

数学 meets データ分析

2016/11/6

原田 慧

KEI HARADA

自己紹介

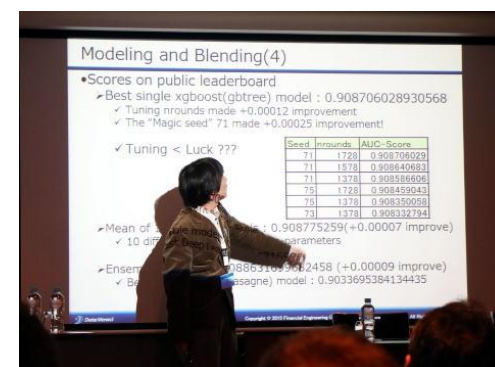
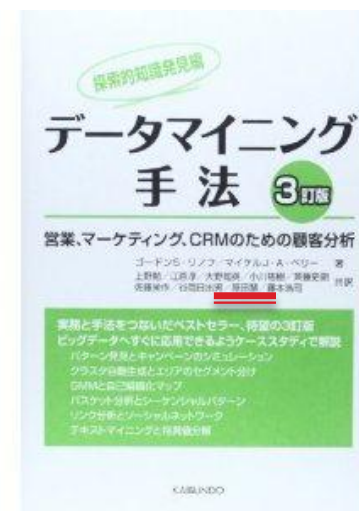
▶ いわゆる「データサイエンティスト」

- 某データ分析企業勤務
- 佐野くんの先輩
- @hoxosh (タコの絵のひと)



▶ 数理学博士から就職

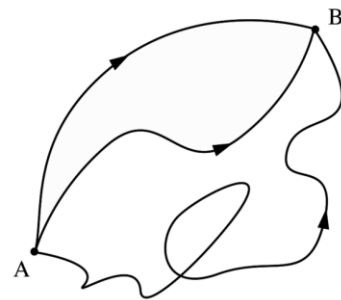
- 何でも不器用にこなし、特に数学が得意
- データ分析の本をちょっと翻訳した
- KDDCUP2015というデータ分析コンペで世界2位！
 - ITLeadersというところに連載した
- <http://it.impressbm.co.jp/articles/-/12969>
- 人狼知能ミニ大会というもので優勝した
- <http://aiwolf.org/>



学生時代の研究テーマ(1)

- ▶ 大義名分は「Feynman経路積分の数学的な定式化」
 - すごく振動する関数(の関数)を、AからBに至るすべてのpathについて平行移動不変な積分をするとシュレディンガー方程式の基本解ができる

$$\exp \left\{ i \int_{t_a}^{t_b} L(t) dt / \hbar \right\} = \exp \{ i S(t_b, t_a) / \hbar \}$$



- ▶ どこが難しいか(面白いか)
 - 1. 振動するものに数学はとても弱い
 - 2. Pathの空間は無限次元空間で、いろんなものがコンパクトでなくて詰む、例えばLebesgue測度がない
絶対可積分は甘え コンパクト近傍は甘え

学生時代の研究テーマ(2)

- 1. 振動するものに数学はとても弱い、だけちょっと説明
- 例題1: $S1 = -1 + 1 - 1 + 1 - 1 + 1 - 1 + \dots = ?$
- 例題2: $S2 = -1 + 2 - 3 + 4 - 5 + 6 - 7 + \dots = ?$
- 学生のとくに閃いたことは、 $(0.25, 0.50, 0.25)$ と合成積をとると、 $S1 = -0.5, S2 = -0.25$
- $S3 = 1 + 2 + 3 + 4 + 5 + 6 + 7 + \dots$ とおくと、
 $S2 + S3 = 4 + 8 + 12 + 16 + \dots = 4 * S3$
よって、 $S3 = -1/12$
ゼータの計算ができた！

今日のお話の要約

- ▶ 「データサイエンティスト」はセクシーらしいね
 - あと「人工知能」すごいらしいね
 - ▶ 数学はデータ分析の役に立つ
 - プログラミングもできるとさらに強い
 - ▶ データ分析ができるとこういう楽しい問題がある
- ▶ 結論：数学の人がデータ分析やるとたぶん楽しい
 - だからみなさんもやってみてください

データサイエンティストはセクシー？

- ▶ これは違う、自信を持って言える
 - やってることはかなり地味
 - 経営層から見たらもしかしたらセクシーかもしれないけど
 - ビジネスを数字で、客観的・合理的にみて、改善するというの
は、まあ本来当たり前のこと
 - データサイエンティストはそのとき使う武器がデータ・データ
分析であるだけ
 - 集計されたExcelではなくて、その元のデータを自由にさわる
 - 世の中の需要はかなりあると思う
 - ちょっとデモ

人工知能は人類を滅ぼす？

▶ これはわからない

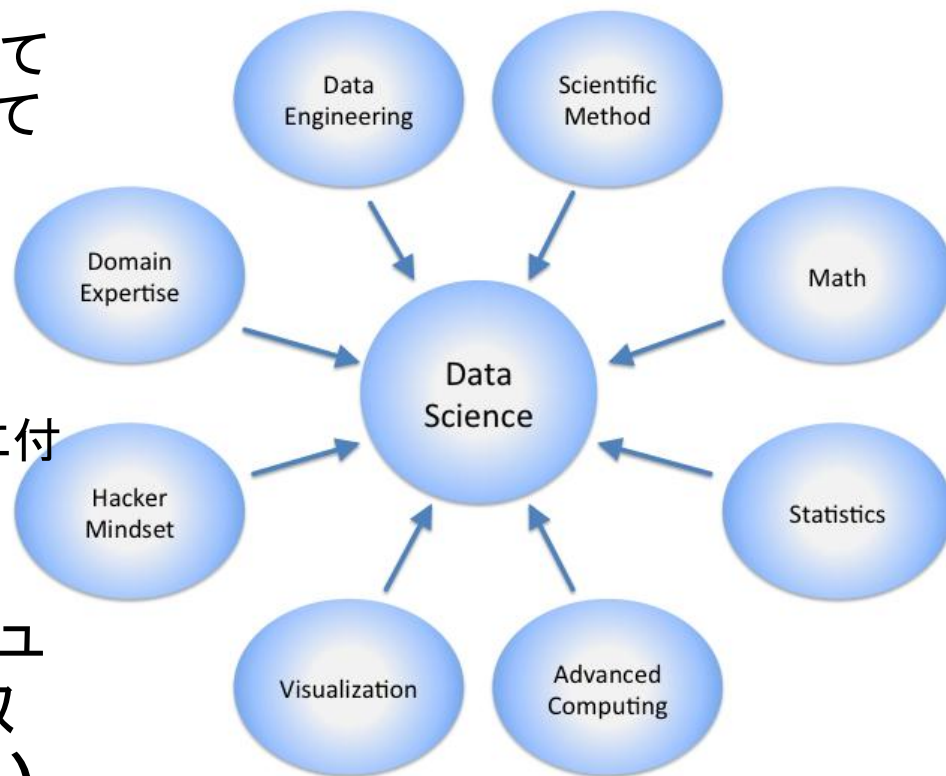
- 個人的には、もしすごい人工知能ができたら、ものすごく無関心・消極的に滅ぼされるかとも思っている
- いま新聞をにぎわせているAIはそういうものではない
 - どうやら人は自分にはよくわからないプログラムを「AI」と呼びたくなるらしい

▶ 強いAIと弱いAI

- 明確な定義はないけど、特定のことに強いのが弱いAI
 - データ分析の結果できるものは弱いAI
- 知性・自我・意識とかをもっているようにしか見えないのが強いAI
- 作る立場からすると、さわっているうちにどんどん弱いAIに見えてくるというジレンマ・・・

数学とデータ分析

- ▶ 数学は多少は必要
 - それ以上に、数学者が持っている素養がデータ分析にとってもマッチする
 - 特にHacker MindsetとScientific Method
 - これらはサラリーマンが身に付けるのはとても大変
- ▶ プログラムができて、コミュニケーションがちゃんと取れて、数学ができると強い



データ分析の技術から楽しい問題(1)

▶ マリオをクリアする人工知能(?)

- もともと親会社にいたtomomoto_factoryさんのネタ
 - DeepMindという会社がAtariでやったことで有名
 - AlphaGoの会社
- 画像とマリオの残機を入力して、ゼロからスタートして何度も何度もプレイするうちに強化学習によってだんだん賢くなるものを作りたい
 - 目標は1-4クリア
- ゼロの状態のデモ

データ分析の技術から楽しい問題(2)

▶ 人狼をする人工知能

- 人狼ゲームは人狼陣営(嘘つき)と人間陣営に分かれて、人狼は村人を騙す、村人は嘘つきを探すゲーム

▶ 人狼ゲームの特徴

- チェスや将棋、囲碁のような完全情報ゲームではない
- 対戦相手とのコミュニケーションが必要
 - 例えば合意形成
- 「騙される」って実はすごい知性
 - 行動をもとに推定するだけなら意外と簡単
- デモ