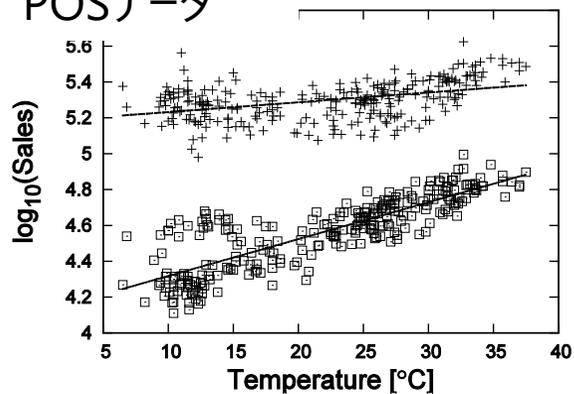
津波被災地域の避難行動のGIS解析



道路ネットワークの関節点解析



POSデータ



NRI デジタルで携わっている業務

- 現在は具体的な業務やそれに関わる技術開発に従事

業務改善（統計分析、モデル開発）

- 不動産会社における広告・物件評価分析
- 人材情報サービス会社の営業員/音声分析
- 生命保険会社での解約予兆分析

産学連携（アルゴリズム開発）

- 連携候補先大学の調査
- 共同研究のディレクション

研究開発（モデル開発）

- 広告クリエイティブ評価モデル構築
 - CM効果推定モデル
- ナレッジグラフ推論/説明モデル構築
 - 解釈可能 AI の基礎技術開発
- 自然言語処理
 - 文書校正、ニューラル翻訳

対外発表・人材開発

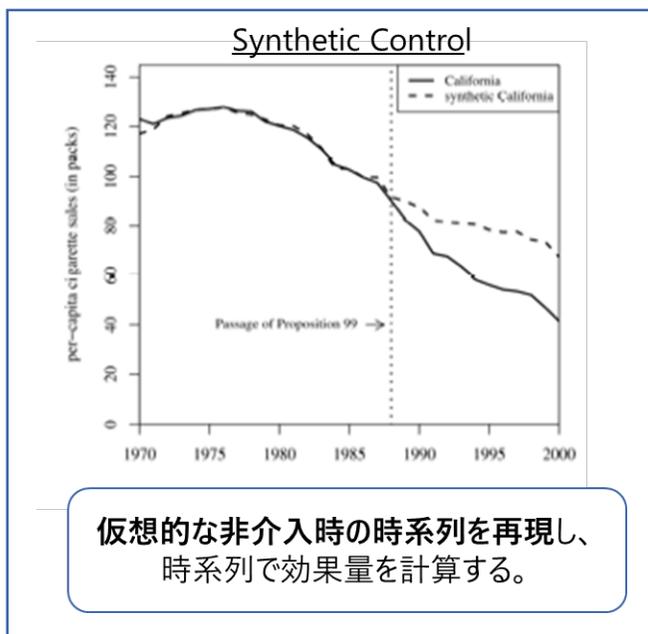
- 研修・講演での講義
- 社外向けフォーラムでの発表
- 論文・寄稿

自己紹介

■ アナリティクス案件

- モデルの導入とその導入効果検証などの分析

海外大学との因果推論の研究

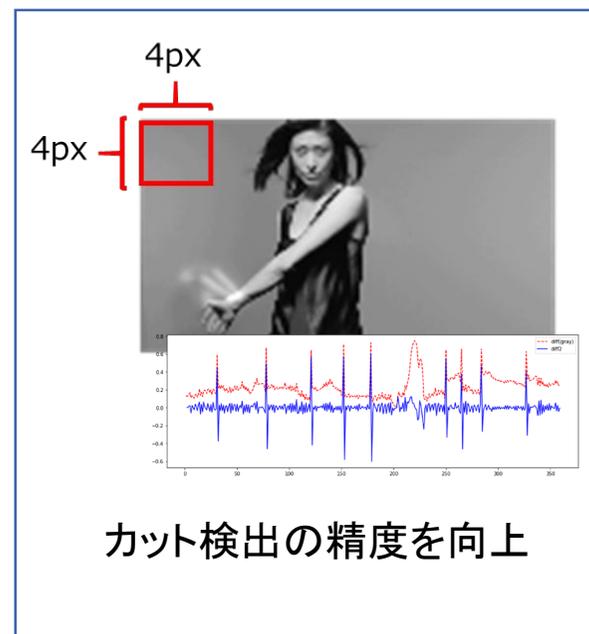


推論エンジン(自然言語処理)

ヘレンは、ヘレンの部屋で死んだ。
dead(Helen,HelenRoom)
窓は閉まっていた。
!windowopen(HelenRoom)
庭は広い
locatbig(Garden)
部屋は広くない
!locatbig(HelenRoom)
通気口がある。
tuukikou(HelenRoom,Roylce)
何者かは小さい
small(strange)

...

CM動画解析 & 効果解析



「因果推論」のガイダンス

■ 本講義のねらい

- 基本的な統計分析を身に付けた後の、一段深いデータの見方
- 因果推論の考え方や使うもの
- 現場で使われることの多い、「効果検証」という側面を中心に因果推論を扱っていく。

■ 方法

- Python
 - Pandas等、データ分析関連ライブラリ
- 使用データ
 - https://github.com/kabigon-ds/uec_de
 - ↑のJupyterNotebookとcsvをサンプルとして利用します。
 - Github、Googleコラボ等での配布
 - 講義中にスクリプトを説明していく演習を交えます。

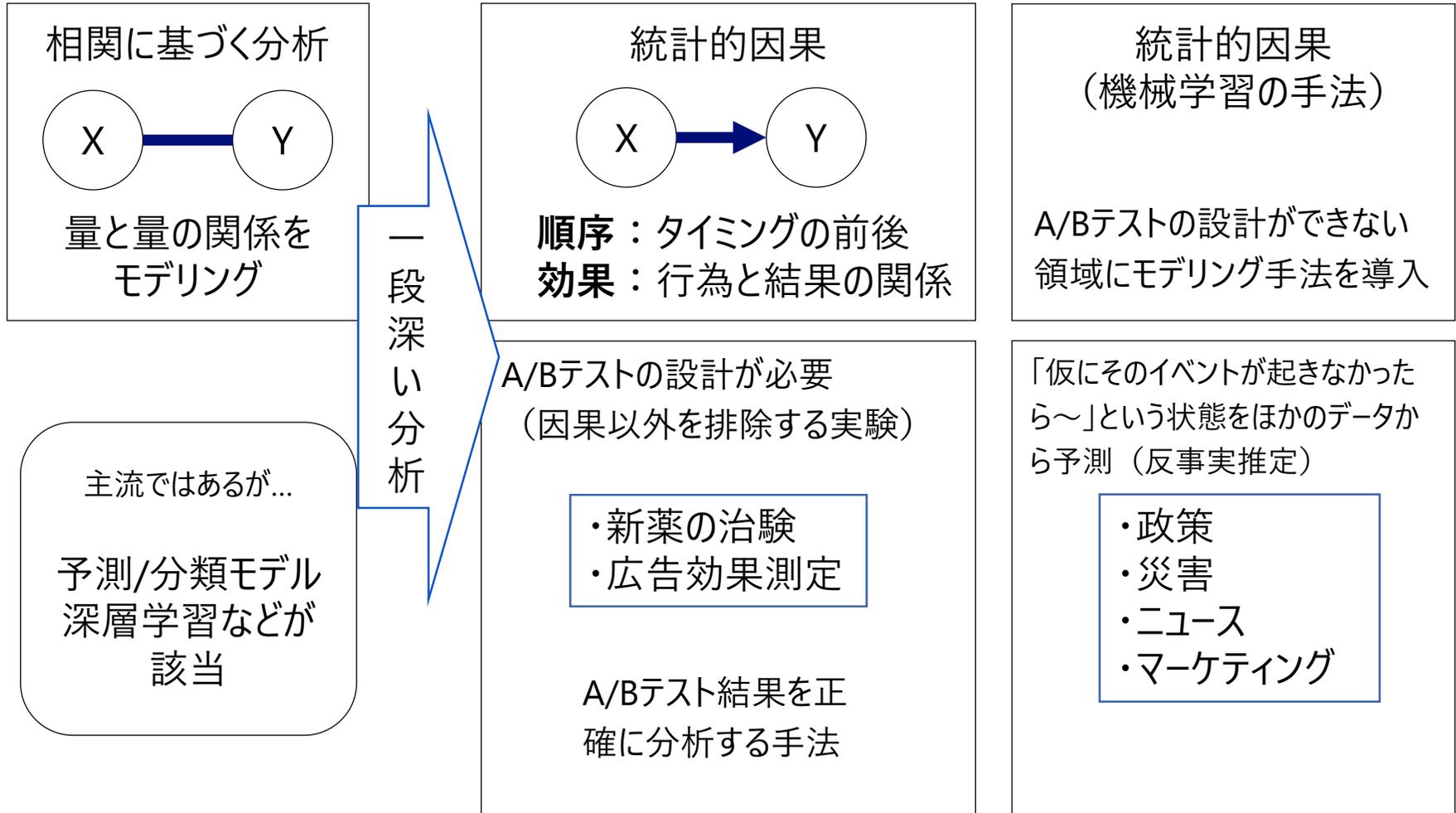
■ できるようになってほしいこと

- データ分析を行うときに「因果」に気を付けて分析を行えるようにする
- 主要なケースについては、因果効果を適切に測定できるようになる。

因果分析の概観

■ 因果分析とは、

「ある時点のイベントが、対象データに対して、どう影響したか大きさを測定すること」



因果推論分野の動向

- マーケティング施策等の効果分析を中心に、因果推論を中心とした効果分析手法の発展が期待されている。
- 特に、データの複雑化や高速化が進み、機械学習やAIの手法が、因果推論に適用され始めている。
- 深層学習と因果推論を組み合わせ、AIの出力に対して解釈可能性や説明可能性を持たせ、AI技術の高度化に繋げる試みも行われている。

昨今の「因果推論」領域での取り組み

| 組織 | 取り組み |
|---------------------------------|--|
| 理化学研究所 革新知能統合研究センター (AIP) | 2017年度センター創設に伴い、因果推論チームを創設 http://www.riken.jp/en/research/labs/aip/generic_tech/cause_infer/ |
| IBM Research | 2019年AI業界で注目している3つのトレンドについての発言：相関主体の分析が因果関係へと代わっていく。 AIに説明性・解釈可能性を与える重要な技術になるだろう。 https://www.ibm.com/blogs/think/jp-ja/ai-year-review/ |
| Uber | 因果推論用のpythonライブラリをオープンソースとして公開。 https://github.com/uber/causalml |
| CyberAgent AI tech studio | GAN（AIの一種）で因果推論手法を開発 https://adtech.cyberagent.io/research/archives/2283 |
| NEC X | NECのシリコンバレー子会社の3つの注力技術領域のひとつが「因果構造探索」。 https://tech.nikkeibp.co.jp/atcl/nxt/column/18/00141/032300045/ https://jpn.nec.com/press/201806/images/2001-01-01.pdf |
| Sony Computer Science Lab. | AIによる要因分析サービス「CALC 3.0」をリリース ～ 因果情報に基づく介入・施策効果をシミュレーションする新機能を搭載 ～ https://www.sony CSL.co.jp/press/prs20191106/ |

The text is framed by two decorative swooshes. The top swoosh is a gradient bar transitioning from blue on the left to red on the right. The bottom swoosh is a solid blue bar.

Share the Next Values!