

# VISUALIZATIONS FOR MACHINE LEARNING

(An Introduction)

HARI SUBRAMONYAM

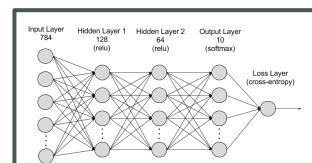


<https://xkcd.com/1838/>

1

What is the zip-code written on the mail? (Handwriting Recognition)

0 0 0 0 0 0 0 0 0 0 0 0 0 0 0  
1 1 1 1 1 1 1 1 1 1 1 1 1 1 1  
2 2 2 2 2 2 2 2 2 2 2 2 2 2 2  
3 3 3 3 3 3 3 3 3 3 3 3 3 3 3  
4 4 4 4 4 4 4 4 4 4 4 4 4 4 4  
5 5 5 5 5 5 5 5 5 5 5 5 5 5 5  
6 6 6 6 6 6 6 6 6 6 6 6 6 6 6  
7 7 7 7 7 7 7 7 7 7 7 7 7 7 7  
8 8 8 8 8 8 8 8 8 8 8 8 8 8 8  
9 9 9 9 9 9 9 9 9 9 9 9 9 9 9

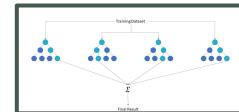
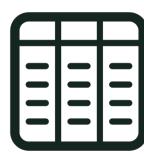
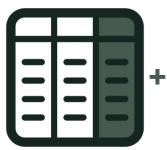


Fold 1	Training set	Testing set	$\epsilon_1$
Fold 2	Training set	Testing set	$\epsilon_2$
Fold 3	Training set	Testing set	$\epsilon_3$
Fold 4	Training set	Testing set	$\epsilon_4$

How much will a flight from CA to NYC cost in December ? (Flight Price Prediction)



Airline  
Date of Journey  
Source  
Destination  
Route  
Total Stops  
Departure Time  
Arrival Time  
Pilot Name  
Price  
Duration  
Departure Gate



Fold 1	Testing set	Training set	$\epsilon_1$
Fold 2	Training set	Testing set	$\epsilon_2$
Fold 3	Training set	Testing set	$\epsilon_3$
Fold 4	Training set	Testing set	$\epsilon_4$

DATA  
COLLECTION &  
PROCESSING

FEATURE  
SELECTION

MODELING

EVALUATION

Machine Learning Pipeline

Liu et al. (2017)

2

# In what ways can **visualizations** support machine learning tasks?

3

## **Understanding Data**

What is the quality of the data?

Is the data representative?

What features are available in the dataset?

Is the test set representative of the data set as a whole?

Is the data labeled correctly?

## **Model Training**

Why is the model behaving the way it is?

What is the structure of the model?

How can I refine the model?

Is this the best technique for modeling the learning task?

How does the model output change with changes in model parameters?

## **Model Evaluation**

How accurate is the model output?

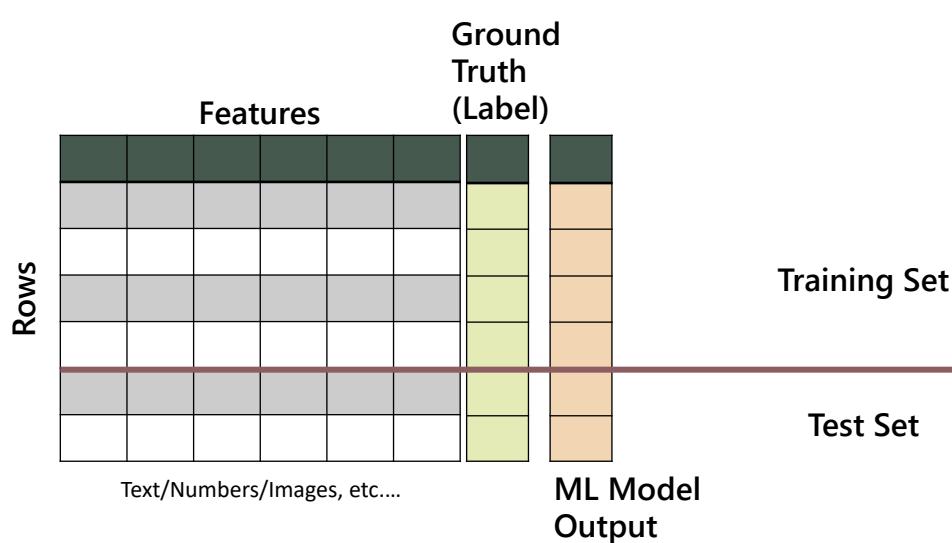
How can I explain the model output?

Is the model fair?

4

# Visualizations for Understanding ML Datasets

5



6

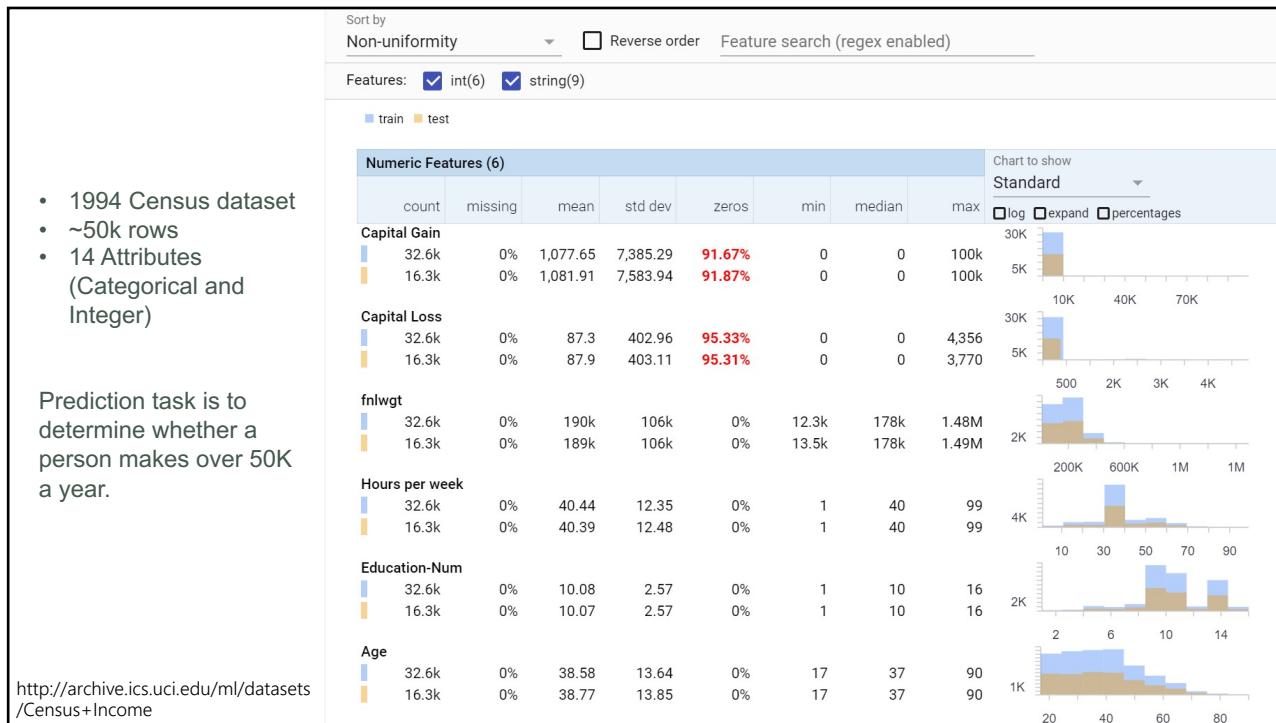
- Understanding Data Characteristics before Modeling
- Selecting Features for Modeling
- Debugging Data based on Model Outputs

7

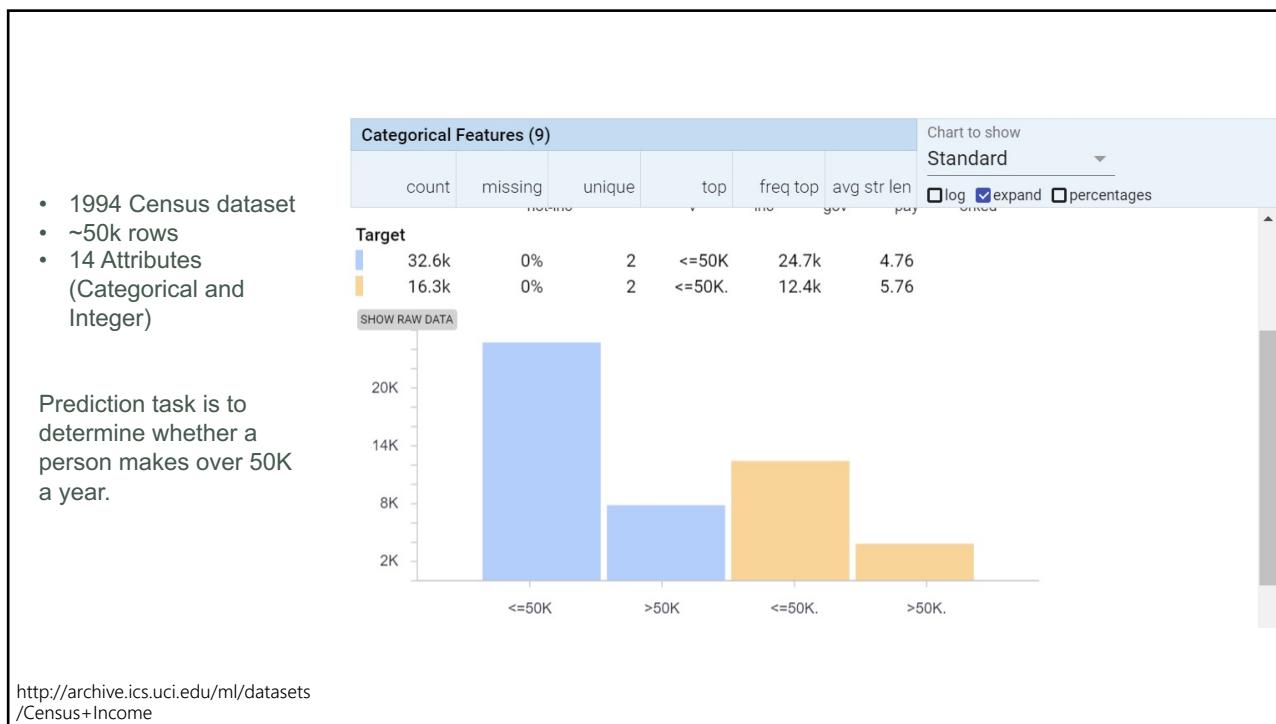
## FACETS



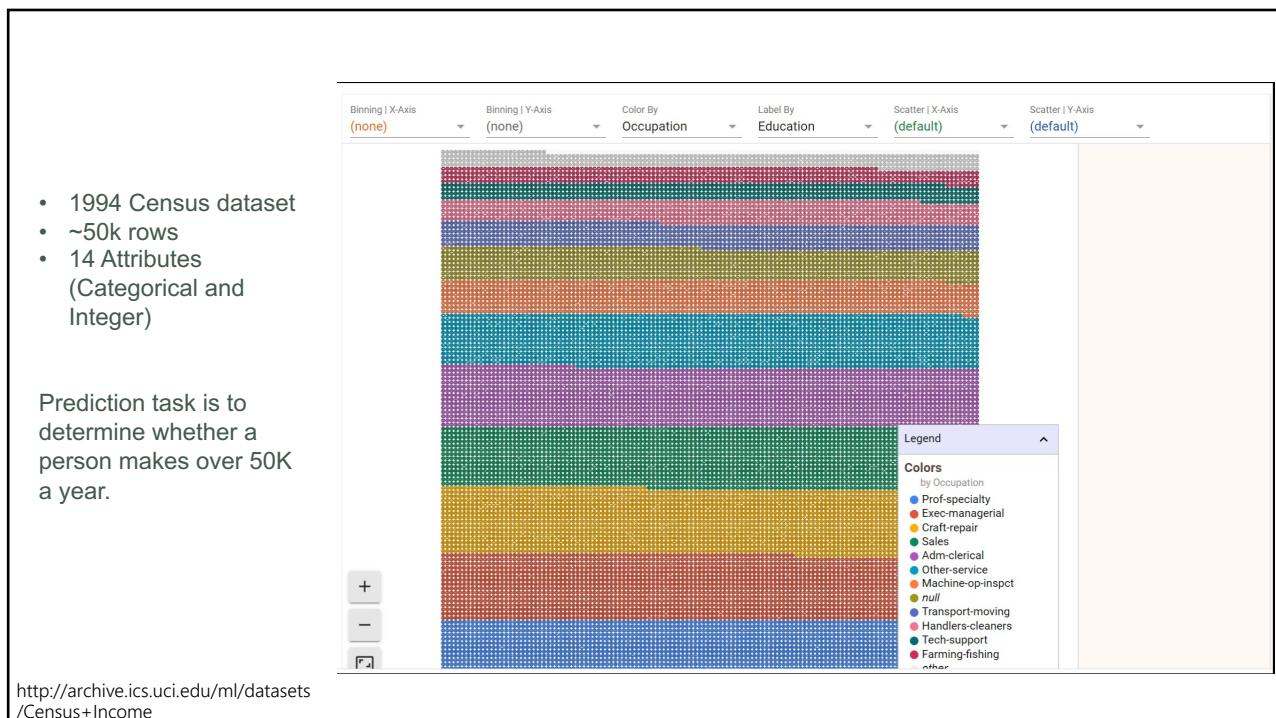
8



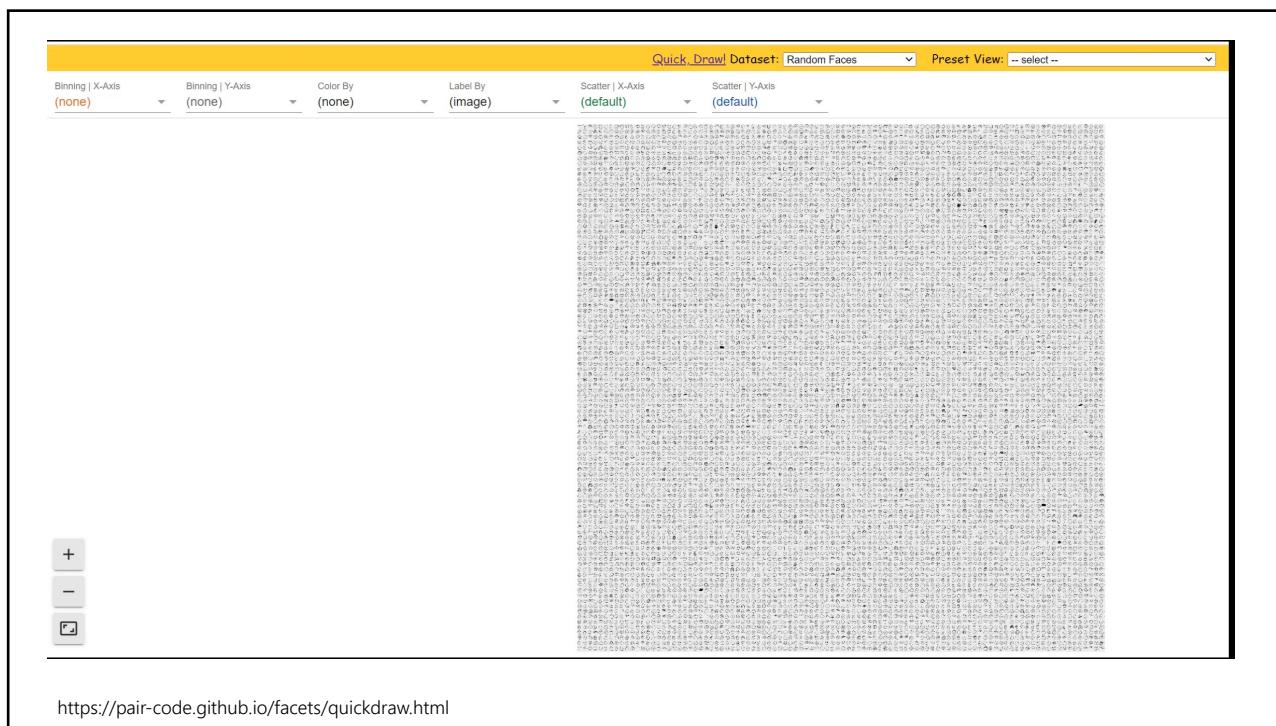
9



10



11



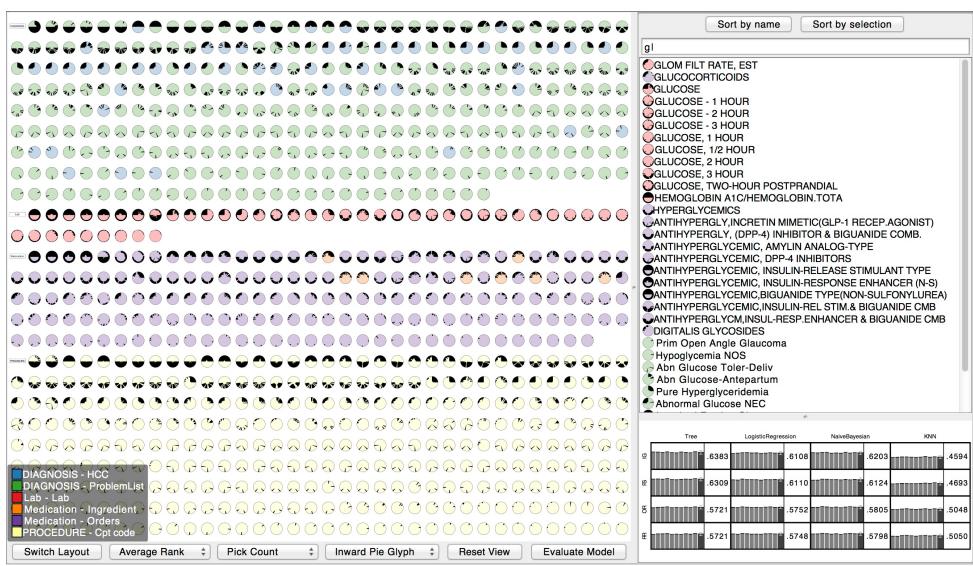
12



<https://qz.com/994486/the-way-you-draw-circles-says-a-lot-about-you/>  
<https://medium.com/analytics-vidhya/analyzing-sketches-around-the-world-with-sketch-rnn-c6cbe9b5ac80>

13

## INFUSE (INteractive FeatUre SElection)

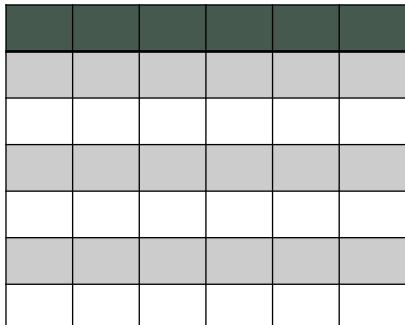


Krause, Perer, and Bertini (2014)

14

**Domain Task:** Predict if a patient is at risk of developing diabetes.

858 Features



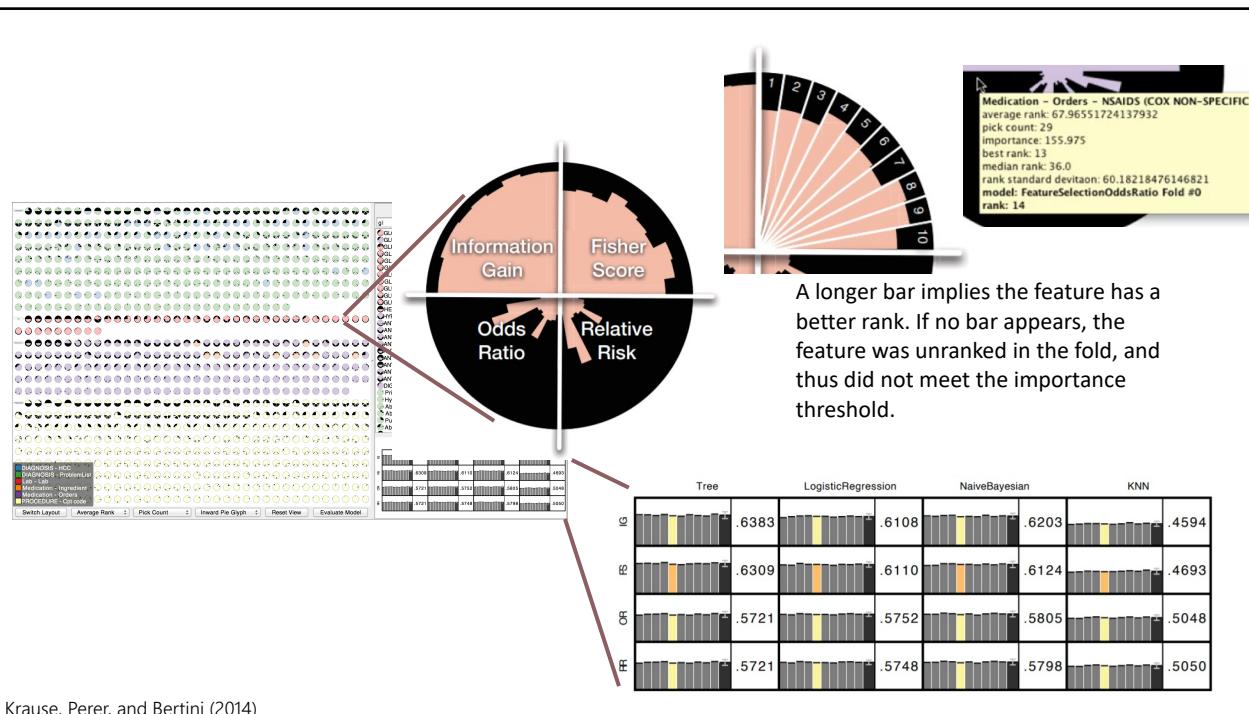
Diabetes Dataset

**ML Task 1:** Comparison of feature selection algorithms. (4 Algorithms )

**ML Task 2:** Comparison of classification algorithms. (4 classifiers)

**ML Task 3:** Manual selection and testing of feature sets

15



16

# INFUSE: Interactive Feature Selection for Predictive Modeling of High Dimensional Data

Josua Krause, Adam Perer, Enrico Bertini

VAST 2014

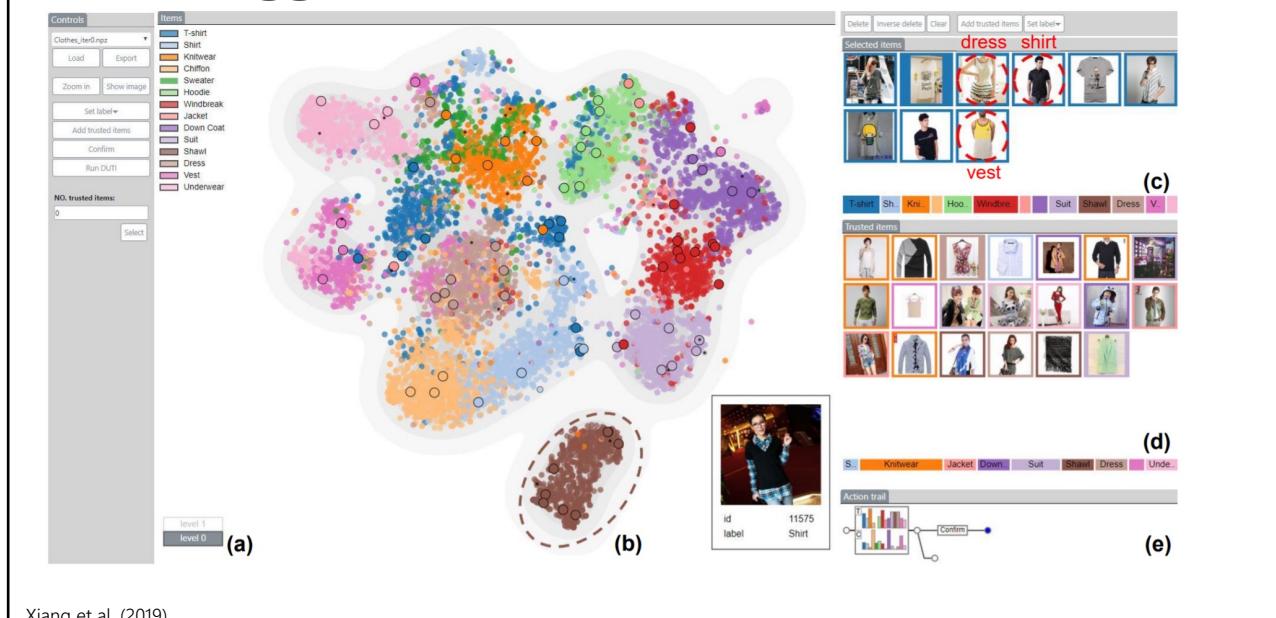
Krause, Perer, and Bertini (2014)

17

How expressive/effective is this visualization?

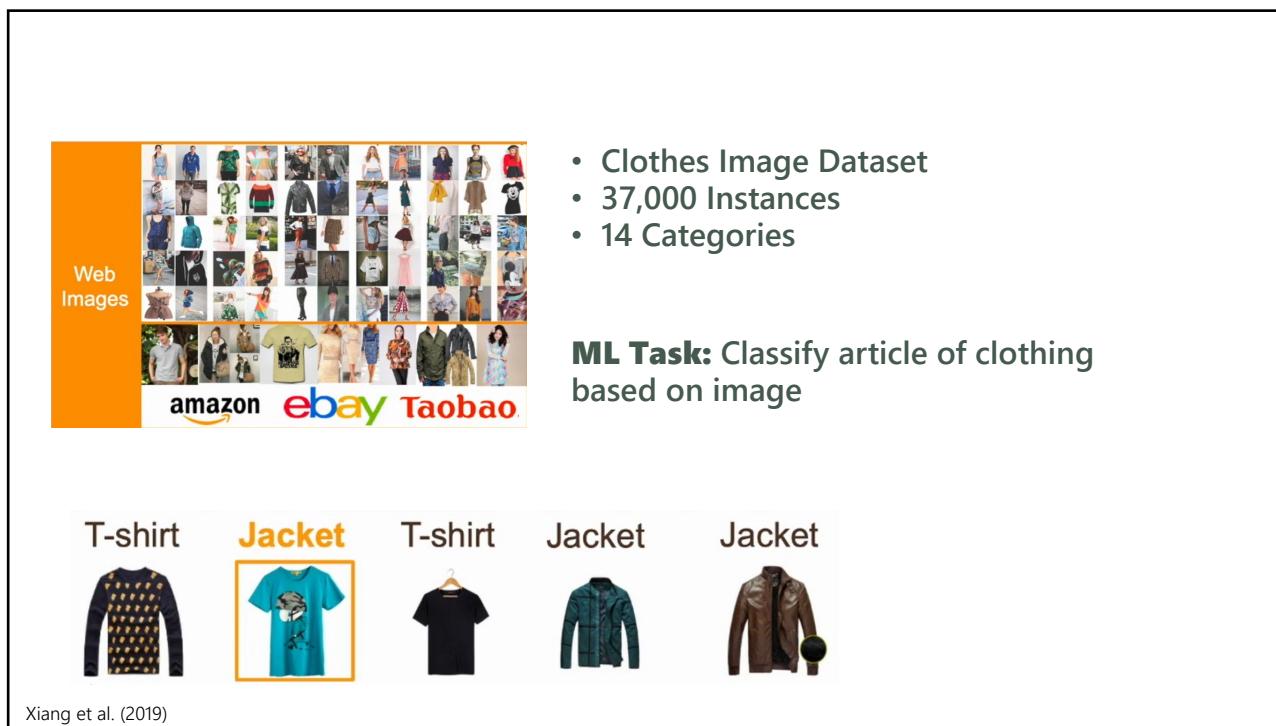
18

## DataDebugger



Xiang et al. (2019)

19



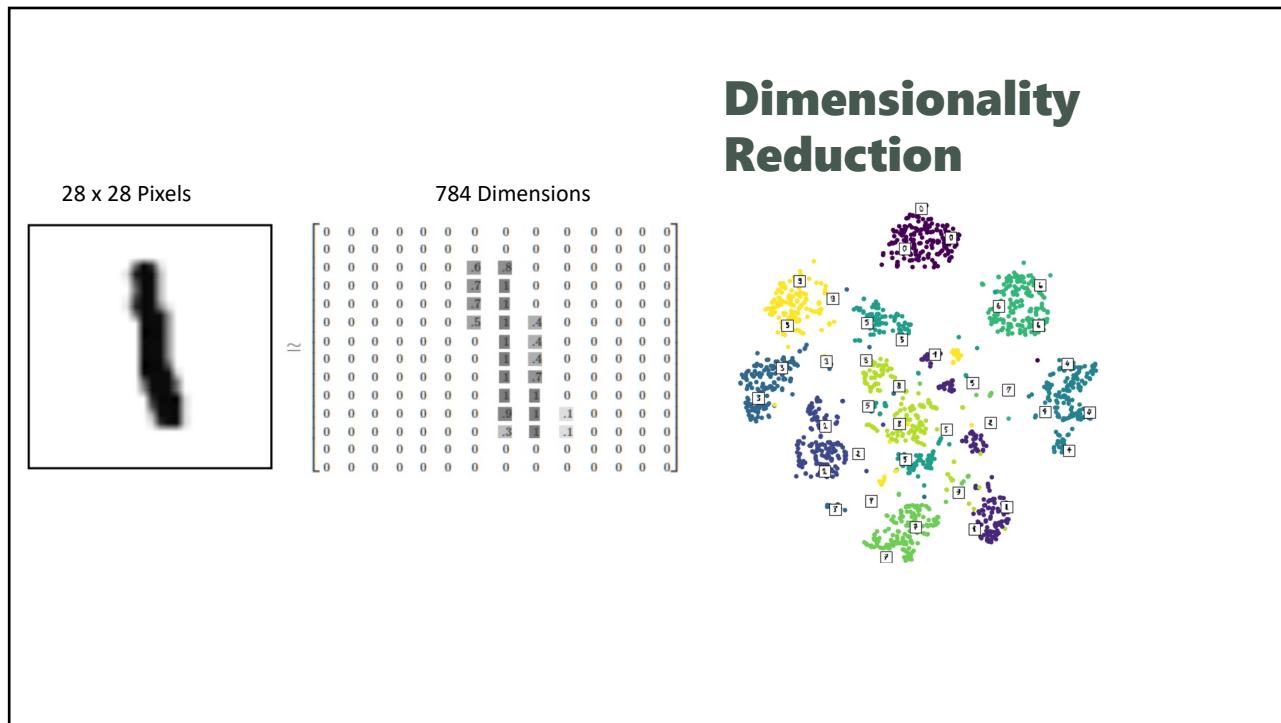
Xiang et al. (2019)

20

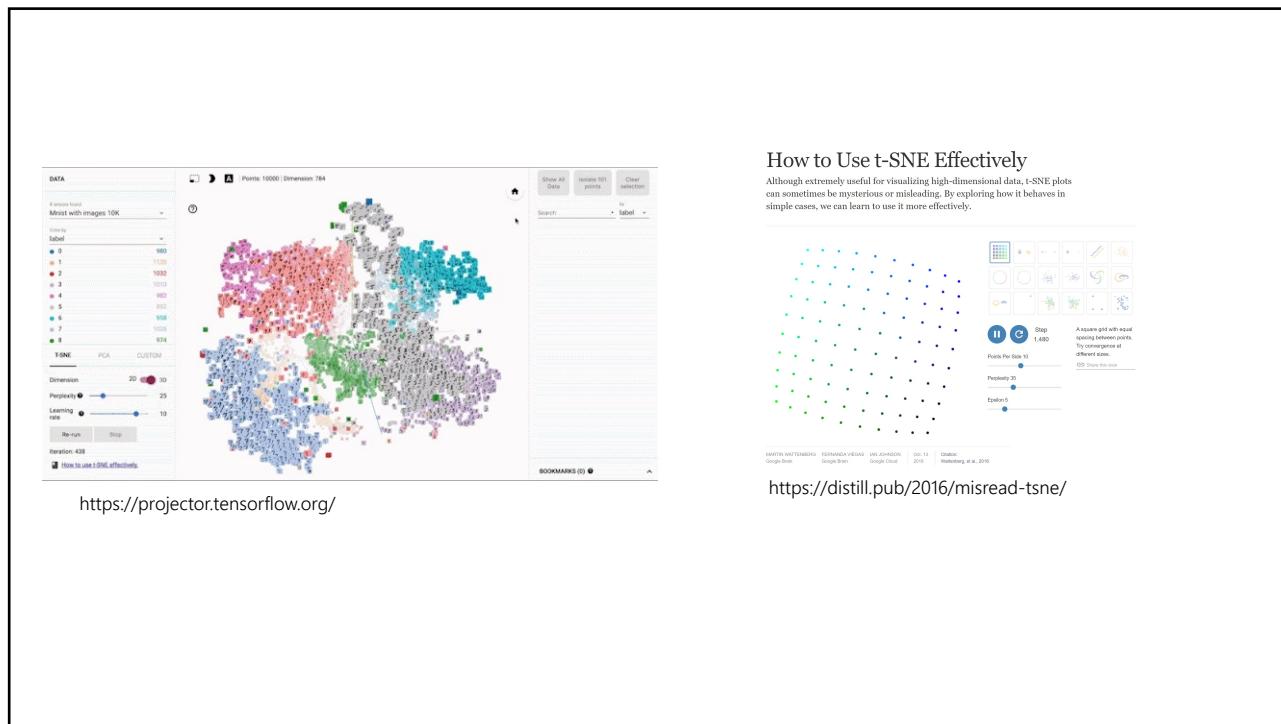
- Clothes Image Dataset
- 37,000 Instances
- 14 Categories

**ML Task:** Classify article of clothing based on image

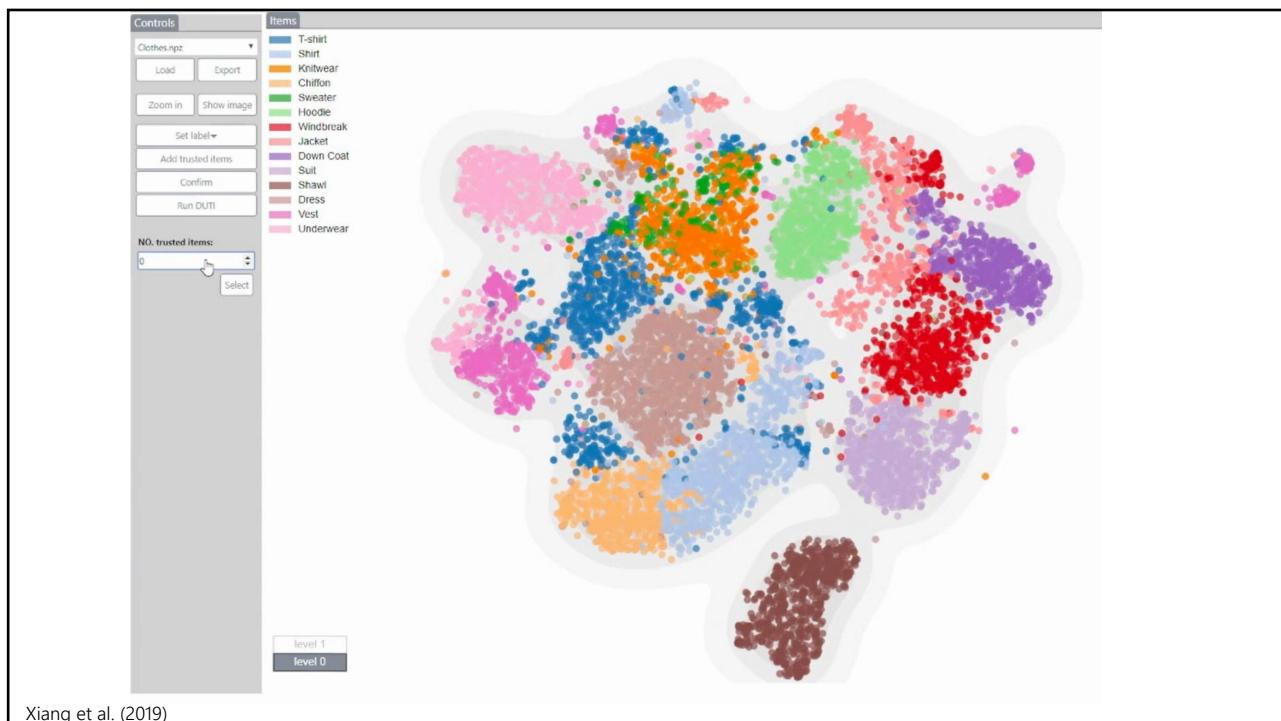




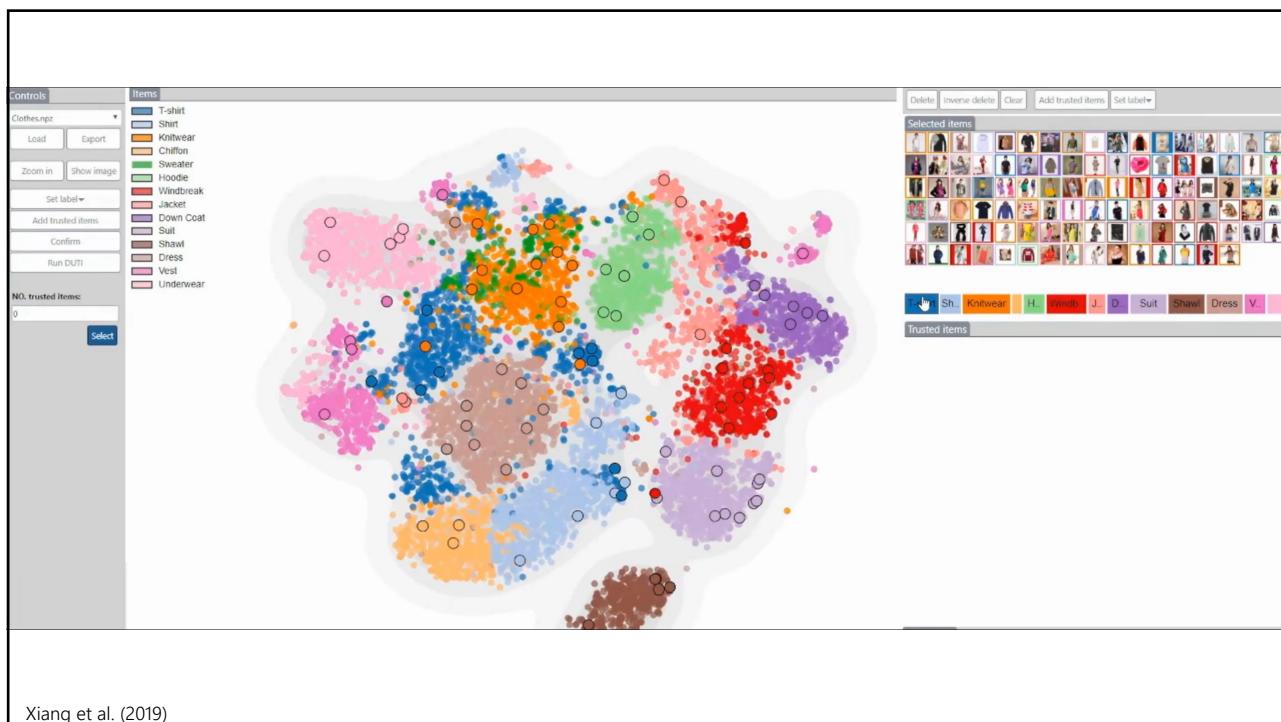
21



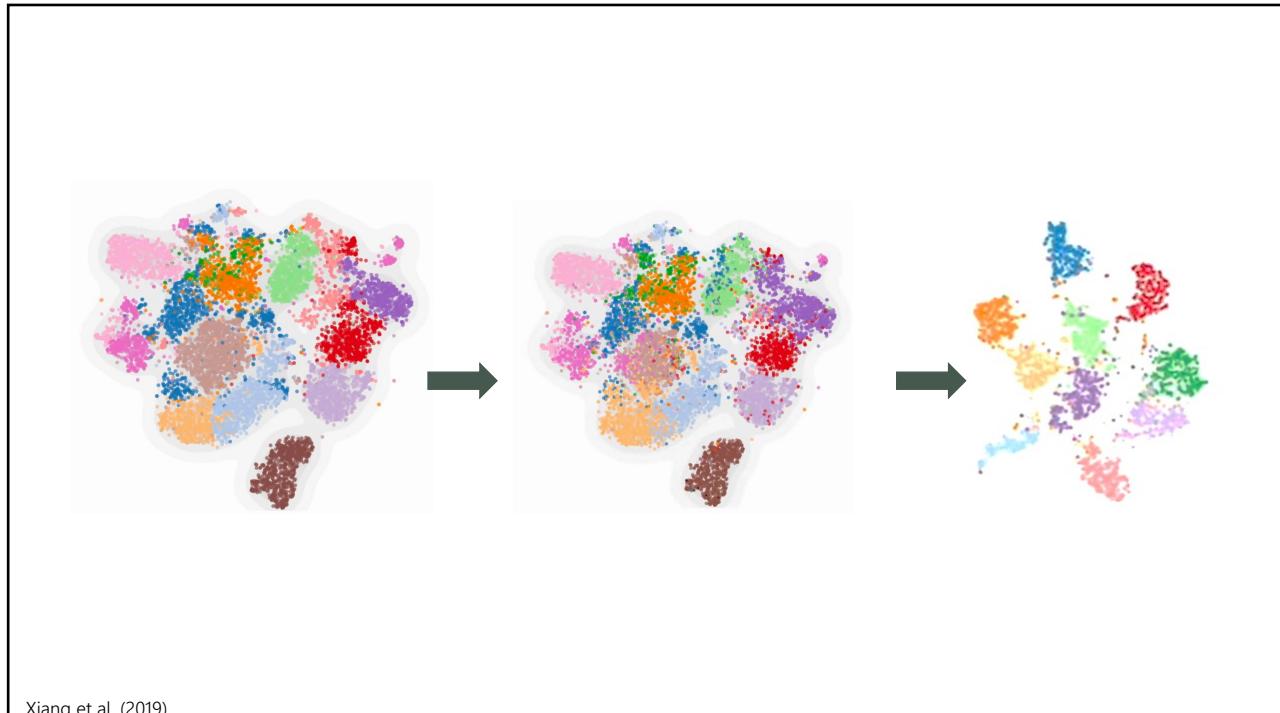
22



23



24

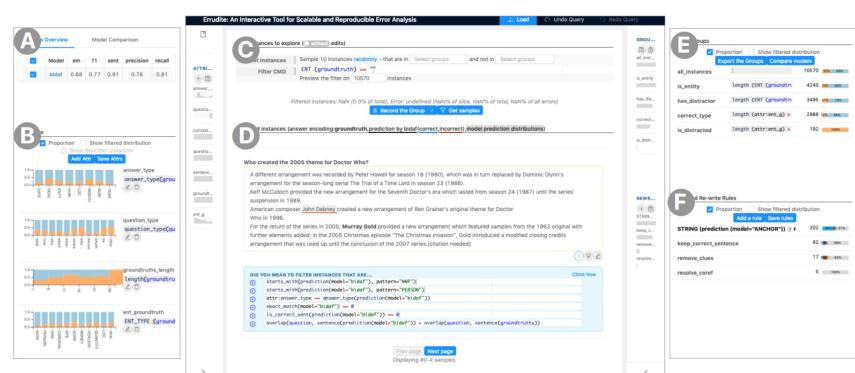


25

## Errudite

### Error Analysis by:

1. Expressive grouping of error instance
2. Counterfactual evaluation



Wu et al. (2019)

26

Find instances to explore (without edits)

Get Instances | Sample 10 instances worst predicted by bidaf that are in Select groups and not in Select groups  
Filter CMD | Preview the filter on 10570 instances

Filtered instances: 10570 (100.0% of total), Error: 3378 (32.0% of slice, 32.0% of total, 100.0% of all errors)  
Record the Group Get samples

Returned instances (answer encoding:groundtruth,prediction by bidaf(correct,incorrect),model prediction distributions)

What year did Börte's give birth to Jochi?

As previously arranged by his father, Temüjin married Börte of the Onggirat tribe when he was around 16 in order to cement alliances between their respective tribes. Soon after Börte's marriage to Temüjin, she was kidnapped by the Merkits and reportedly given away as a wife. Temüjin rescued her with the help of his friend and future rival, Jamukha, and his protector, Toghrul Khan of the Kerait tribe. She gave birth to a son, Jochi (1185–1226), nine months later, clouding the issue of his parentage. Despite speculation over Jochi, Börte would be Temüjin's only empress, though he did follow tradition by taking several morganatic wives.

When did Wei Yilin die?

The Chinese medical tradition of the Yuan had "Four Great Schools" that the Yuan inherited from the Jin dynasty. All four schools were based on the same intellectual foundation, but advocated different theoretical approaches toward medicine. Under the Mongols, the practice of Chinese medicine spread to other parts of the empire. Chinese physicians were brought along military campaigns by the Mongols as they expanded towards the west. Chinese medical techniques such as acupuncture, moxibustion, pulse diagnosis, and various herbal drugs and elixirs were transmitted westward to the Middle East and the rest of the empire. Several medical advances were made in the Yuan period. The physician Wei Yilin (1277–1347) invented a suspension method for reducing dislocated joints, which he performed using anesthetics. The Mongol physician Hu Sihui described the importance of a healthy diet in a 1330 medical treatise.

Wu et al. (2019)

## Error Instances

rrudit: An Interactive Tool for Scalable and Reproducible Error Analysis

Metrics Overview Model Comparison

Model	em	f1	sent	precision	recall
bidaf	0.68	0.77	0.91	0.78	0.81

Proportion Show filtered distribution Show Rewritten instances Add Attr Save Attrs

utes

Proportion Show filtered distribution Show Rewritten instances Add Attr Save Attrs

answer\_type answer\_type<groundtruth>

question\_type question\_type<question>

context\_length length(context)

question\_length length(question)

Find instances to explore (without edits)

Get Instances | Sample 10 instances worst predicted by bidaf that are in Select groups and not in Select groups  
Filter CMD | Preview the filter on 10570 instances

Filtered instances: 10570 (100.0% of total), Error: 3378 (32.0% of slice, 32.0% of total, 100.0% of all errors)  
Record the Group Get samples

Returned instances (answer encoding:groundtruth,prediction by bidaf(correct,incorrect),model prediction distributions)

What year did Börte's give birth to Jochi?

As previously arranged by his father, Temüjin married Börte of the Onggirat tribe when he was around 16 in order to cement alliances between their respective tribes. Soon after Börte's marriage to Temüjin, she was kidnapped by the Merkits and reportedly given away as a wife. Temüjin rescued her with the help of his friend and future rival, Jamukha, and his protector, Toghrul Khan of the Kerait tribe. She gave birth to a son, Jochi (1185–1226), nine months later, clouding the issue of his parentage. Despite speculation over Jochi, Börte would be Temüjin's only empress, though he did follow tradition by taking several morganatic wives.

When did Wei Yilin die?

The Chinese medical tradition of the Yuan had "Four Great Schools" that the Yuan inherited from the Jin dynasty. All four schools were based on the same intellectual foundation, but advocated different theoretical approaches toward medicine. Under the Mongols, the practice of Chinese medicine spread to other parts of the empire. Chinese physicians were brought along military campaigns by the Mongols as they expanded towards the west. Chinese medical techniques such as acupuncture, moxibustion, pulse diagnosis, and various herbal drugs and elixirs were transmitted westward to the Middle East and the rest of the empire. Several medical advances were made in the Yuan period. The physician Wei Yilin (1277–1347) invented a suspension method for reducing dislocated joints, which he performed using anesthetics. The Mongol physician Hu Sihui described the importance of a healthy diet in a 1330 medical treatise.

Wu et al. (2019)

Errudite: An Interactive Tool for Scalable and Reproducible Error Analysis

**Metrics Overview**

Model	em	f1	sent	precision	recall
bidaf	0.68	0.77	0.91	0.78	0.81

**Model Comparison**

**Find instances to explore (0 unapplied edits)**

Get Instances | Sample 10 instances worst predicted by bidaf - that are in Select groups and not in Select groups

Filter CMD | STRING (groundtruth) in STRING (prediction (model="ANCHOR")) and F1 (model="ANCHOR") == 0

Preview the filter on 10570 instances

Returned Instances (answer encoding groundtruth,prediction by bidaf(correct,incorrect),model prediction distributions)

**Attributes**

- Proportion
- Show filtered distribution
- Show Rewritten instances

Add Attr Save Attrs

answer\_type

question\_type

question\_type(question\_type)

context\_length

length(context)

question\_length

length(question)

**JOHN MAYOW DIED IN WHAT YEAR?**

John Mayow died in what year?  
In the late 17th century, Robert Boyle proved that air is necessary for combustion.  
English chemist John Mayow (1614-1679) refined this work by showing that fire requires only a part of air that he called spiritus nitroaereus or just nitroaereus.  
In one experiment he found that placing either a mouse or a lit candle in a closed container over water caused the water to rise and replace one-fourteenth of the air's volume before extinguishing the subjects.  
From this he surmised that nitroaereus is consumed in both respiration and combustion.

**WHAT YEAR DID BÖRTE'S GIVE BIRTH TO JOCHI?**

As previously arranged by his father, Temüjin married Börte of the Chinggis tribe when he was around 16 in order to cement alliances between their respective tribes.  
Soon after Börte's marriage to Temüjin, she was kidnapped by the Merkits and reportedly given away as a wife.  
Temüjin rescued her with the help of his friend and future rival, Jamukha, and his protector, Toghrul Khan of the Kerait tribe.  
She gave birth to a son, Jochi (1185-1226), nine months later, clouding the issue of his parentage.  
Despite speculation over Jochi, Börte would be Temüjin's only empress, though he did follow tradition by taking several morganatic wives.

**WHEN DID WEI YILIN DIE?**

Prev page Next page  
Displaying #0-10 samples.

**GROU...**

**REWR...**

Wu et al. (2019)

29

The figure displays the Errudit interface, which is designed for analyzing errors in large datasets. The top navigation bar includes 'Load', 'Undo Query', 'Redo Query', and a 'GROUP...' button. The main area is divided into several sections:

- Metrics Overview:** Shows a table of metrics for different models. The table includes columns for Model, metric, and value.
- Model Comparison:** Shows a table comparing 'Model' and 'saaa' across various metrics.
- Find Instances to explore (30000 edits):** A search interface for finding specific instances. It includes a 'Get Instances' section with a 'Filter CMD' input: `attr.count_g > 5`. Below it, a preview shows 121512 instances. A message indicates 16490 instances (13.6% of total) have errors.
- Returned instances:** Buttons for 'Correct prediction' and 'Incorrect prediction'.
- Attributes:** A section for filtering by attributes. It includes checkboxes for 'Proportion' and 'Show filtered distribution', and buttons for 'Add Attr.' and 'Save Attr.'. Three charts are shown:
  - answer\_type:** Distribution of answer types: response (~0.4), PNP (~0.3), and others (~0.3).
  - question\_type:** Distribution of question types: QP (~0.4), FQ (~0.3), and others (~0.3).
  - question\_length:** Distribution of question lengths: length(Question) (~0.4), length(Question) (~0.3), and others (~0.3).
- What color is the sheets?**: A section showing a photo of a person sitting at a desk with a red and white striped sheet. It lists ground truth colors: striped (~3), white (~2), stripes (~1), white striped (~1), multicolored (~1), white with blue, yellow, and red stripes (~1), and multi-color (~1). Filter buttons include 'saaa:blue' and 'vqcounting:white'.
- What is on the sandwiches?**: A section showing a photo of two sandwiches on a plate. It lists ground truth ingredients: meat, cheese, lettuce (~1), tomato and lettuce (~1), lettuce, cheese and tomato (~1), ground truth vegetables (~1), ham, lettuce, tomatoes (~1), ground truth lettuce, tomatoes, cabbage (~1), ground truth meat tomato (~1), ground truth bh (~1), ground truth ham and lettuce (~1), and avocado (~1). Filter buttons include 'saaa:lettuce' and 'vqcounting:onions'.

Wu et al. (2019)

30

## Programming by Demonstration

American composer [John Debney](#) created a new arrangement of Ron Grainer's original theme for Doctor Who in 1996.

For the return of the series in 2005, **Murray Gold** provided a new arrangement which featured samples from the 1963 original with further elements added; in the 2005 Christmas episode "The Christmas Invasion", Gold introduced a modified closing credits arrangement that was used up until the conclusion of the 2007 series.[citation needed]

```
DID YOU MEAN TO FILTER INSTANCES THAT ARE...
⊕ starts_with(prediction(model="bidaf"), pattern="NNP")
⊕ starts_with(prediction(model="bidaf"), pattern="PERSON")
⊕ attr:answer_type == answer_type(prediction(model="bidaf"))
⊕ exact_match(model="bidaf") == 0
⊕ is_correct_sent(prediction(model="bidaf")) == 0
⊕ overlap(question, sentence(prediction(model="bidaf"))) > overlap(question, sentence(groundtruths))
```

Wu et al. (2019)

31

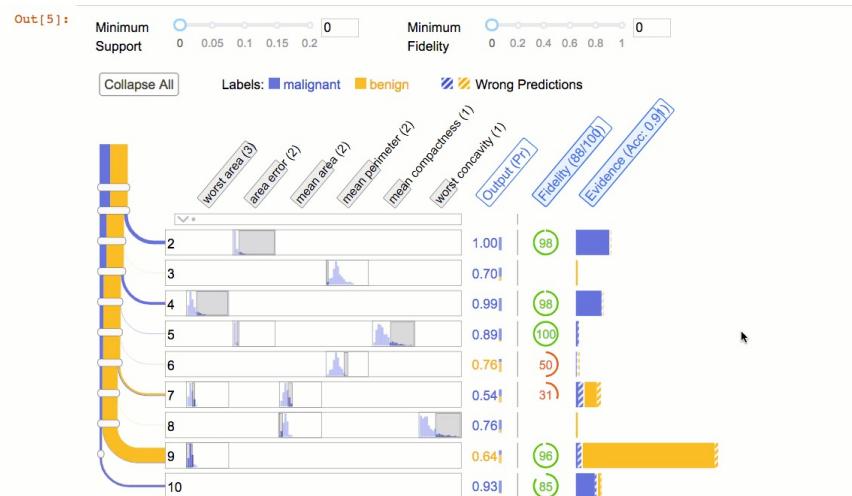
How does this approach compare to conventional GUI input elements ?

32

# Visualizations for Modeling

33

## RuleMatrix



Ming, Qu, &amp; Bertini (2019)

34

## Domain Questions

What knowledge has the model learned?

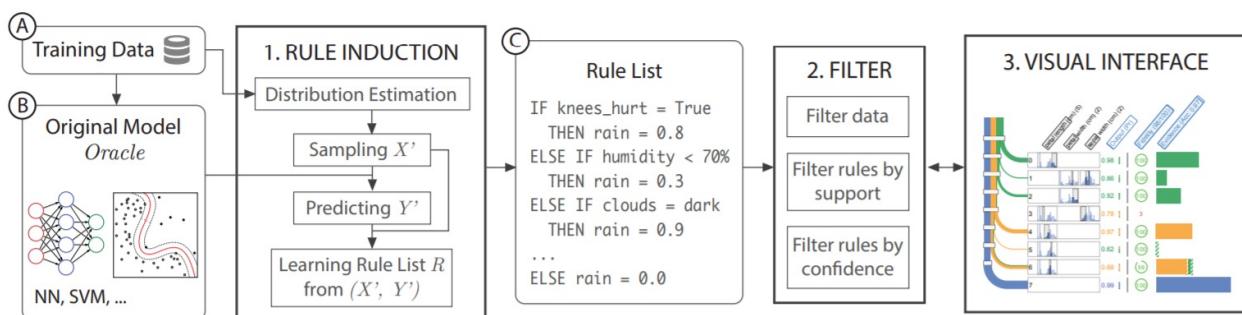
How certain is the model for each piece of knowledge?

What knowledge does the model utilize to make a prediction?

When and where is the model likely to fail?

Ming, Qu, & Bertini (2019)

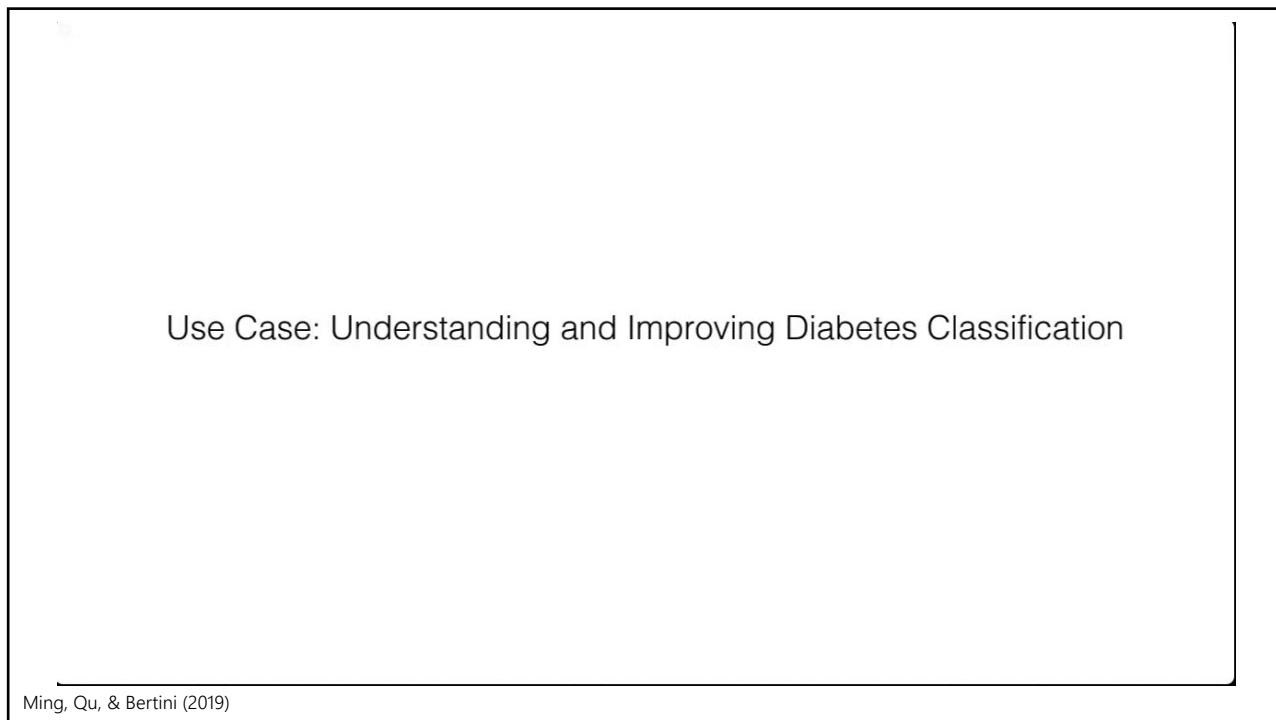
35



36



37



38

## SMILY (Similar Medical Images Like Yours)

refine-by-region      refine-by example      refine-by concept



Cai et al. (2019)

39

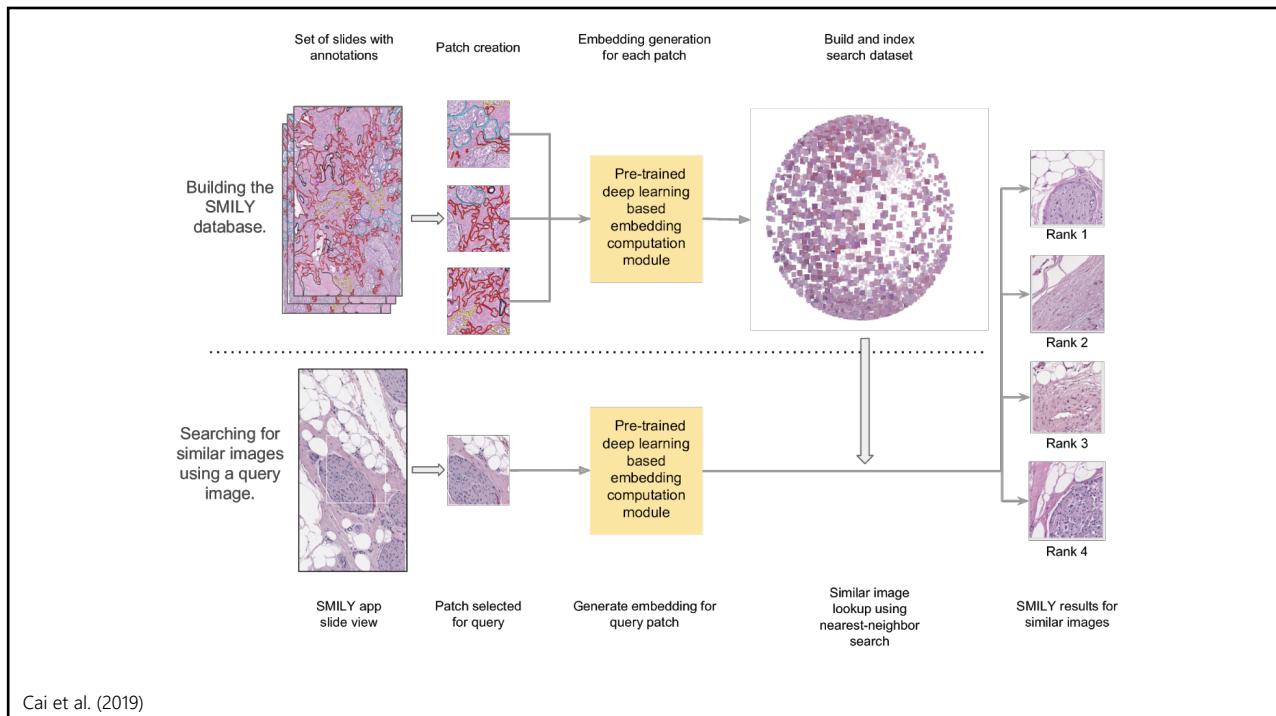
## Domain Task

Pathologists need to retrieve visually similar medical images from past patients (e.g. tissue from biopsies) to reference when making a medical decision with a new patient.

Control which types of similarity matter

Cai et al. (2019)

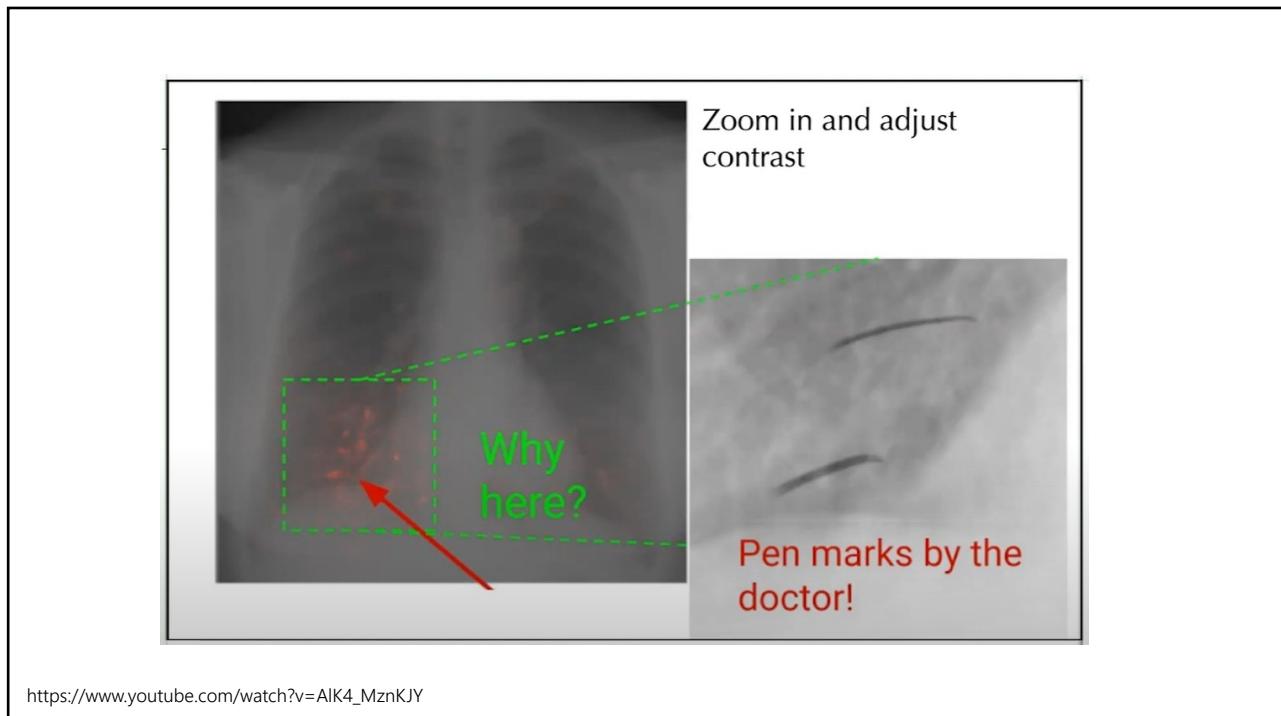
40



41



42

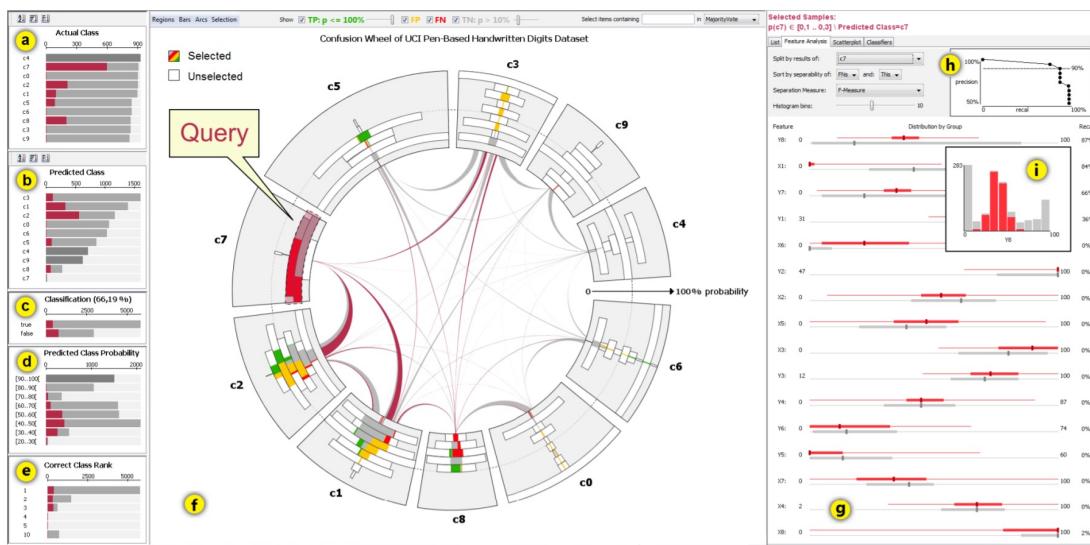


43

## Visualizations for Model Evaluation

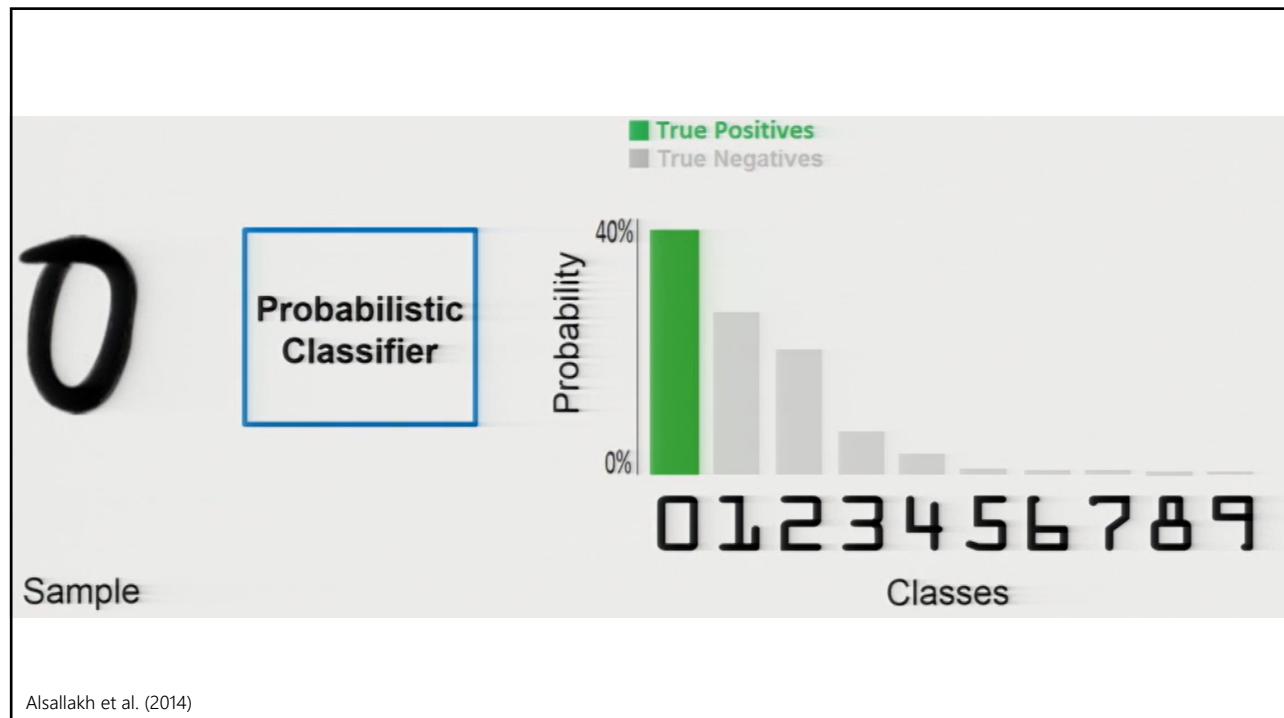
44

## Confusion Wheel

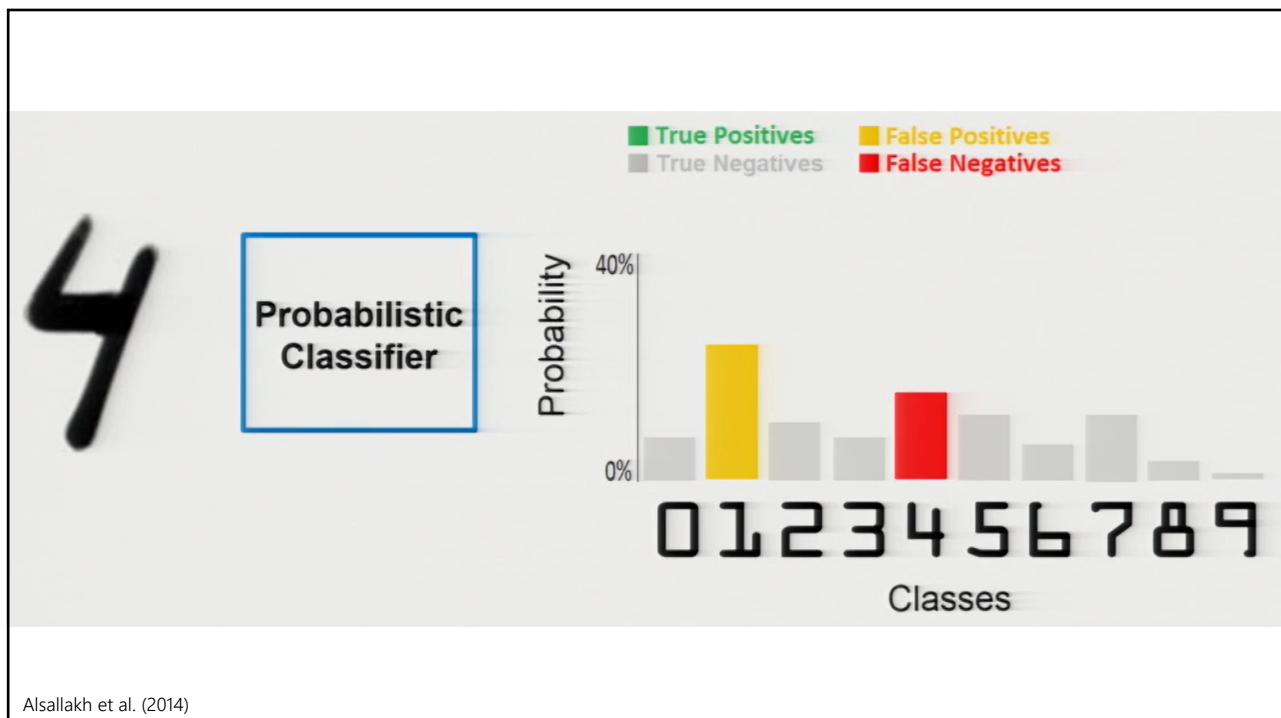


Alsallakh et al. (2014)

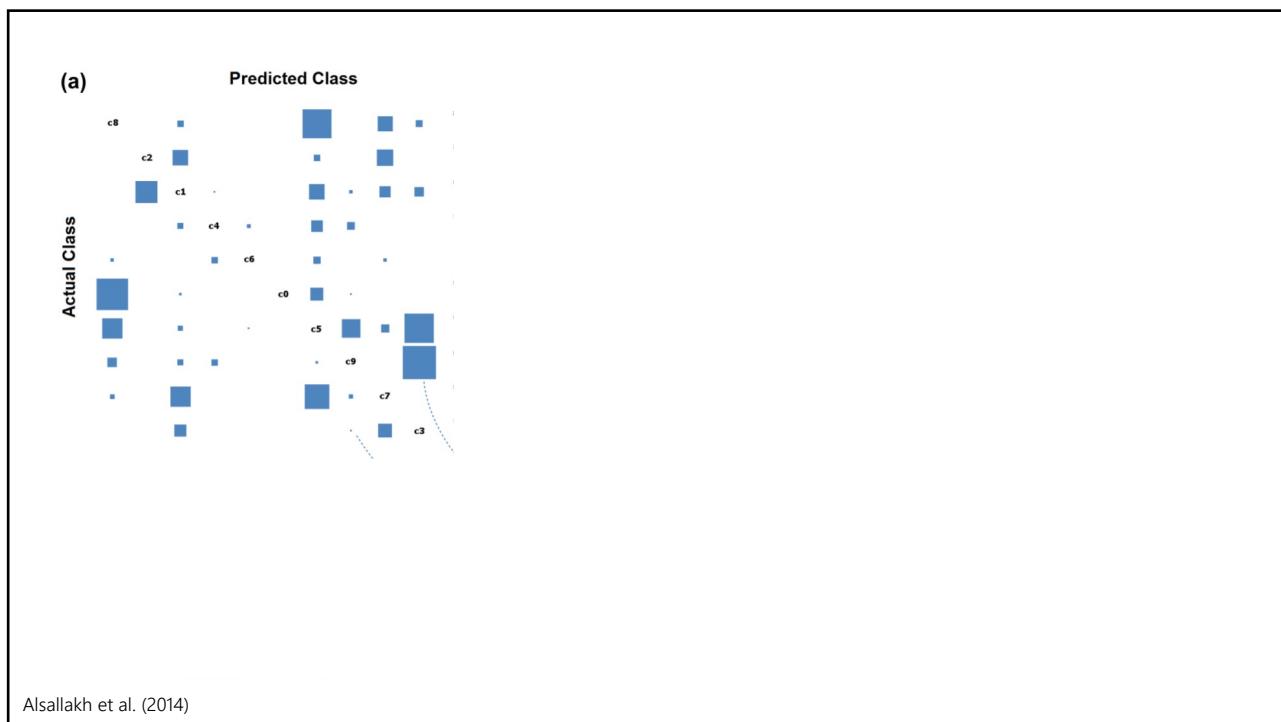
45



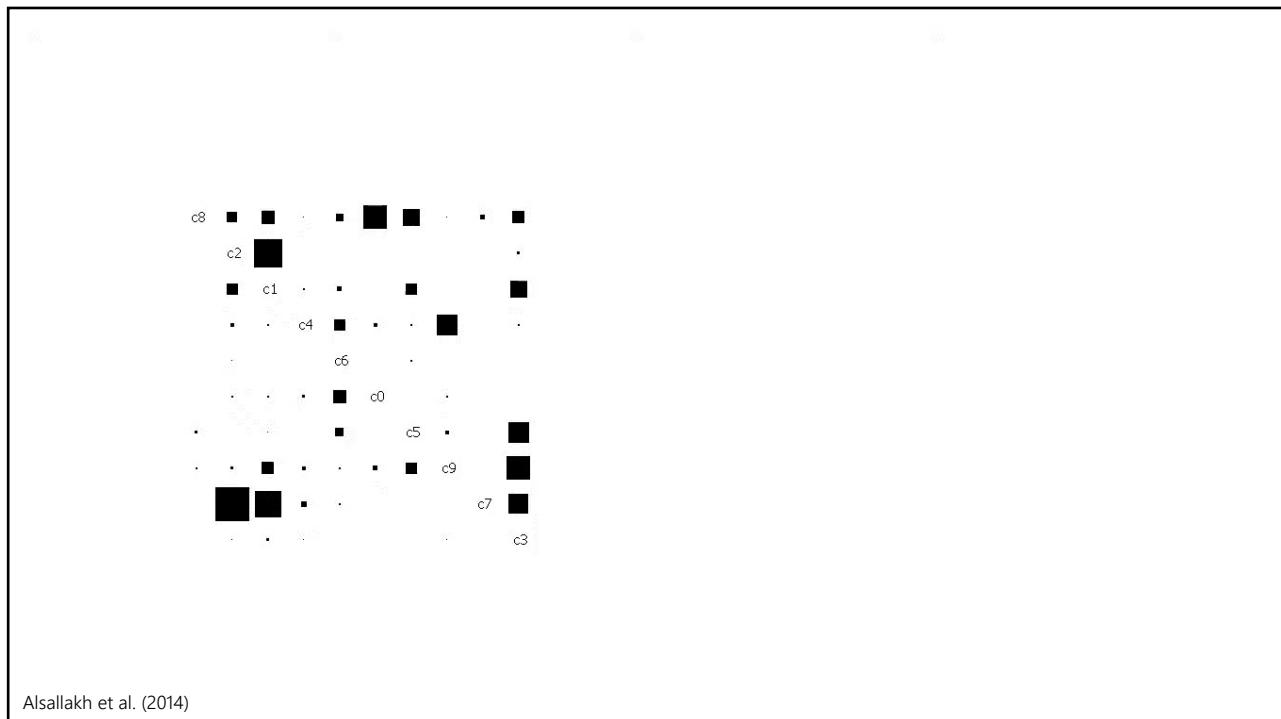
46



47



48

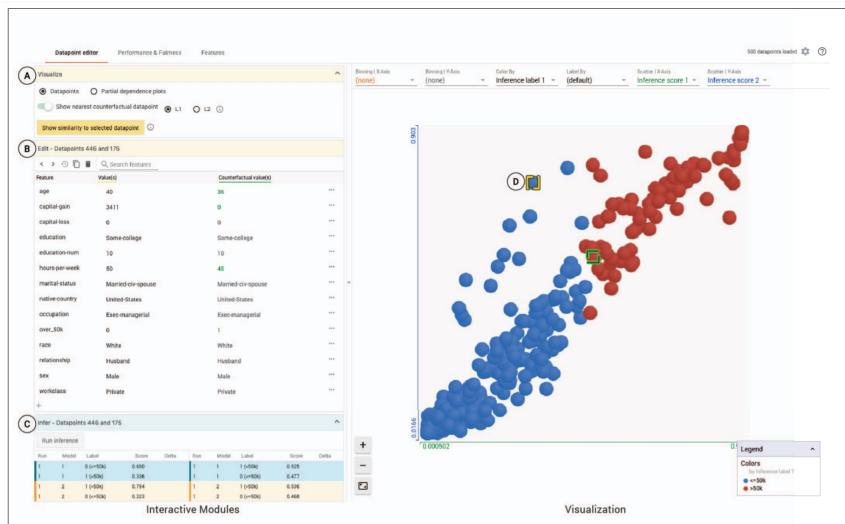


49

In what other ways can we visualize this data?

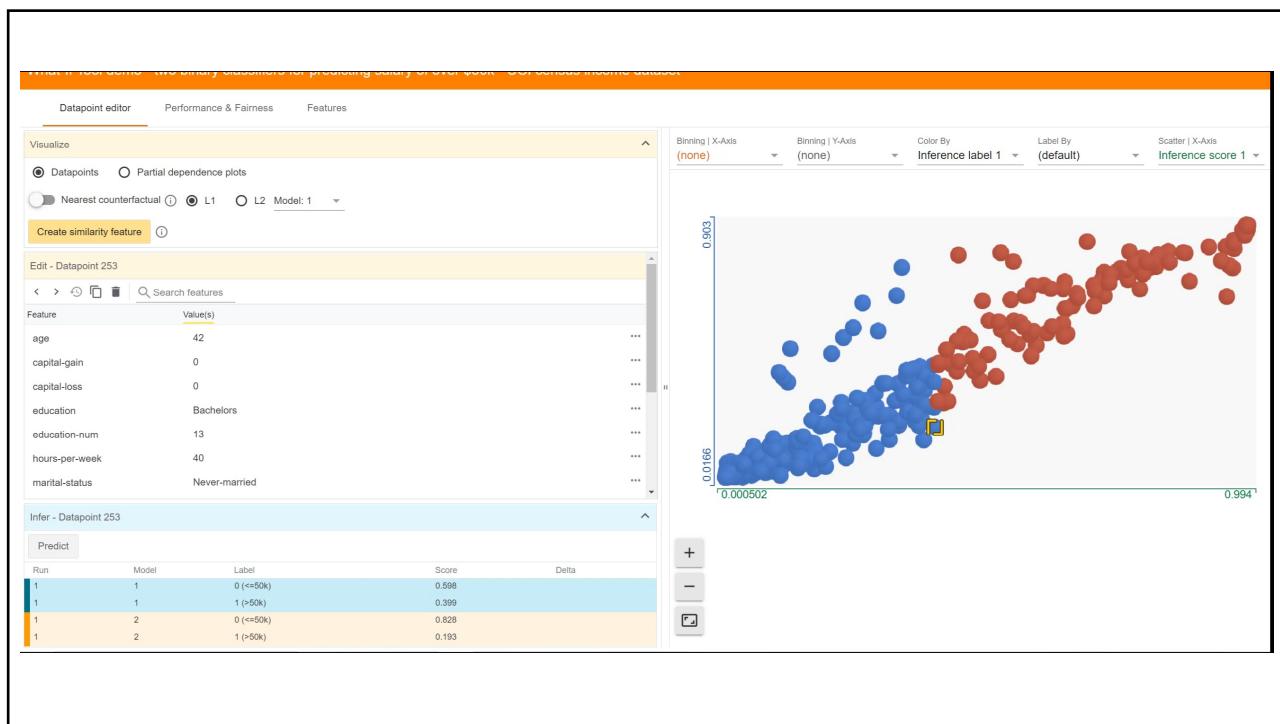
50

## The What-If Tool



Wexler et al. (2019)

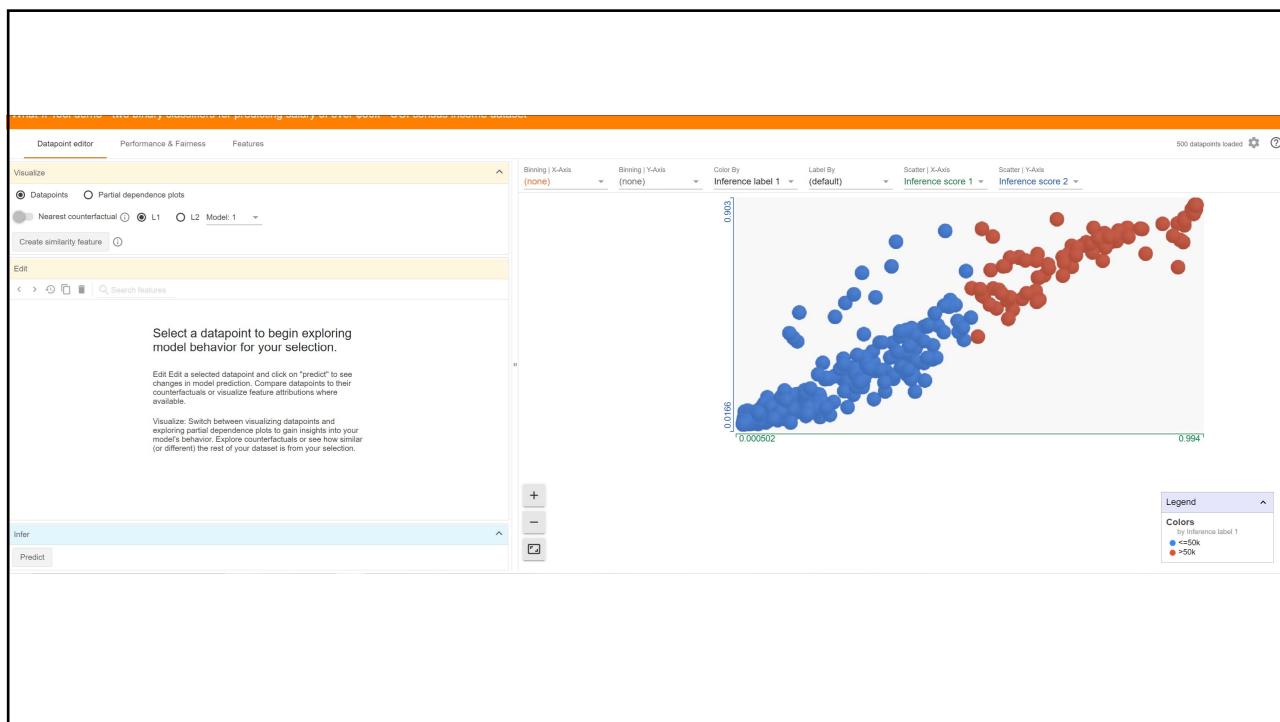
51



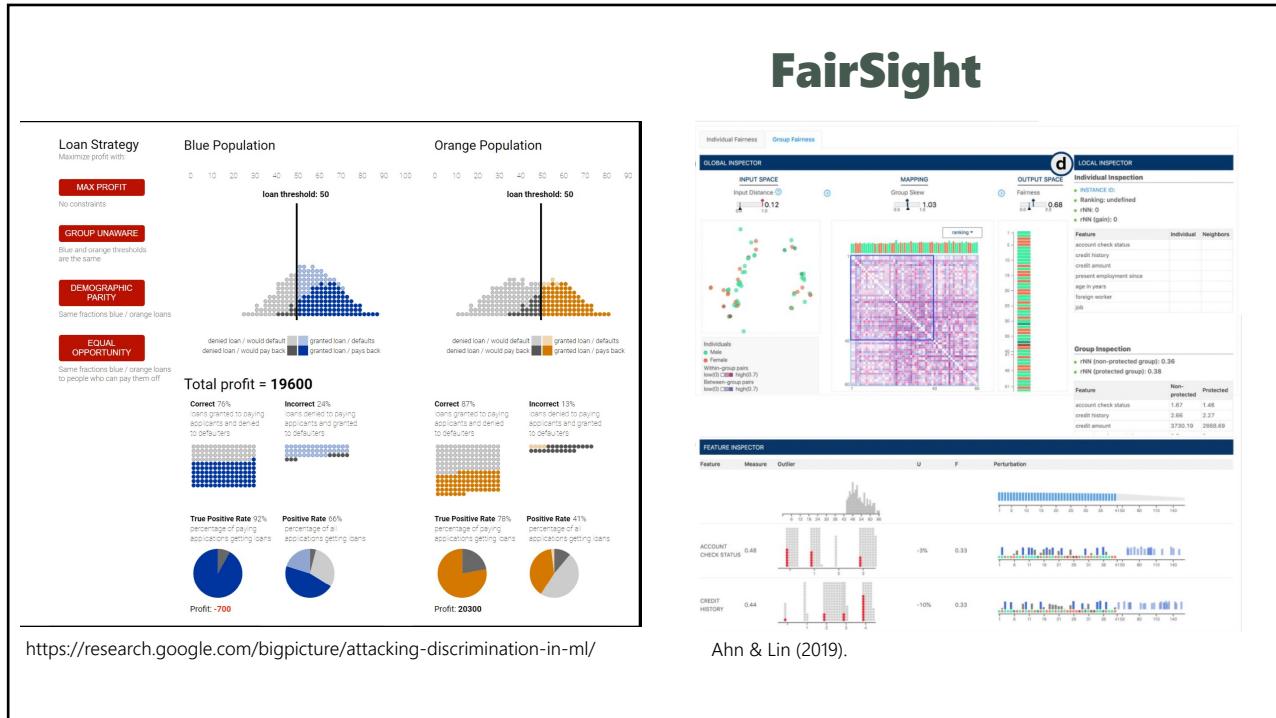
52



53



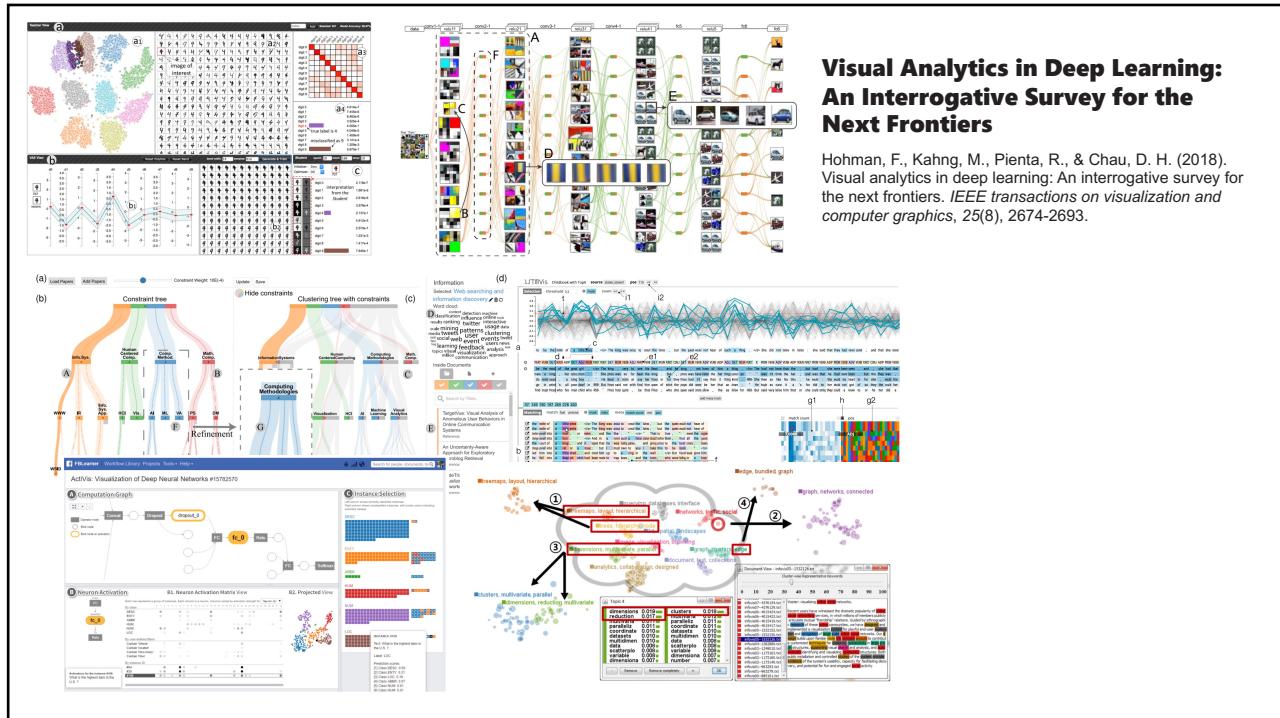
54



55

What is different/unique about visualizing ML data?

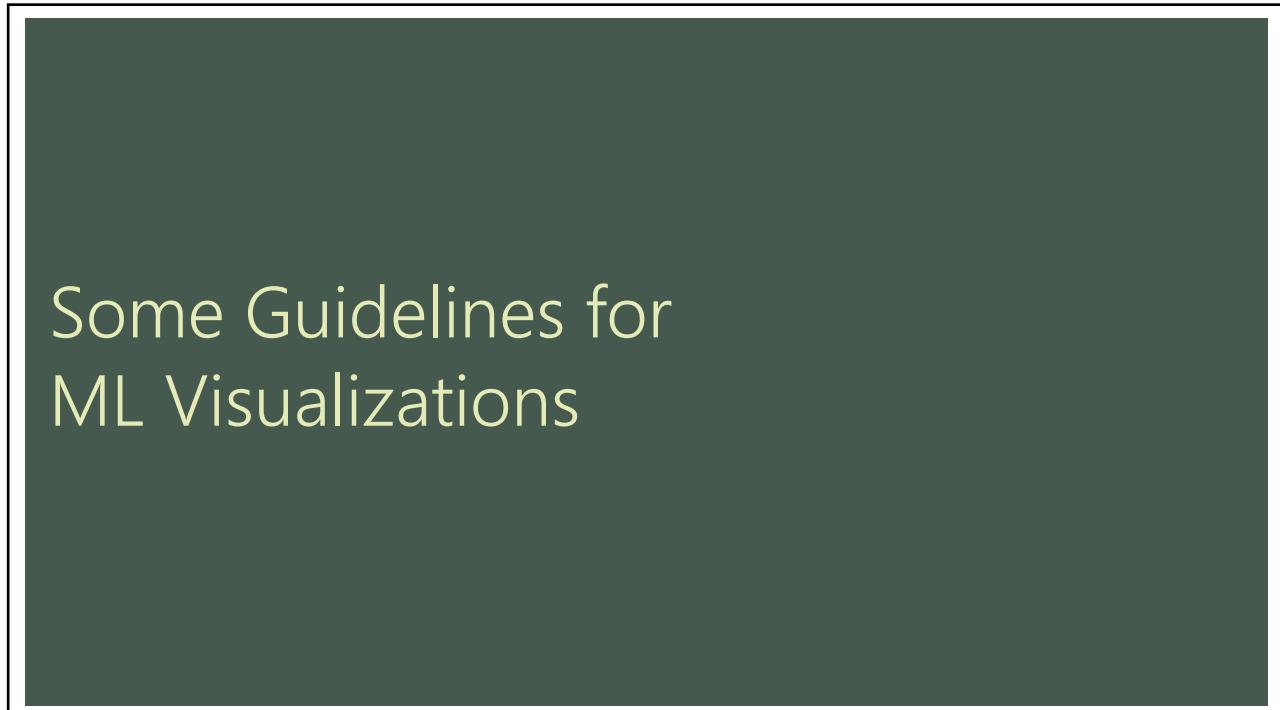
56



57

## Visual Analytics in Deep Learning: An Interrogative Survey for the Next Frontiers

Hohman, F., Kahng, M., Pienta, R., & Chau, D. H. (2018). Visual analytics in deep learning: An interrogative survey for the next frontiers. *IEEE transactions on visualization and computer graphics*, 25(8), 2674-2693.

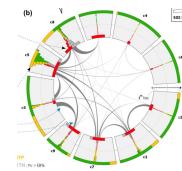


58

1. Visualizations should align with user expertise
2. Provide effective data representations for the task
3. Support understanding of model uncertainty
4. Exploit interactivity and promote rich interactions
5. Support expressive inputs

- Model Developers and Builders
- Model Users
- Domain Experts
- Non-Experts
- Learners/Students

Confusion matrix			
	Predicted Yes	Predicted No	Total
Actual Yes	13.2% (66)	9.8% (49)	23.0% (115)
Actual No	5.2% (26)	71.8% (359)	77.0% (385)
Total	18.4% (92)	81.6% (408)	
2	Predicted Yes	Predicted No	Total
Actual Yes	14.0% (70)	9.0% (45)	23.0% (115)
Actual No	5.4% (27)	71.6% (350)	77.0% (385)
Total	19.4% (97)	80.6% (403)	

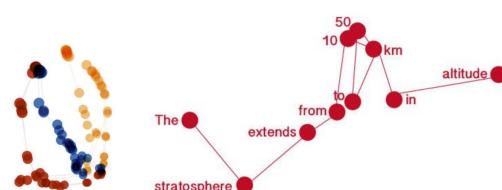


Dudley & Kristensson (2018)

59

1. Visualizations should align with user expertise
2. Provide effective data representations for the task
3. Support understanding of model uncertainty
4. Exploit interactivity and promote rich interactions
5. Support expressive inputs

- Debugging and Improving Models
- Comparing and Selecting Models
- Interpretability and Explainability
- Teaching ML Concepts



"You mean to imply that I have nothing contrary. I can supply you with every dinner parties," warmly replied Chich spoke to prove his own rectitude and animated by the same desire.

Kutuzov, shrugging his shoulders, rep smile: "I meant merely to say what I

Dudley & Kristensson (2018)

60

1. Visualizations should align with user expertise
2. Provide effective data representations for the task
3. Support understanding of model uncertainty
4. Exploit interactivity and promote rich interactions
5. Support expressive inputs

Uncertainty is an inevitable feature of data-driven models in most real-world applications.



**Jessica Hullman**  
Northwestern University  
**Theories of inference for visual analysis**  
December 3, 2021



Dudley & Kristensson (2018)

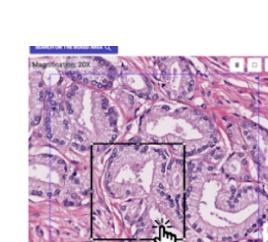
61

1. Visualizations should align with user expertise
2. Provide effective data representations for the task
3. Support understanding of model uncertainty
4. Exploit interactivity and promote rich interactions
  - Editing data points
  - Evaluating Hypotheses
  - Constructing Explanations
5. Support expressive inputs

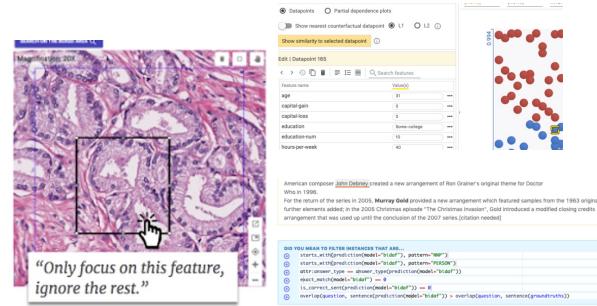


Dudley & Kristensson (2018)

62

1. Visualizations should align with user expertise
  2. Provide effective data representations for the task
  3. Support understanding of model uncertainty
  4. Exploit interactivity and promote rich interactions
  5. Support expressive inputs
  - Direct Manipulation
  - Query-by-demonstration

"Only focus on this feature, ignore the rest."



Dudley & Kristensson (2018)

63

## **Additional Resources**

# A visual introduction to machine learning

<http://www.r2d3.us/visual-intro-to-machine-learning-part-1/>

## How to Use t-SNE Effectively

Although extremely useful for visualizing high-dimensional data, t-SNE plots can sometimes be mysterious or misleading. By exploring how it behaves in simple cases, we can learn to use it more effectively.

<https://distill.pub/2016/misread-tsne/>

# The Building Blocks of Interpretability

Interpretability techniques are normally studied in isolation

We explore the powerful interfaces that arise when you combine them—and the rich structure of this combinatorial space.

<https://distill.pub/2018/building-blocks/>

## But what is a Neural Network?

An overview of what a neural network is, introduced in the context of recognizing hand-written digits.

<https://www.3blue1brown.com/topics/neural-networks>

# **Visual Analytics in Deep Learning: An Interrogative Survey for the Next Frontiers**

Hohman, F., Kahng, M., Pienta, R., & Chau, D. H. (2018). Visual analytics in deep learning: An interrogative survey for the next frontiers. *IEEE transactions on visualization and computer graphics*, 25(8), 2674-2693.