

機械学習・AIの技術的進展と最新事例の紹介

東京工業大学 工学院 経営工学系
中田和秀

日本ファイナンス学会 特別公開セッション

「オルタナティブデータと機械学習が拓く新たなファイナンス実務と研究」

2020年12月5日

目次

1. 説明可能なAI
2. 反実仮想機械学習
3. 自然言語モデル BERT
4. 量子コンピュータ （時間があれば）

目次

1. 説明可能なAI
2. 反実仮想機械学習
3. 自然言語モデル BERT
4. 量子コンピュータ (時間があれば)

説明可能なAI –概要–

□ 「説明可能なAI」とは

- 深層学習（機械学習）の予測は一般に「ブラックボックス」
- 実務では、**モデルの解釈**や**判断根拠の説明**も必要であることが多い。
 - 疾病診断モデルの「医師/患者への判断根拠の説明」
 - 金融商品提供者の「顧客への商品推薦理由の説明」

→ **「説明可能なAI」**（eXplainable AI, XAI）の重要性が高まっている。

□ 「説明可能なAI」の動向

- 「説明可能なAI」論文数の増加（右図）
- 「人工知能技術戦略実行計画（案）」（内閣府、H.30）の「信頼できるAI」へのアプローチとして取り上げられる。



図の引用：[1]

説明可能なAI –種類–

□ 「説明可能なAI」の代表的アプローチ

1. Interpretable Models

- 最初から可読性/解釈性の高い（シンプルな）モデルを設計

2. Model Induction

- 精度の高いブラックボックスモデルを構築後、それを分析することで解釈性や判断の根拠を得る

3. Deep Explanation

- 深層学習（Deep Learning）が説明可能になるように構築
- 「画像の注目領域のハイライト」などが有名
（右図：説明文生成の事例[2]）



□ 次ページより、2・3について紹介

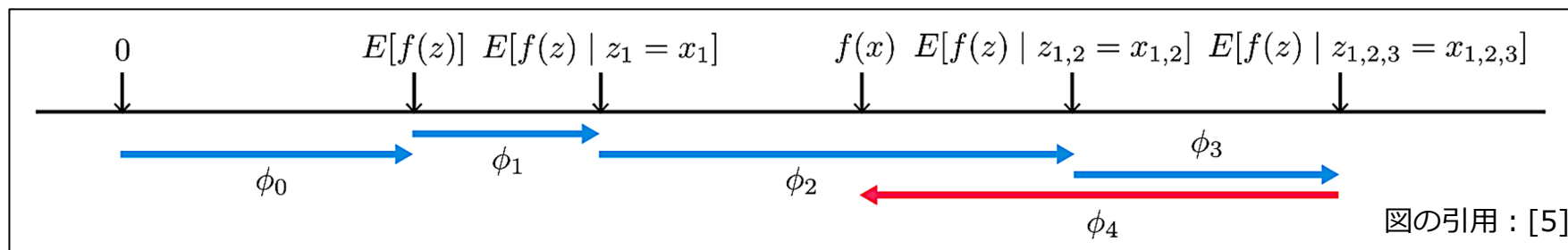
説明可能なAI – Model Induction–

□ 大域的な説明

- 複雑なモデルを、単一の決定木やルールモデルなどで近似する方法。
- Making Tree Ensembles Interpretable [3] :
ツリーアンサンブルモデルを、精度を保ちつつ単純なモデルに変換する。

□ 局所的な説明

- 複雑なモデルについて、局所的な入力に対する**予測の根拠**を説明する。
- 入力ごとに「根拠」が変わるため、より厳密な分析が可能になる。
- LIME[4]、SHAP[5]（下図）が有名。
各変数（図では0～4）について、予測への寄与度を計算する。

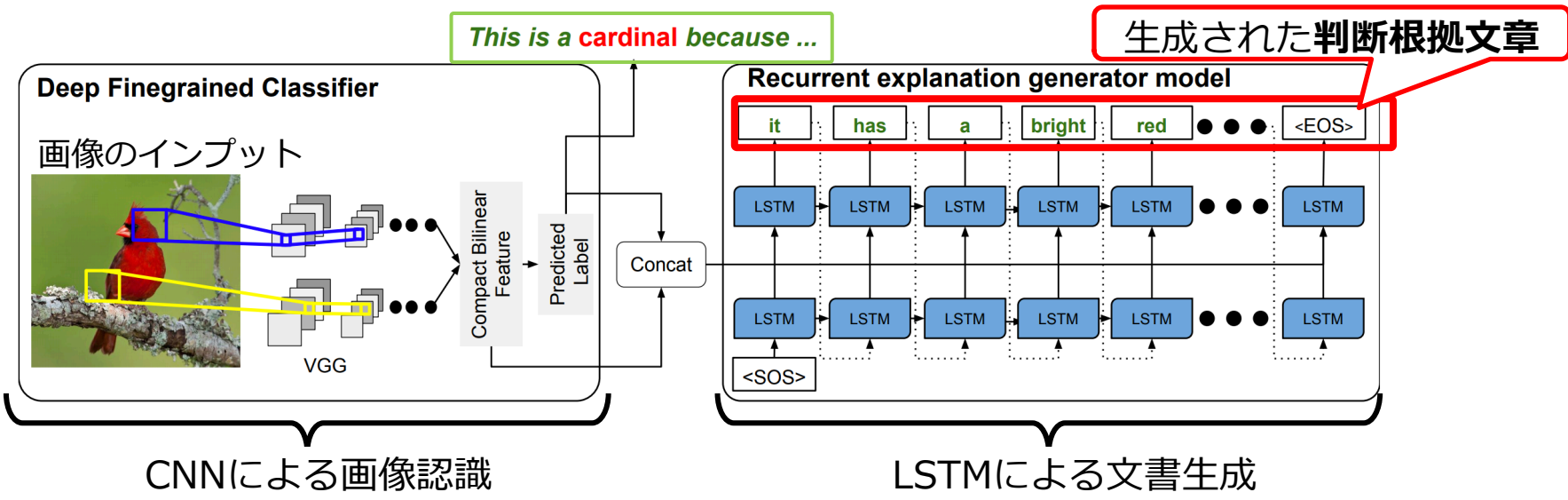


説明可能なAI –Deep Explanation–

□ 画像認識における事例

Generating Visual Explanations [6]

■ 画像認識モデルの判断根拠を、文章で説明するモデル



■ 出力例



This is a Bronzed Cowbird because ...

Definition: this bird is **black** with **blue** on its wings and has a long **pointy beak**.

Description: this bird is **nearly all black** with a short **pointy bill**.

Explanation-Label: this bird is **nearly all black** with **bright orange eyes**.

Explanation-Dis.: this is a **black bird** with a **red eye** and a **white beak**.

Explanation: this is a **black bird** with a **red eye** and a **pointy black beak**.

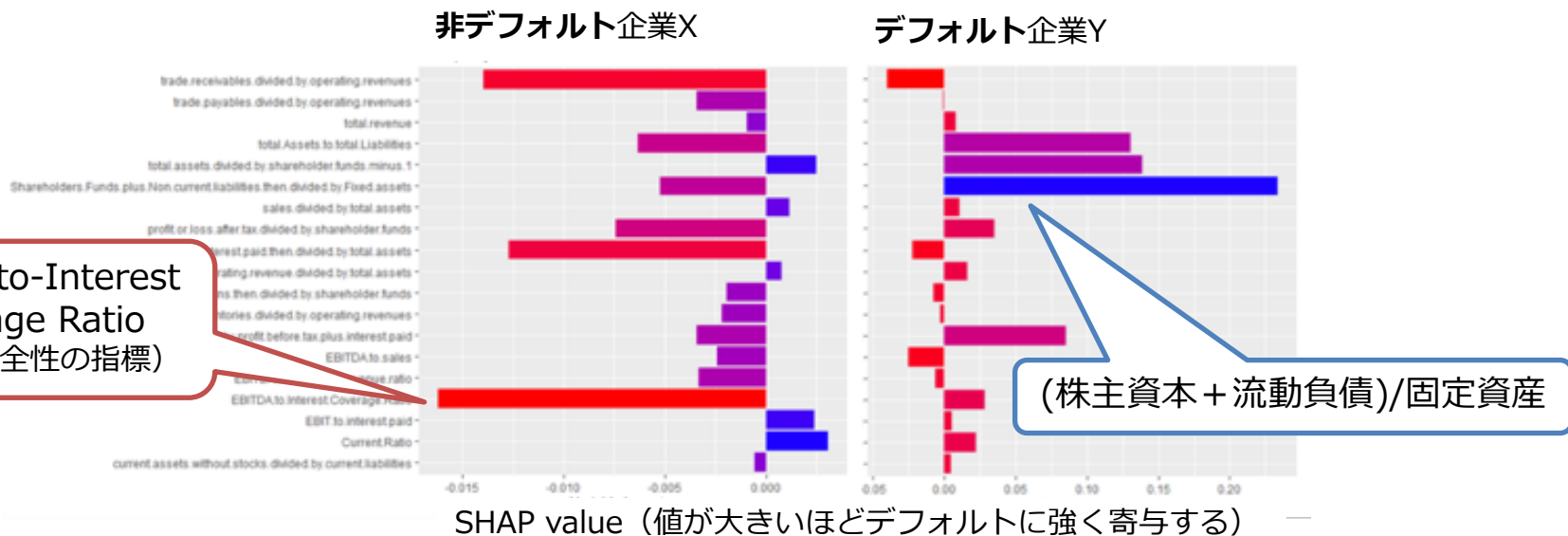
提案モデルは既存手法が言及していない属性・特徴について説明している

説明可能なAI -金融分野における応用事例-

金融×機械学習×SHAPの事例

Explainable Machine Learning in Credit Risk Management [7]

- 15,000社の中小企業の貸借対照表から、企業の1年後のデフォルトを、高精度で複雑なモデル「XGBoost」を用いて予測する。
- その際**SHAP**によって、予測に寄与した説明変数を抽出する。



参考文献

- [1] Amina Adadi, Mohammed Berrada. "Peeking inside the black-box: A survey on Explainable Artificial Intelligence (XAI)." *IEEE Access* 6: 52138-52160, 2018.
- [2] Kelvin Xu, Jimmy Ba, Ryan Kiros, Kyunghyun Cho, Aaron Courville, Ruslan Salakhutdinov, Richard Zemel, Yoshua Bengio. "Show, attend and tell: Neural image caption generation with visual attention." *Proceedings of the 32th International conference on machine learning*. 2015.
- [3] Satoshi Hara, Kohei Hayashi. "Making tree ensembles interpretable: A Bayesian model selection approach." *Proceedings of the 21th International conference on artificial intelligence and statistics*, 2018.
- [4] Marco Tulio Ribeiro, Sameer Singh, Carlos Guestrin. "" Why should I trust you?" Explaining the predictions of any classifier." *Proceedings of the 22nd ACM SIGKDD international conference on knowledge discovery and data mining*. 2016.
- [5] Scott Lundberg, Su-In Lee. "A unified approach to interpreting model predictions." *Advances in neural information processing systems* Vol.30 (NIPS 2017). 2017.
- [6] Lisa Anne Hendricks, Zeynep Akata, Marcus Rohrbach, Jeff Donahue, Bernt Schiele, Trevor Darrell. "Generating visual explanations." *Computer Vision – ECCV 2016*. Springer, 2016.
- [7] Niklas Bussmann, Paolo Giudici, Dimitri Marinelli, Jochen Papenbrock. "Explainable Machine Learning in Credit Risk Management." *Computational Economics*: 1-14, 2000.

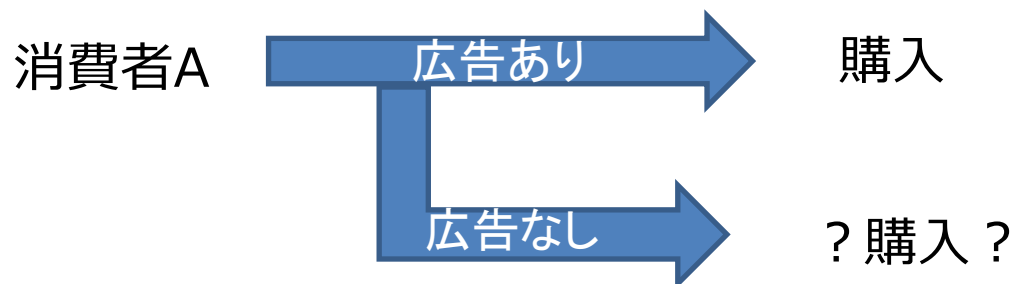
目次

1. 説明可能なAI
2. 反実仮想機械学習
3. 自然言語モデル BERT
4. 量子コンピュータ (時間があれば)

反実仮想機械学習 –概要–

□反実仮想（反事実）とは

反実仮想とは、実際には起こっていない状況を想定すること。



現実社会には反実仮想を必要とするタスクが多く存在する。

□反実仮想機械学習の分野

- Individual Treatment Effect (ITE) Prediction
介入が目的変数に与える個別的因果効果を予測する
- Off-Policy Evaluation (OPE)
ある方策で集まったデータを使って仮想的な方策を評価
- Unbiased Recommender Learning
蓄積されたログデータのバイアスの影響を取り除く

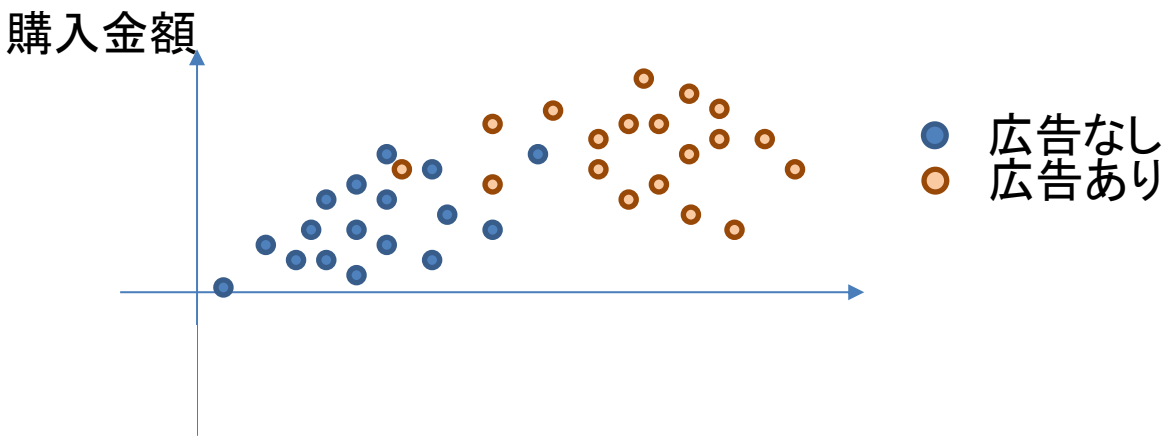
反実仮想機械学習

広告の有無がユーザの購買に与える影響を考える。

ユーザ	広告あり	広告なし
1	150	?
2	?	0
3	?	150
4	200	?
5	300	?

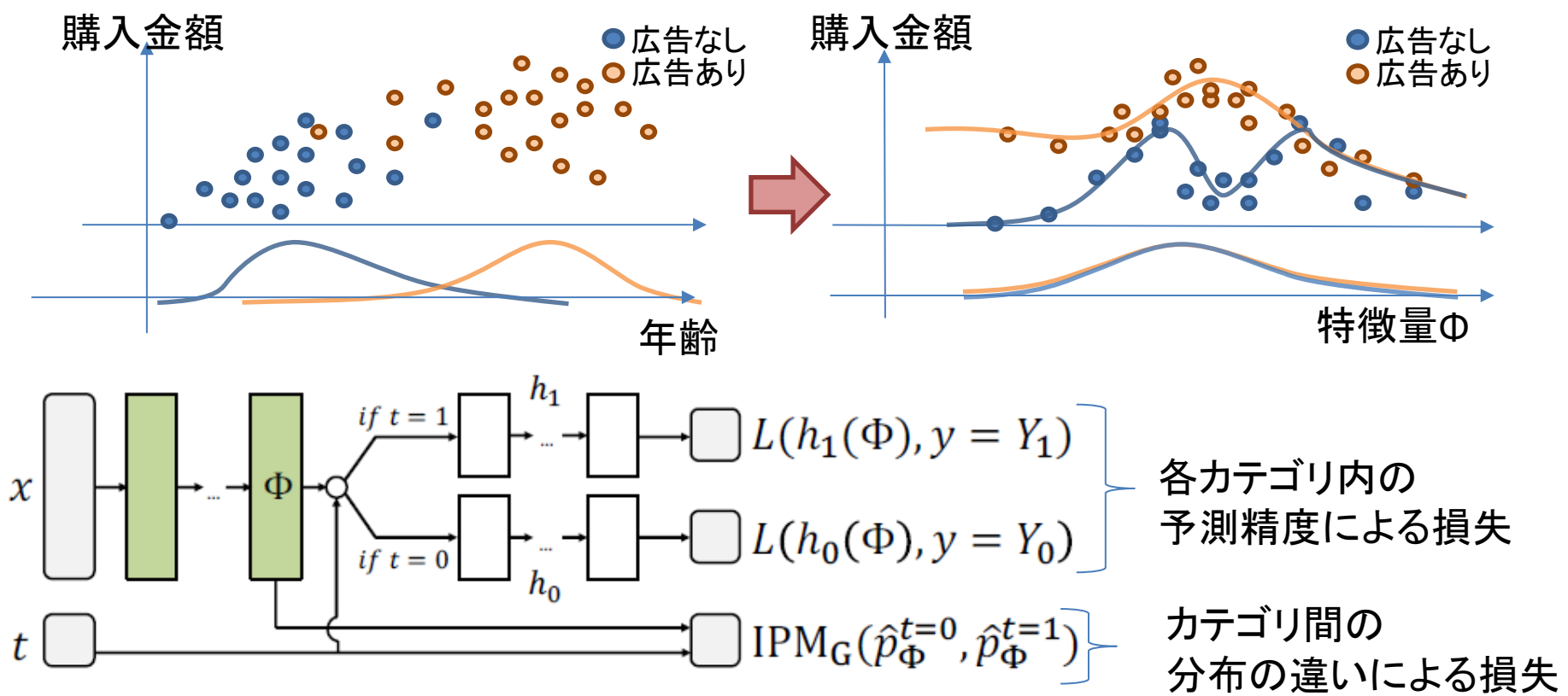
反実仮想

この時、反実仮想特有の問題として広告の与えられた人と与えられなかった人の特徴が異なる点が問題である。



反実仮想機械学習

例えば、分布が同一の新たな特徴量 Φ を作成し、その特徴量をもとにそれぞれの予測モデルを学習する。



上記モデルは医療、経済分野のデータで有効性が確認されている。[1].

□Uberの事例

Uberは一部のユーザにクーポンを配り、離反しないようにしている。

Uber

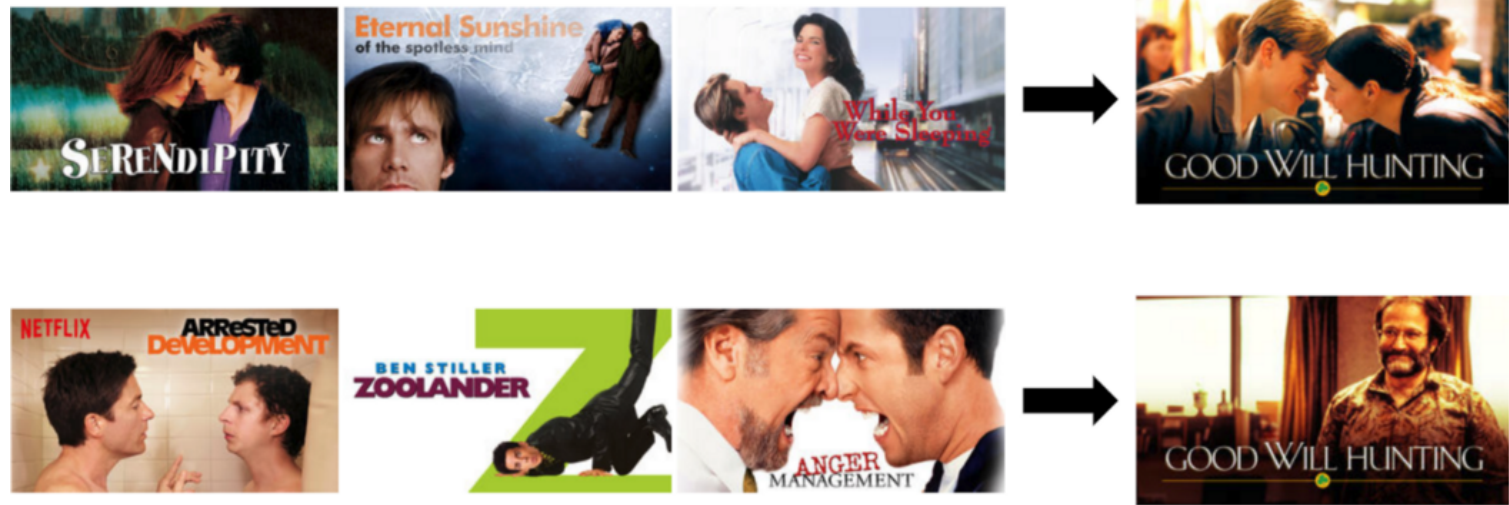
ユーザ	クーポンあり	クーポンなし
1	✕	?
2	?	○
3		✕
4	○	?
⋮	⋮	⋮

クーポンを提供するユーザのコストなどを加味した選択の最適化に、
ITE Predictionの技術が用いられている。 [2]

反実仮想機械学習 - Off-Policy Evaluation -

□ Netflixの事例

Netflixでは同じ作品でも人によって異なるArtworkが表示される。



表示するArtworkの最適化を機械学習で行われているが、モデルの評価にOff-Policy Evaluationの技術が使われている。 [3]

参考文献

- [1] Uri Shalit, Fredrik D. Johansson, David Sontag. “Estimating individual treatment effect: generalization bounds and algorithms”, arXiv preprint arXiv:1606.03976, 2017
- [2] Shuyang Du, James Lee, Farzin Ghaffarizadeh. “Improve User Retention with Causal Learning”, *Proceedings of Machine Learning Research*, 104:34-49, 2019
- [3] THE NETFLIX TECH BLOG “Artwork Personalization atNetflix”
<https://netflixtechblog.com/artwork-personalization-c589f074ad76>

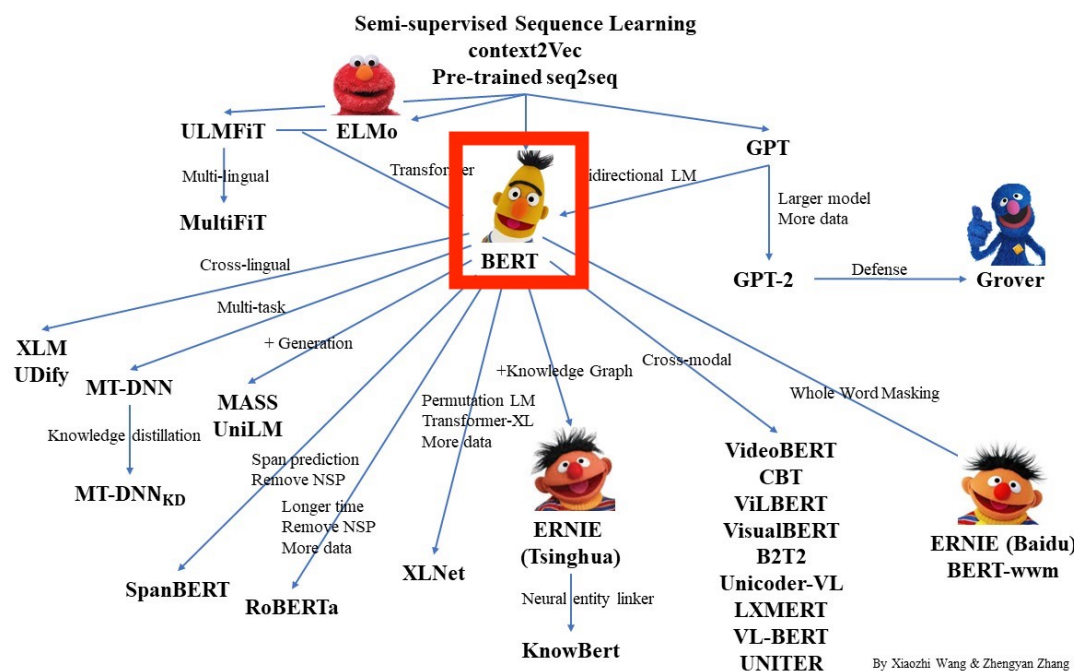
目次

1. 説明可能なAI
2. 反実仮想機械学習
3. 自然言語モデル BERT
4. 量子コンピュータ (時間があれば)

BERT -概要-

BERT(**B**idirectional **E**ncoder **R**epresentations from **T**ransformers)

- 2018年10月にGoogleのグループが発表した自然言語処理のモデル [1]
- 意味判定, 質疑応答などの様々なタスクで当時の最先端技術のスコアを大幅に更新。一部のタスクでは**人間を超えるスコアを記録**



BERT関連の研究の系統図。
BERTが発表されて以来、多くの
自然言語処理の研究がBERTを元
にしている

引用:
<https://github.com/thunlp/PLMpapers>

By Xiaozhi Wang & Zhengyan Zhang @THUNLP

BERTの学習

事前学習と転移学習の2段階で学習を行うことで幅広いタスクに対応

□ 事前学習

- 様々なタスクへの**応用を前提とした学習**
- 文章内の単語の穴埋めタスク, 2文が連続するかを判定するタスクで学習を行う
- 学習コストは高いが、学習済のモデルは汎用的に使える(データ:13GB, パラメータ:3.4億)

□ 転移学習

- 事前学習済みのBERTの最終層を目的のタスクに合わせた構造に付け替えて学習
- 新しいモデルをゼロから学習するより**少ないコストで学習を行える**(コスト:データ, 計算量)

彼は昨日買い物に_____。

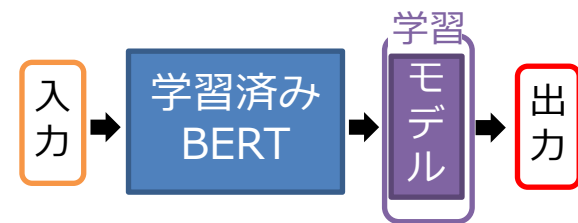
正解:行った

彼は昨日買い物に行った。
牛乳を2L買った。

正解:連続

彼は昨日買い物に行った。
ペンギンは空を飛べない。

正解:不連続



事前学習タスク(上図)と
転移学習のイメージ(下図)

BERTのネットワーク構造

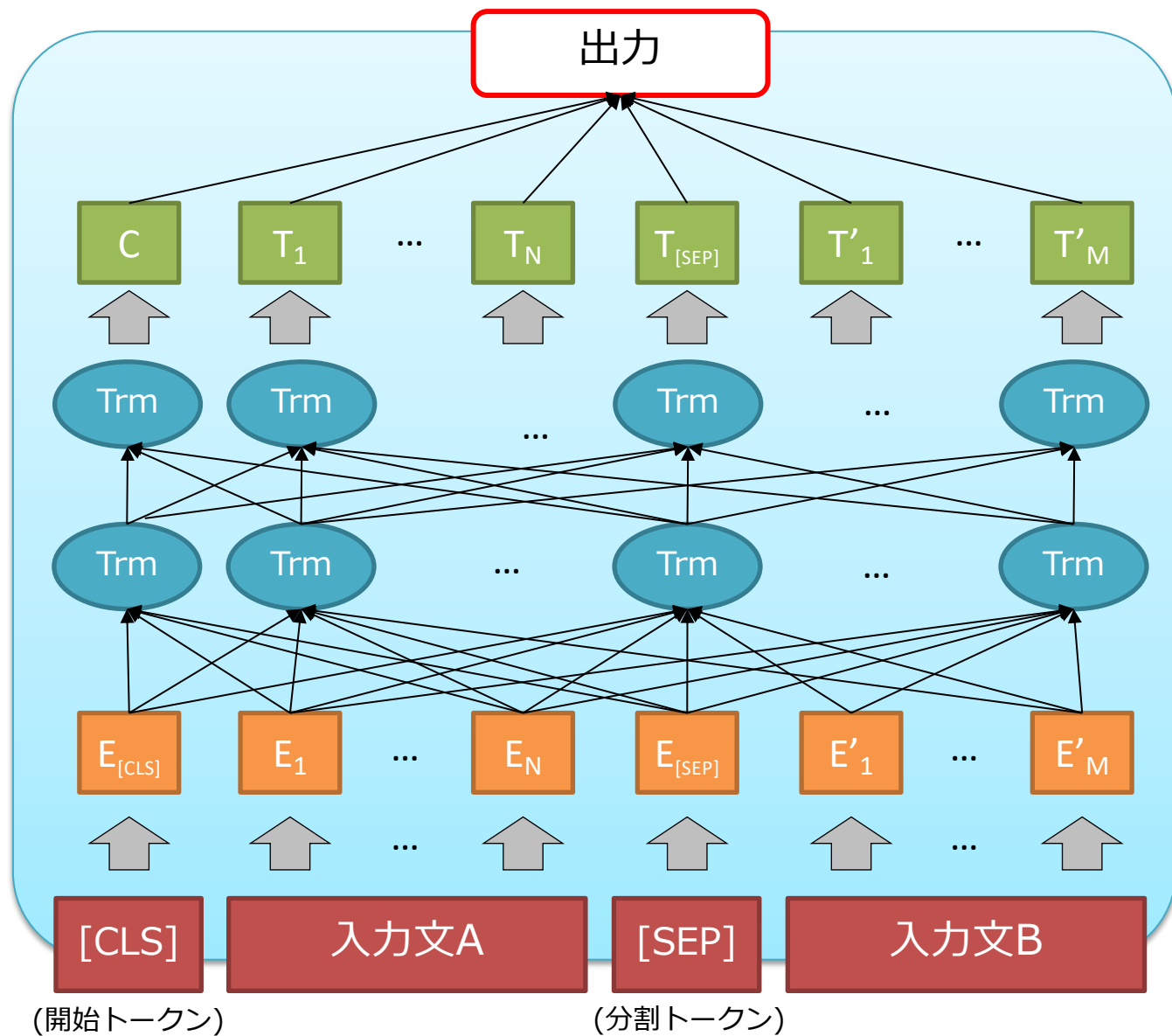
双方向の
ネットワーク



Transformer
(self-attention)
を利用



性能と計算の
効率性を両立



応用事例 1 : Google検索

- 2019年10月にGoogle検索のアルゴリズムにBERTが組み込まれる [2]
- 長い検索フレーズの微妙なニュアンスや文脈の理解度が大幅に上がる
- 長文で複雑になりやすい**音声検索の利便性が向上**

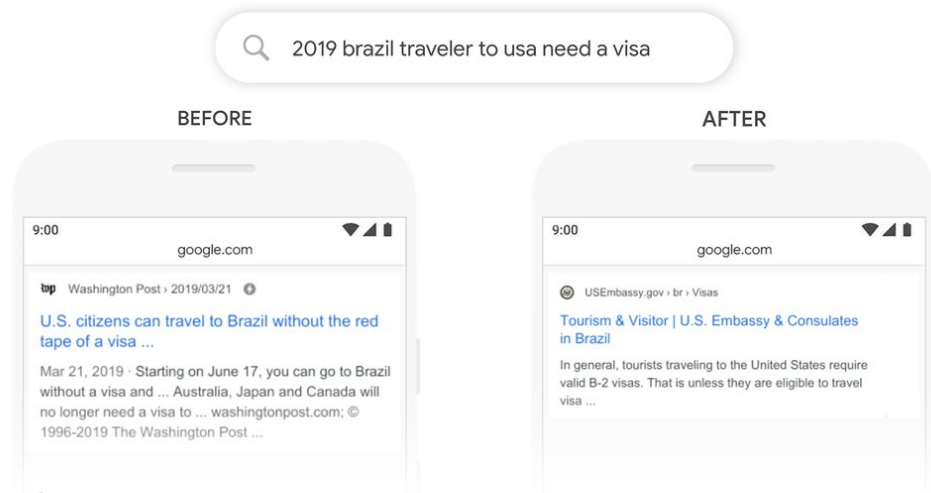
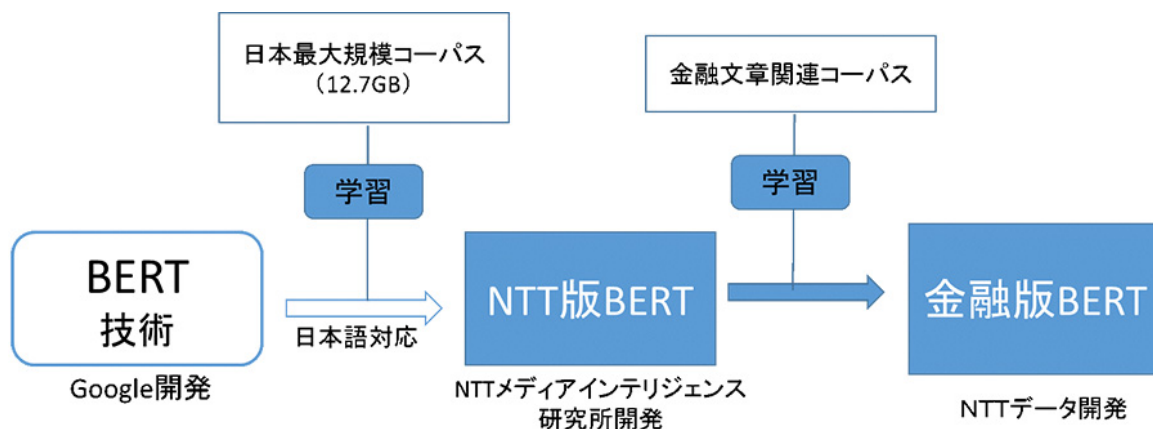


図: “2019 米国へのブラジル旅行者 ビザは必要”という検索フレーズに対して、以前は「ブラジルに旅行する米国人」についての記事を表示していたが、BERT導入により文脈を正しく理解し、「アメリカに旅行するブラジル人」についての記事を表示するように

引用:<https://blog.google/products/search/search-language-understanding-bert> [4]

応用事例 2 : 金融版BERT

- NTTデータが金融分野に特化したBERTを2020年に開発。2021年度中にサービス提供を開始 [3]
- まず日本語の大規模なコーパス(文章を集めたもの)で学習を行い、金融分野に特化させるために金融文書で追加学習
- 日報からの情報抽出, 稟議書の記載内容チェック, 財務情報からのリスク抽出, FAQの回答自動引き当て, チャットボットによる問い合わせ対応等への活用を想定



引用:
<https://www.nttdata.com/jp/ja/news/release/2020/071000/>

参考文献

- [1] Jacob Devlin, Ming-Wei Chang, Kenton Lee, Kristina Toutanova “BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding” arXiv preprint arXiv:1810.04805, 2018.
- [2] <https://blog.google/products/search/search-language-understanding-bert>
- [3] <https://www.nttdata.com/jp/ja/news/release/2020/071000/>

目次

1. 説明可能なAI
2. 反実仮想機械学習
3. 自然言語モデル BERT
4. 量子コンピュータ

量子コンピュータ -概要-

量子コンピュータ：
量子力学の原理を用いたアナログコンピュータ
ゲート型とアニーリング型に分かれる

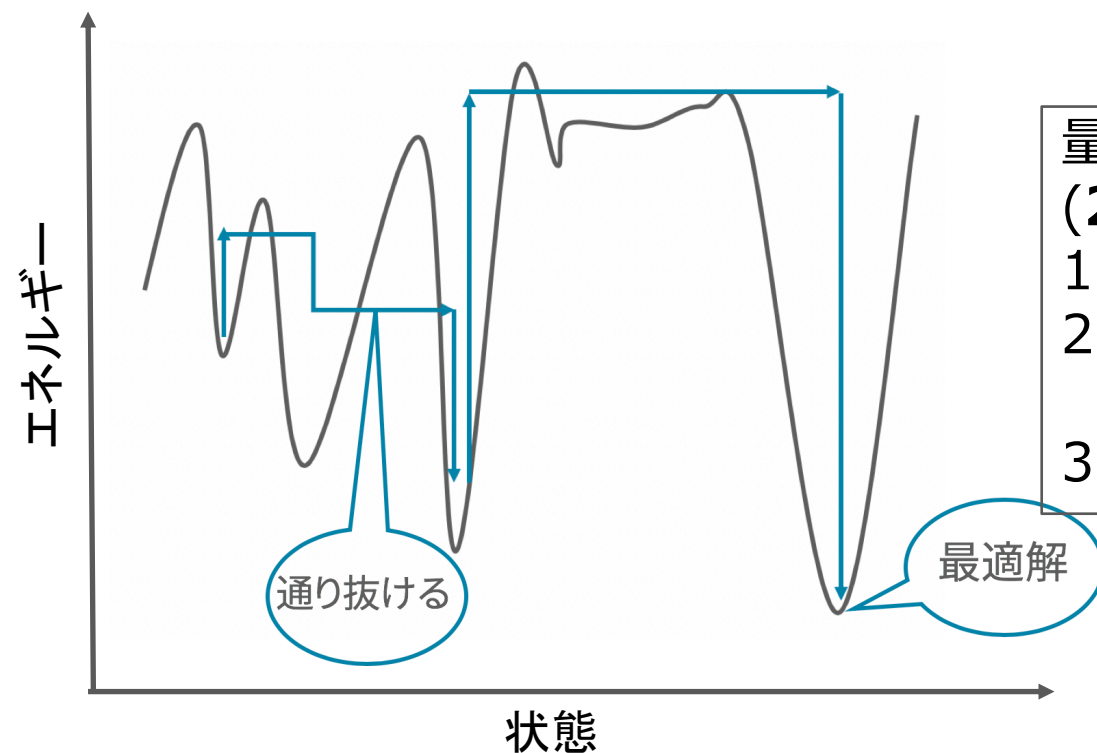
	量子ゲート方式	量子アニーリング方式
用途	全ての計算	組合せ最適化, 機械学習など
特徴	高速化を保証したアルゴリズムがあり	ヒューリスティック, ノイズに強い
量子ビット数	70程度	5640
開発企業	IBM,Googleなど	D-Wave Systems, NEC

ゲート方式はノイズの影響で現状では実用的な計算が難しい.

量子アニーリング方式では, 組合せ最適化や機械学習の小規模な問題に対して, 実用的な計算能力を獲得しつつある.

量子アニーリングの原理と組合せ最適化

エネルギー関数の山を通り抜けて、最適解に到達



量子アニーリングの手順
(**20 μ s**ほど)

1. 横磁場を強くかける
2. 徐々に弱める
(アニーリング)
3. エネルギーが低い状態に到達

物理的なエネルギー関数をうまく調節することで、
解きたい組合せ最適化問題に対応させる。

組合せ最適化を解くまでの手順

問題を
定式化

QUBOで
定式化

$$\begin{array}{ll} \text{minimize} & \mathbf{x}^T Q \mathbf{x} \\ \text{s.t.} & \mathbf{x} \in \{0, 1\}^N \end{array}$$

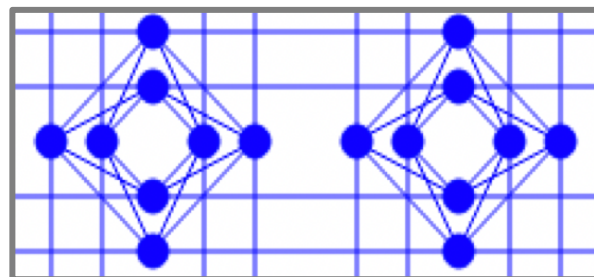
エネルギー
関数に変形

イジング
モデルへ

$$E = \sum_{i < j} J_{ij} \sigma_i \sigma_j + \sum_i h_i \sigma_i \quad (\sigma = \pm 1)$$

ハードウェア
上に表現

ハードウェア
グラフへ



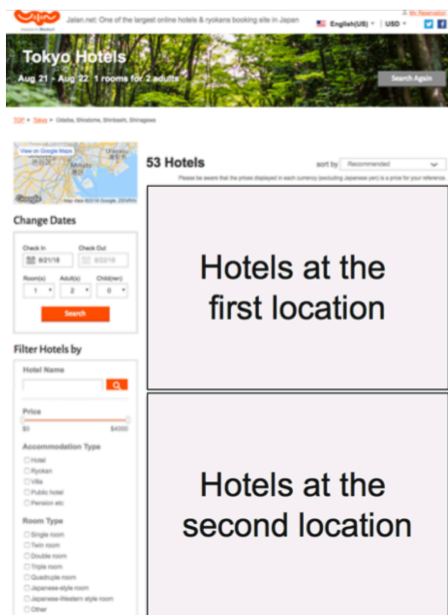
量子アニーラ
に投入

最適化

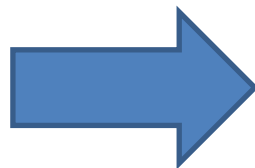


事例1(組合せ最適化)

- ホテル予約サイトにおけるWebページ内の表示ホテルの配置最適化
- 推定売上を最大化しつつ、類似したホテルの並びへのペナルティを与える



最適化



1. B. North City hotel
2. G. South Budget hotel
3. A. North City hotel
4. E. North Budget hotel
5. H. South Budget hotel
6. C. North City hotel
7. F. South Budget hotel
8. D. North City hotel

ホテル予約サイト
(じゃらん)

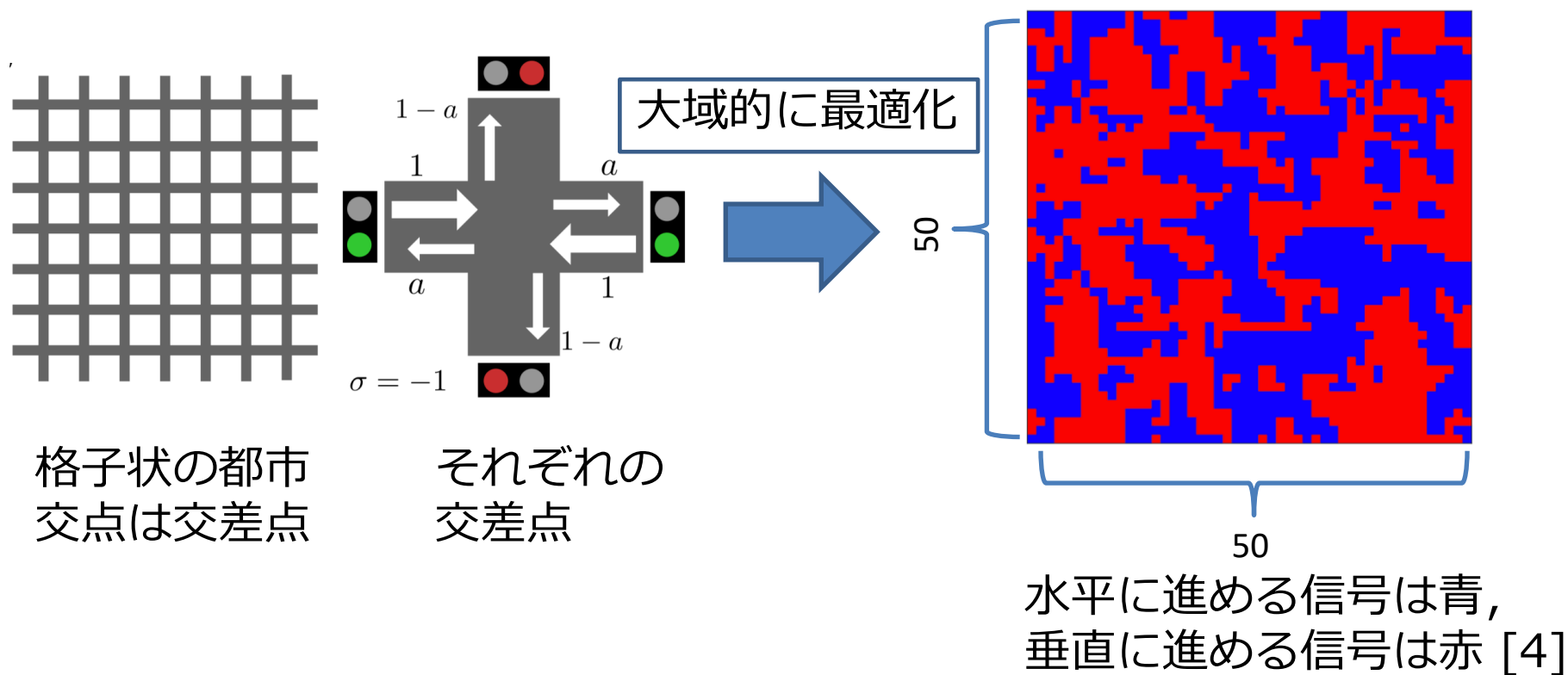
異なる種類の人気ホテルが並ぶ[3]

類似するアイテムが並ぶことを避けながら、人気のアイテムの配置に成功

事例2(組合せ最適化)

信号の最適化

50×50の道路を持つ都市で、交通量と信号の切り替わり頻度を考慮して最適化



交通量には仮定を置きながらも大域的な最適化に成功 [4]

事例3(機械学習)

画像の特徴抽出

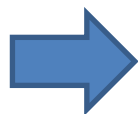
Wを古典計算機で, Hを量子アニーリングで, ハイブリッドに最適化

$$V \approx WH$$

V :大量の画像

W :特徴を表す行列

H :特徴を選択する01行列



[5] 5行7列の画像のマトリックスは、学習された特徴.
左の2つの画像は、元の画像(上)と再構成(下)を示す.
再構成は、緑色で囲まれた特徴を合計することで得られる.

多くの画像から、共通する特徴を抽出し、画像の再構成に成功

参考文献

- [1] Jun Cai, William G. Macready, Aidan Roy. "A practical heuristic for finding graph minors." *arXiv preprint arXiv:1406.2741* (2014).
- [2]
https://www.dwavesys.com/sites/default/files/Advantage_Datasheet_v9_0.pdf
- [3] Naoki Nishimura, Kotaro Tanahashi, Koji Suganuma, Masamichi J. Miyama, Masayuki Ohzeki. "Item listing optimization for e-commerce websites based on diversity." *Frontiers in Computer Science* 1 (2019): 2.
- [4] Daisuke Inoue, Akihisa Okada, Tadayoshi Matsumori, Kazuyuki Aihara, Hiroaki Yoshida. "Traffic Signal Optimization on a Square Lattice using the D-Wave Quantum Annealer." *arXiv preprint arXiv:2003.07527* (2020).
- [5] Daniel O'Malley, Velimir V. Vesselinov, Boian S. Alexandrov, Ludmil B. Alexandrov. "Nonnegative/binary matrix factorization with a d-wave quantum annealer." *PloS one* 13.12 (2018): e0206653.

おわりに

本講演で紹介したAI技術

1. 説明可能なAI
2. 反実仮想機械学習
3. 自然言語モデル BERT
4. 量子コンピュータ

近年すごい勢いで機械学習・AI技術は発展しており、
今後ファイナンス分野の様々な場面での応用も期待できる