

Human-Computer Interaction perspective on building ML applications

HCI分野の“for ML”論文の紹介

加藤 淳 (@arcatdmz)

国立研究開発法人 産業技術総合研究所 主任研究員

XX for ML 論文読み会 #1



加藤 淳 (@arcatdmz)

<https://junkato.jp/ja>

 **東京大学 五十嵐研究室** **ACM DIS '12, CHI '13**
'09 学士, '11 修士, '14 博士 (情報理工学)

 **Microsoft Research Asia** **ACM UIST '12**
'12/1-4 Research Intern and Fellow

 **Microsoft Research Redmond** **ACM PLDI '13**
'12/6-9 Research Intern

 **Adobe Research Seattle**
'13/8-11 Research Intern

 **産業技術総合研究所** **ACM CHI '15, DIS '17, Multimedia '18, ...**
'14/4- 研究員 '18/10- 主任研究員 (主務)

 **アーチ株式会社**
'18/7- 技術顧問 (兼任)

加藤 淳 (@arcatdmz)

<https://junkato.jp/ja>

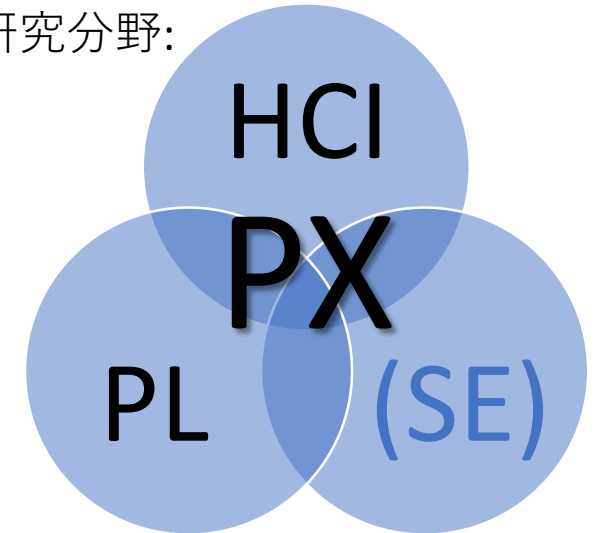
- **Human-Computer Interaction分野の学会運営**

- [ACM CHI 2017](#): Associate Chair (≡ Program Committee)
- [ACM UIST 2018](#), [2017](#), [2014](#): Program Committee
- [ACM IUI 2018](#), [2019](#): Program Committee (≡ Reviewer)

- **Programming Experience分野(?)の学会運営**

- LIVE workshops (co-located w/ [ECOOP 2016](#), [SPLASH 2017](#), [2018](#)): Steering Committee, Program Committee
- PX workshops (co-located w/ [ECOOP 2016](#), [SPLASH 2017](#), [<Programming> 2017](#), [2018](#)): Program Committee
- [<Programming> 2018](#) conference: Program Committee

研究分野:



小特集

プログラミング・ エクスペリエンスの 新潮流

—言語設計から産業応用まで—

SIGPX <http://sigpx.org>

2016-

参加者層

Human-Computer Interaction, ソフトウェア工学, プログラミング言語など
プログラミング
の研究者/エンジニア

主旨

国際会議などの
集合して情報共有する場
プログラミング体験に関して誰でも
相談できる窓口としての役割

主宰

加藤（産総研） **増原**（東工大）

第5回: 2018年12月3日(月)

東京大学本郷キャンパス理学部7号館



HCI分野の“for ML”論文:

そもそもHCI分野の国際会議ってどんな？

- 総合トップ: **ACM CHI**

ACM CHI Conference on Human Factors in Computing Systems

- 技術寄り: **ACM UIST**

ACM Annual Symposium on User Interface Software and Technology

—— トップ会議の壁 ——

- ちょっとデザイナー寄り、最近伸びてる: **ACM DIS**

ACM Conference on Designing Interactive Systems

- 賢いシステム: **ACM IUI**

International Conference on Intelligent User Interfaces

- プログラミング言語やAPI設計など: **IEEE VL/HCC**

IEEE Symposium on Visual Languages and Human-Centric Computing

HCI分野の“for ML”論文:

本日はご紹介する論文

- Disclaimer: 大半が2018年以前刊行です...
- Gestalt [UIST '10]
- Label-and-Learn [IUI '17]
- Machine Learning for Makers [DIS '17]
- DS.js [UIST '17 Best Paper Honorable Mention]
- How Non-Experts Actually Build Models [DIS '18]

Gestalt: Integrated Support for Implementation and Analysis in Machine Learning

UIST 2010

- Kayur Patel et al., University of Washington & Microsoft Research
- 機械学習のワークフローを支援できるような統合開発環境
- これまでの開発環境はコード重視だがデータも重視
- いわゆるモダンな統合開発環境が研究対象になった黎明期

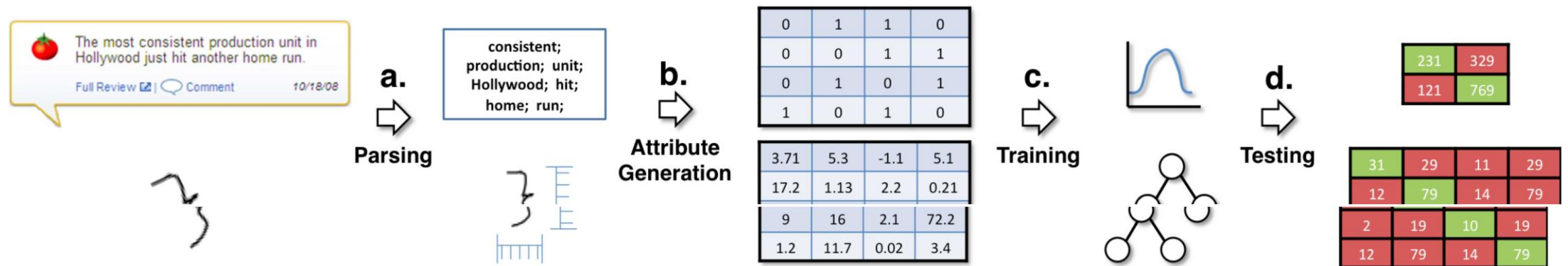


図1: 機械学習の典型的なワークフロー (感情分析とジェスチャ認識で同じワークフロー)

Gestalt: Integrated Support for Implementation and Analysis in Machine Learning

UIST 2010

- 開発環境とビジュアライザを行ったり来たりしなくて済む
- ワークフロー全体が支援される

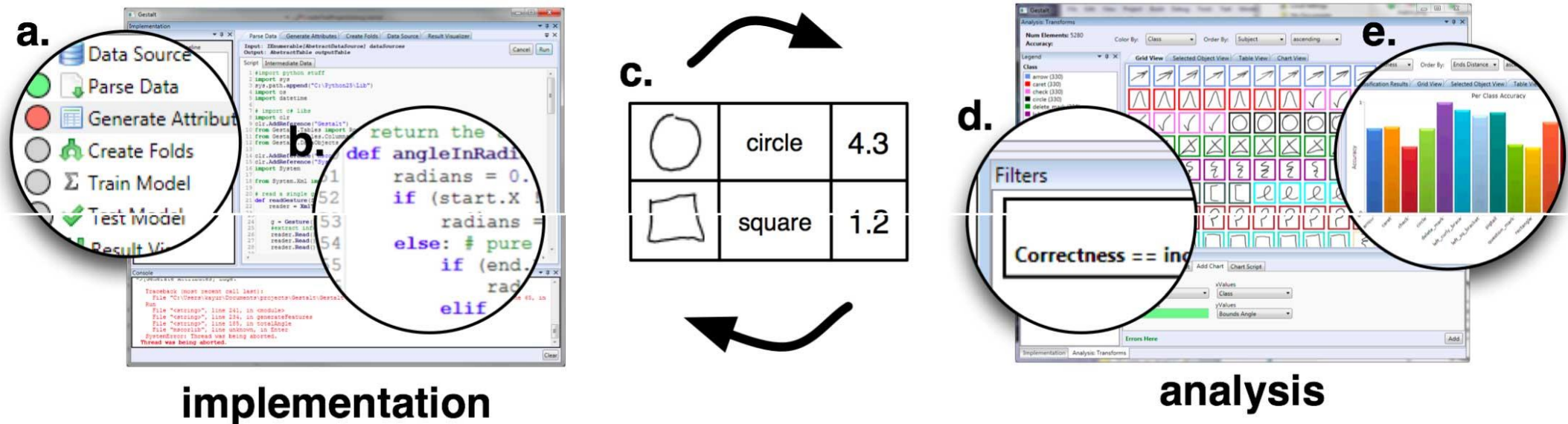


図2: 提案手法 (a. パイプライン b. コード c. データ構造 d. フィルタ、並べ替え、色付け e. 可視化)

Gestalt: Integrated Support for Implementation and Analysis in Machine Learning

UIST 2010

- 既存研究は以下がほとんど
 - Domain-specific tools: 1. 画像処理などドメイン限定 2. 非プログラマ対象
 - Disconnected general purpose tools: e.g., 可視化のみ支援するTableau
- 本研究はConnected general purpose tools: e.g., MATLAB
- そこで、MATLABのような環境をベースラインに評価実験
 - ワークフロー全体が見えてるのすごくいい
 - 可視化をカスタマイズできるのもいい
 - インタラクティブに結果とコードを行き来できるのもいい
 - 小規模ユニットテストしたいユーザがいた
 - 個々データが独立な前提➡格子状の可視化は時系列データ等に不向き

Google Colaboratory

- Gestaltを作ったKayurが博士修了後にGoogleに就職して作った
- Jupyter Colaboratory [2014]がベース
 - <https://github.com/jupyter/colaboratory/graphs/contributors>
 - この頃からJupyterにはポテンシャルがあった
 - Literate programming（文芸的プログラミング）の再興
- ちなみに
 - KayurはGoogle Colaboratoryのリリリース直前にAppleに移籍しました
 - Appleで機械学習インタラクション系の研究インターン募集しています

Label-and-Learn: Visualizing the Likelihood of Machine Learning Classifier's Success During Data Labeling

IUI 2017

- Sun et al., University of Waterloo
- インタラクティブな機械学習手法を使うためデータにラベリングしている段階で、いかに有用な情報を提示するか？
- 既存手法:
 - 1) 現段階の結果のみ見せて、さらにラベリングできるようにする
 - 2) ラベリング完了状態からモデルをいじれるようにする(Gestalt)かモデル同士比較しやすくする
- もっといろんな情報出そう

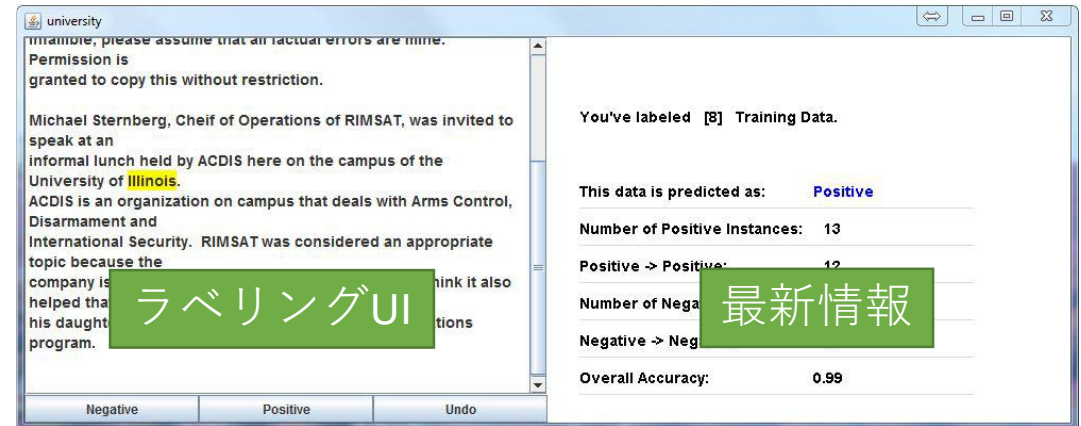
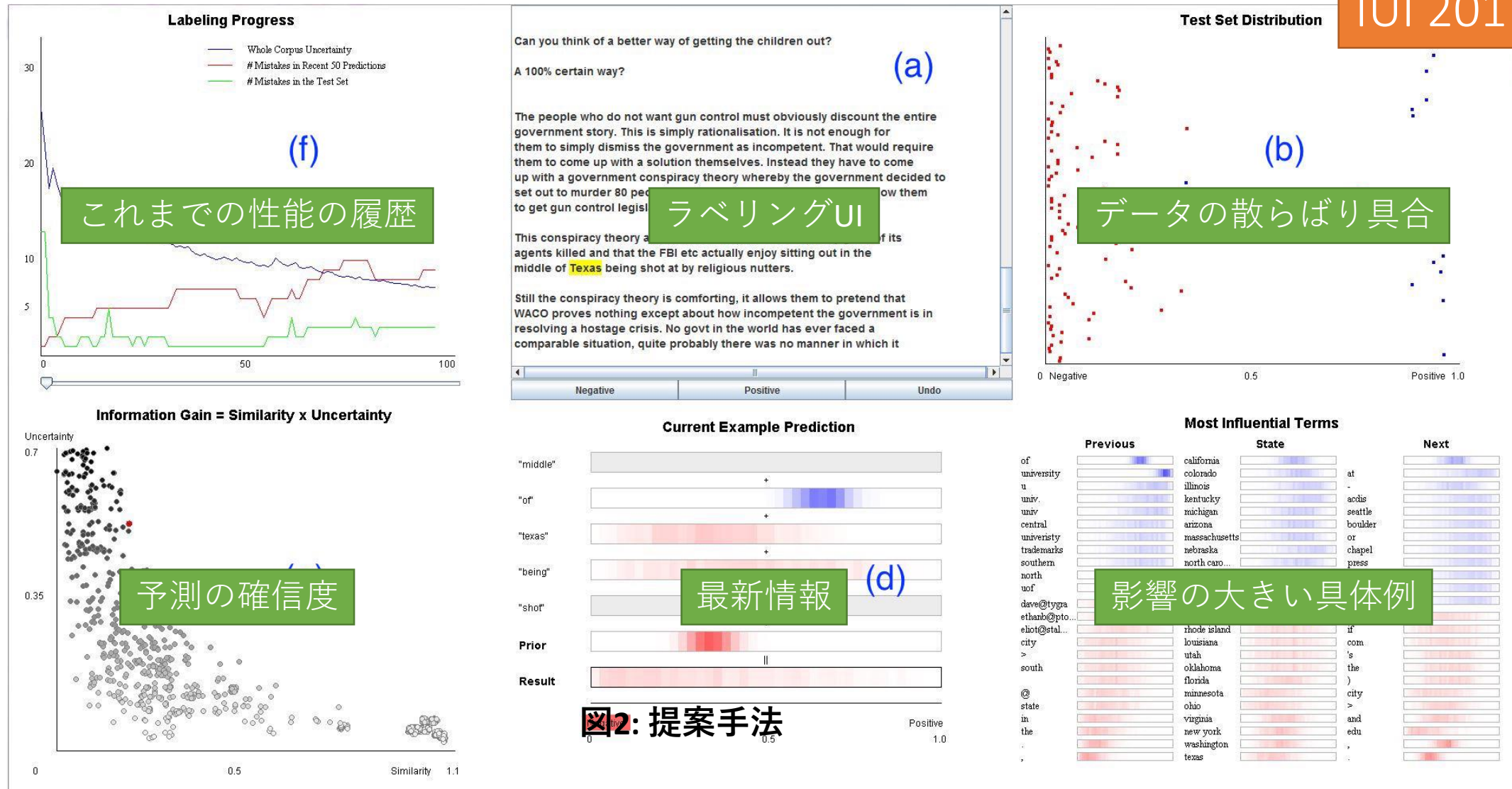


図3: 典型的な既存手法 (最新の統計値だけ表示)

Label-and-Learn: Visualizing the Likelihood of Machine Learning Classifier's Success During Data Labeling

IUI 2017



Label-and-Learn: Visualizing the Likelihood of Machine Learning Classifier's Success During Data Labeling

IUI 2017

- Naïve Bayesで固有表現抽出した例にしか使えないんじゃない？
- 一般性はどれくらいあるのか
- 同じポリシーの可視化を他のモデルでもやってみました

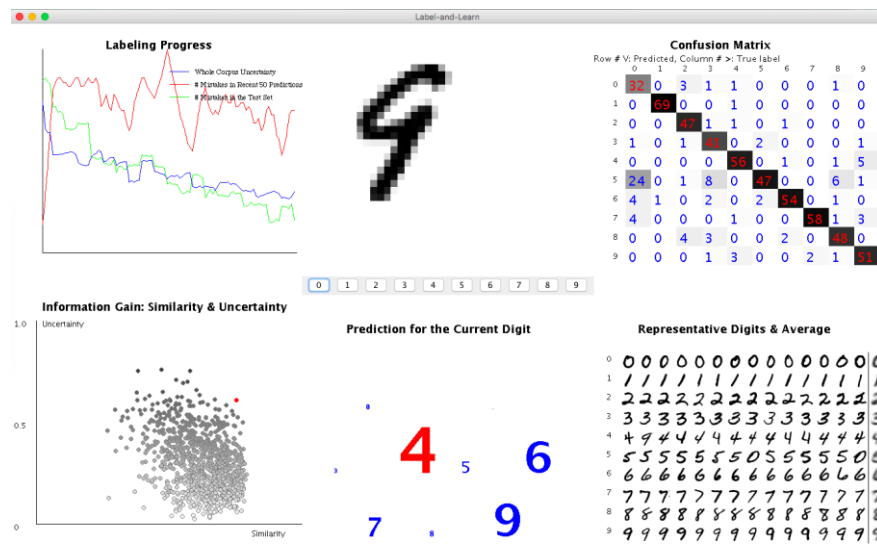
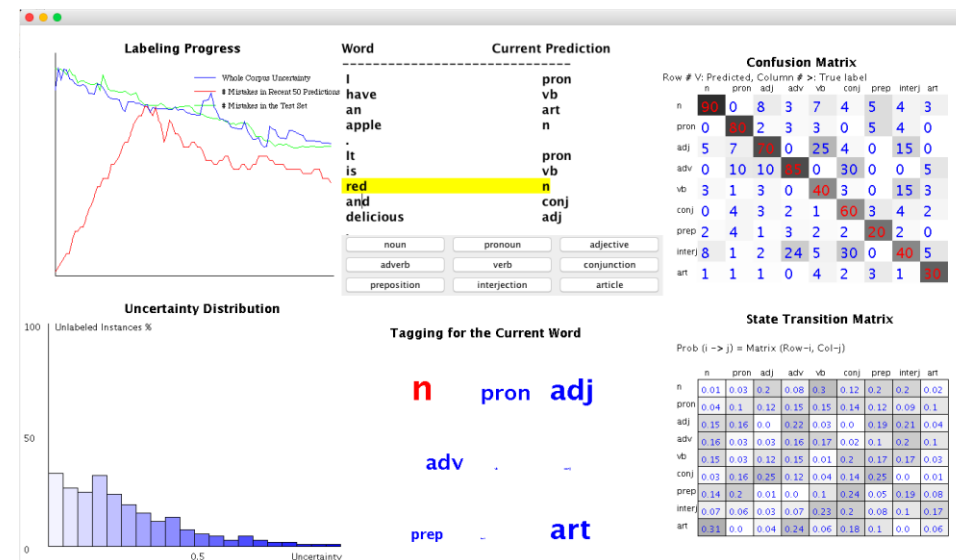


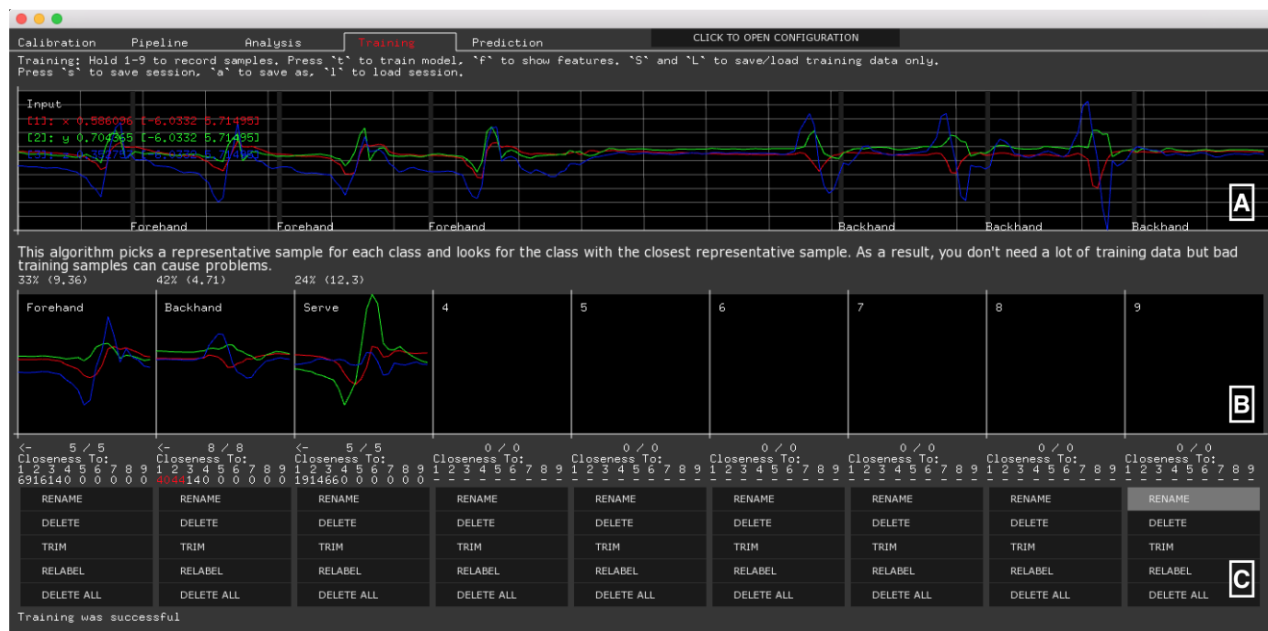
図6:
(左) 深層学習で
文字認識
(右) HMMで品
詞タグ付け



Machine Learning for Makers: Interactive Sensor Data Classification Based on Augmented Code Examples

DIS 2017

- Mellis et al., UC Berkeley
 - First authorのDavid MellisはArduinoを作った一人
 - Last authorはPhysical Computing支援研究で第一人者のBjoern Hartmann
- プロ向けには機械学習用APIを提供
- 素人向けに、APIを使って作られたパイプラインのパラメタ調整と学習をインタラクティブに行えるUI (右図)を提供

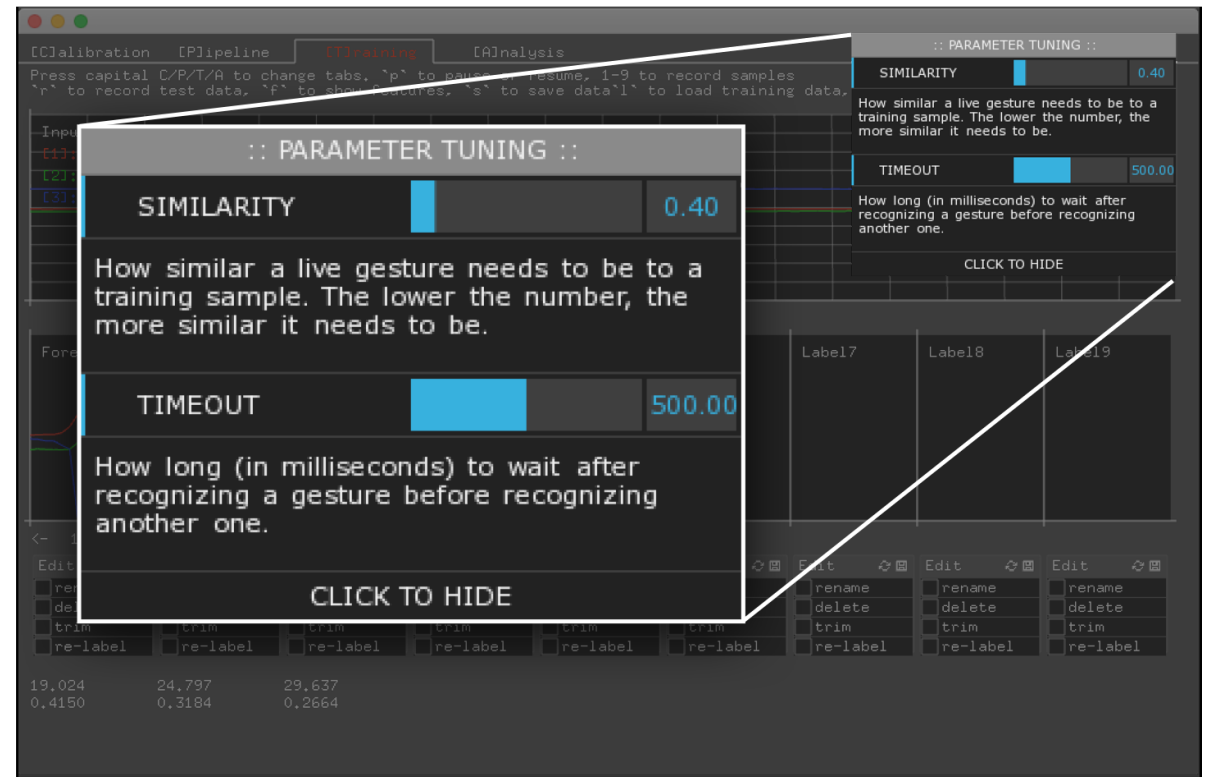


Machine Learning for Makers: Interactive Sensor Data Classification Based on Augmented Code Examples

DIS 2017

- GitHub: <https://github.com/damellis/ESP/wiki>

- 素人はプロが用意したUIでパラメタ調整できる
- エッジデバイスでの機械学習支援かと期待したけど、実際はPC/Macが母艦になる必要があった
- 提案されているインタラクションはどれも割と既出 (Juxtapose, Live Tuning, ...)



DS.js: Turn Any Webpage into an Example-Centric Live Programming Environment for Learning Data Science

UIST 2017

- Zhang et al., UC San Diego
 - Online PythonTutorで有名なPhilip Guoの研究室の成果
 - Best Paper Honorable Mention

- HTML TableやCSVファイル(へのリンク)などをデータセットとして利用
- 即席ライブプログラミング環境

(a) Load DS.js!

(b) Population in the world increased from 1990 to 2008 with 1,423 million and 27% growth. Measured by persons, the increase was highest in India (290 million) and China (192 million). Population growth was highest in Qatar (174%) and United Arab Emirates (140%).^[80]

Rank	Country	Population (thousands) 1990	Population (thousands) 2010
1	World	5,306,425	6,895,889
2	China	1,145,195	1,341,335
3	India	873,785	1,224,614
4	United States	253,339	310,384
5	Indonesia	184,346	239,871
6	Brazil	149,650	194,946
7	Pakistan	111,845	173,593
8	Nigeria	97,552	158,423
9	Bangladesh	105,256	148,692
10	Russia	148,244	142,958
11	Japan	126,536	122,251

(c) - ... and generate data visualizations

(c) - write code to manipulate data tables

```
1 t6; // This table is denoted as t6
2 t6.filtered = t6.where('Population (thousands) 1990', x => x < 200000);
3 t6.filtered.barplot(1, 3);
4
```

(d) - table API suggestions

In-Context Operations

- set('Population (thousands) 1990', x => x)
- get_column('Population (thousands) 1990')
- select_columns('Population (thousands) 1990')
- drop_columns('Population (thousands) 1990')
- rename_column('Population (thousands) 1990', 'new_label')
- where('Population (thousands) 1990', x => true)
- sorted('Population (thousands) 1990')
- groupby('Population (thousands) 1990')

DS.js: Turn Any Webpage into an Example-Centric Live Programming Environment for Learning Data Science

UIST 2017

- 既存研究は一般的な統合開発環境やMATLAB、RStudio
- 教育用に使いやすいよう、さまざまな工夫が詰まっている
 - ブラウザで見ているデータをすぐ使えるbookmarkletになっている
 - データをDSTableクラスでラップして学習用APIを提供
 - コードと可視化結果が密に連携するような実装
 - tableに限らず柔軟にデータ取得してコードに貼り込めるUI
 - コードはクエリパラメタとして共有可能

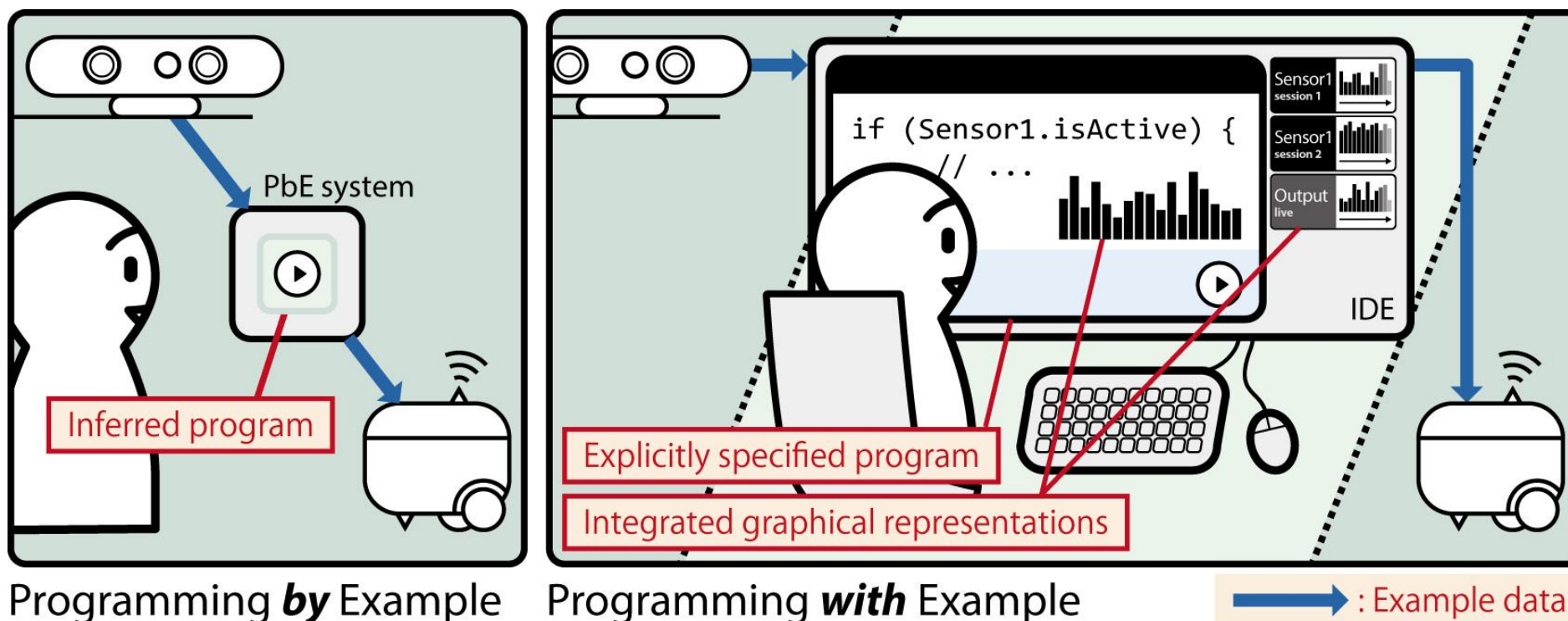
- b.)
- Alice
 - Bob
 - Carol
 - Doug
 - Eli
 - London
 - New York
 - Paris
 - Tokyo

RANK	PLAYER	POSITION	TEAM	SALARY(M)
1	Kobe Bryant	SF	Los Angeles Lakers	25
2	Joe Johnson	SF	Brooklyn Nets	24.894863
3	LeBron James	SF	Cleveland Cavaliers	22.9705
4	Car	In-Context Operations	New York Knicks	22.875
5	Dw	Global Operations	Houston Rockets	22.359364
6	Chris Bosh	PF	Los Angeles Clippers	22.19273
7	Chris Paul	PG	Oklahoma City Thunder	21.468695
8	Kevin Durant	SF	San Antonio Spurs	20.158622
9	Derrick Rose	PG	Cleveland Cavaliers	20.093064
10	Dw	Global Operations	Miami Heat	20
11	LaMarcus Aldridge	PF	Memphis Grizzlies	19.689
12	Kevin Love	PF	Memphis Grizzlies	19.689
13	Marc Gasol	PF	Memphis Grizzlies	19.688

Programming with Examples

IEEE Computer 49(7), 2016

- Kato et al., <https://junkato.jp/programming-with-examples/>
- コードよりデータが大事な時代の開発環境論



Grounding Interactive Machine Learning Tool Design in How Non-Experts Actually Build Models

DIS 2018

- Yang et al., CMU HCII & Microsoft Research
- 素人（データサイエンスや統計などの学位を持たず必要に応じて機械学習を利用した問題解決を行っている人）を調査した
- とくにインタラクティブな機械学習ツール（iML Tools）対象
- 24人（と雇われプロ）を対面調査、オンラインで98人調査
- とりあえずGUIで簡単にすればいいだろうという風潮に喝

Grounding Interactive Machine Learning Tool Design in How Non-Experts Actually Build Models

DIS 2018

Non-Experts			Supporting Experts*
Profession	Example ML Problem	count	count
Professional Software Engineer	Bug report classifier	4	2
Project Manager	User feedback classifier	2	2
Manager	HR Policy Q&A bot	2	1
Business Analytics	Predictive machine maintenance	1	2
Artist	Emotion classifier for wearables	1	
Botanist	Predictive plant nutrient management	1	1
Academic Researcher	Sensor signal classifier	1	
Clinical Researcher	Prognostic classifier	1	
Mechanical Engineer	Insurance risk estimate	1	1

Grounding Interactive Machine Learning Tool Design in How Non-Experts Actually Build Models

DIS 2018

- 「素人」の取り組み方
 - だいたい最初からデータを持っていて、そこから何らかの洞察を得たいと思っている（一般的なアプリを作りたいわけではない）
 - 機械学習アプリ像が最初から見えているケースは稀だが、その場合は、手動だったり経験則に基づく手法を置き換えたいと思っている
- 機械学習うまくいかなかったらどうするか
 - 洞察が得られればいい・部分的でも動けばいいので、ゴールを動かす
 - 違うアルゴリズムとかモデルを試しがち
- 機械学習をどう捉えているか
 - 何か入れると何か出てくるマッピングだと思っている
 - モデルがおかしいとはあまり思わない（e.g., 植物画像を一枚突っ込んだら病気かどうか出てきてほしいと思って頑張ってラベル付けする）

Grounding Interactive Machine Learning Tool Design in How Non-Experts Actually Build Models

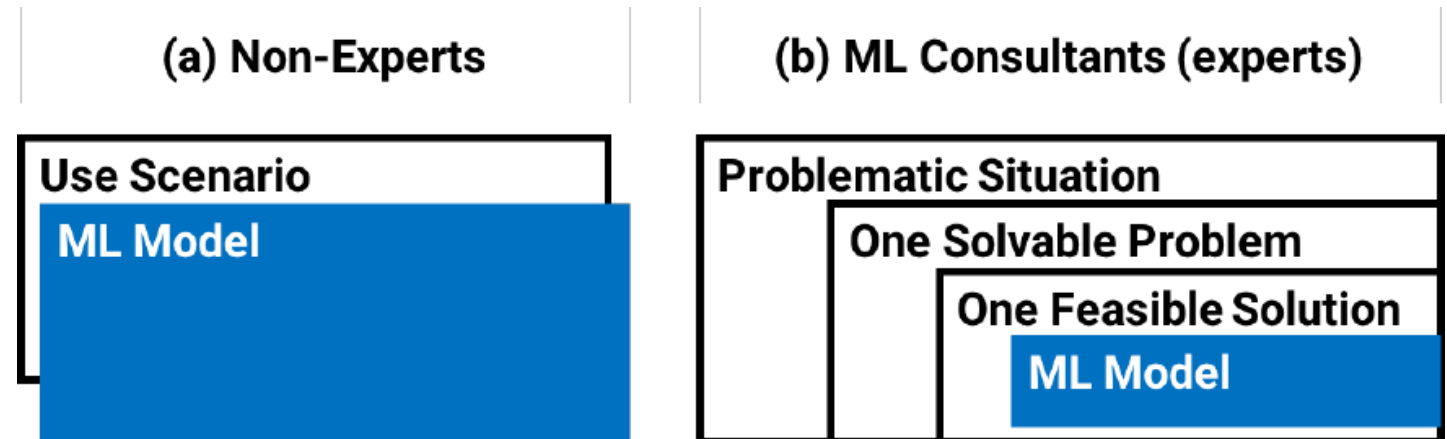
DIS 2018

- 「素人」に信頼されるために
 - モデルをアップデートするたびにできるだけその意味を説明する
 - ときに説明しやすさを優先してパフォーマンスを犠牲にする
 - 徐々に作っていく過程を見せると信頼してもらえる
- そもそも「素人」はどんな人？
 - 問題領域に関する知識はコンサルより当然深いので、使えそうな素性や適切な可視化手法については議論をリードする
 - 実はプログラミングはできる（ほぼソフトウェアエンジニア対象）
 - とはいえ機械学習で殆どコード書かない（前処理だけなど<100行）
 - とにかくサンプルコードをコピペするしドキュメントは読まない

Grounding Interactive Machine Learning Tool Design in How Non-Experts Actually Build Models

DIS 2018

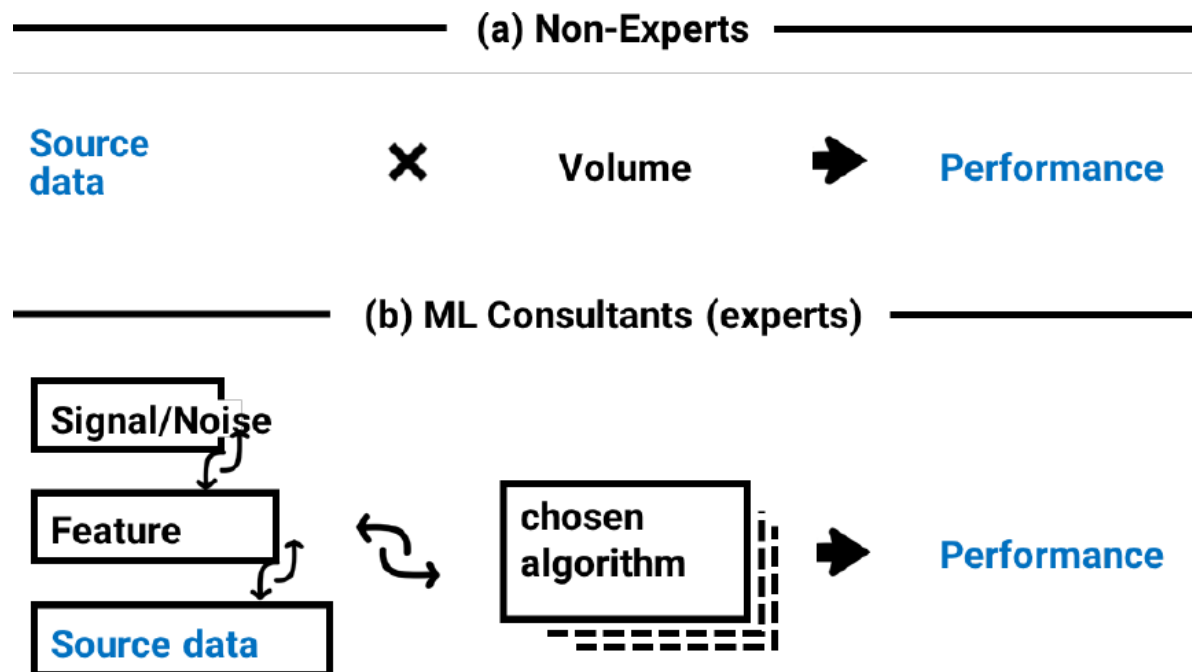
- 「素人」はどんな人？の続き
 - ちょっとでもドキュメントちゃんと読んでればモデル作るのが本当になくなるのに
 - コピペ大好きなので意外とテキストベースのコーディングをする
 - プロのコンサルのほうで毎回モデルを組み替えたりするのでGUIツールを駆使する傾向がある
 - プロは二か月かけるところ、素人は一瞬でモデルに満足するので性能上がらない



Grounding Interactive Machine Learning Tool Design in How Non-Experts Actually Build Models

DIS 2018

- そもそもプロと素人で性能の判断指標が違う
- 素人はデータいっぱい用意すれば性能上がると思ってる
- 素人の手元にデータがある状況から、いかに正しくモデルを作っていけるか支援する必要がある



Grounding Interactive Machine Learning Tool Design in How Non-Experts Actually Build Models

DIS 2018

(a) Typical Existing iML Tools and Observed Non-Expert Usage



Display data;
List data processing options.

List feature engineering options;
List algorithm options.

Display performance matrix.



Did not read through data;
Often did not pre-process data.

Did not understand "feature";
Used one algorithm throughout,
Or tried all algorithms available.

Did not know how to improve
performance; Proceeded to deploy
model once reaching good enough
percentage accuracy.

(b) Test-Driven Machine Teaching :: Interaction Flow



Display data;
Check quality of the data;
Require user to select test cases;
Suggest neglected data regions in the selection.



Actively examine data to elicit a test set;
Develop a holistic understanding of the data;

- Examine whether data instances are important or representative;
- Gradually formulate their ML goals, priorities and expected algorithm behaviors;

Suggest/guide feature set design;
Suggest/guide algorithm picks.
based on data characteristics and user
priorities captured from the test cases.

Make informed feature and
algorithm picks.

- Over time, learns principles and best practices of choosing algorithms.

Communicate model behavior
via test cases;
Suggest debugging strategies.

Examine model performance
based on its behavior on
concrete data instances;
Revise features and algorithms
under guidance and
strategically.

HCI分野の“for ML”論文:

本日はご紹介した論文

- Disclaimer: 大半が2018年以前刊行です...
- Gestalt [UIST '10]
- Label-and-Learn [IUI '17]
- Machine Learning for Makers [DIS '17]
- DS.js [UIST '17 Best Paper Honorable Mention]
- How Non-Experts Actually Build Models [DIS '18]

{ Programmers | Software Engineers }
are { People | Users }, too.

— Prof. Brad Myers



SIGPX <http://sigpx.org>

2016-

参加者層

Human-Computer Interaction, ソフトウェア工学, プログラミング言語など
プログラミング
の研究者/エンジニア

主旨

国際会議などの
集合して情報共有する場
プログラミング体験に関して誰でも
相談できる窓口としての役割

主宰

加藤（産総研） **増原**（東工大）

第5回: 2018年12月3日(月)

東京大学本郷キャンパス理学部7号館



Appendix

今回カバーできなかった2018年刊行論文

Interactions for Untangling Messy History in a Computational Notebook

- Mary Beth Kery and Brad A. Myers. VL/HCC, 2018.
- Best Paper

Trends and Trajectories for Explainable, Accountable and Intelligible Systems: An HCI Research Agenda

- CHI 2018

The Exploratory Labeling Assistant: Mixed-Initiative Label Curation with Large Document Collections

- UIST 2018