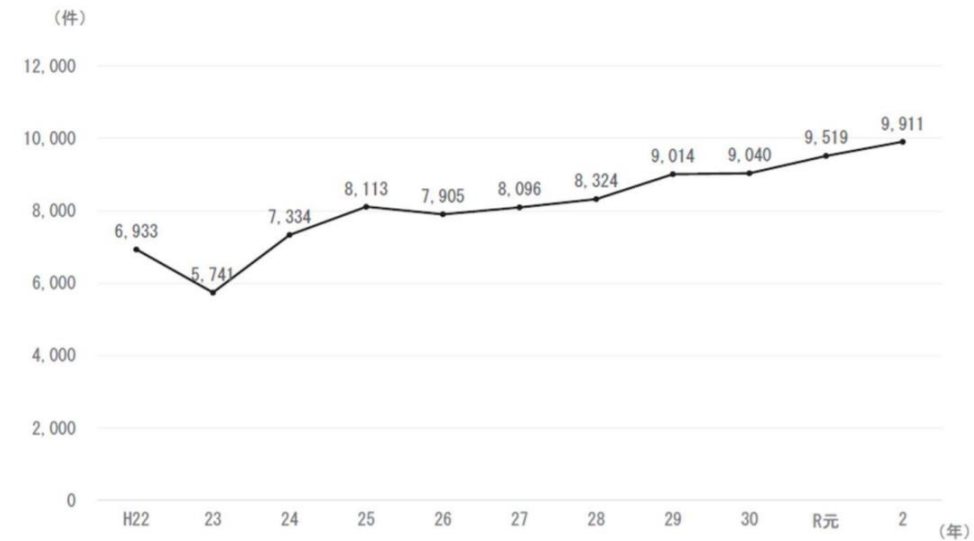


# インシデント損害金額を推定する ウェブサイトの開発と評価

菊池研4年 伊藤 充司

# 研究背景

- 近年、クレジットカードなどの個人情報の流出が増加している  
(2020年の上場企業で約100件)



(出典：警察庁)

- 個人情報漏洩した時の損害額推定のためのモデルが複数ある

# モデル比較

	JO[1]	Romanosky[2]	山田[3]
損害賠償額	○	○	×
セキュリティ対策	×	○	○
お詫び料	×	○	○
ソース	JNSA	Advisen	JNSA+決済短信

[1]山田道洋, 菊池浩明, 松山直樹, 乾孝治, "個人情報漏洩の損害額の新しい数理モデルの提案"

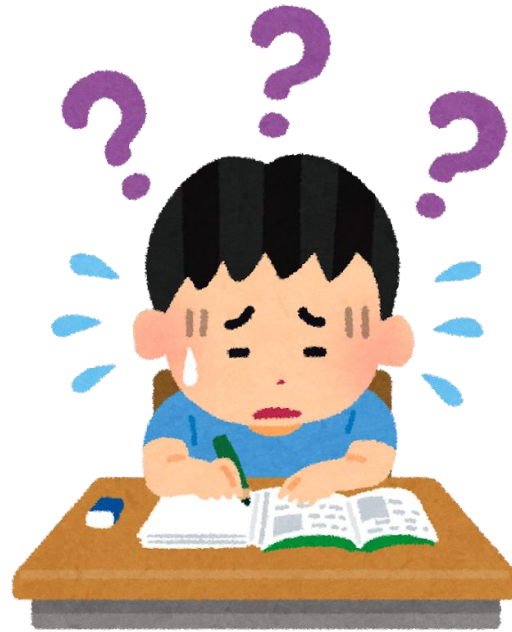
[2] Romanosky, S.: Examining the costs and causes of cyberincidents, Journal of Cybersecurity

[3]情報セキュリティインシデントに関する調査報告書別紙

([https://www.jnsa.org/result/incident/data/2016incident\\_survey\\_attachment\\_ver1.0.pdf](https://www.jnsa.org/result/incident/data/2016incident_survey_attachment_ver1.0.pdf))

# 問題点

モデルの計算式は条件や面倒な計算などがあり扱いにくい



# モデルの計算式

例) 山田モデル

$$\begin{aligned} & \log(\text{損害金}) \\ &= \beta_0 + \beta_1 \times \log(x_1) + \beta_2 \times \log(x_2) \\ & \quad + \beta_3 \times x_3 + \dots + \beta_{16} \times x_{16} \end{aligned}$$

係数		Estimate	定義域
定数		-3.9632	
log(被害人数)	$\log(x_1) \quad \beta_1$	0.0379	
log(売上高)	$\log(x_2) \quad \beta_2$	0.9904	
故意	$x_3 \quad \beta_3$	0.6261	0,1
事後対応度	$x_4 \quad \beta_4$	N/A	0,1
経済的ランク	$x_5 \quad \beta_5$	0.1590	1,2,3
精神的ランク	$x_6 \quad \beta_6$	0.0128	1,2,3
本人特定容易度	$x_7 \quad \beta_7$	0.2079	1,3,6
業種	不動産業, 物品賃貸業	-0.0773	0,1
	建設業	-1.4450	
	情報通信業	-0.1350	
	林業	-0.4030	
	電気・ガス・熱供給・水道業	-0.9330	
	生活関連サービス業, 娯楽業	-1.0040	
	卸売行, 小売業	-0.4550	
	医療, 福祉	-0.6319	
	宿泊業, 飲食サービス業	-0.4607	
	製造業	-0.7577	
	教育, 学習支援業	-0.0654	
	学術研究, 専門・技術サービス業	-0.1173	
	金融業, 保険業	-1.7570	
運輸業, 郵便業	-0.8893		
氏名	$x_9 \quad \beta_9$	-0.6231	0,1
住所	$x_{10} \quad \beta_{10}$	-0.5169	0,1
電話番号	$x_{11} \quad \beta_{11}$	-0.5337	0,1
生年月日	$x_{12} \quad \beta_{12}$	-0.2348	0,1
性別	$x_{13} \quad \beta_{13}$	0.2624	0,1
職業	$x_{14} \quad \beta_{14}$	0.1453	0,1
メールアドレス	$x_{15} \quad \beta_{15}$	-0.3845	0,1
ID/PASS	$x_{16} \quad \beta_{16}$	-0.2810	0,1

# 研究目的

- 知識のない一般の人でも簡単に被害金額の推定を行えるウェブサイトの開発
- 2021年のインシデントに適用して各モデルの有用性の評価

# 研究目的

- **知識のない一般の人でも簡単に被害金額の推定を行えるウェブサイトの開発**
- 2021年のインシデントに適用して各モデルの有用性の評価

# 開発したウェブサイト①

### 被害金額推定

被害人数は何人ですか？（人）

---

インシデントの起こった年の売上高はいくらですか？（百万円）

---

故意？

事故

故意

---

事後対応度？

適切

不適切

## 入力

- 入力項目は被害人数や売上高など18項目



# 開発したウェブサイト②

## 被害金額推定

各モデルにおいて推定される被害金額は以下の通りです。

各モデル	推定損害額
山田モデル	2,148,245.52 万円
JOモデル	160,314,000.00 万円
Romanoskyモデル (一部省略版)	28,146.42 万円

戻る

## 出力

- 3つのモデルを整理して出力している

例) ベネッセホールディングスが2014年に起きた情報漏洩

# デモ

- ベネッセホールディングス

2014/6/27 に判明した情報漏洩について損害金額を推定



The screenshot shows the Benesse Customer Service website. The header includes the Benesse logo and the text "ベネッセお客様本部" and "より確かな安心・信頼に向けて". A navigation menu contains links for "TOP", "情報セキュリティ強化の取り組み", "不審な勧誘への対応について", "ベネッセこども基金", and "お客様本部について". The main content area shows a breadcrumb trail: "お客様本部TOP > お客様本部について > 事故の概要". A sidebar on the left has a menu with "お客様本部の想い", "事故の概要" (selected), and "今までの取り組み". The main text area is titled "事故の概要" and contains a sub-section "事故発覚の経緯" with a blue background. The text below reads: "2014年6月27日、株式会社ベネッセコーポレーションは、お客様からの問い合わせによりお客様の個人情報情報が社外に漏えいしている可能性を認識し、緊急対策本部を設置するとともに、これらの問い合わせで提供された情報を手がかりとして社内調査を開始しました。"

<http://windy.mind.meiji.ac.jp/~atsu110/Yamada/index.html>

# 使用する事件の詳細

- 実際の損害額：2,583,100万円
- 被害人数：2,895万人
- 売上高：46,326,400万円
- 情報漏洩経路：内部犯による持ち出し
- 法的措置：受けている
- 漏洩した情報：氏名、住所、電話番号、生年月日、性別、メールアドレス、出産予定日

調べた主なサイト

[https://www.benesse-hd.co.jp/ja/about/release\\_20140709.pdf](https://www.benesse-hd.co.jp/ja/about/release_20140709.pdf) (お詫びと説明)

<https://pdf.irpocket.com/C9783/Q8aV/le8C/mNjf.pdf> (決算短信)

# 意見

- 売上高はどこのもどれを見るべきなのかが分かりにくい
- どのモデルをどう参考にすべきなのか
- 社会的責任度の知名度の判断がしにくい

## 社会的責任度？

一般より高い

一般的

適正な取扱いを確保すべき個別分野の業種（医療、金融・信用、情報通信など）、公的機関 **知名度**のある企業は一般より高いに分類

# 研究目的

- 知識のない一般の人でも簡単に被害金額の推定を行えるウェブサイトの開発
- **2021年のインシデントに適用して各モデルの有用性の評価**

# 評価

## 方法

- 実際の損害金額から誤差が50%以内なら正解として正解率計算  
例) 1000万円が正解の場合  
500~1500万円を正解とする
- 一般的に誤差の計算として使われるRMSEを計算

# 評価に使うデータ

評価に使うデータは、JNSAが集計した2005年から2016年までのと自分で集計した2021年ので比較していく

期間	レコード数	企業数
2005-2016	15569	8853
2021	143	108

# 正解率の比較

各モデル	2005-2016[%]	2021[%]
山田	27.6	21.2
JO	6.2	6.3
Romanosky	9.2	8.5



# RMSEの比較

各モデル	2005-2016[万円]	2021[万円]
山田	86049	20922
JO	91128	15960
Romanosky	32680	15925

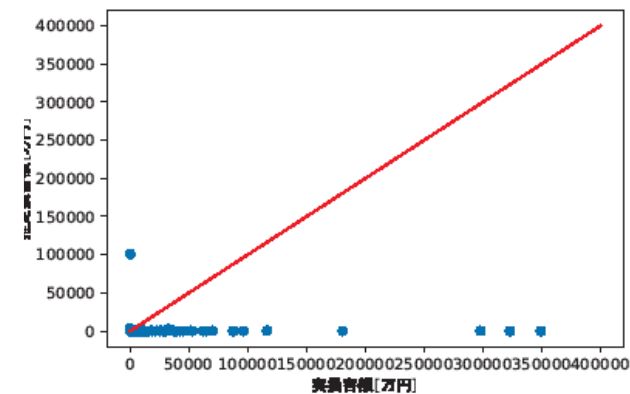
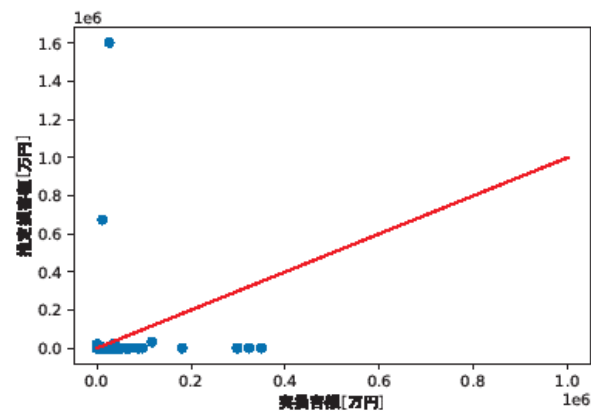
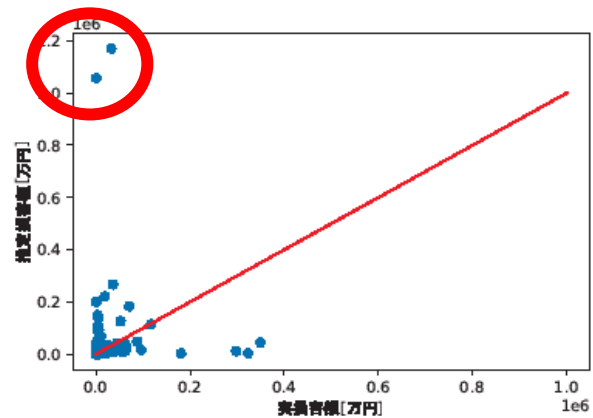
# 各モデルの散布図

山田

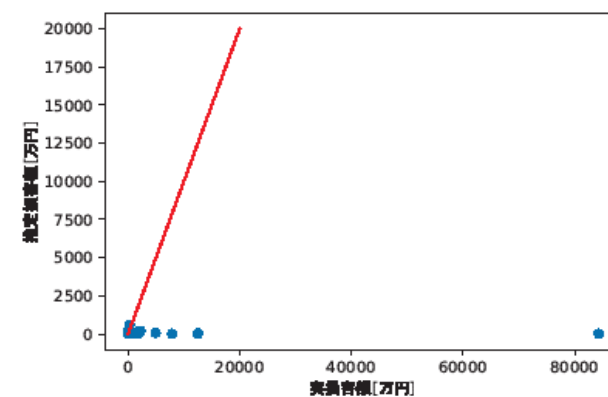
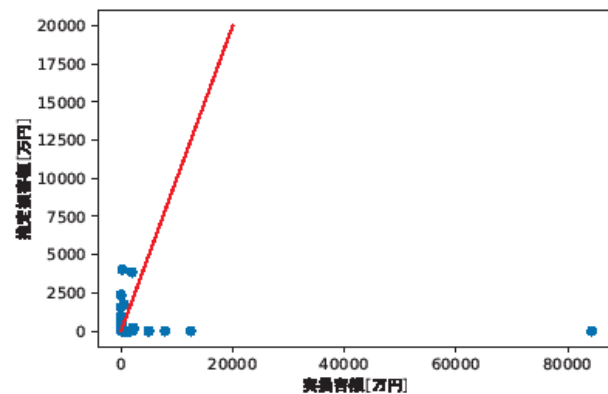
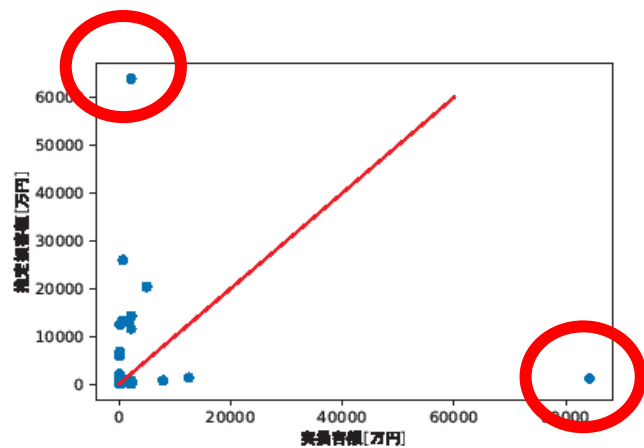
JO

Romanosky

2005-2016



2021



## まとめ

- ウェブサイトを使うことで各モデルを簡単に比較できるようになった

# 課題

- ウェブサイトを使ってもらった感想をもとに入力項目、出力結果ともにわかりやすくできることがある  
(各モデルの説明や決算短信のどこを参照すべきかの例など)