



2020年7月29日

東北大学 知のフォーラム

第5回 実験家のためのデータ駆動科学オンラインセミナー
人間と社会のデータ科学

超スマート社会と生活者理解

東北大学大学院経済学研究科

石垣 司



2019年7月30-31日
東北大学経済学部オープンキャンパス

東北大学経済学部
理系入試
2020年度入学者より導入

3回 (AO前期・一般入試前期・後期) の理系入試のチャンス
日本の経済学部で唯一 完全理系型の入試科目
新たな社会 Society 5.0 を先導するエコノミストを育成

Faculty of Economics

理系入試始まる！
来るべき近未来社会 Society 5.0 を
創り出す皆さんへ

昨年度のオープンキャンパスでの講演資料

経済学部で理系入試？（背景）

- Society 5.0: データを社会に還元し価値を生む超スマート社会



- Society 5.0 ではデータが“資源”
– AI技術、ビッグデータ活用があらゆる産業で浸透

内閣府HPより

第3次AIブームの本質は“データ”

- Deep Learning **ブームの歴史的背景** (超短縮版)

	[Data]	[Hardware]	[Method]
1990's	ICTの発展		
2000's	ICTの普及	GPUの転回	Deep化の再発見
2010's	ビッグデータ	GPUの普及	Deep化技術の発展

現在の第3次AIブーム



将棋
プロ棋士(2013) 名人(2017)

囲碁
元世界ランク1位(2016) 現役世界ランク1位(2017)



画像認識
人間の精度を超える(2015)



<https://pocketalk.jp/>
自動翻訳



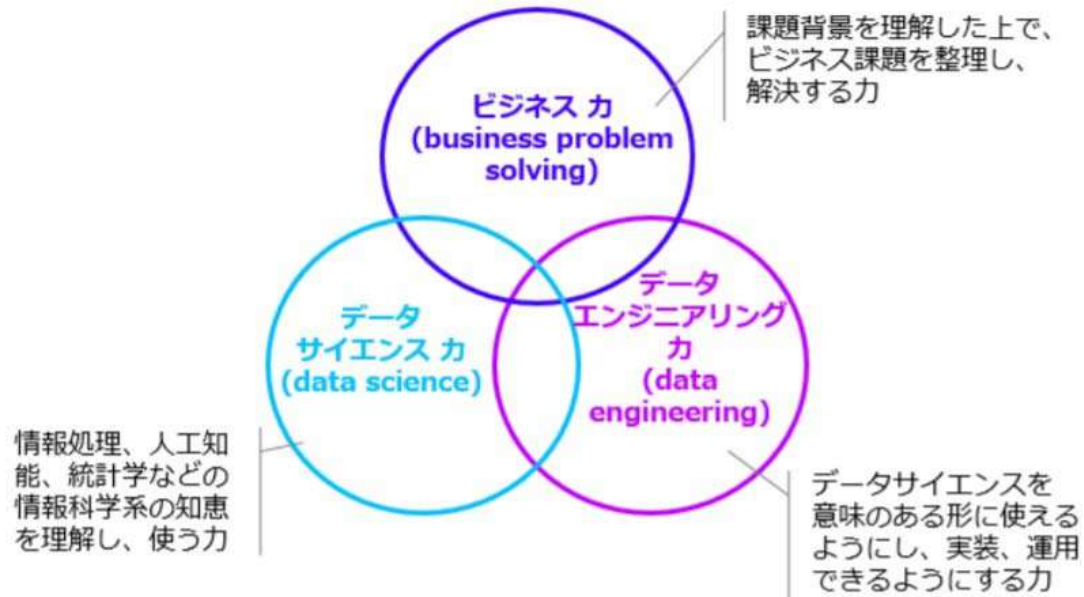
無人走行バス
地域を支える新しい交通手段。
高齢者の移動もより便利に。
内閣府HPより
完全自動走行への期待

- 技術的本質 ⇒ ビッグデータ + データ学習技術**

理学部や工学部の仕事でしょ？

- **必要とされる経済・経営がわかるデータ科学人材**
 - データ科学には数学の力が必要
 - しかし、数学やプログラミングスキル“のみ”では不十分

データサイエンティストに求められるスキルセット



データサイエンティスト協会HPより

Advantage 1

5年で修士号が取得可能なプログラム

東北大学経済学部では **学部4年** + **修士1年** の **5年** で修士号取得が可能な4つの一貫プログラムを用意しました

- データ科学
- 日本の経済・経営
- 高齢社会の地域経済政策
- 高度グローバル人材コース

Advantage 2

ビジネス力を備えた**データサイエンティスト**の養成

- 現代社会では、**文系・理系の知識をともに備えたIT人材**が求められています
- **データを分析してビジネスに役立てるアナリスト、コンサルタント、マーケター**などの活躍の場が広がっています

経済・経営学の知識を修得

経済学、経営学、会計学、統計学、経済史

Society 5.0^{*}
活躍できる人材

修士課程

統計・計量分析特別演習
データ科学特別演習
中級計量経済学特論
上級計量経済学特論
Applied Data Sciences

3・4年次

計量分析
ファイナンス
オペレーションズ・リサーチ
応用実証分析

2年次

数理統計
経済経営数学
データサイエンス

1年次

経済経営数学基礎
統計学入門

理系の知識を活用

生活者理解による超スマートサービス

- **超スマート社会** (Society5.0, 第5期科学技術基本計画)
 - 「必要なもの・サービスを、必要な人に、必要な時に、必要なだけ提供」、するためには？
- **十人十色のサービス受容者(生活者)の理解**
 - **多種多様なデータの活用 ⇒ スマートなサービス提供へ**
 - “無駄な資源、時間、労力の最小化” & “便益、満足度の最大化”



非スマートな画一的サービスに

超スマートサービス技術

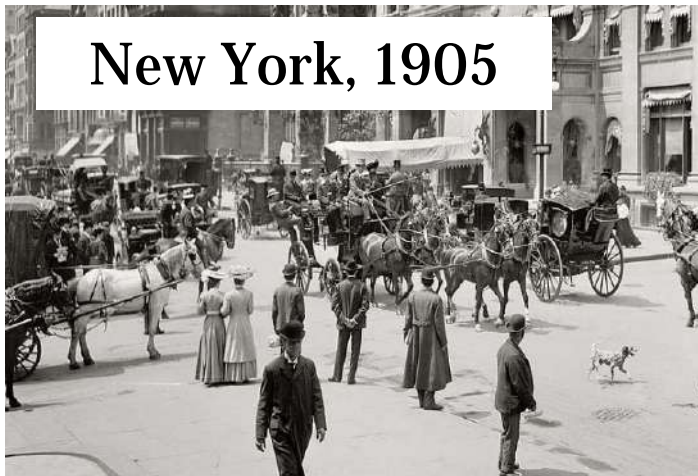
- 一人一人の特性・嗜好の理解
- プライバシーの重視
- Big Data 活用



スマートサービス実現の土壌

生活者のニーズを埋める

- **ニーズ(欠乏状態)は聞いても分からない**(スマホが無くて不便だった?)



https://commons.wikimedia.org/wiki/File:5th-Ave-New-York-N.Y._1905.jpg

1769蒸気自動車、1885ガソリン自動車
(貴族・富裕層の趣味)

T型フォード発売
1908



簡素な構造
運転が簡単



https://commons.wikimedia.org/wiki/File:Ave_5_NY_2_fl.bus.jpg

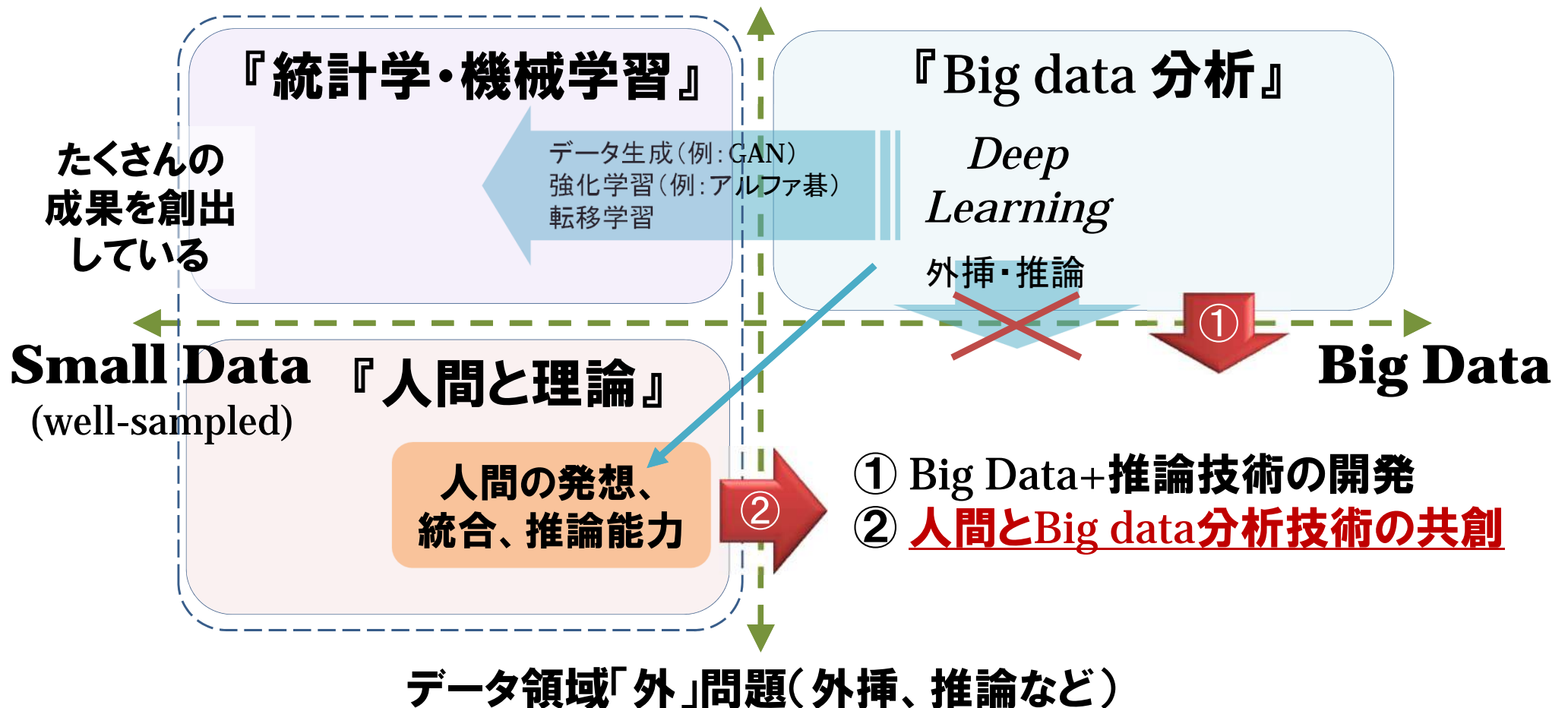
- データ・調査無くして「**“無い”状態**」の把握は困難
- **しかし、ニーズ探索は外挿・推論が必要**

『帰納型のAIは、近くのサンプル(過去の経験)を参考に“みようみまね”をしている「内挿マシン」にすぎない。よって、データ集合が占める領域から遠く離れた空間(未体験ゾーン)の予測、つまり外挿は非常に不得手である。』[樋口知之, 情報管理, vol.59, 2016]

生活者データからニーズを埋めるために

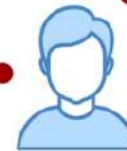
- 現在の技術的背景での(平凡な...)帰結

データ領域「近傍」問題(内挿、認識・検出など)



データから個人のニーズを埋める

- Big Data × Deep Learningの推薦システムを例に
 - 行動変容(需要の喚起)を引き起こすことができるか？



Accurate Recommendation

  **Game of Thrones: Season 2**
 96%
 2012, HBO, 10 episodes

- Accurate; Closely related to history items.
- But **already well-known** to user (same series).

推薦されなくても見る

Novel Recommendation

  **Chernobyl: Miniseries**
 96%
 2019, HBO, 5 episodes

- Related to history items; High rating.
- Less Popular; **Unlikely to be known** by user.
- **Help items discovery.**

推薦されたから見た

ニーズを埋める～推薦システムを例に #1

- 高い推薦精度で、かつ、ありきたりな推薦を避ける
 – マルチカテゴリ商品の同時学習



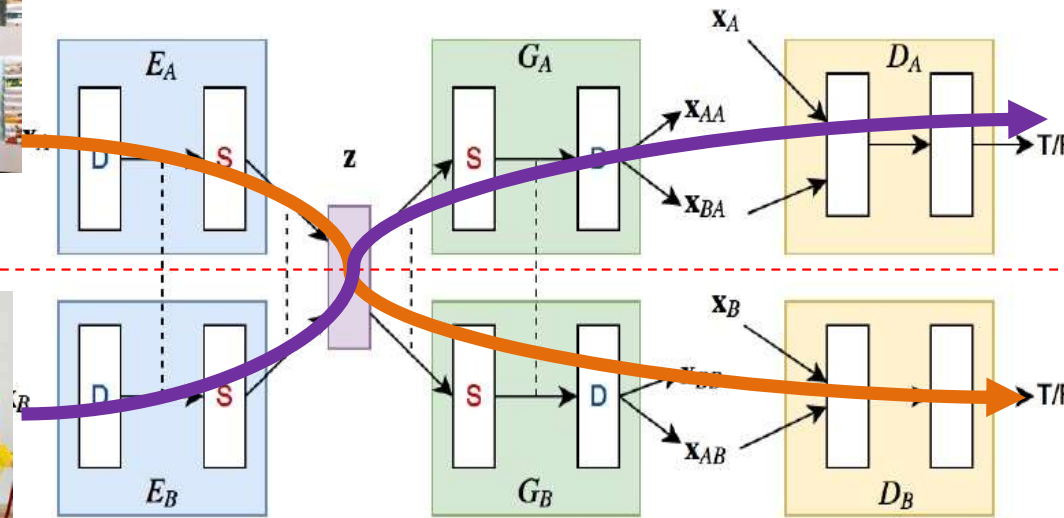
Linh Nguyen君
(Ph.D. student)



洋服カテゴリ



パーソナルケア製品カテゴリ



洋服の購買履歴から
パーソナルケア製品
の嗜好を推定

パーソナルケア製品の
購買履歴から
洋服の嗜好を推定

- Cycle VAE-GAN -
Deep Learningの一種

高い推薦精度で異なるカテゴリの商品の嗜好を推定

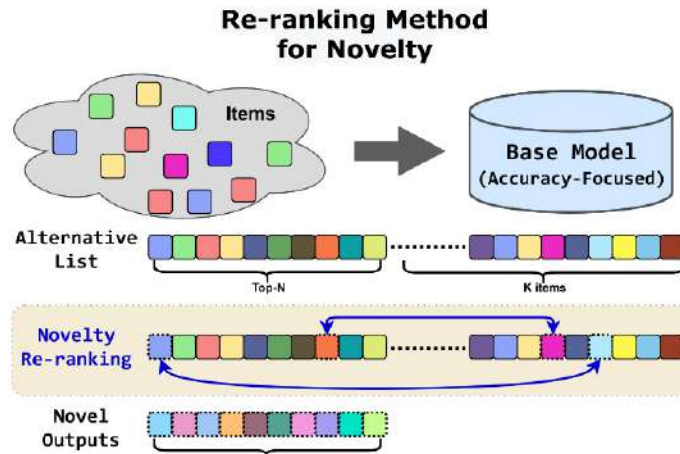
ニーズを埋める～推薦システムを例に #2

推薦精度と“目新しさ”の両立

– 1-stage かつ end-to-end アルゴリズムの開発



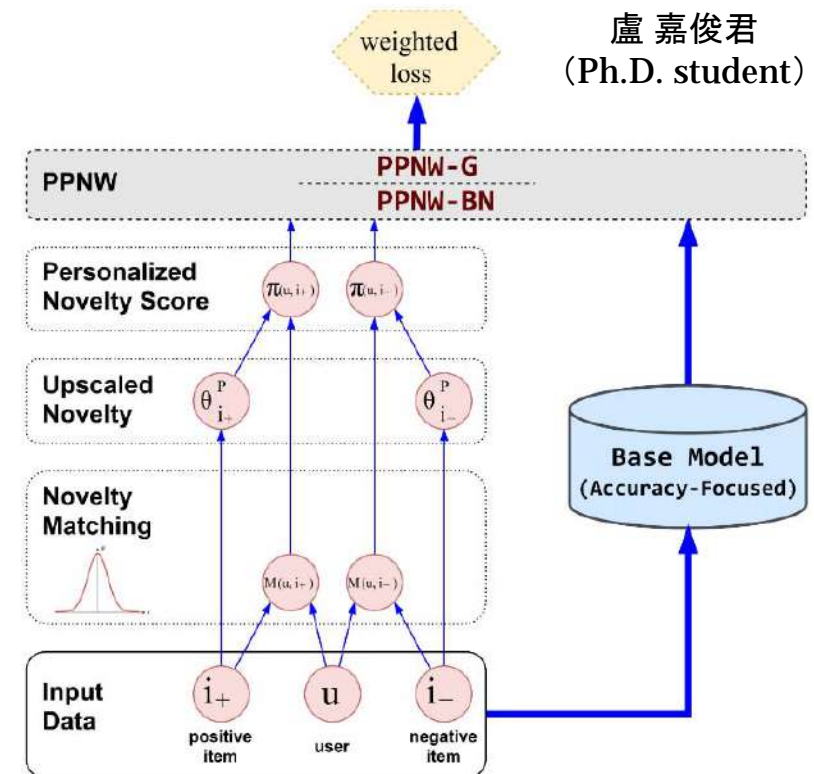
盧 嘉俊君
(Ph.D. student)



Limitations of most Loss Functions

- Most loss functions are **incapable of distinguishing** 2 types of negative item, “unknown negative items” and “disliked negative items”.

Note that in recommender system, we want to suggest those “**interesting but unknown items**”.



高い推薦精度で新規性の高い商品の嗜好を推定

スマートな意思決定サポートへ向けて

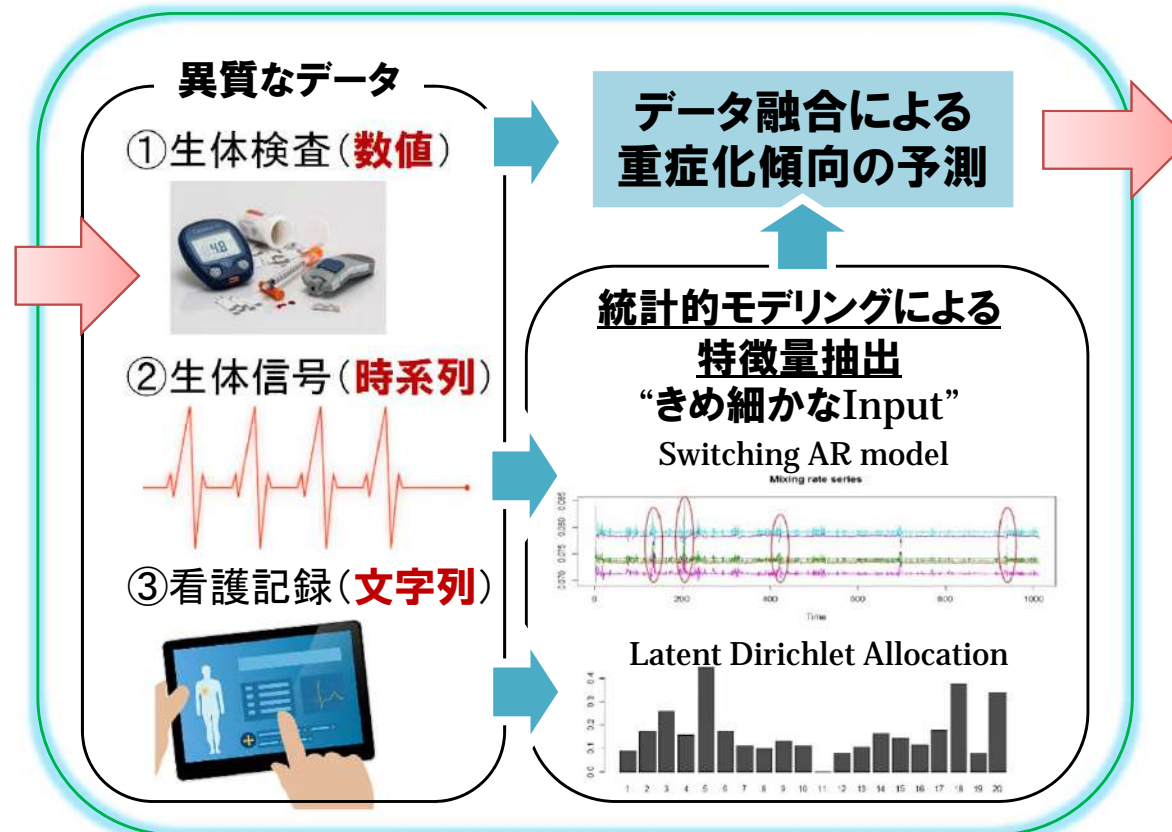
- **集中治療室内患者の早期転帰予測**
 – 東北大学病院との共同研究(症例数137)



院生
募集中

石塚治也君
(H30 修士課程修了)

重症化しやすい
患者を早期に
予測したい



**必要な労力を
必要な患者へ**

予測精度Good

- 従来法と比べてAUCの値が0.1ポイント以上の大きな改善
- 定性的データである看護記録の利用が予測精度向上に大きく貢献
- 人的制約の厳しい医療スタッフの負担緩和のため、重症化傾向の患者をAIが警戒し医師へ報告するシステムを目指している

石塚治也、石垣司、小林直也、工藤大介、中川敦寛、異質なデータを統合した敗血症患者の転帰予測システム、電子情報通信学会論文誌D, Vol.J101-D, No.3, pp.481-493, 2018 (学生論文特集秀逸論文)

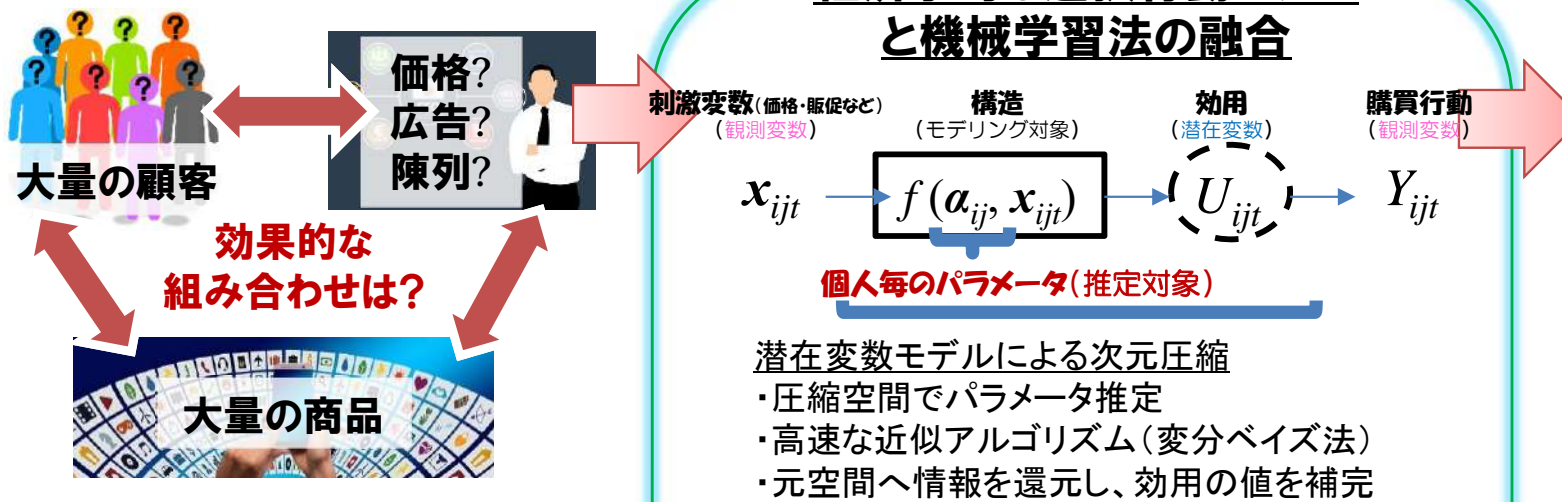
H. Ishizuka, T. Ishigaki, N. Kobayashi, D. Kudo, A. Nakagawa, In-hospital Mortality Prediction for Trauma Patients Using Cost-sensitive MedLDA, Data Science and Service Research Discussion Paper, No.79, 2018

超スマートマーケティングを目指して

“顧客&商品&マーケティング施策”の発見

– Big Dataの本質的情報不足を克服 (パラメータ数 >> データ数)

ベイズモデリングによる
経済学的な選択行動モデル
と機械学習法の融合



必要な資源を
効果的な顧客へ

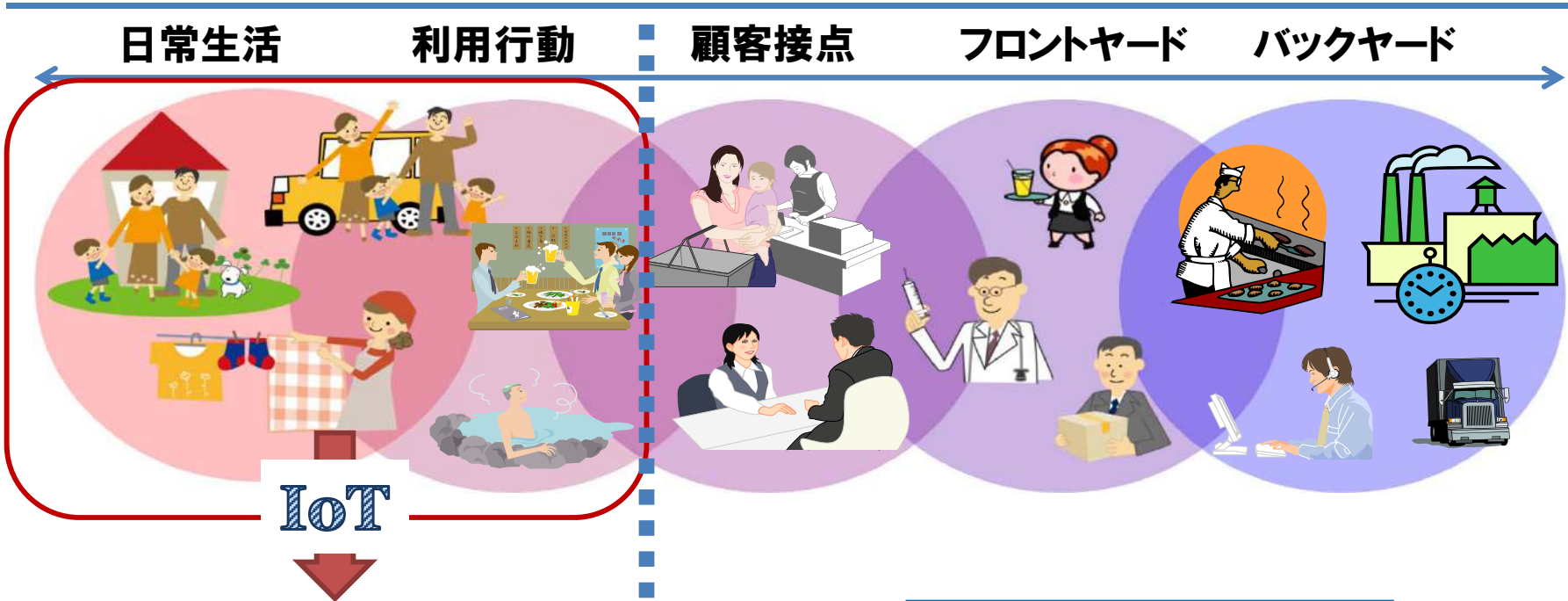
異質性の発見

- スーパーマーケットの購買データに適用 (1600人×500商品×3施策=2,400,000パラメータ。(3か月分のデータ数94,000))
- マーケティング施策の“効果あり”と“効果なし”をマーケティング施策毎に把握
- 店舗全体規模の最適化へ

[例] 1000人×100商品×5施策 = 500,000組合せ(パラメータ) (>>購買件数)



おわりに



これからのビッグデータ

- コネクテッドカー
- ウェアブルセンサ
- スマート家電, etc.

今までのビッグデータ

- 購買履歴データ
- 診察画像データ
- 品質管理データ
- 金融取引データ
- コールセンター
- 発注データ
- 交通ICデータ
- 接客データ
- 機器利用データ, etc.

- モノ・サービスの利用場面の様々なデータが利用可能な時代へ突入
- サービス提供から生活・利用まで俯瞰できる視座の育成
- 超スマート社会ではビッグデータの質がより重要に