

応用基礎レベル教材について

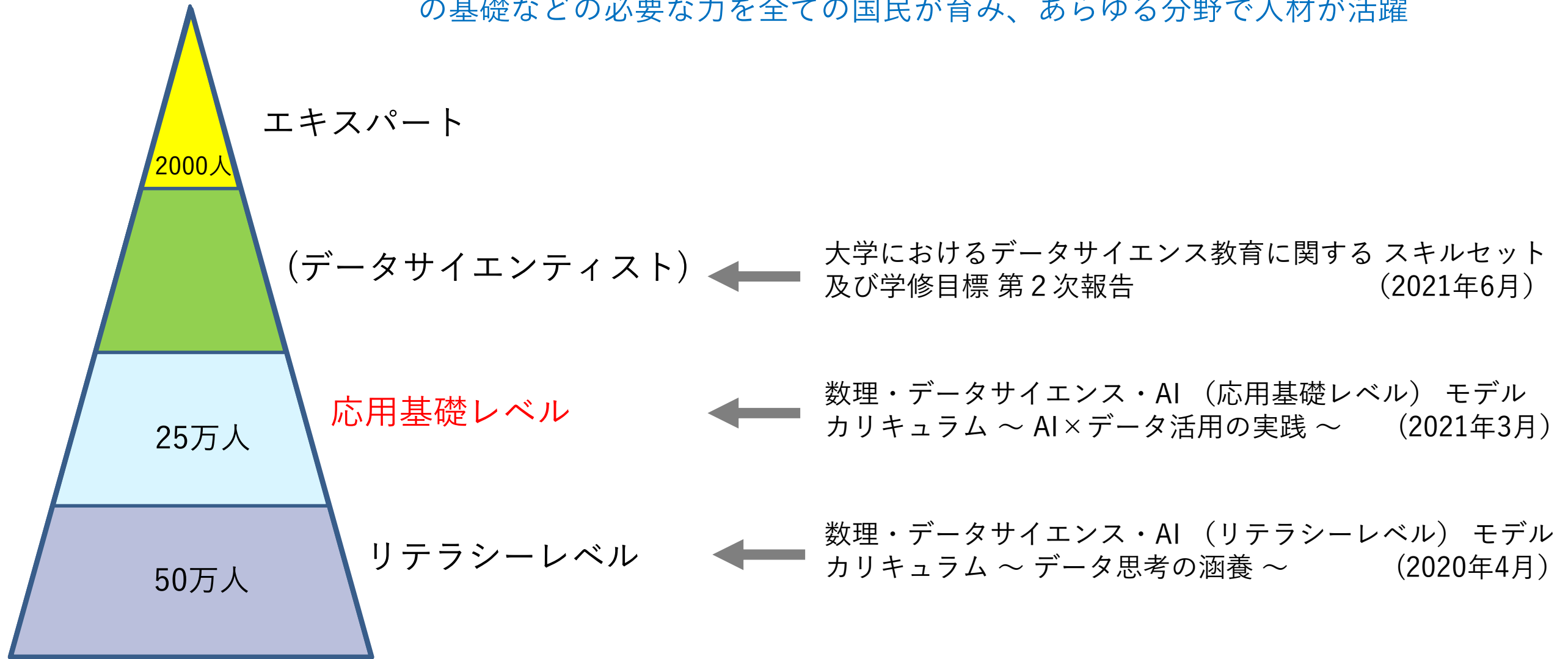
東京大学 数理・情報教育研究センター

北川 源四郎

育成レベルとコンソーシアムが公開したモデルカリキュラム

AI戦略2019

デジタル社会の「読み・書き・そろばん」である「数理・データサイエンス・AI」の基礎などの必要な力を全ての国民が育み、あらゆる分野で人材が活躍



応用基礎レベル モデルカリキュラムの構成

数理・データサイエンス・AI（応用基礎レベル）モデルカリキュラム ～ AI×データ活用の実践 ～

3. AI基礎

3-1. AIの歴史と応用分野（☆）

3-2. AIと社会（☆）

3-3. 機械学習の基礎と展望（☆）

3-4. 深層学習の基礎と展望（☆）

3-5. 認識

3-6. 予測・判断

3-7. 言語・知識

3-8. 身体・運動

3-9. AIの構築と運用（☆）

1. データサイエンス基礎

1-1. データ駆動型社会とデータサイエンス（☆）

1-2. 分析設計（☆）

1-3. データ観察

1-4. データ分析

1-5. データ可視化

1-6. 数学基礎（※）

1-7. アルゴリズム（※）

2. データエンジニアリング基礎

2-1. ビッグデータとデータエンジニアリング（☆）

2-2. データ表現（☆）

2-3. データ収集

2-4. データベース

2-5. データ加工

2-6. ITセキュリティ

2-7. プログラミング基礎（※）

☆はコア学修項目， ※は数理・データサイエンス・AIを学ぶ上で基盤となる学修項目

リテラシーレベル選択項目と応用基礎レベル学修項目の対応

リテラシーレベル モデルカリキュラム

導入	1. 社会におけるデータ・AI利活用
基礎	2. データリテラシー
心得	3. データ・AI利活用における留意事項
選択	4-1. 統計および数理基礎
	4-2. アルゴリズム基礎
	4-3. データ構造とプログラミング基礎
	4-4. 時系列データ解析
	4-5. テキスト解析
	4-6. 画像解析
	4-7. データハンドリング
	4-8. データ活用実践(教師あり学習)
	4-9. データ活用実践(教師なし学習)

応用基礎レベル モデルカリキュラム

1. データサイエンス基礎	1-1. データ駆動型社会とデータサイエンス 1-2. 分析設計 1-3. データ観察 1-4. データ分析 1-5. データ可視化 1-6. 数学基礎 1-7. アルゴリズム
2. データエンジニアリング基礎	2-1. ビッグデータとデータエンジニアリング 2-2. データ表現 2-3. データ収集 2-4. データベース 2-5. データ加工 2-6. ITセキュリティ 2-7. プログラミング基礎
3. AI 基礎	3-1. AIの歴史と応用分野 3-2. AIと社会 3-3. 機械学習の基礎と展望 3-4. 深層学習の基礎と展望 3-5. 認識 3-6. 予測・判断 3-7. 言語・知識 3-8. 身体・運動 3-9. AIの構築と運用

東京大学 MIセンターにおける教材開発

- コンソーシアムでは数理・データサイエンス・AIの教材ポータルを設置し、各大学の教材を集約

<http://www.mi.u-tokyo.ac.jp/consortium/e-learning.html>

- コンソーシアム活動の一環として、東京大学 数理・情報教育研究（MI）センターではモデルカリキュラムに対応した教材を開発・公開

http://www.mi.u-tokyo.ac.jp/6university_consortium.html

- リテラシーレベル教材： 2020年5月
- 応用基礎レベル教材： 2021年6月

この教材の特徴

- モデルカリキュラム(応用基礎レベル)に**完全準拠**
 - 全てのキーワード（知識・スキル）を解説
 - キーワードとの対応表を用意
- MIセンターの総力を結集して作成
- 一部は講義スライドのほかに補助教材も提供
- 教育目的なら**ページ単位で自由**に改変・利用・再配布可能

CC BY-NC-SA(表示－非営利－継承)ライセンス

応用基礎レベル教材作成ワーキンググループ

全体統括： 駒木 文保 教授（センター長）

編集調整： 北川 源四郎 特任教授

斎藤 洋 教授

執筆者

久野 遼平 講師 1章1節, 3章9節

島田 尚 准教授 1章2節

河合 玲一郎 教授 1章3節

北川 源四郎 特任教授 1章4節

寒野 善博 教授 1章4節, 6節

小林 亮太 准教授 1章5節

荻原 哲平 准教授 1章6節

小池 祐太 准教授 1章6節

小林 浩二 講師 1章7節, 2章2節

斎藤 洋 教授 2章1節, 3節

森 純一郎 准教授 2章3節, 3章1節, 7節

佐藤 重幸 助教 2章4節, 5節

小林 克志 特任准教授 2章6節, 7節

山肩 洋子 准教授 3章2節, 5節

二反田 篤史 助教* 3章3節, 4節

郡山 知樹 講師 3章6節, 8節

* 現在, 九州工業大学 大学院情報工学研究院 知能情報工学研究系 准教授

3章8節の一部は下記の2名に執筆を依頼した。

柏野 善大 東京大学 先端科学技術研究センター 助教

宮寄 哲郎 東京大学大学院 情報理工学系研究科

システム情報学専攻 助教

CCライセンス (Creative Commons Licenses)

<https://creativecommons.org/licenses/by-nc-sa/4.0/deed.ja>



表示 - 非営利 - 継承 4.0 国際
(CC BY-NC-SA 4.0)

あなたは以下の条件に従う限り、自由に：

共有 — どのようなメディアやフォーマットでも資料を複製したり、再配布できます。

翻案 — マテリアルをリミックスしたり、改変したり、別の作品のベースにしたりできます。

あなたがライセンスの条件に従っている限り、許諾者がこれらの自由を取り消すことはできません。

あなたの従うべき条件は以下の通りです。



表示 — あなたは 適切なクレジットを表示し、ライセンスへのリンクを提供し、変更があったらその旨を示さなければなりません。これらは合理的であればどのような方法で行っても構いませんが、許諾者があなたやあなたの利用行為を支持していると示唆するような方法は除きます。



非営利 — あなたは営利目的でこの資料を利用してはなりません。



継承 — もしあなたがこの資料をリミックスしたり、改変したり、加工した場合には、あなたはあなたの貢献部分を元の作品と同じライセンスの下に頒布しなければなりません。

追加的な制約は課せません — あなたは、このライセンスが他の者に許諾することを法的に制限するようないかなる法的規定も技術的手段も適用してはなりません。

応用基礎レベル教材サイトへのアクセスの方法

● コンソーシアムHPから



教材ポータルサイトへのリンク

教材ポータルサイト



東大教材サイトへのリンク

● 東大 数理・情報教育センターHPから



東大教材サイトへのリンク

数理・情報教育研究センター
Mathematics and Informatics Center

 東京大学
The University of Tokyo

[トップ](#)
ENGLISH

センター概要 メンバー 大学間コンソーシアム 産業界コンソーシアム 関連講義 e-learning教材 社会人教育

数理・データサイエンス教育強化拠点コンソーシアム

MIセンターは、2017年度政府予算に盛り込まれた「数理・データサイエンス教育の強化」事業の東京大学における実施主体です。同事業で選定された6大学（北海道大学、東京大学、滋賀大学、京都大学、大阪大学、九州大学）のセンターのコンソーシアムの幹事校として、大学、産業界、研究機関等と幅広くネットワークを形成し、地域や分野における先進的教育モデルの拠点として 実践的な教育の普及に努めます。

[\[コンソーシアムホームページ\]](#)

数理・データサイエンス・AIの活用事例動画

本動画集は数理・データサイエンス・AIリテラシーレベル教材の導入となるような活用事例を収集したものです。数理・データサイエンス・AIリテラシーレベル教材との対応は下の教材一覧をご覧ください。

[人工知能（AI）の活用事例動画一覧](#)

数理・データサイエンス・AI教材

本教材は数理・データサイエンス教育強化コンソーシアムが公開したリテラシーレベル及び応用基礎レベルのモデルカリキュラムの全国展開に向け、東京大学 数理・情報教育センターが開発したものです。モデルカリキュラムに完全準拠したものであり、全ての大学、高専において、学生への教育にご利用いただけます。

[本教材の利用についての詳細](#)

応用基礎レベル教材の詳細目次

◆ [教材目次詳細版](#)

◆ [キーワードと教材の対応表](#)

下に移動

下に移動

章	節	教材	実習用補助教材
1. データサイエンス基礎	1-1. データ駆動型社会とデータサイエンス	スライド 	
	1-2. 分析設計	スライド 	
	1-3. データ観察	スライド 	
	1-4. データ分析	スライド 	補助資料
	1-5. データ可視化	スライド 	補助教材
	1-6. 数学基礎(前半)	スライド 	
	1-6. 数学基礎(後半)	スライド 	
2. データエンジニアリング基礎	1-7. アルゴリズム	スライド 	
	2-1. ビッグデータとデータエンジニアリング	スライド 	
	2-2. データ表現	スライド 	
	2-3. データ収集	スライド 	
	2-4. データベース	スライド 	
	2-5. データ加工	スライド 	
	2-6. ITセキュリティ	スライド 	
3. AI基礎	2-7. プログラミング基礎	スライド 	
	3-1. AIの歴史と応用分野	スライド 	
	3-2. AIと社会	スライド 	
	3-3. 機械学習の基礎と展望	スライド 	
	3-4. 深層学習の基礎と展望	スライド 	
	3-5. 認識	スライド 	
	3-6. 予測・判断	スライド 	
	3-7. 言語・知識	スライド 	
	3-8. 身体・運動	スライド 	
	3-9. AIの構築・運用	スライド 	

目次詳細版とキーワードー教材対応表

目次詳細版

1. データサイエンス基礎		
1-1. データ駆動型社会とデータサイエンス (☆)	データ駆動型社会、Society 5.0	1章1節 15-20
	データサイエンス活用事例 (仮説検証、知識発見、原因究明、計画策定、判断支援、活動代替など)	1章1節 4-14
	データを活用した新しいビジネスモデル	1章1節 8
1-2. 分析設計 (☆)	データ分析の進め方、仮説検証サイクル	1章2節 4
	分析目的の設定	1章2節 5
	様々なデータ分析手法 (回帰、分類、クラスタリングなど)	1章2節 14-24
	様々なデータ可視化手法 (比較、構成、分布、変化など)	1章2節 9-17,24
	データの収集、加工、分割/統合	1章2節 6-8
	分析目的に応じた適切な調査 (標本調査、標本誤差)	1章2節 25-30
	サンプルサイズ的设计	1章2節 28-30
1-3. データ観察	ランダム化比較試験、実験計画法	1章2節 27,30
	データの集計、比較対象の設定、クロス集計表	1章3節 4-7, 21
	データのバラツキ、ヒストグラム、散布図	1章3節 8-19
1-4. データ分析	データの特異点、相違性、傾向性、関連性	1章3節 5,7-9, 11-18
	単回帰分析、重回帰分析、最小二乗法	1章4節 6,8,18
	ロジスティック回帰分析、最尤法	1章4節 22,28,30
	時系列データ、時系列グラフ、周期性、移動平均	1章4節 34-35,38-40
	クラスター分析、デンドログラム	1章4節 47,50
	パターン発見、アソシエーション分析、リフト値	1章4節 56, 58, 60, 62
	主成分分析、次元削減	1章4節 66,71
	連続最適化問題、組み合わせ最適化問題	1章4節 96-97
	ナップサック問題、巡回セールスマン問題	1章4節 101, 106
1-5. データ可視化	可視化目的 (比較、構成、分布、変化など) に応じた図表化	1章5節 4-18
	1〜3次元の図表化 (棒グラフ、折線グラフ、散布図、積み上げ縦棒グラフ、箱ひげ図、散布図行列、ヒートマップなど)	1章5節 5-22
	適切な縦軸、横軸候補の洗い出し	1章5節 22
	不必要な誇張表現、強調表現がもたらす影響	1章5節 23, 24
	ビッグデータの可視化	1章5節 25-27
	関係性の可視化 (ネットワーク構造、グラフ構造、階層構造)	1章5節 28-32
	地図上の可視化、地理情報システム (GIS)	1章5節 33, 34
	挙動・軌跡の可視化	1章5節 35
	ダイナミックな可視化、リアルタイム可視化	1章5節 34

キーワードー教材対応表

注：太字はコア学習科目

キーワード (知識・スキル)	章一節	ページ
1〜3次元の図表化	1-5	5-18
1 変数関数の積分法	1-6 (後半)	122, 124
1 変数関数の微分法	1-6 (後半)	107
2 変数関数の積分法	1-6 (後半)	139
2 変数関数の微分法	1-6 (後半)	135
Accuracy	3-6	13
AIアシスタント	3-9	11
AI技術の活用領域の広がり	3-1	47-48
AIクラウドサービス	3-1	50-51
AIシステムの開発	3-9	4-7
AIシステムの品質	3-9	21
AIと知的財産権	3-2	45
AIとロボット	3-8	5,20
AIに関する原則/ガイドライン	3-2	6
AIの開発環境と実行環境	3-9	8-10
AIの開発基盤 (大規模並列GPUマシンなど)	3-9	16
AIの学習と推論	3-9	4
AIの計算デバイス (GPU、FPGAなど)	3-9	16
AIの公平性	3-2	30-31
AIの社会実装	3-9	15
AIの社会的受容性	3-2	25
AIの信頼性	3-2	43
AIの説明可能性	3-2	43-44
AIの歴史	3-1	4-24
AI倫理	3-2	11
API	2-3	23
AUC (Area Under the Curve)	3-6	15
ER図	2-4	15
FTP	2-3	21
Hadoop	2-5	22

スライド教材の構成

章	節	教材	実習用補助教材
1. データサイエンス 基礎	1-1. データ駆動型社会とデータサイエンス	スライド PDF	
	1-2. 分析設計	スライド PDF	
	1-3. データ観察	スライド PDF	
	1-4. データ分析	スライド PDF	補助資料
	1-5. データ可視化	スライド PDF	補助教材

1. タイトル
2. 概要（学修目標）
3. 本教材の目次
4. 内容
5. . . .

1-1 データ駆動型社会と データサイエンス

東京大学 数理・情報教育研究センター
2021年4月30日

東京大学 数理・情報教育研究センター 久野遼平 2021 CC BY-NC-SA

1

概要

- 本節では
 - データ駆動型社会においてデータサイエンスを学ぶことの意義を理解する
 - データを活用した一連のプロセスを体験し、データ利活用の流れ（進め方）を理解する（仮説検証、知識発見、原因究明、計画策定、判断支援、活動代替など）
 - 来るべき未来社会について理解を深める
- ことを目標とします。

東京大学 数理・情報教育研究センター 久野遼平 2021 CC BY-NC-SA

2

本教材の目次

1. データサイエンス活用事例	4ー7
2. データを活用した新しいビジネスモデル	8ー11
3. 進むAIの社会実装とデータ駆動社会	12ー15
4. Society 5.0とスーパーシティ	16ー24

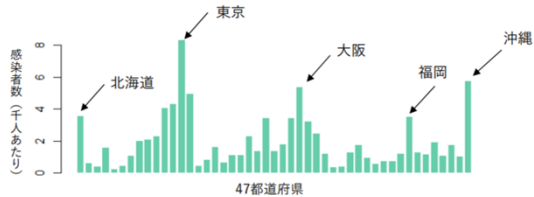
東京大学 数理・情報教育研究センター 久野遼平 2021 CC BY-NC-SA

3

スライドの例

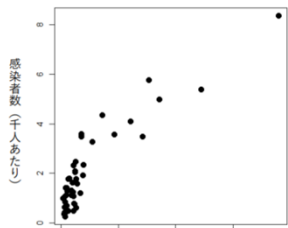
説明変数との関係を見る

下の棒グラフは左から北海道、青森、・・・、沖縄の順に47都道府県の人口千人あたりの新型コロナウイルスの累積感染者数を示しています。都道府県によって感染した人の割合が大きく異なることがわかります。



このような都道府県（以下では単に県と書きます）による感染者の違いを各県の県庁所在地の人口密度と関連づけて説明することを考えてみましょう。

実際に、各県の県庁所在地の人口密度を横軸に、千人当たりの感染者数を縦軸にとって**散布図**を描いてみると、右のように人口密度と感染者数には明確な関係がみえます。



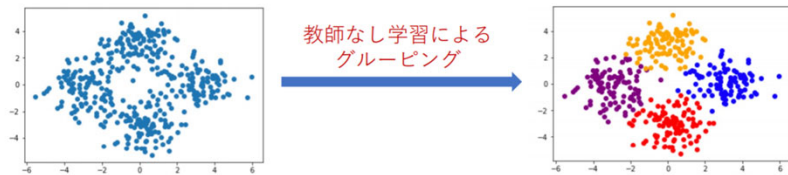
東京大学 数理・情報教育

教師なし学習

訓練データは特徴ベクトルのあつまり $x_1, \dots, x_n$ です。

教師値が与えられない点が教師あり学習と異なります。

教師なし学習によるグルーピング：データ同士の類似性に基づくグルーピング（クラスタリング）は教師なし学習の代表例です。



クラスタリングの例

顧客セグメンテーション→ 顧客情報や利用状況によるクラスタリング

店舗クラスタリング→ 店舗や周辺情報からクラスタリング

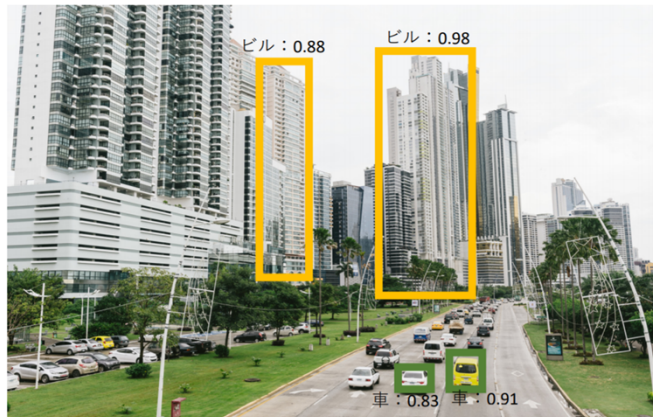
東京大学 数理・情報教育研究センター 二反田篤史 2020 CC BY-NC-SA

27

事例：物体検出

画像認識より一段むずかしい問題設定。

どこに何が写っているかを検出します。

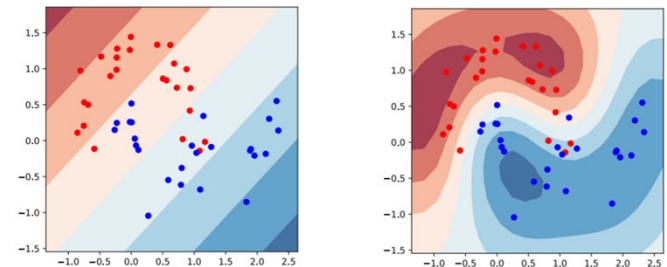


「どこに」 + 「なにが」

東京大学 数理・情報教育研究センター 二反田

カーネルSVM

- 線形SVMは分離境界が直線になるため、データによっては必ずしもうまく予測ができません。
- カーネルSVMは、データ点同士の類似度を表すカーネル関数を用いる手法で、複雑な識別境界を実現可能です。



線形SVM

RBFカーネルSVM

2次元入力変数を2クラス(赤or青)で2クラス分類した例. 赤い領域は赤のクラスに、青の領域は青のクラスに分類されます。

東京大学 数理・情報教育研究センター 郡山知樹, 2021 CC BY-NC-SA

10

章	節	教材	実習用補助教材
1. データサイエンス 基礎	1-1. データ駆動型社会とデータサイエンス	スライド 	
	1-2. 分析設計	スライド 	
	1-3. データ観察	スライド 	
	1-4. データ分析	スライド  補助資料	
	1-5. データ可視化	スライド 	補助教材

1. Rによる計算コード
2. データ (csvファイル)

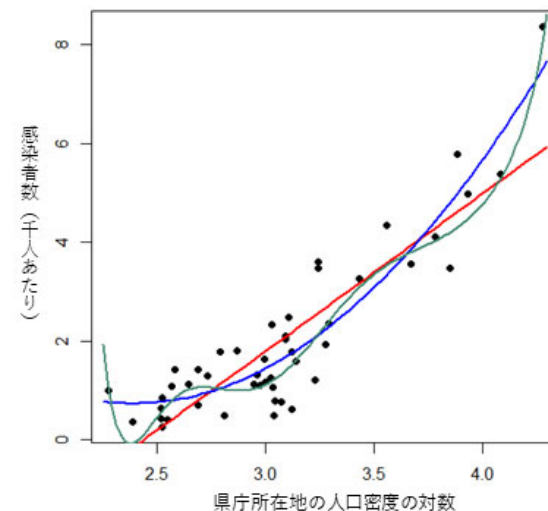
多項式回帰

```
##### 多項式回帰 #####
plot( x,y,pch=16,xlim=c(2.3,4.3) )
newdata <- seq( 2.0,4.5, length=100 )

# 直線回帰
fit <- lm( y ~ x )
beta <- coef(fit)
pred <- beta[1] + beta[2]*newdata
points( newdata,pred, type="l", col="red",lwd=2 )

# 2次回帰
fit <- lm( y ~ poly(x,2,raw=TRUE) )
beta <- coef(fit)
pred <- beta[1] + beta[2]*newdata + beta[3]*newdata^2
points( newdata,pred, type="l", col="blue",lwd=2 )

# 8次回帰
fit <- lm( y ~ poly(x,8,raw=TRUE) )
beta <- coef(fit)
pred <- beta[1]
for (i in 1:8){
  pred <- pred + beta[i+1]*newdata^i
}
points( newdata,pred, type="l", col="aquamarine3",lwd=2 )
```



アンケートのお願い

今後、適宜、修正・改善の予定
アンケートにご協力ください

応用基礎レベル教材についてのアンケート

応用基礎レベル教材の利用につきまして、今後の参考のため、アンケートにご回答ください。メールアドレス（任意）をご記入いただいた大学には、教材の更新情報などをご連絡させていただく予定です。
利用の状況や、ご意見をお気軽にご記入ください。後日、追加のご意見やご要望も歓迎いたします。

*必須

貴校名 *

回答を入力

興味のある章・節 *

回答を入力

想定される利用法 *

- ☐ 節全体を授業で利用
- ☐ ページ単位で授業で利用
- ☐ 授業の参考にする
- ☐ 学生の自習用
- ☐ その他: _____

教材を見てのご意見

回答を入力

その他ご要望事項

回答を入力