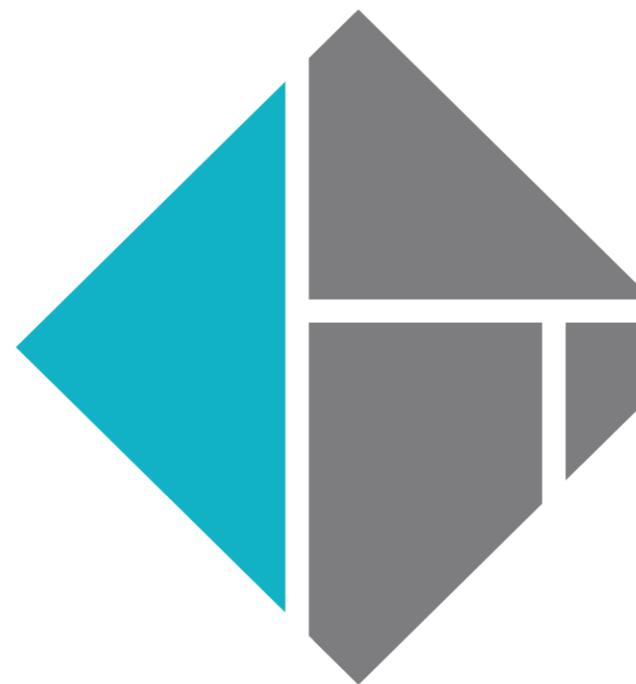


北陸信越工学教育協会シンポジウム
AI技術の事業利用と技術者の育成

2021年2月19日

株式会社インテック
先端技術研究所
青木功介





株式会社インテック
先端技術研究所 所属
シニアスペシャリスト

1992年 富山大学理学部 大学院修了
同年 インテックに入社。
同年 インテック・システム研究所に出向。

現在まで、コンピュータグラフィックス、
画像処理、ヒューマンインタフェースの研究開発に従事。

2010年 色情報を利用した画像処理の研究により、博士(工学)取得。
2015年 アスベスト迅速検査により環境賞(環境大臣賞)を受賞。

社名	株式会社 インテック INTEC Inc.
URL	https://www.intec.co.jp
設立	1964年1月11日
本社	富山・東京
売上高	1,151億98百万円（2019年3月期）
社員数	3,717名（2020年6月1日現在）
事業概要	技術研究、ICTコンサルティング、ソフトウェア開発、システム・インテグレーション、ネットワークサービス、アウトソーシングサービス
本社所在地	本社 富山県富山市牛島新町5-5
	東京本社 東京都江東区新砂1-3-3



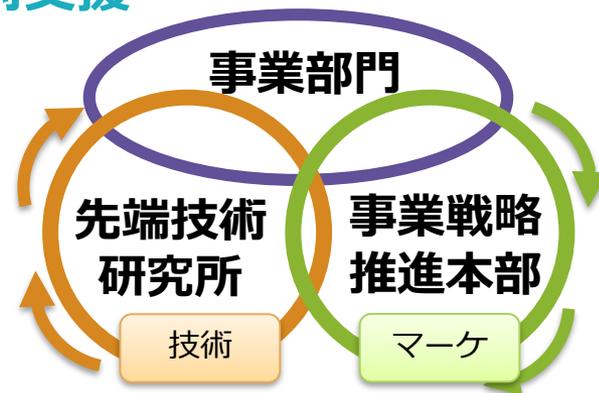
行政、金融、医療、流通、製造業等、様々な企業向けシステムの構築を行っている会社。

■ 研究テーマ

AIの活用
ブロックチェーンの活用
IoTの活用
ビジネスクリエーション

■ 研究所の役割

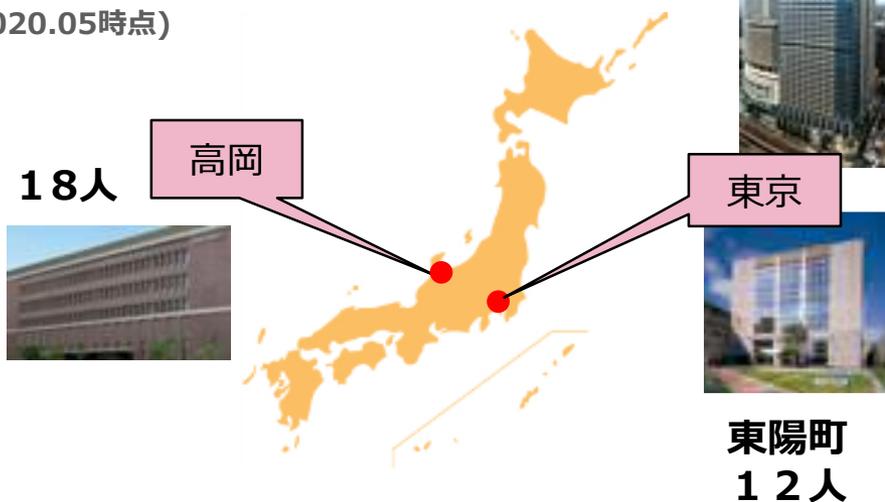
新規事業に結び付く研究開発
各事業部の技術支援



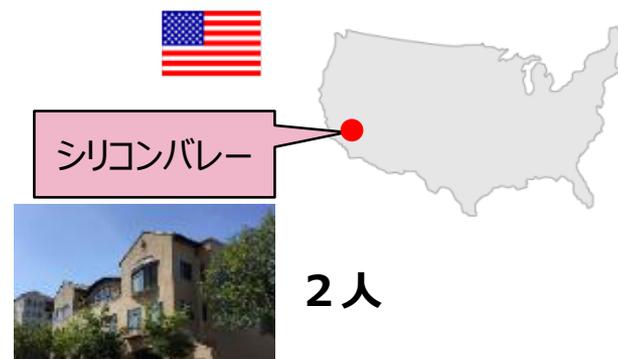
■ 研究拠点

44人

(2020.05時点)



丸の内
12人



1. 画像処理とAI技術
 2. AI技術を用いた事例紹介
(事例) 航空写真からの家屋変化領域検出システム
AI技術事業化の問題点
 3. 社内でのAI技術者の育成
 - (1) 専門技術者の育成
 - (2) シチズン・データ・サイエンティストの育成
- ・ まとめ

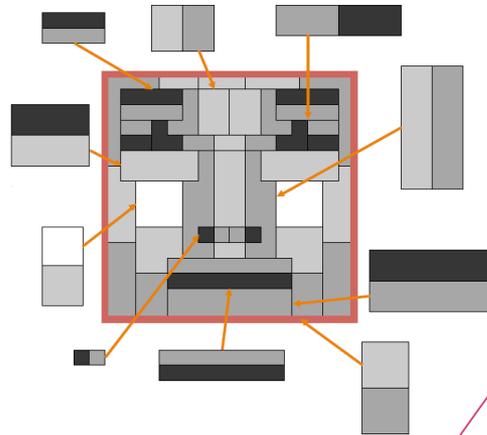
1. 画像処理とAI技術



既存の画像認識

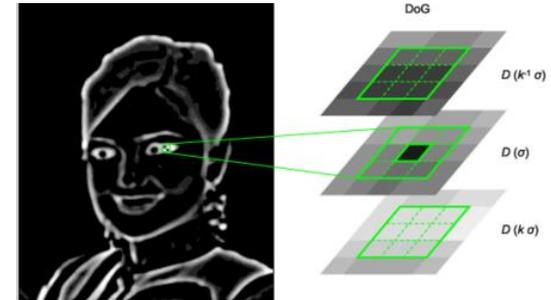
Haar-like

引用 : https://www.slideshare.net/MPRG_Chubu_University/ss-32258845

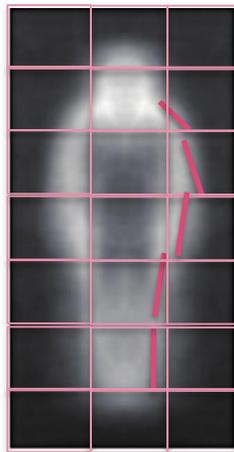


2領域間の明度差を
特徴量とする。

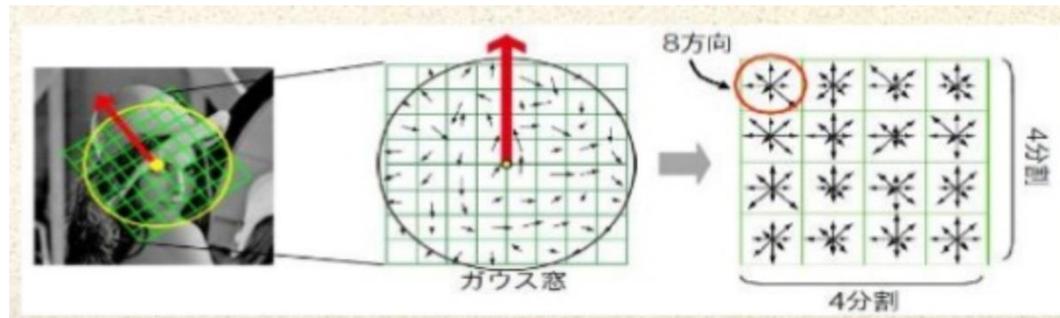
SIFT



HOG



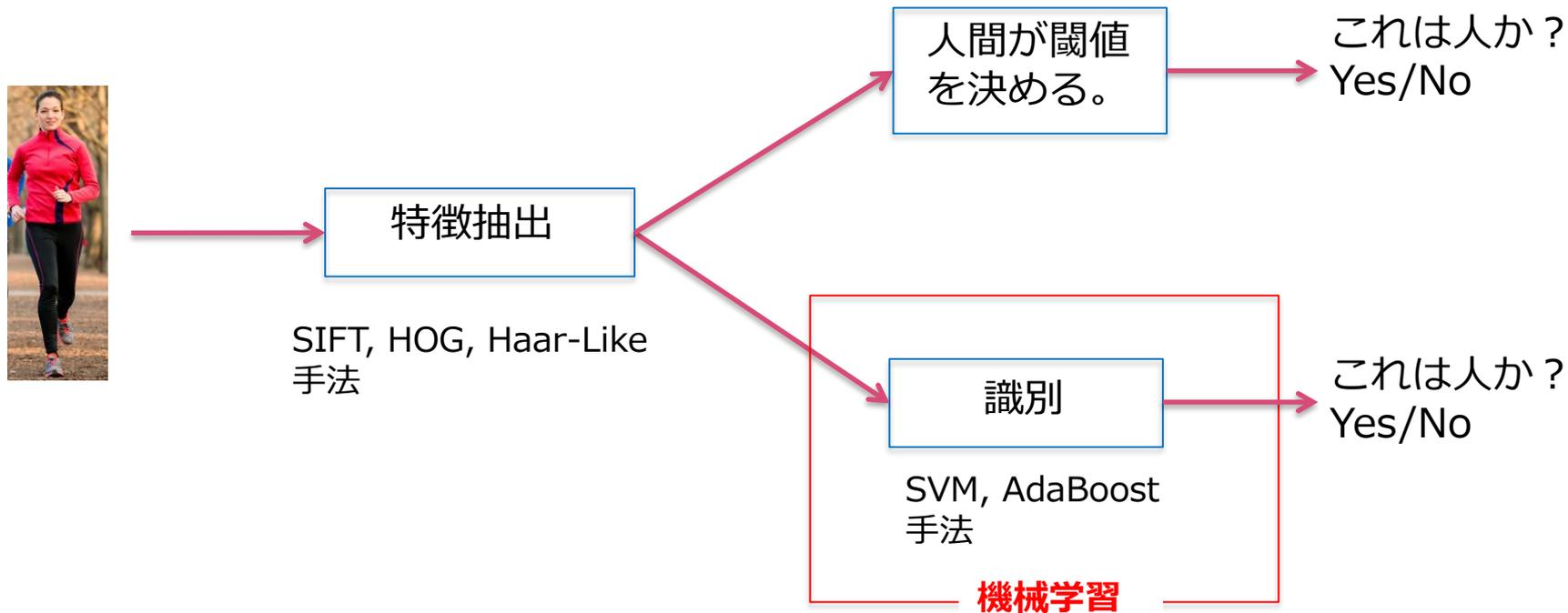
エッジ部分の
明度変化量を
特徴量とする



引用 :
https://pythonhosted.org/sift_pyocl/sift.html
<http://www.slideshare.net/lawmn/siftsurf>

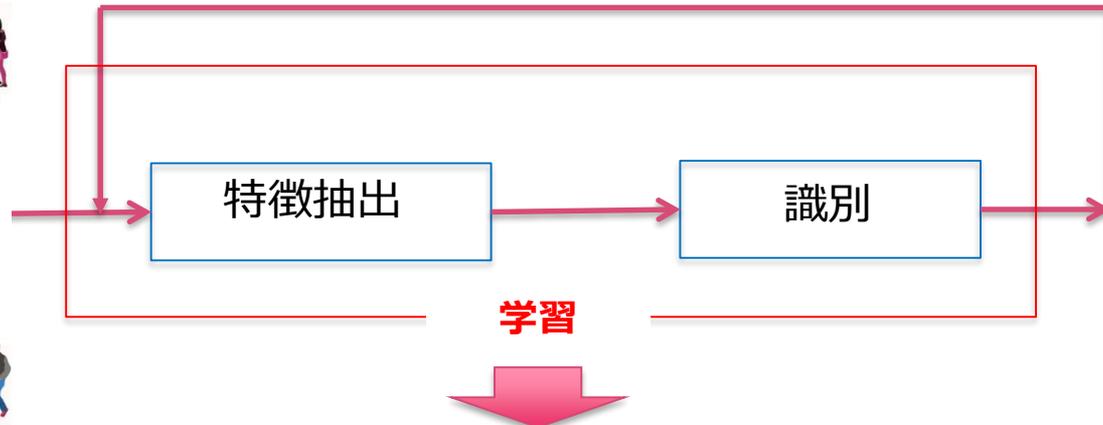
引用 : https://www.slideshare.net/MPRG_Chubu_University/ss-32258845

既存の画像認識

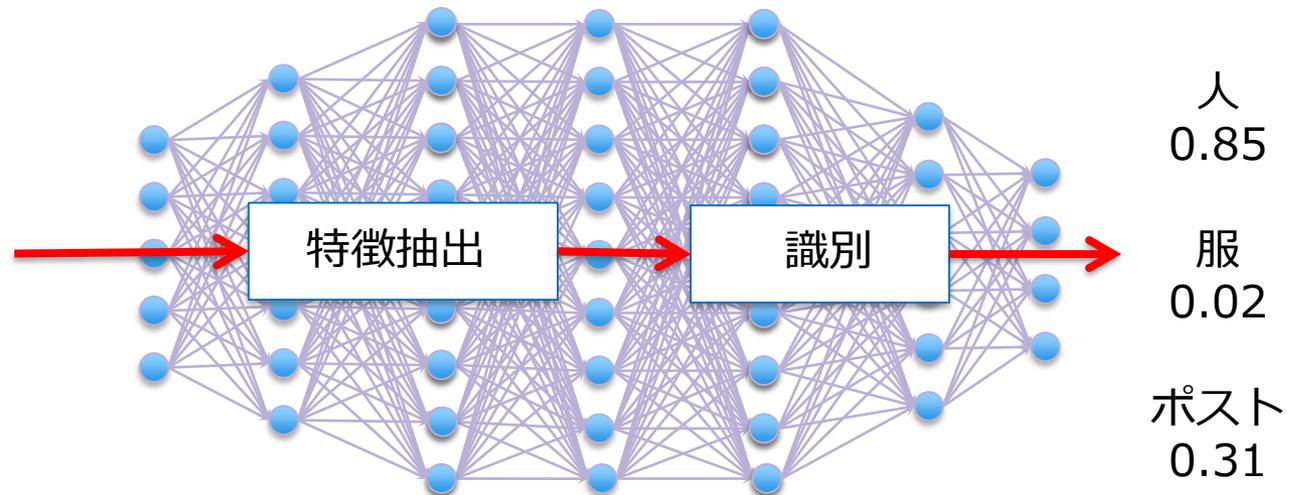


Deep Learningによる画像認識

大量の人の画像

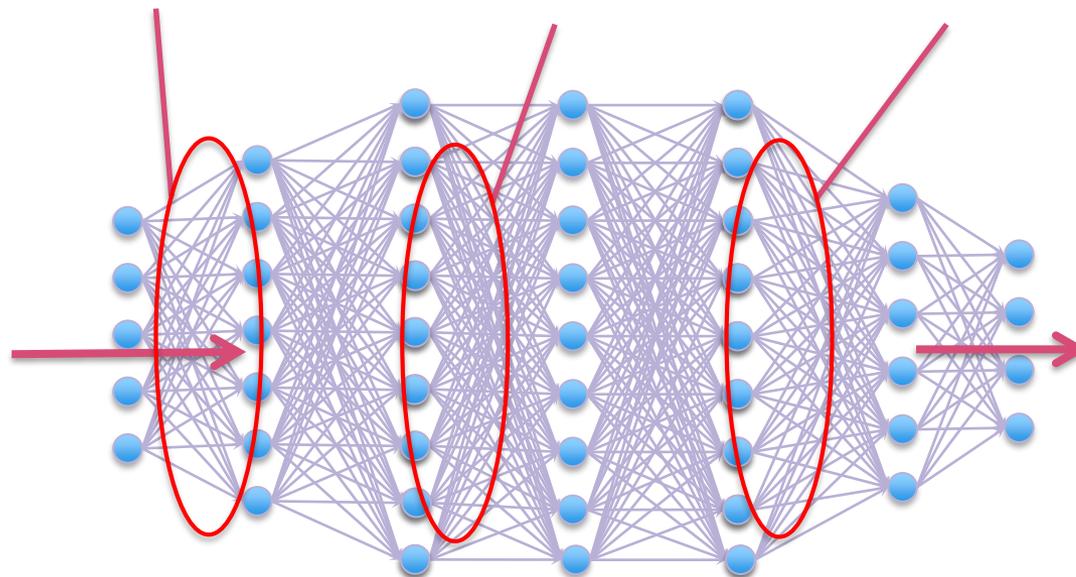


Deep Learning Network



写真、イラスト：Corporate+

Deep Learningによる画像認識



人
0.85
服
0.02
ポスト
0.31

人物写真：Corporate+

画像分類コンテストの結果



写真： pixabayより

ILSVRC(ImageNet Large Scale Visual Recognition Challenge)

既存の画像処理は

人が特徴や識別値を見つけてプログラミング



AIを使った画像の認識は

大量の画像をコンピュータが学習

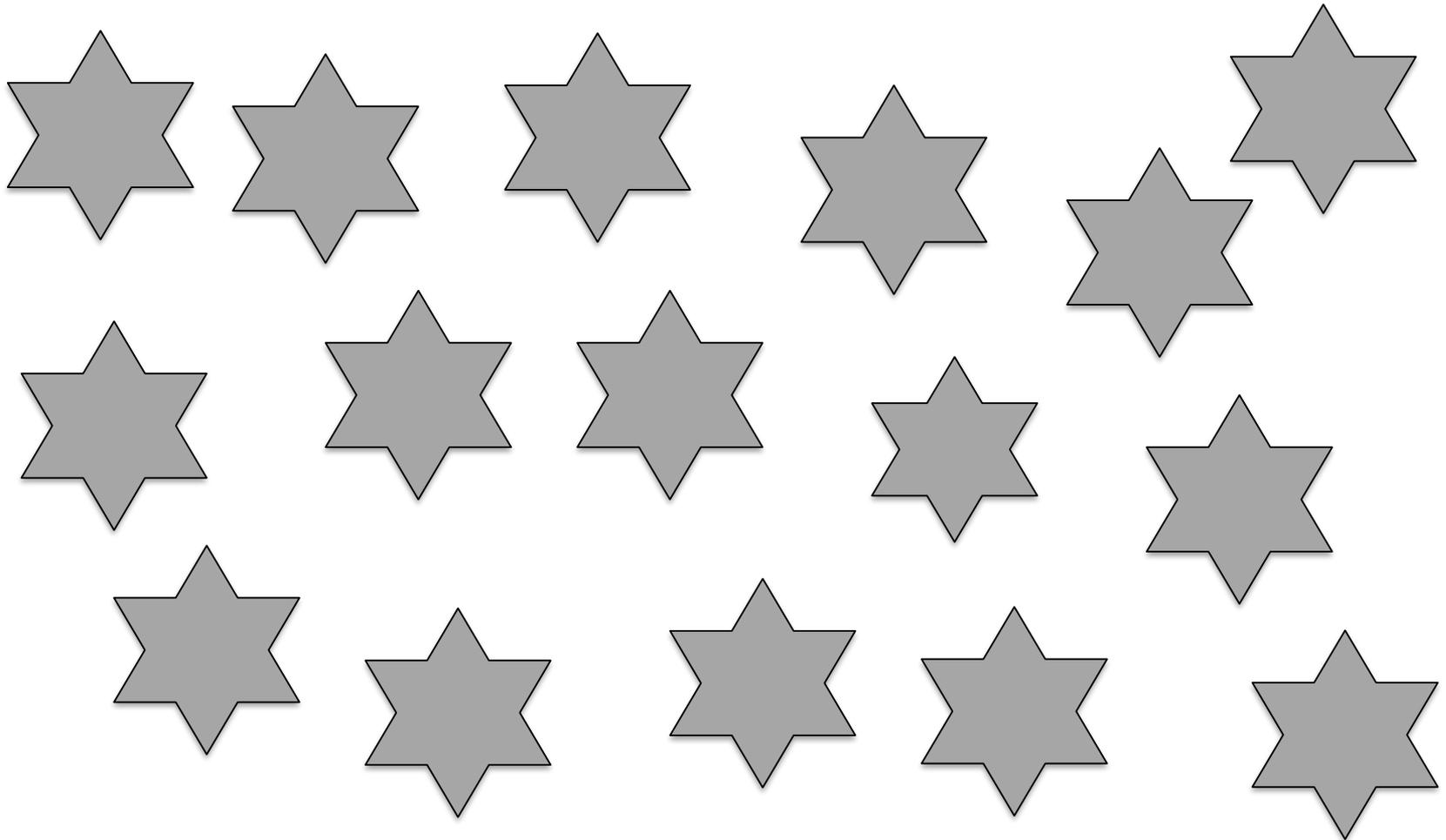
→ より複雑な特徴にも対応

AIが得意 Vs. 既存の画像処理は苦手



人は何をもって鳥と判断しているのか？

AIが苦手 Vs. 既存画像処理は得意



明確にサイズが分かっている、違いを見つける場合。

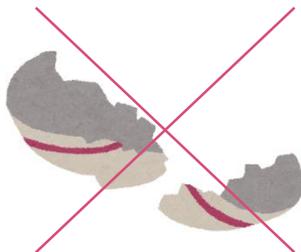
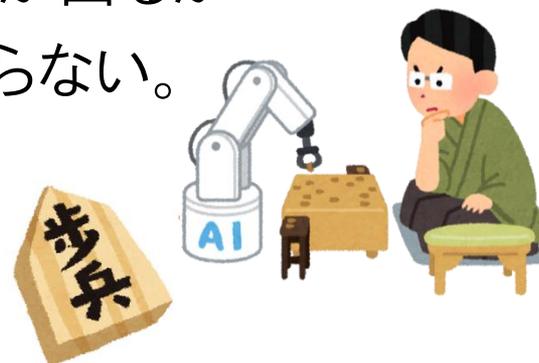
現在のAIの問題点

☆ 明示的な条件の場合、既存の画像処理が有利。

☆ 特定の条件下では、人を超える認識率が出るが
ただしこの条件の外では使いものにならない。

☆ 教え込むのが大変。
大量の教師データが必要。
予想外の学習をしてしまうことがある。

☆ 教え込んだ以外のことはできない。



2. AI技術を用いた事例紹介

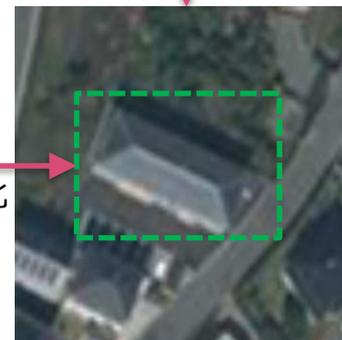
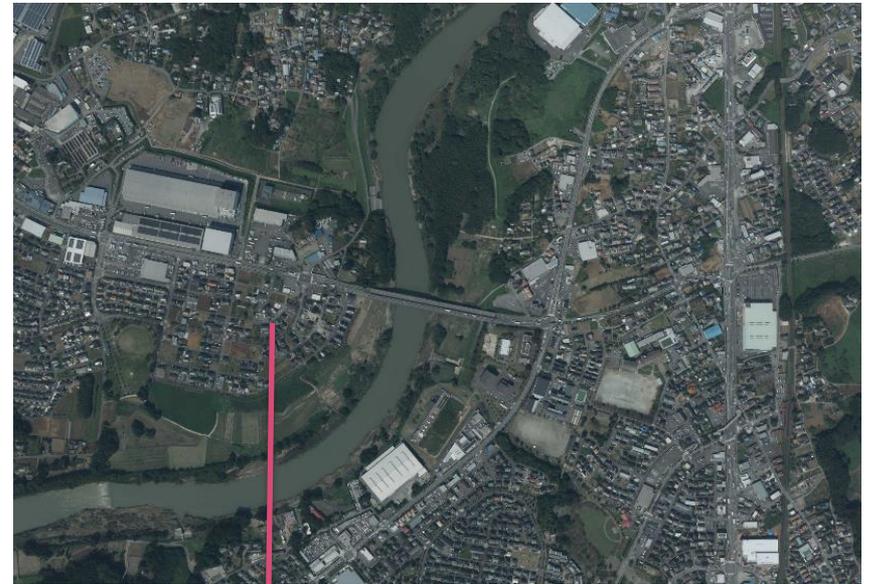


航空写真からの家屋変化領域の検出

旧航空写真



新航空写真



家屋に
変化
があった
領域

固定資産税

市町村が固定資産の所有者に対して課税する税金

国土交通省国土地理院

平成30年2月

「固定資産税調査用空中写真撮影の実態に関する調査業務」報告書より

回答のあった市町村の約73%にあたる1,062市町村では、固定資産の現況調査のための空中写真撮影を実施したことがあると述べている。

①一般的な現場調査方法

- ・現場へ行く
- ・調査員の目で確認

②先進的な自治体の調査方法

- ・異なる年数の航空写真を使用
- ・基本的にヒトの手作業で確認

デメリット：

- ・大量な時間と費用が必要
- ・ヒューマンエラーが発生する可能性がある

第一時期（旧）航空写真



第二時期（新）航空写真



家屋の変化領域の検出を自動化することが望まれている

市町村で数十～数百枚





方針

処理
流れ



(1) 射影変換

航空写真間のズレを補正するために位置変換を行う。
全箇所ですべて画素以内の一致が必要。

(2) 差分抽出

2枚画像の差分を取る。
影や植物の影響を少なくするため、色の差分を強調して取得し、差が大きい箇所のみを候補領域として抽出する。

(3) 変化領域の認識

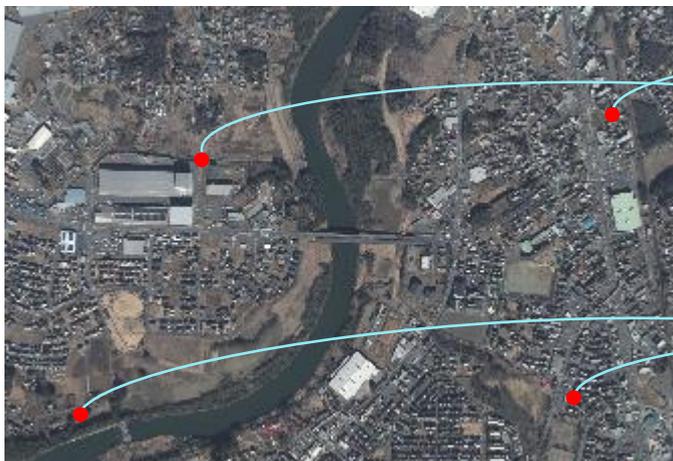
(2) で抽出した候補領域が、家屋かどうかを
ディープラーニングにより判断する。
これにより対象外の領域を省く。

(4) 変化領域の確認

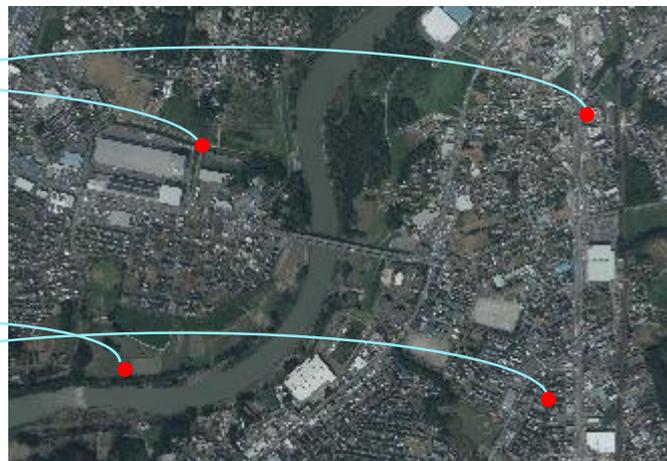
抽出した候補領域を人が見て最終判断する。

4つの対応点を、画像処理による特徴点抽出方法により取得する。

旧航空写真



新航空写真



左上、左下、右上、右下の4象限から対応点を取得する。
この4点が一致するような変換行列を求める。



全域で3画素以内のズレに収まった。

2020年7月特許出願



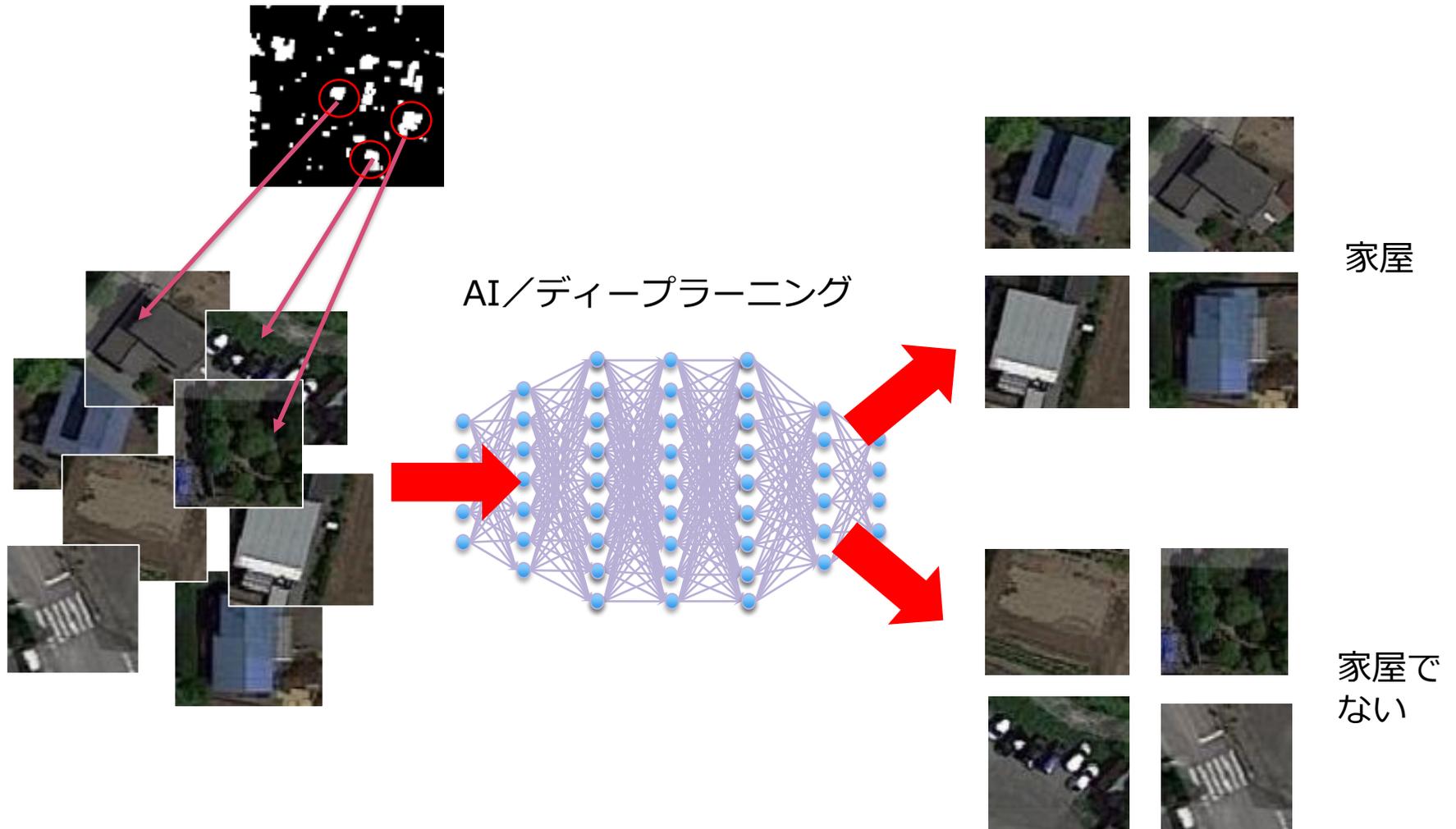


差分領域抽出

RGBではなく、L*a*b*色空間で差分を取る。



差分領域の家屋判定



家屋変化の判定方法

	旧 航空写真	新 航空写真	判定
1	○	○	改築
2	×	○	新築
3	○	×	取り壊し
4	×	×	対象外

ディープラーニングで、
○:家屋と判断
×:家屋でないと判断

例： 1.改築



2.新築



3.取り壊し



①改築

旧



新



②新築

旧



新



③取り壊し

旧



新



処理
流れ

(1) 射影変換

航空写真間のズレを補正するために位置変換を行う。
全箇所ですべて画素以内の一致が必要。

(2) 差分抽出

2枚画像の差分を取る。
影や植物の影響を少なくするため、色の差分を強調して取得し、差が大きい箇所のみを候補領域として抽出する。

(3) 変化領域の認識

(2) で抽出した候補領域が、家屋かどうかを
ディープラーニングにより判断する。
これにより対象外の領域を省く。

(4) 変化領域の確認

抽出した候補領域を人が見て最終判断する。

実際にA市の航空写真を使い評価を行った。

再現率 大→取りこぼしが無い。
適合率 大→過剰検出が無い。

開発手法

家屋の変化領域の検出精度・速度

(1) 再現率 : 94.4%
(検出数67 / 正解数≒71)

(2) 検出速度 : 3分/画像

(3) 適合率 : 23.6%

人による実験

家屋の変化領域の検出精度・速度

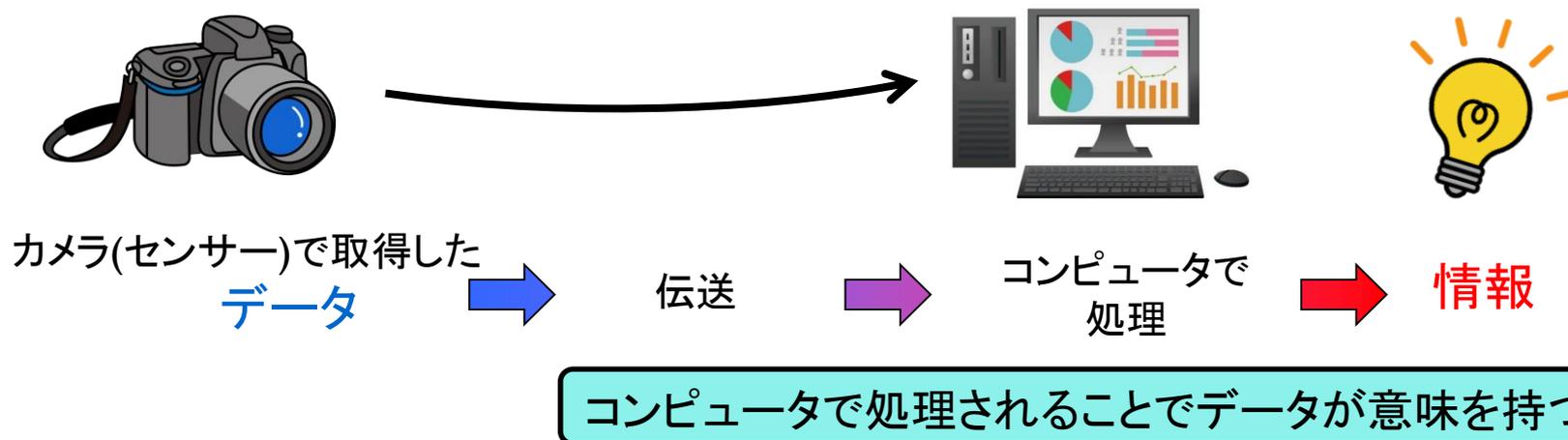
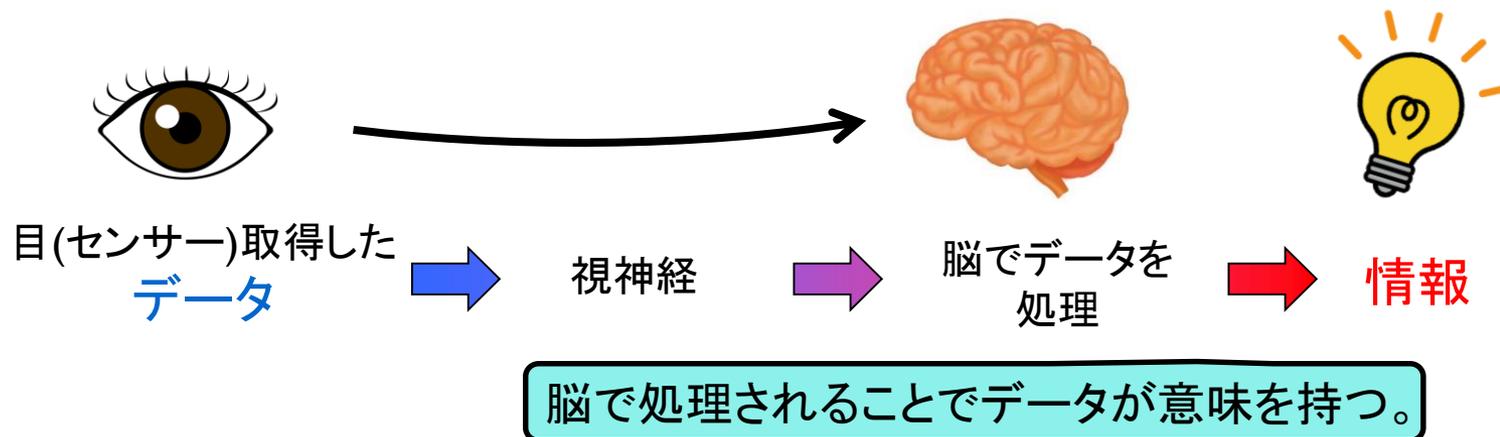
(1) 再現率 : 66.2%
(検出数47 / 正解数≒71)

(2) 検出速度 : 45分/画像

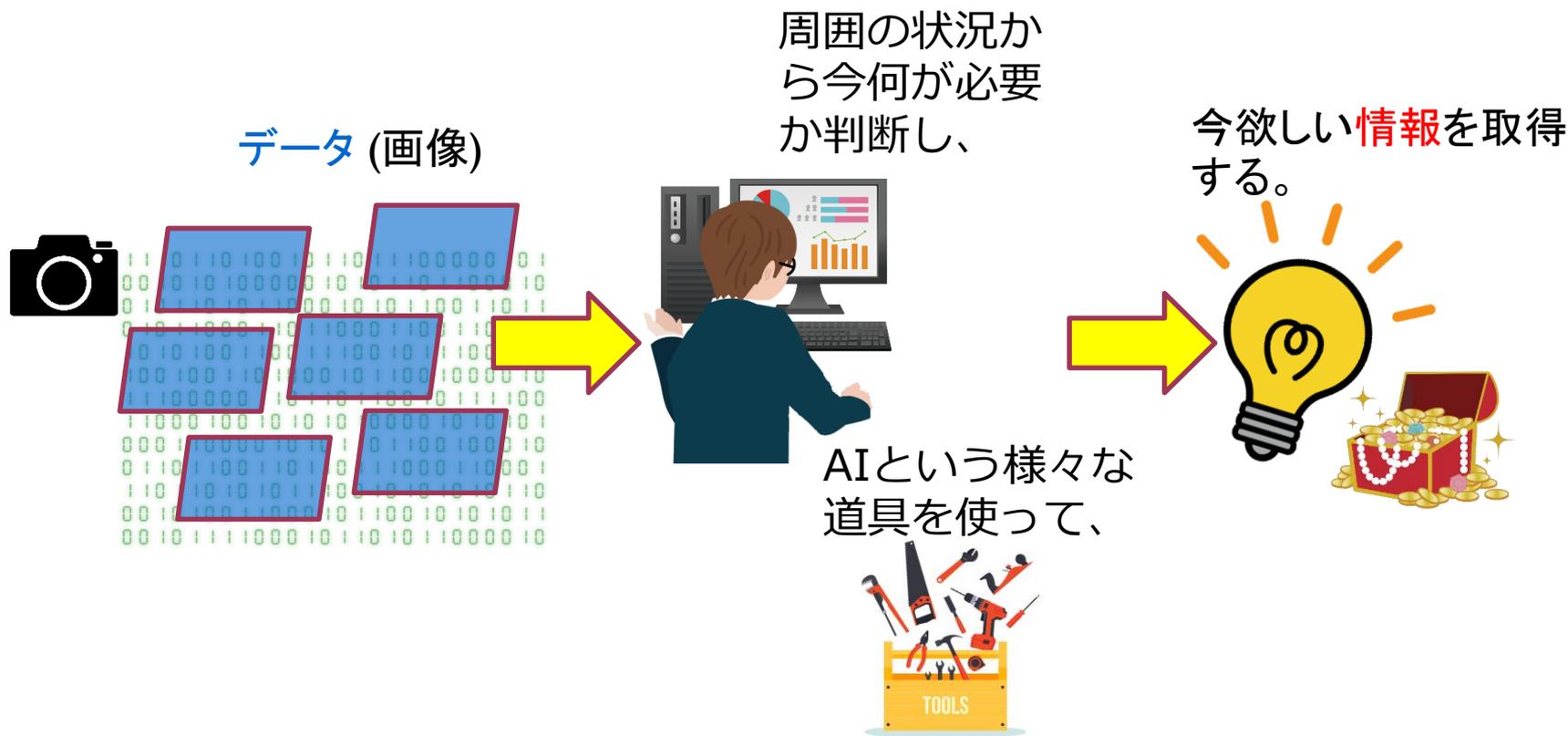
画像処理技術とディープラーニング技術の組み合わせによって、人間の手作業と比べても、高精度・高速な家屋変化領域の検出が実現できた。

3. 社内でのAI技術者の育成





データから情報へ

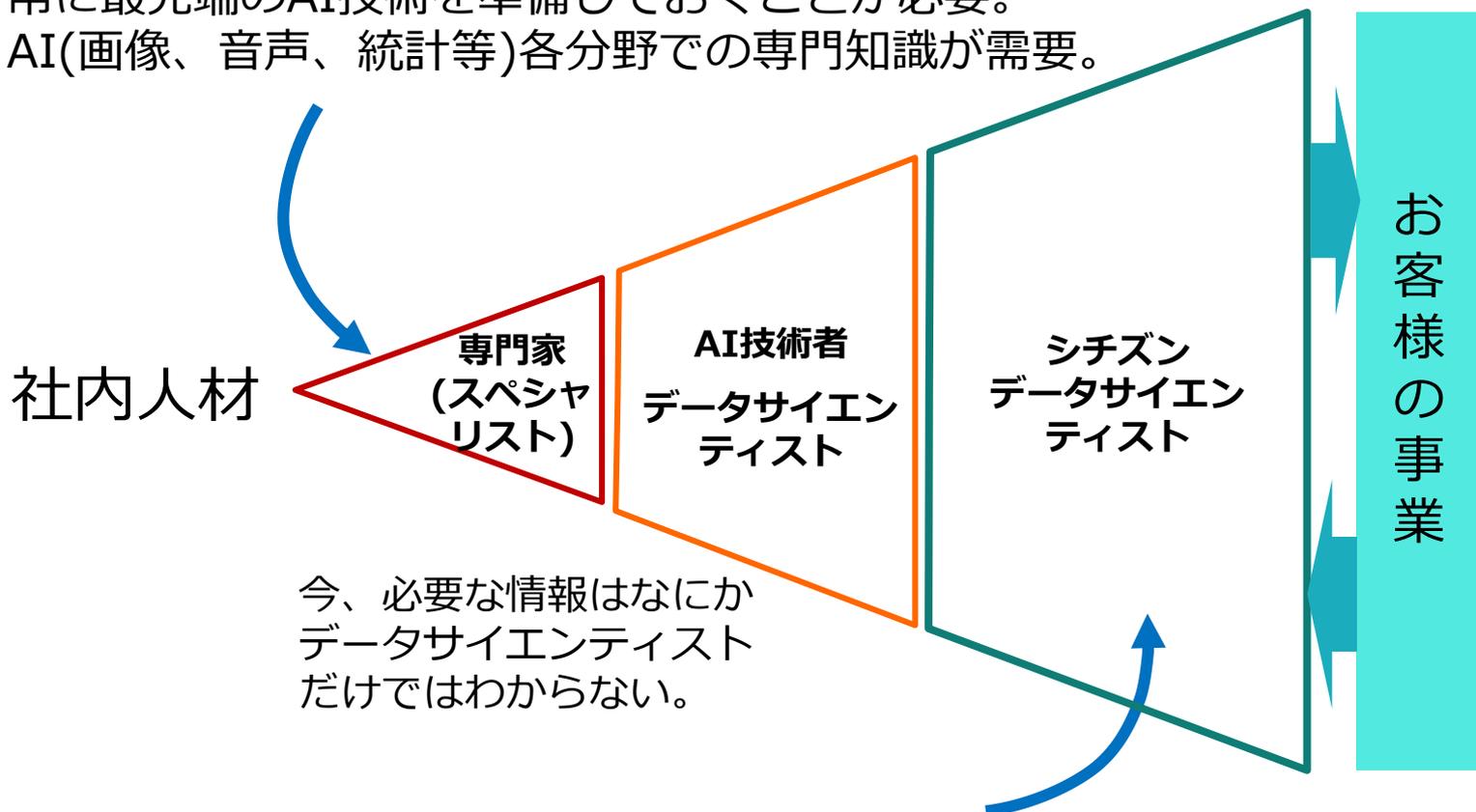


膨大なデータから、必要な情報のみを取り出す。

それを行うのが脳でありデータサイエンティストである。

AI技術者の育成

- 常に最先端のAI技術を準備しておくことが必要。
- AI(画像、音声、統計等)各分野での専門知識が需要。

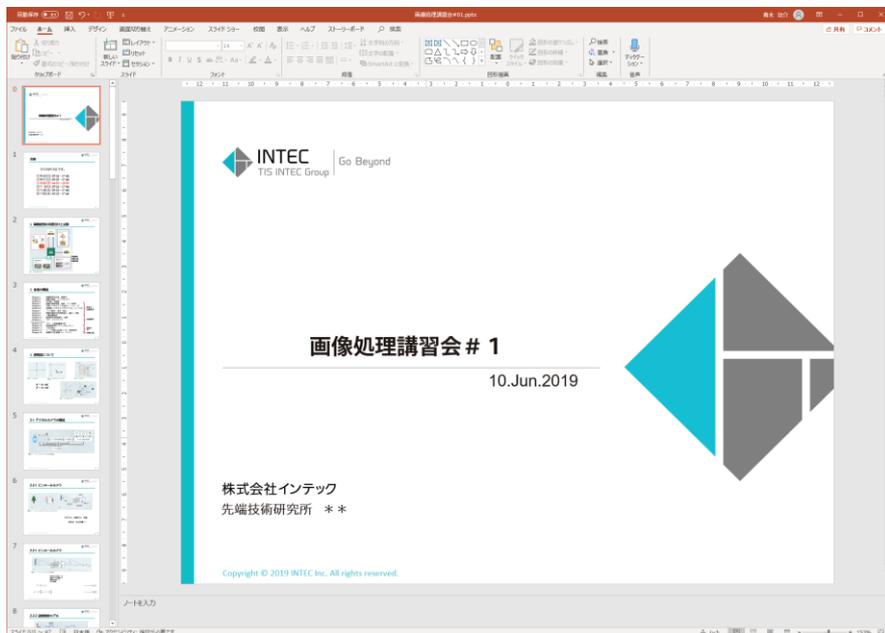


- お客様の事業に近く、お客様が直面しているビジネス課題に関する深い専門知識(ドメイン知識)をもつ人材。
- データサイエンススキルを育成する。

(1) 専門技術者の育成

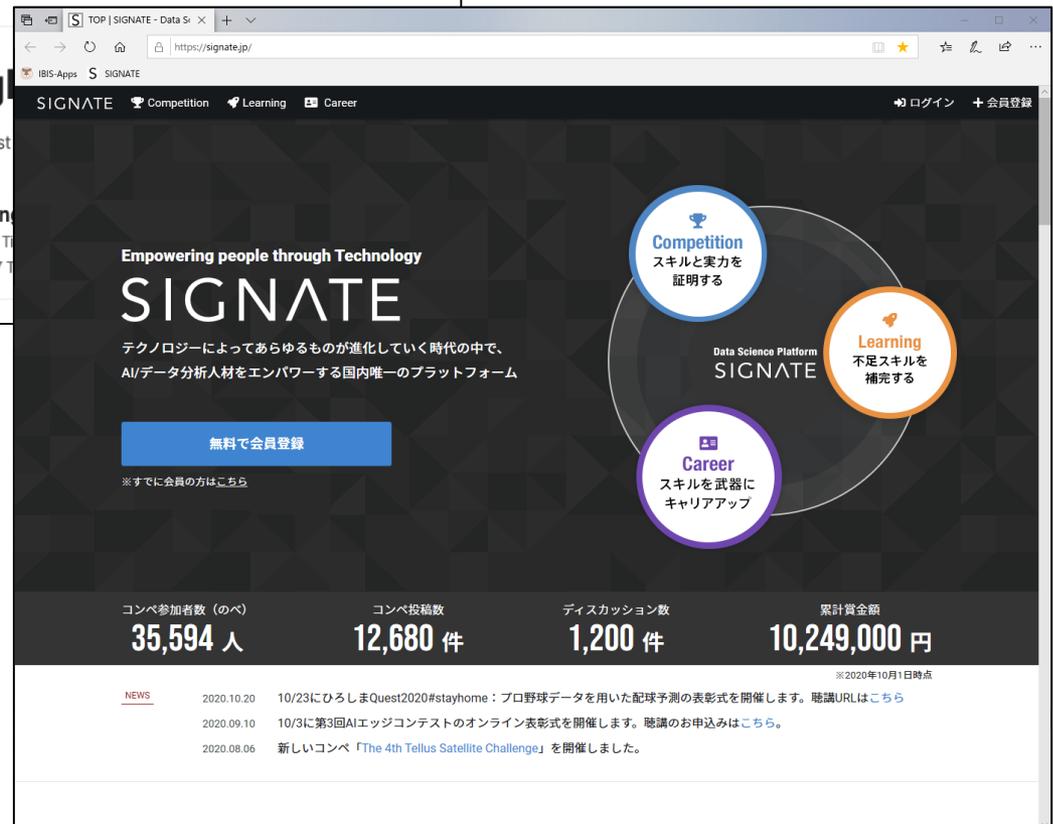
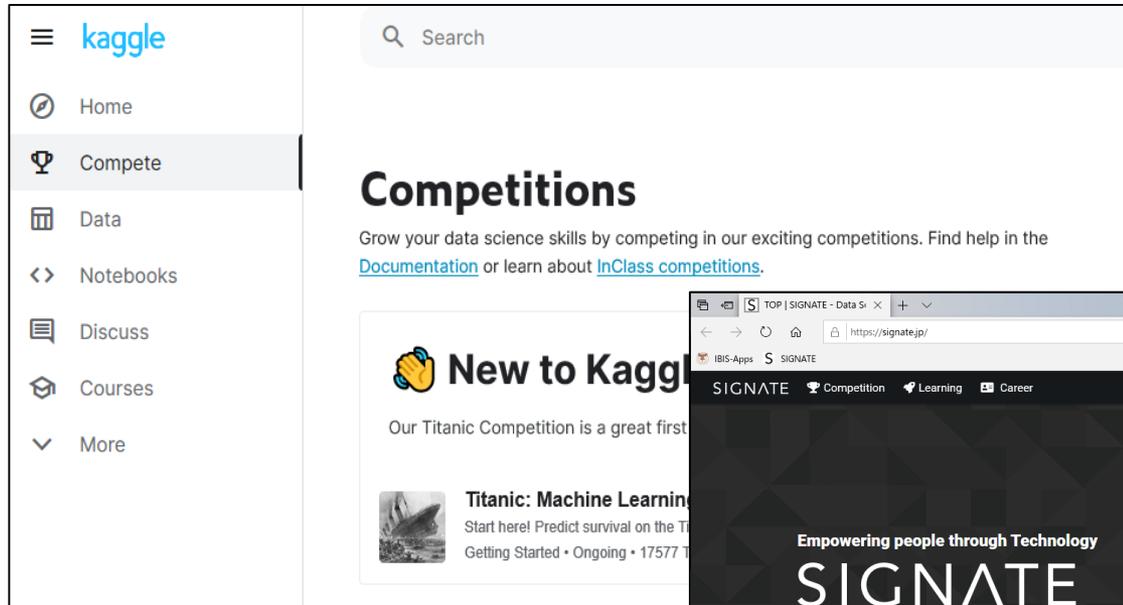
- AI技術者を、より専門的な技術者にするために。
 - 社内勉強会の開催
 - 社外コンテストへの参加





講習会形式
→ 講師が1つの分野を極める

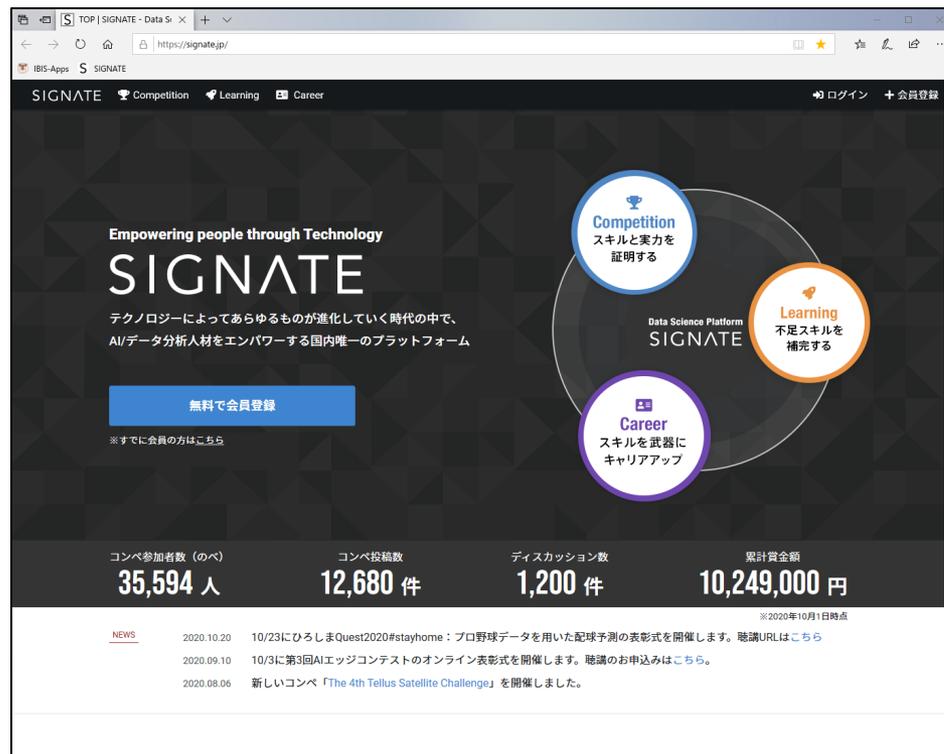
コンテストへの参加 (専門家を育てる)



<https://www.kaggle.com/>
<https://signate.jp/>

コンテストへの参加の意義

- 自分の力を他者と比較できる。
- 実際のデータを使った処理を体験できる。
- 他者の方法から勉強ができる。
- 高成績を出して技術をアピール／公開できる。
- 挑戦する意欲。



<https://signate.jp/>

コンテストへの参加

【SIGNATE コンペティション】

The 4th Tellus Satellite Challenge : 海岸線の抽出 に挑戦

主催：経済産業省

期間：2020/8/6～11/6

参加：803人

提供データ：

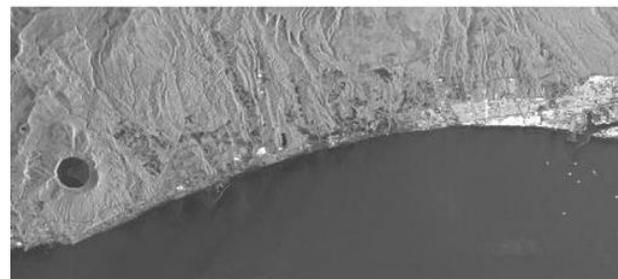
SAR撮影画像(16bit TIFF)

- ・学習用25枚
- ・テスト用30枚
- ・学習用アノテーション座標データ

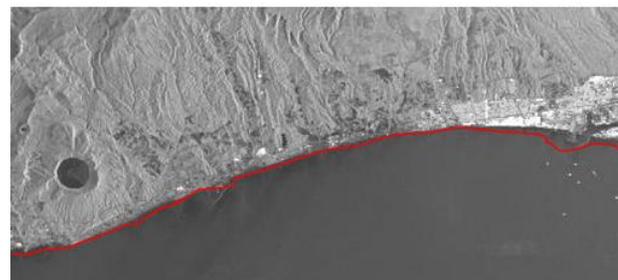
課題：

テスト用30枚についての海岸線を見つける

SAR image

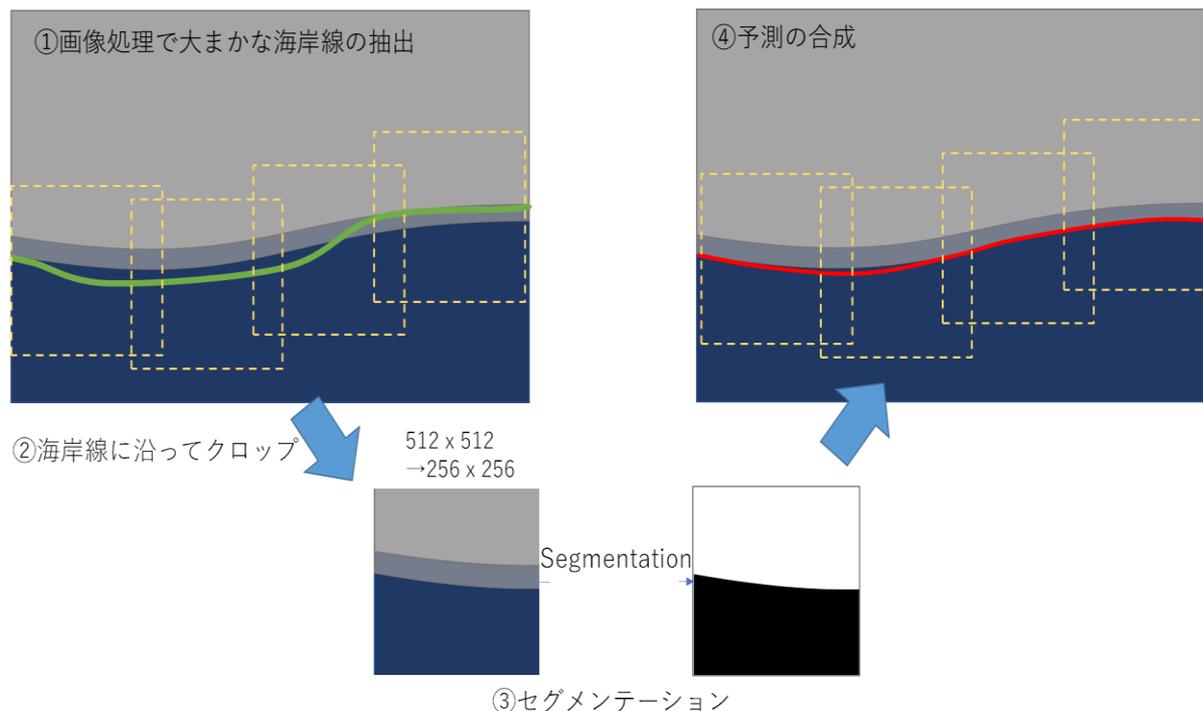


Extract the coastline



出典：<https://signate.jp/competitions/284>

- 16bit画像からのコントラスト変換
- 画像処理によるおおよその領域抽出
- データの増強
- ディープラーニングを使ったセグメンテーション
- . . .



コンテストへの参加

3位入賞



参加 803人中

11月4日時点の暫定順位

順位	チーム名/ユーザー名	暫定評価	最終評価	投稿件数	投稿日時
1	IRAFM-AI	12,788,480	11,184,785	91	2020-10-09 05:40:07
2	Yurara	14,291,184	11,235,125	115	2020-11-06 18:19:10
3	mygle	14,280,712	11,766,797	44	2020-11-06 16:55:11
4	BloodAxe	14,818,092	11,797,105	72	2020-11-06 23:54:13
5	DTSN	15,388,209	12,501,395	122	2020-11-06 16:47:11
6	sanpujany17	15,048,627	12,668,042	56	2020-11-02 17:50:14
7	i-aki.y	18,639,032	12,908,266	12	2020-11-05 21:15:12
8	ishiyuhi	15,898,952			
9	denden_lv	16,167,376			
10	deeshi	16,517,938			
11	nakano	16,915,235			
12	ur	17,102,836			
13	yasuA	17,649,756			
14	aknk	18,241,989			
15	ichiyoda	18,353,167			
16	GreenDoor	18,930,229			
17	KotaNakashima	19,663,258			

順位	チーム名/ユーザー名	暫定評価	最終評価	投稿件数	投稿日時
1	BloodAxe	14,902,369	11,184,785	91	2020-10-09 05:40:07
2	IRAFM-AI	12,788,480	11,235,125	115	2020-11-06 18:19:10
3	aknk	16,294,483	11,766,797	44	2020-11-06 16:55:11
4	mygle	14,048,938	11,797,105	72	2020-11-06 23:54:13
5	DTSN	15,268,951	12,501,395	122	2020-11-06 16:47:11
6	sanpujany17	15,048,627	12,668,042	56	2020-11-02 17:50:14
7	i-aki.y	18,639,032	12,908,266	12	2020-11-05 21:15:12

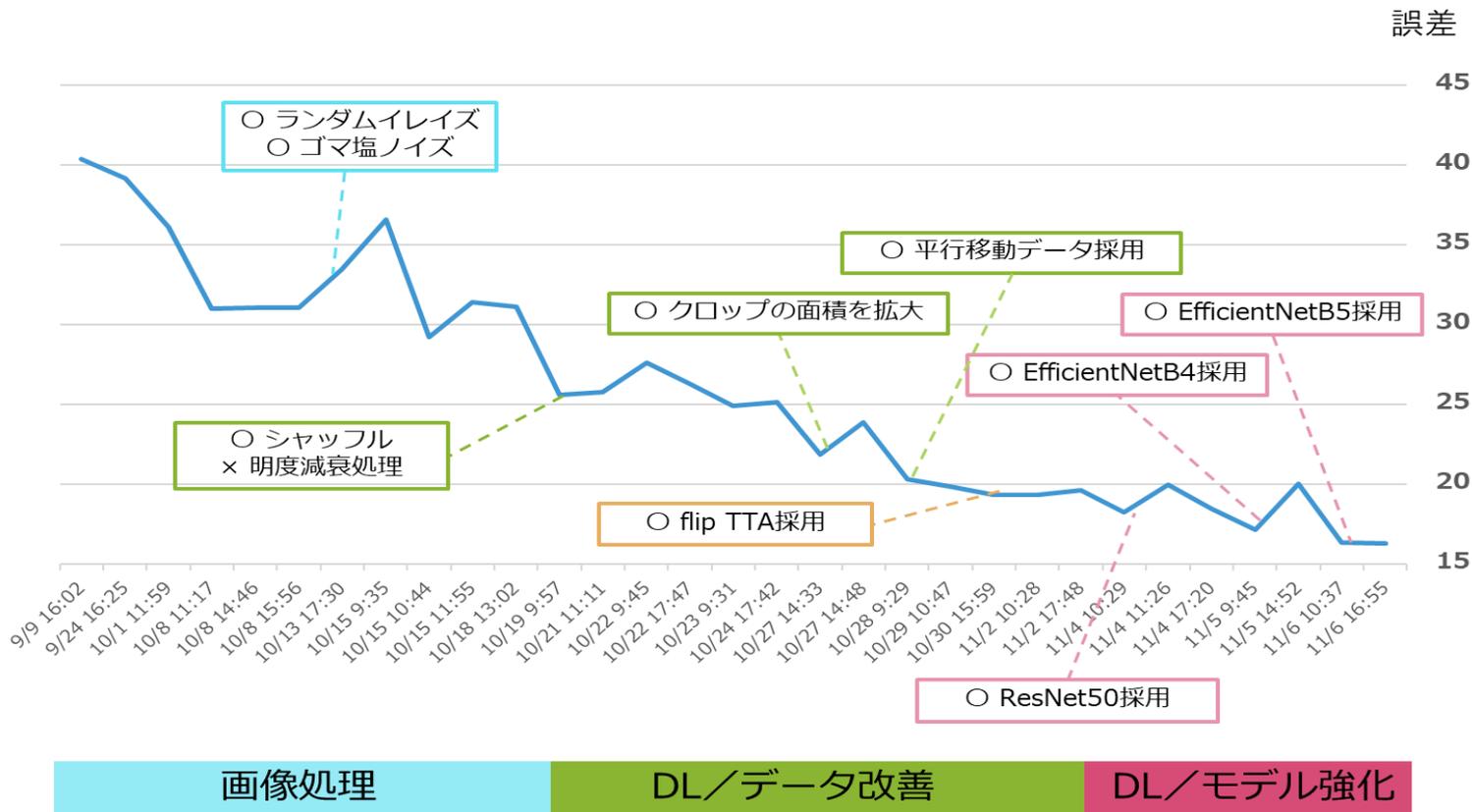
11月7日の最終順位

- ・メンバ全員がとても勉強になった
- ・楽しく参加できた
- ・賞金をいただけた

モチベーションアップ

<https://signate.jp/>

コンテストへの参加／評価の推移



The 4th Tellus Satellite Challenge
実施！～入賞者たちの手法を解説～

<https://sorabatake.jp/18087/>

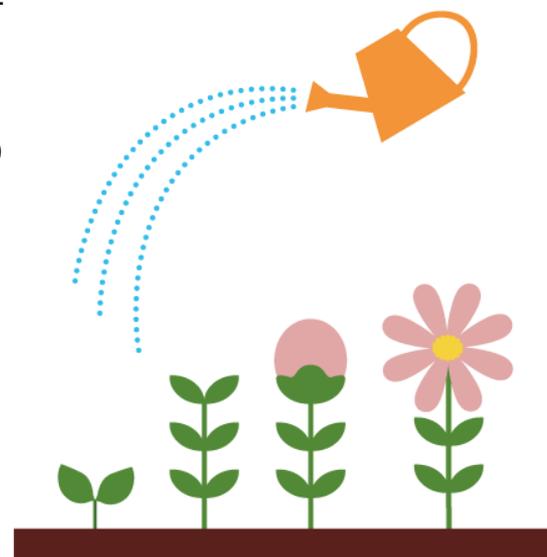
(2) シチズン・データ・サイエンティストの育成

- ・ 実事業の中でデータを活用できる技術者を育成する。

- ワークショップ開催

シチズン・データ・サイエンティストの定義：

機械学習などのデータサイエンスを使って
業績の改善に役立つ知見の獲得や予測モデルの
構築を行おうとするビジネスパーソン



イラスト：<https://www.ac-illustr.com>

機械学習ワークショップ

データ活用を進める上で、データ分析人材を増強することが必須。
社内でのデータ分析人材を育成するため、機械学習ワークショップを実施している。

データサイエンティスト
↓
シチズンデータサイエンティスト

機械学習ワークショップコンテンツ

ハンズオン - データの傾向把握 -

「ターゲット項目に対する各項目の傾向把握」

- 解約の有無に寄与している項目の把握
 - ✓ 日中の利用料金(Day_Charge)と解約の関係は？

⇒ 日中の利用料金における解約有無の100%積み上げ棒グラフ

日中の利用料金が多いと解約の傾向が高い

[演習②]
・ 夕方の利用料金(Eve_Charge)と解約の関係は？

2週間

↓
半年

とめ -

A社の顧客セグメンテーション

客のセグメンテーションを行い、各グループの特徴を把握しました。
「A」が「解約率が高い」ことを分析しました。

顧客の属性/特徴による分類

- ビデオMailの利用が多いグループ
- 解約率が高い
- 通話料金が低いグループ
- 夜間利用が多いグループ

DataSet DSS ハンズオン操作手順書

4. 「score_customers_prepared」画面において、各フィールドは初期値のまま、画面右下の「RUN」ボタンを押下します。

画面下部に「Job succeeded」のメッセージが出力されたら、正常に処理が完了です。

5. 画面上部のアイコンにマウスカーソルを合わせ、表示メニューから「Flow」を選択します。下記のフローが表示されたら、「customers」データセットを分類する「手順は完了」です。

6. 上記フローの右側に位置する「customers_segment」データセットをダブルクリックします。

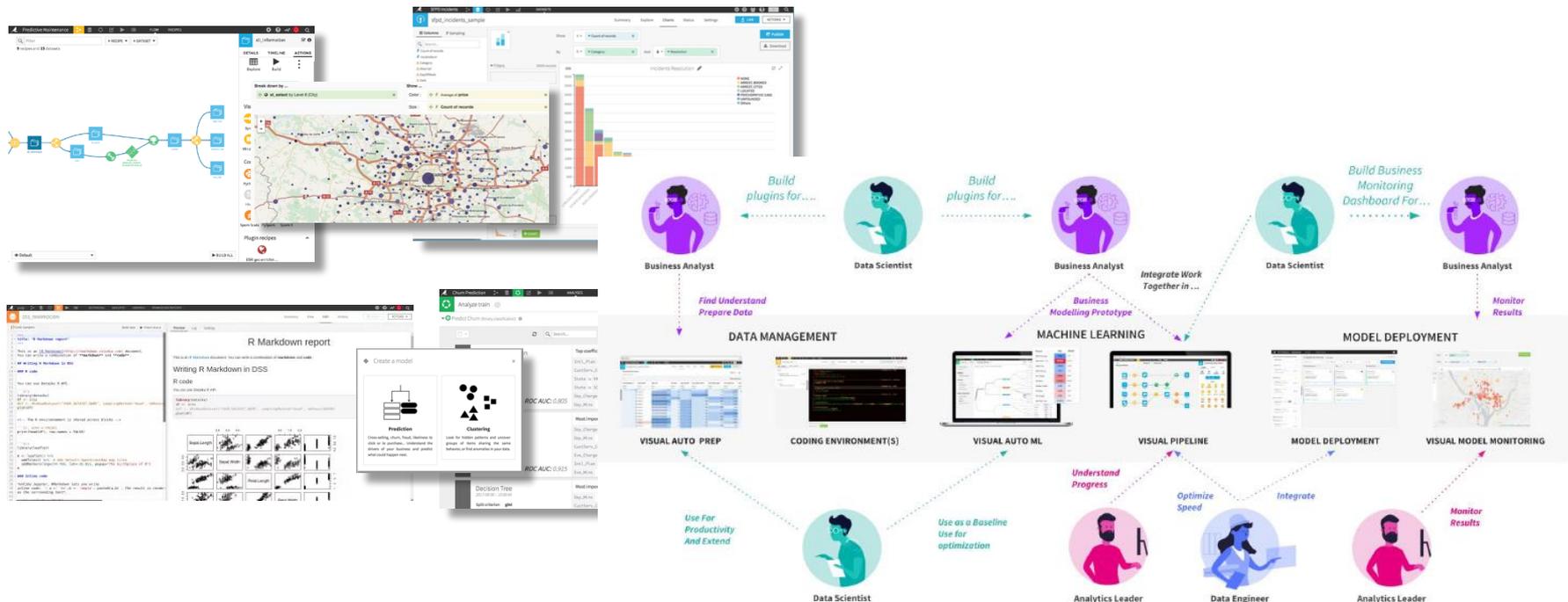
7. 「customers_segment」のデータセット概要 画面において、[cluster_label]カラム(最後のカラム)が追加されており、[cluster_label]カラムにモデル実行結果が表示されていることを確認します。

Copyright © 2019 INTEC Inc. All rights reserved.

機械学習ワークショップ

ハンズオン環境として「Dataiku Data Science Studio」を利用。

- ✓ Dataiku : データ分析が学びやすい環境
 - ✓ <https://www.dataiku.com/>
 - ✓ GUIで分析できる
 - ✓ 同一画面で分析して、エキスパートの知見を共有



各事業間／シチズンデータサイエンティスト同士での情報共有

様々な事業部の担当者が集まって、進捗を確認。

- 私の部では××ステージまで進んだ。
- ○○のところは、このデータを使用した。
- △△の部分は、金融の方ではどうしましたか。
- お客様からも、このような問い合わせがあった。

モチベーションアップ




- ◎ お客様も含めた勉強会やPoCに発展。
- ◎ Qiita等を使い社外への情報発信。



<https://qiita.com/organizations/intec>

まとめ

- ・ 弊社研究所における、AI技術を用いた事例を紹介しました。

(事例) 家屋変化領域検出システム

システムを常に改善できることが必須であり、AIをできるだけシンプルに使うことが重要。

- ・ 社内におけるAI技術者の育成について紹介しました。

ご清聴ありがとうございました

