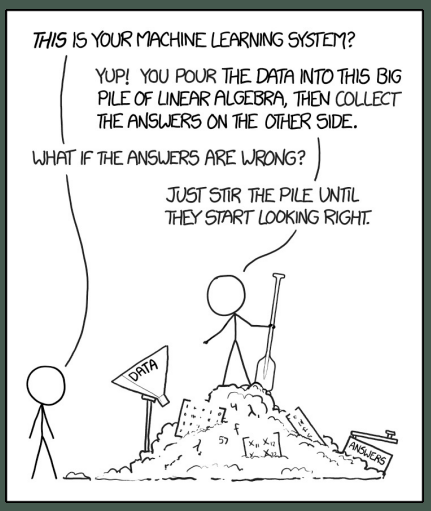
VISUALIZATIONS FOR MACHINE LEARNING

(An Introduction)

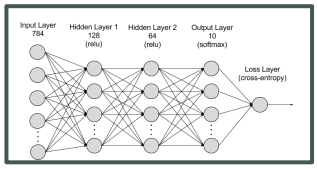
HARI SUBRAMONYAM



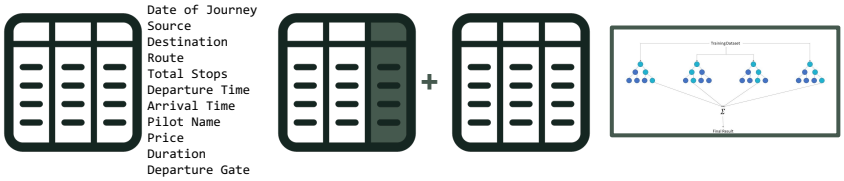
https://xkcd.com/1838/

1

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



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



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 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?

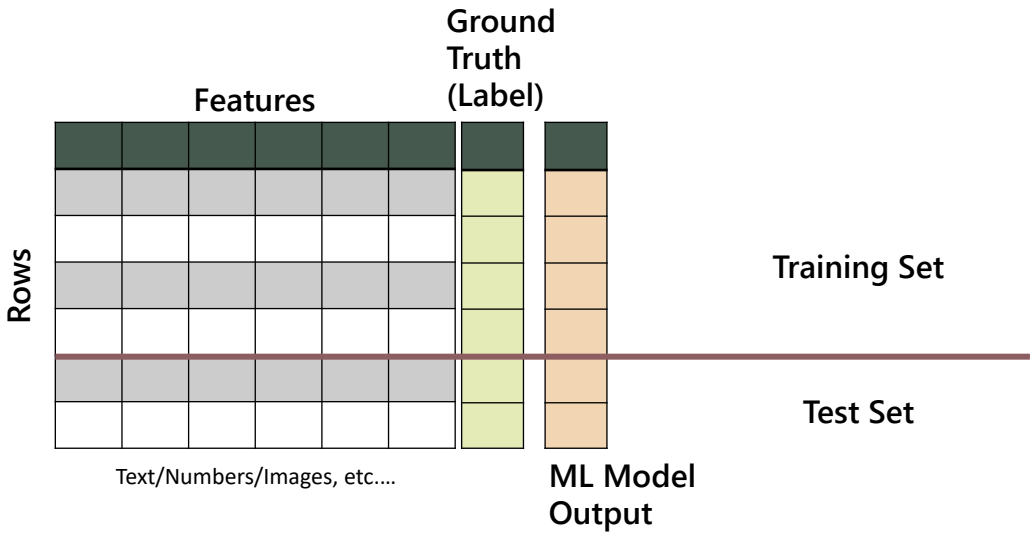
How can I explain the model output?

Is the model fair?

4

Visualizations for Understanding ML Datasets

5

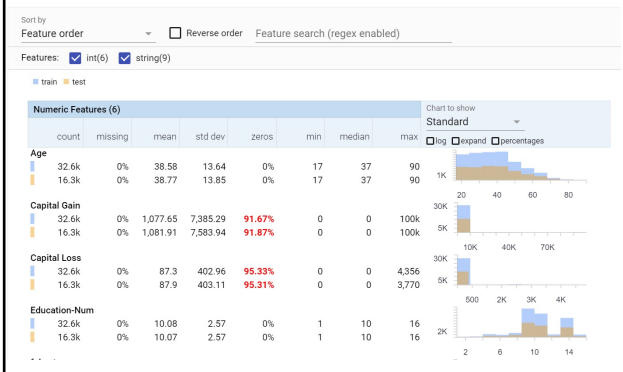


6

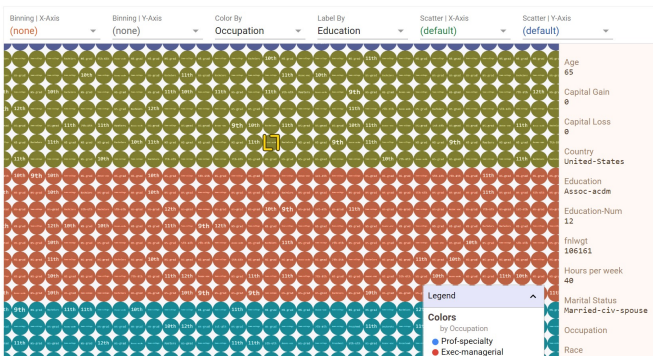
- Understanding Data Characteristics before Modeling
- Selecting Features for Modeling
- Debugging Data based on Model Outputs

7

FACETS



Overview



Dive

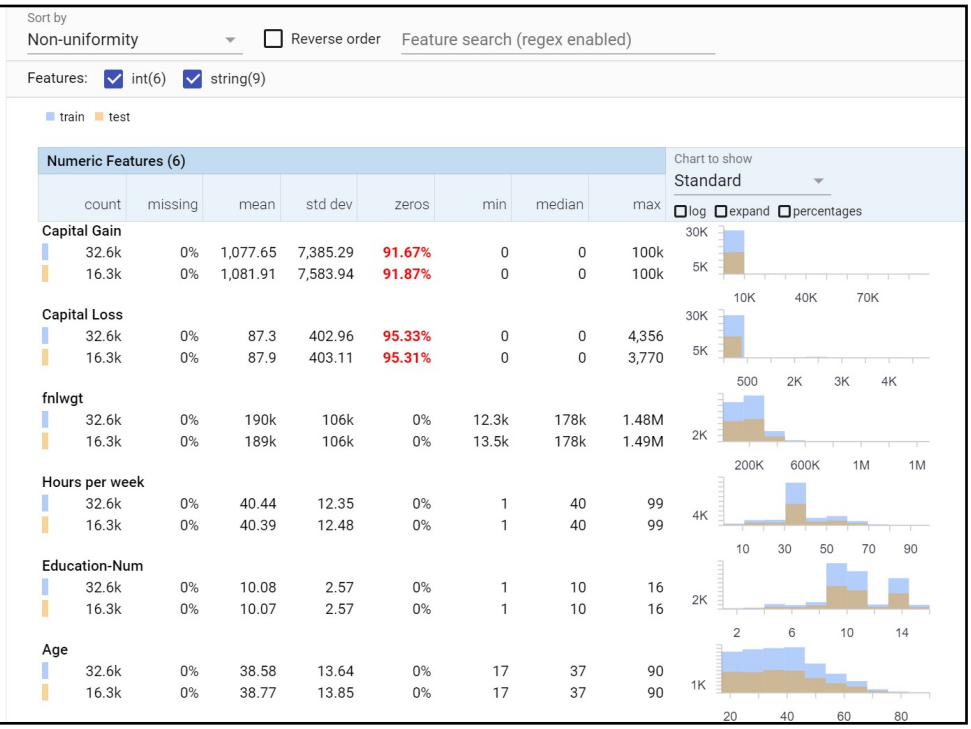
https://pair-code.github.io/facets/

8

- 1994 Census dataset
- ~50k rows
- 14 Attributes (Categorical and Integer)

Prediction task is to determine whether a person makes over 50K a year.

http://archive.ics.uci.edu/ml/datasets/Census+Income

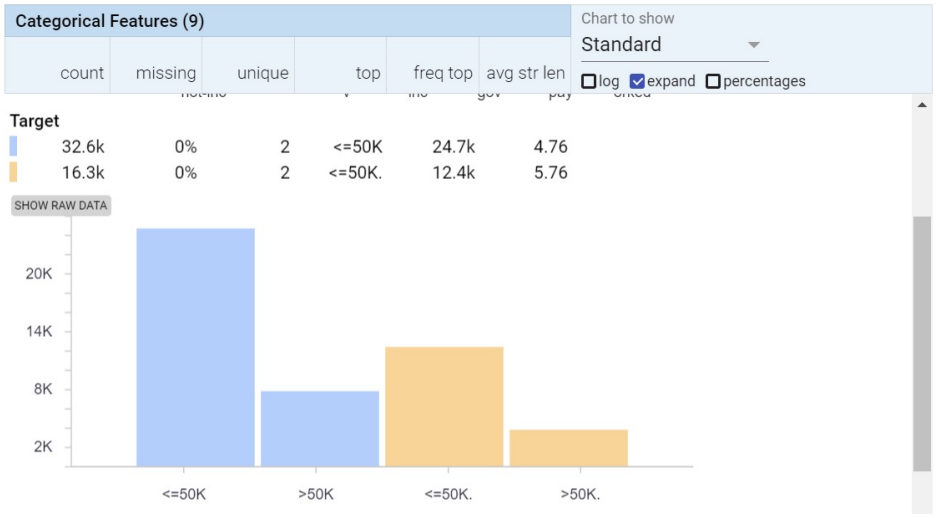


9

- 1994 Census dataset
- ~50k rows
- 14 Attributes (Categorical and Integer)

Prediction task is to determine whether a person makes over 50K a year.

http://archive.ics.uci.edu/ml/datasets/Census+Income



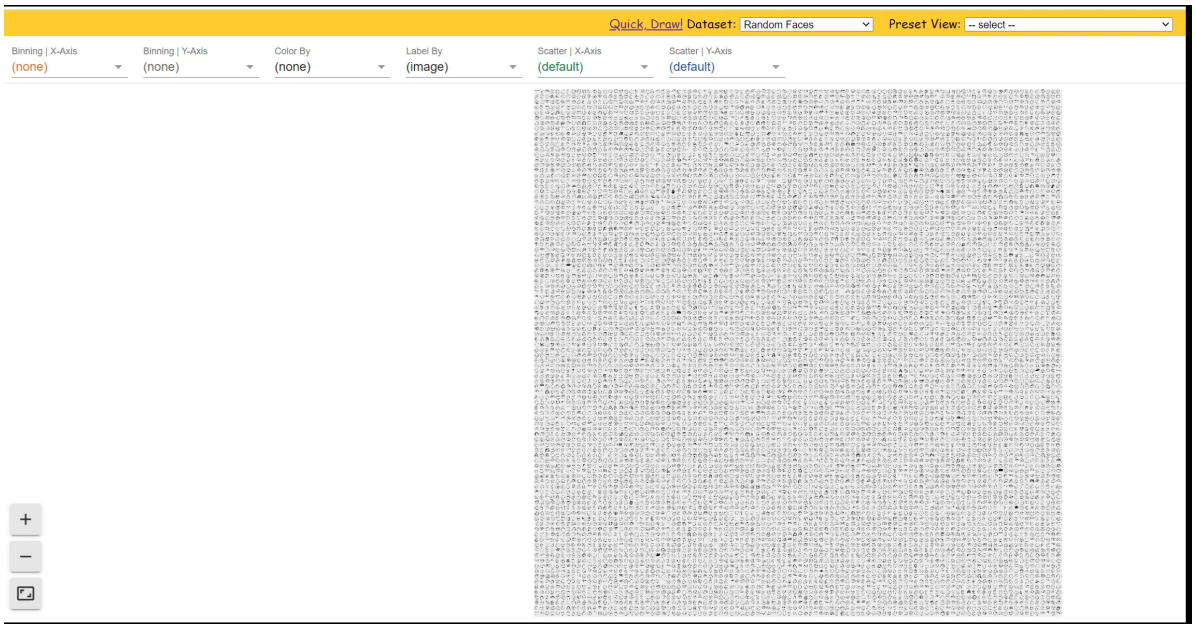
10

- 1994 Census dataset
- ~50k rows
- 14 Attributes (Categorical and Integer)

Prediction task is to determine whether a person makes over 50K a year.

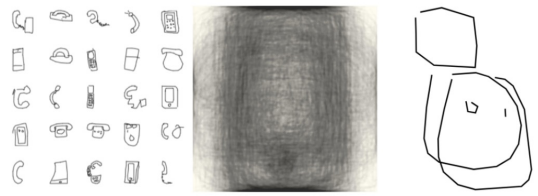
http://archive.ics.uci.edu/ml/datasets/Census+Income

11

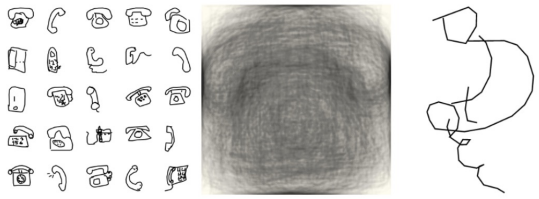


https://pair-code.github.io/facets/quickdraw.html

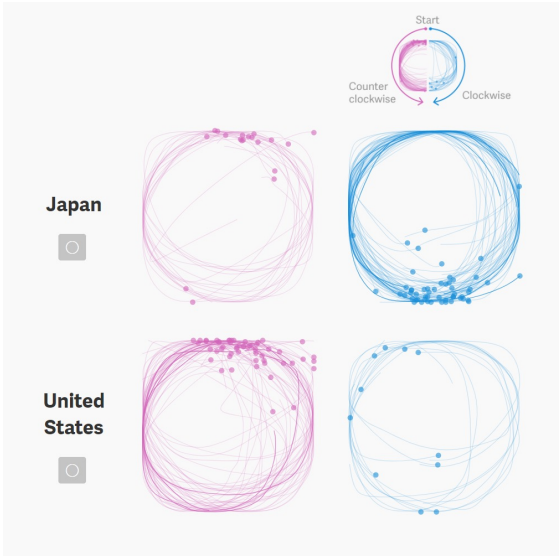
12



Randomly sampled sketches (left), pixel overlay (center) and sketch-rnn average (right) of telephone drawings from around the world.



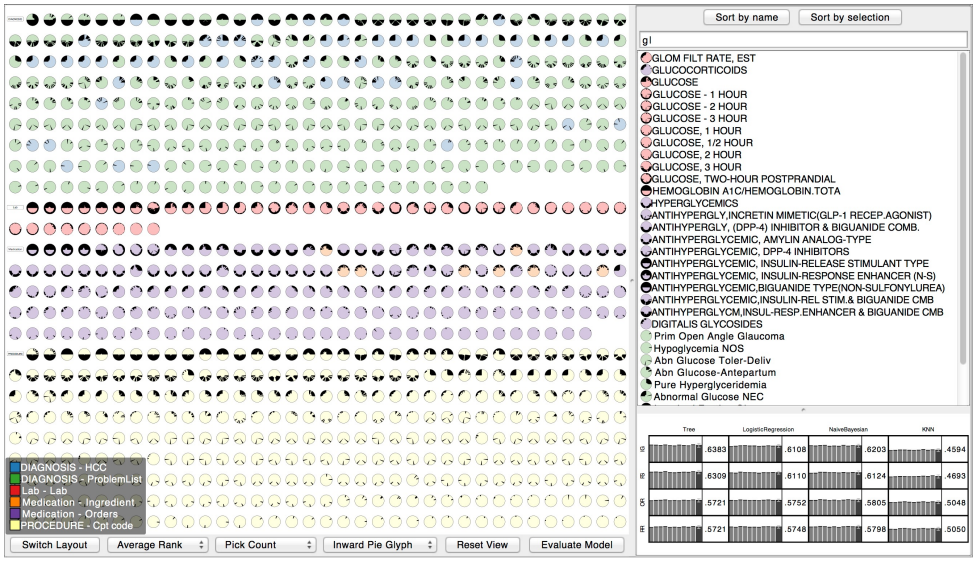
Randomly sampled sketches (left), pixel overlay (center) and sketch-rnn average (right) of telephone drawings from India only.



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)

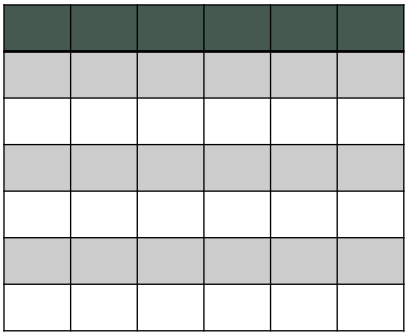


Krause, Perer, and Bertini (2014)

14

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

858 Features



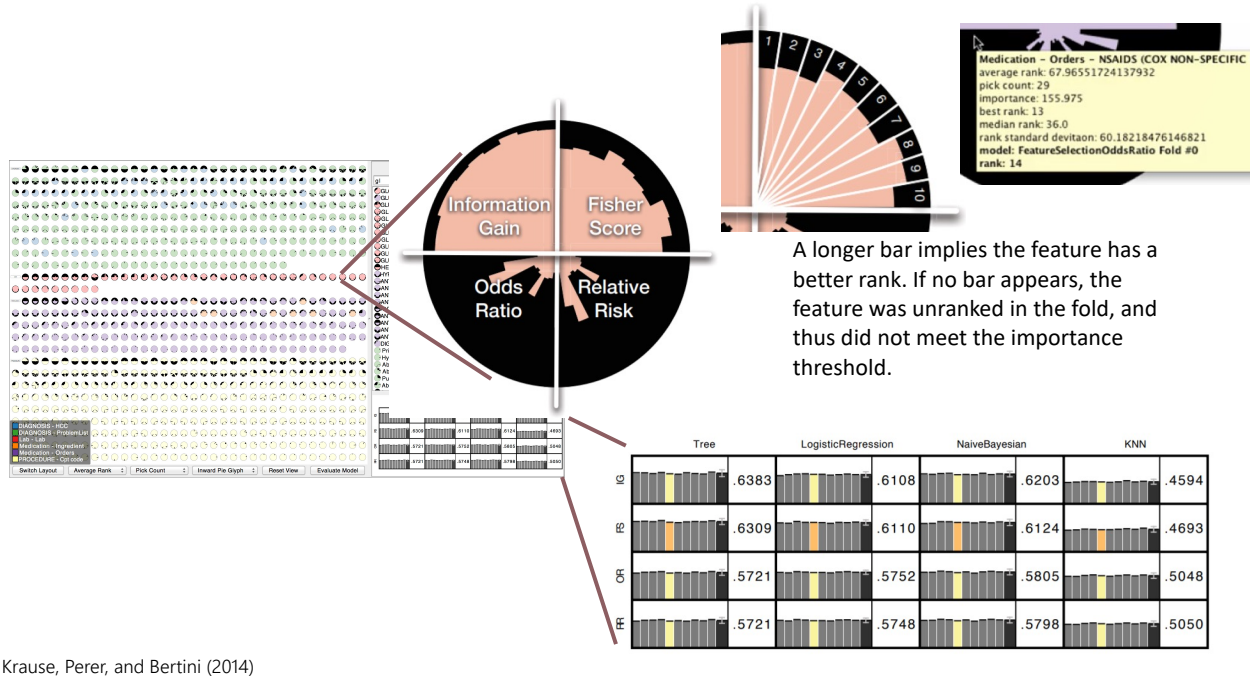
Diabetes Dataset

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

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

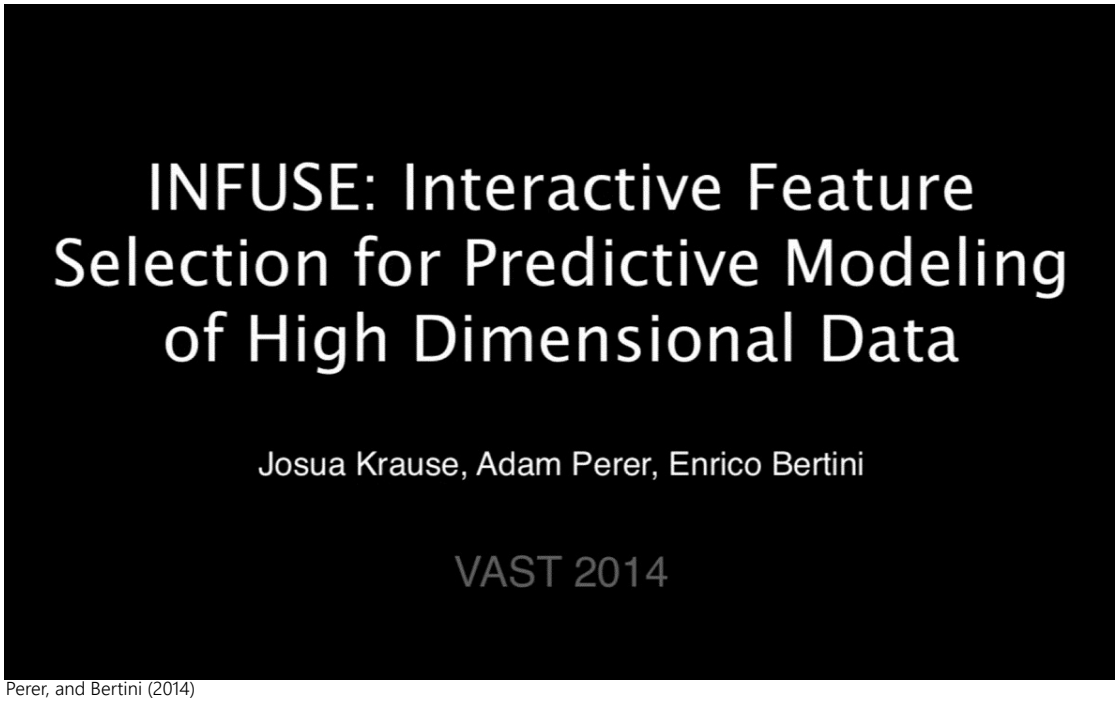
ML Task 3: Manual selection and testing of feature sets

15



Krause, Perer, and Bertini (2014)

16



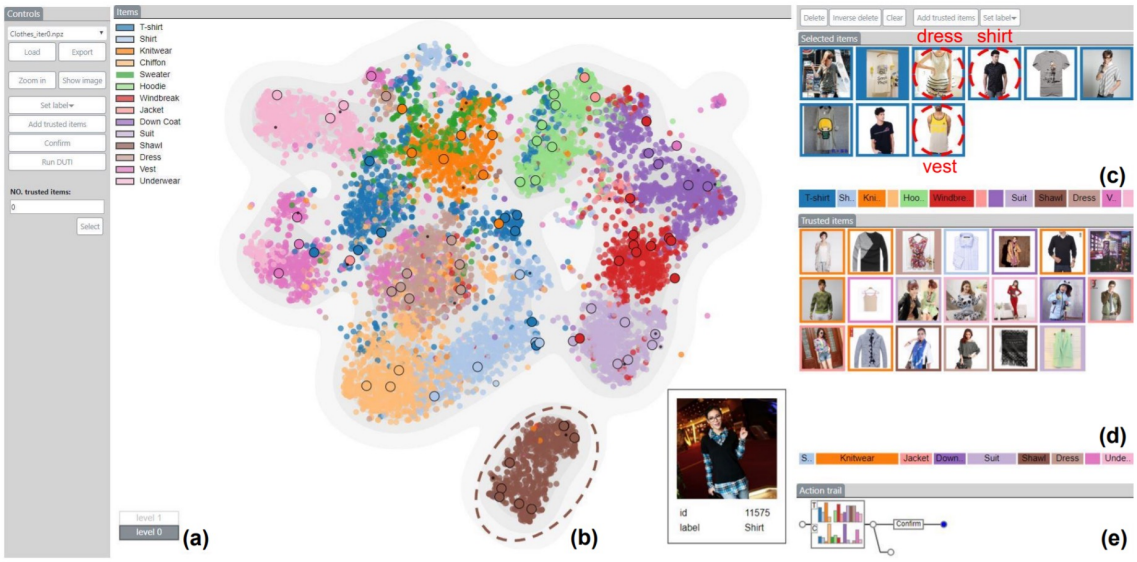
Krause, Perer, and Bertini (2014)

17

How expressive/effective is this visualization?

18

DataDebugger



Xiang et al. (2019)

19

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

ML Task: Classify article of clothing based on image

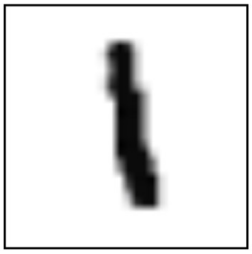
Xiang et al. (2019)

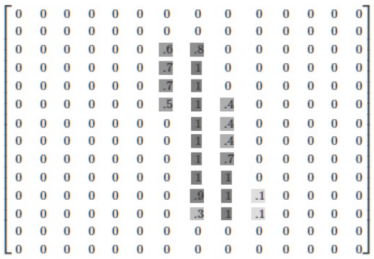
20

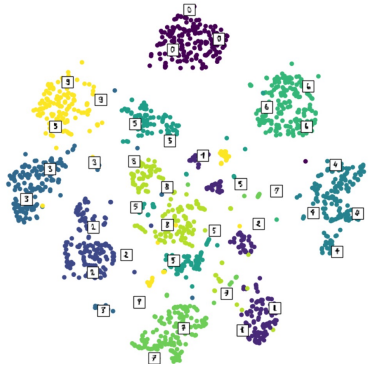
Dimensionality Reduction

28 x 28 Pixels

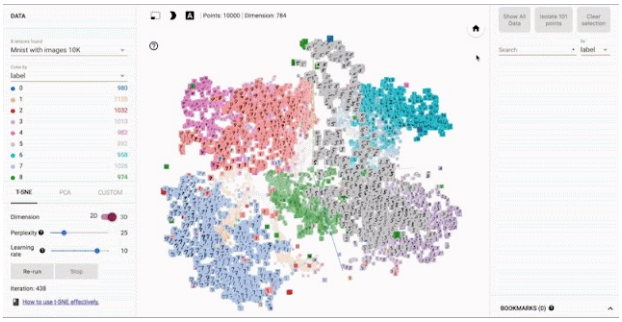
784 Dimensions



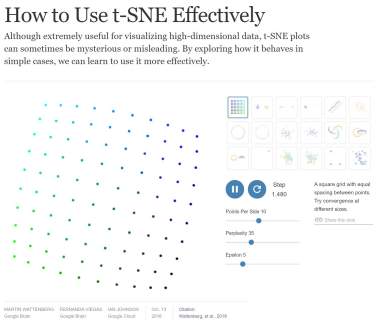




21

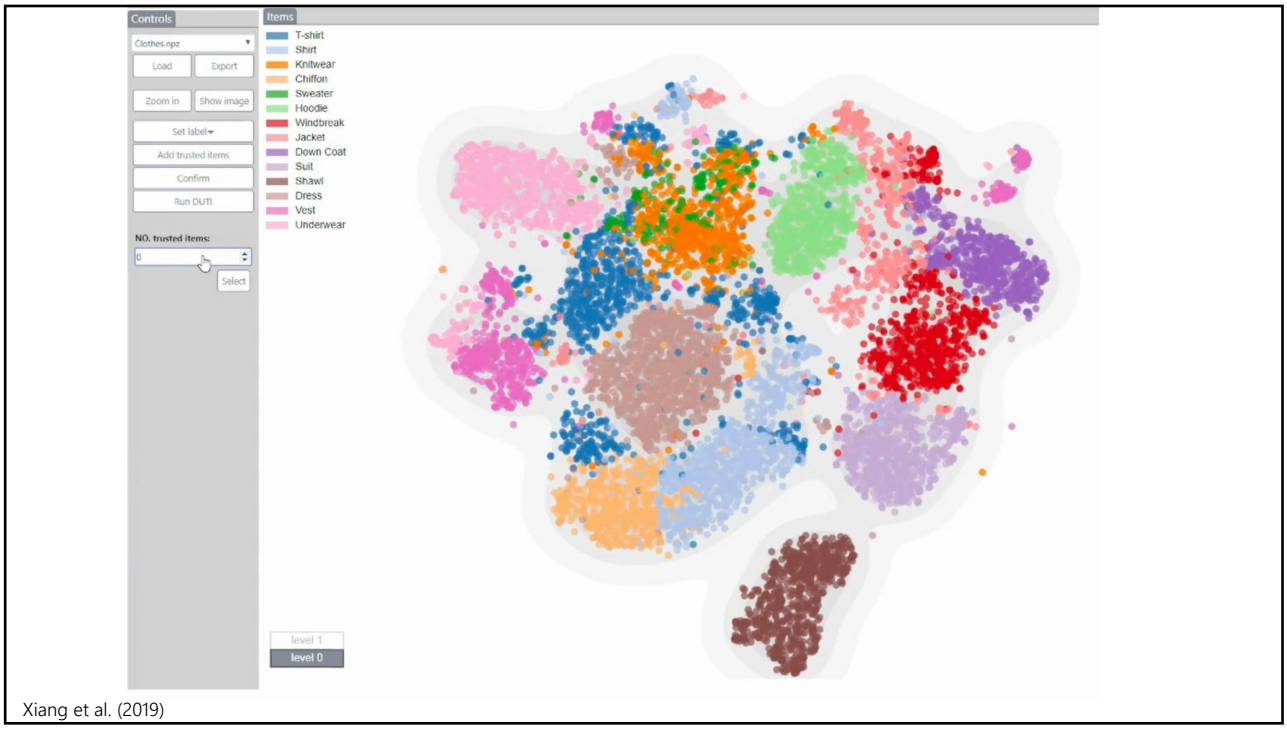


https://projector.tensorflow.org/



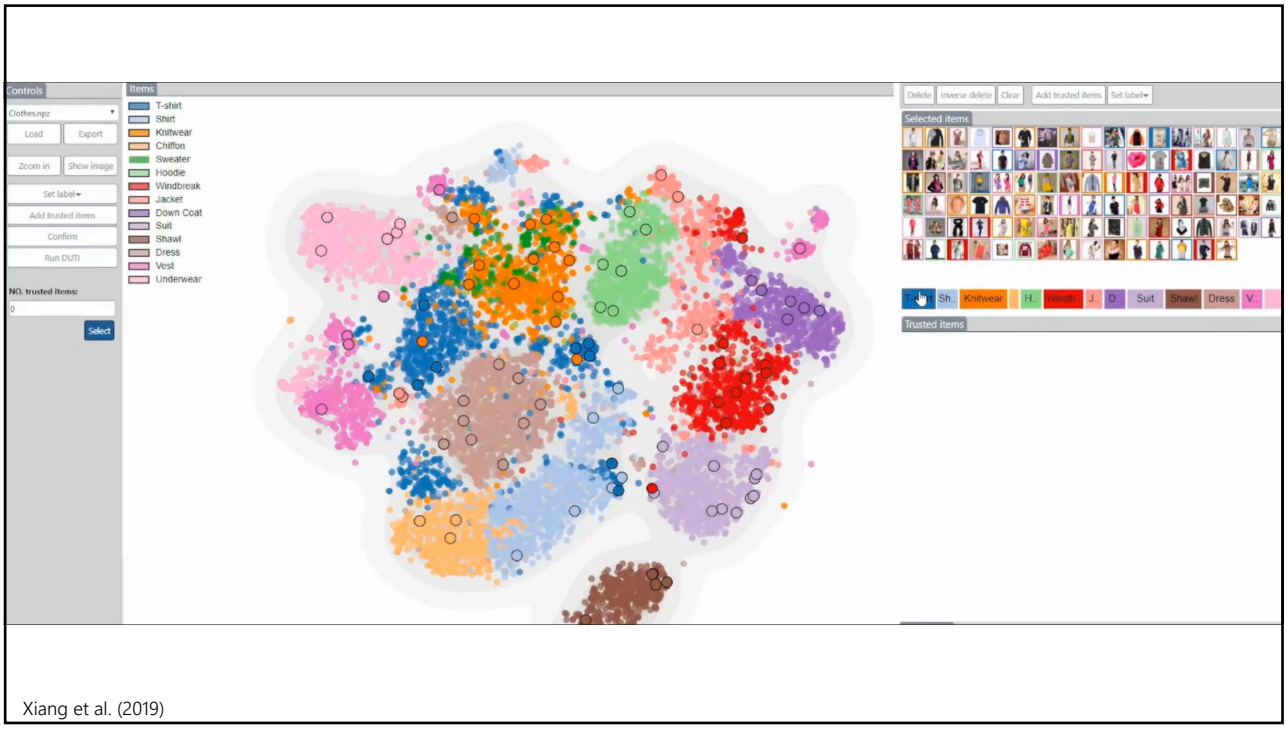
https://distill.pub/2016/misread-tsne/

22



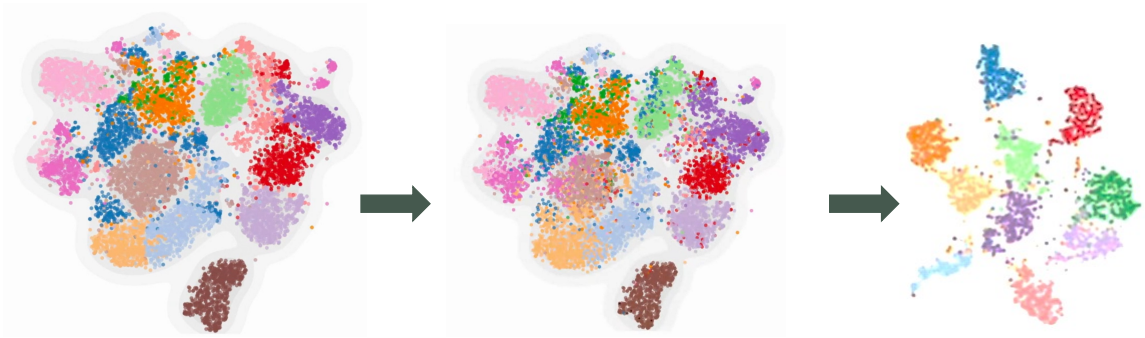
Xiang et al. (2019)

23



Xiang et al. (2019)

24



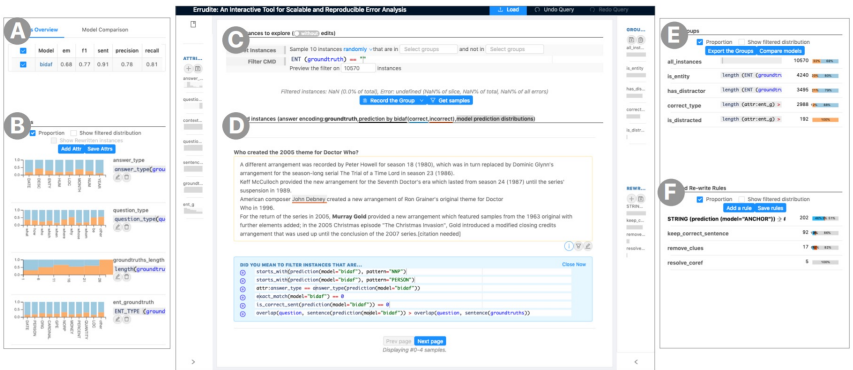
Xiang et al. (2019)

25

Errudite

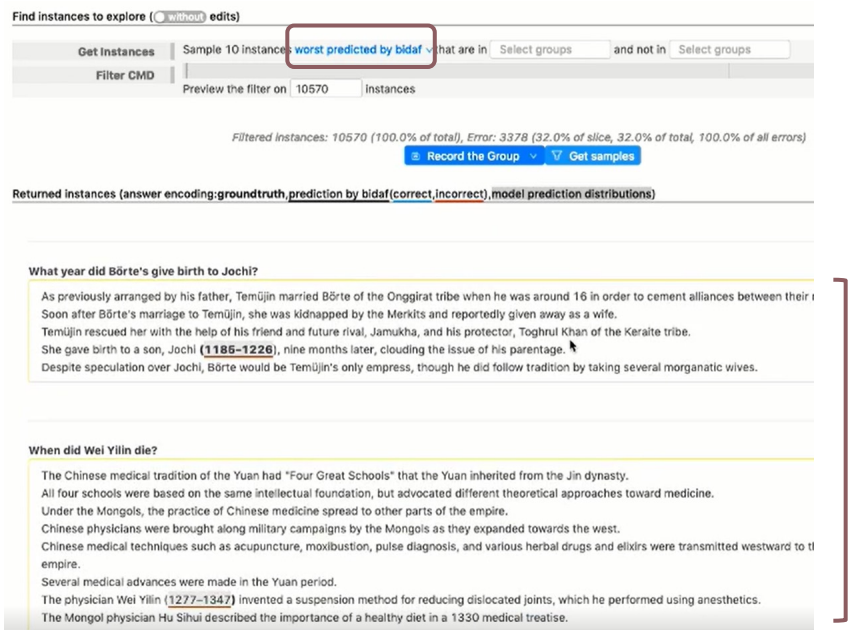
Error Analysis by:

1. Expressive grouping of error instance
2. Counterfactual evaluation



Wu et al. (2019)

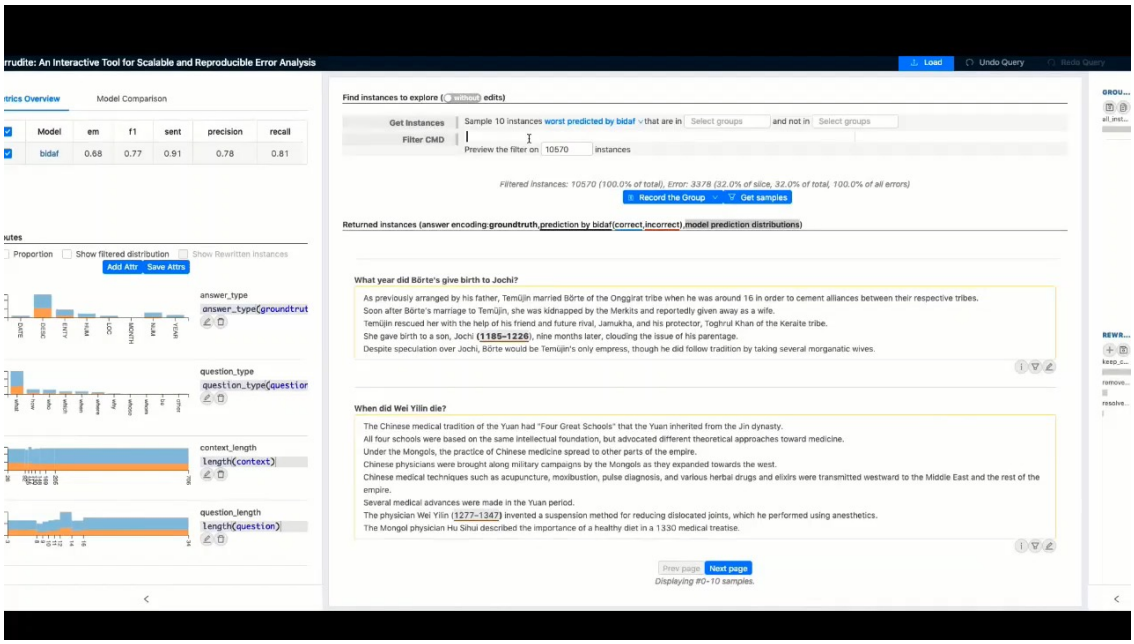
26



Error Instances

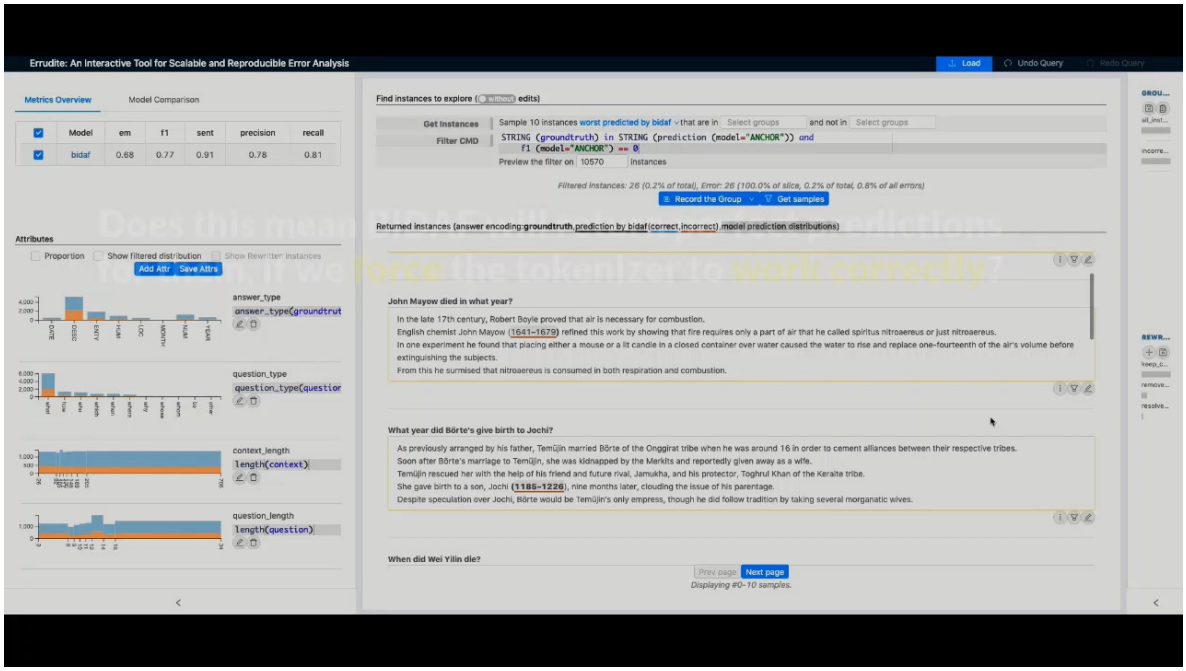
Wu et al. (2019)

27



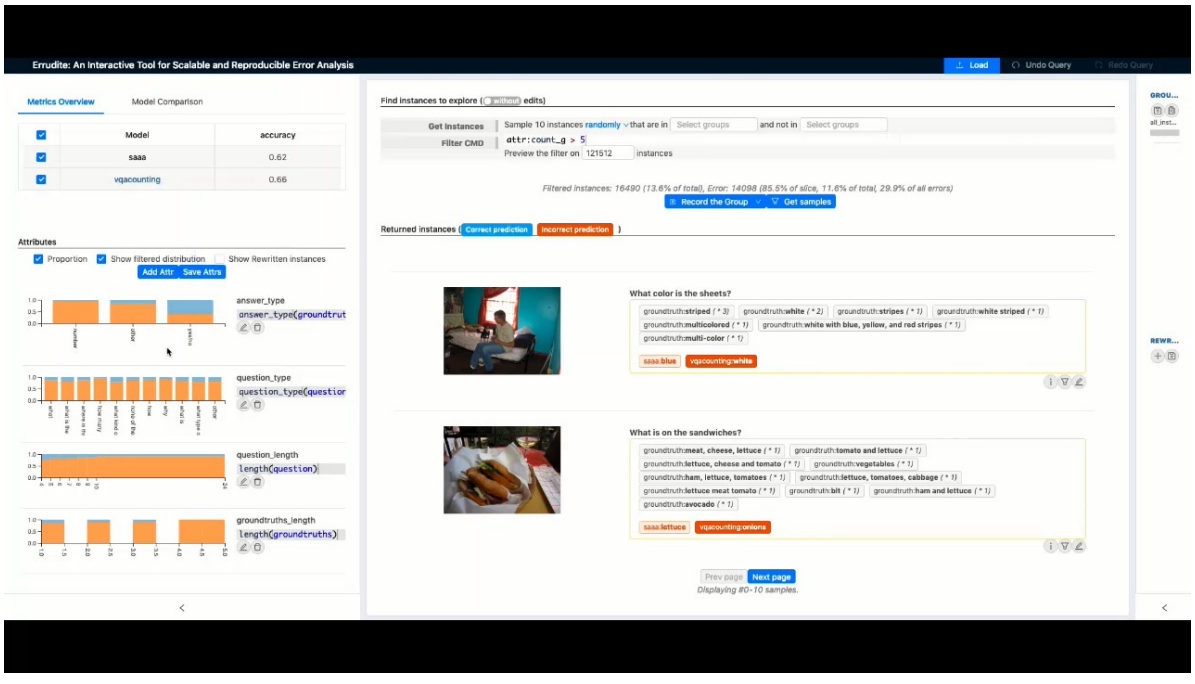
Wu et al. (2019)

28



Wu et al. (2019)

29



Wu et al. (2019)

30

Programming by Demonstration

Wu et al. (2019)

31

How does this approach compare to conventional GUI input elements ?

32

Visualizations for Modeling

33

RuleMatrix

Ming, Qu, & Bertini (2019)

34

Domain Questions

What knowledge has the model learned?

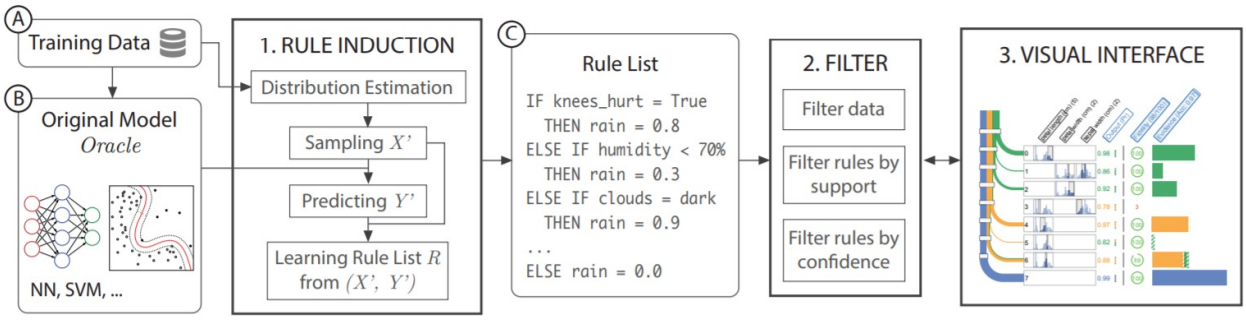
How certain is the model for each piece of knowledge?

What knowledge does the model utilize to make a prediction?

When and where is the model likely to fail?

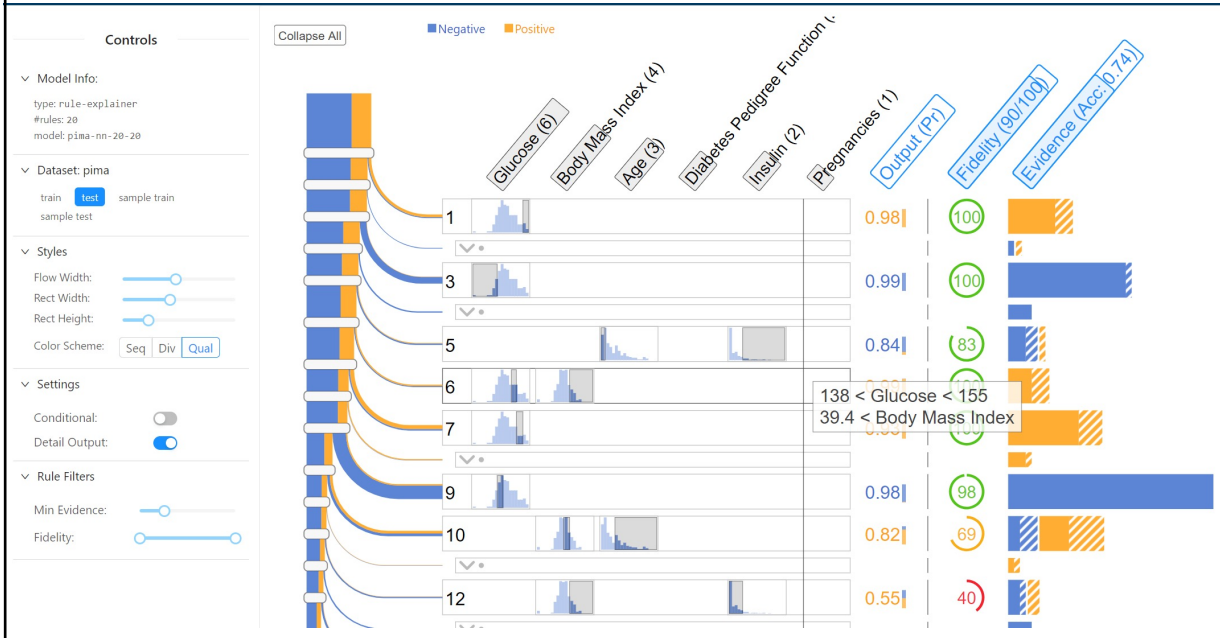
Ming, Qu, & Bertini (2019)

35



Ming, Qu, & Bertini (2019)

36



Ming, Qu, & Bertini (2019)

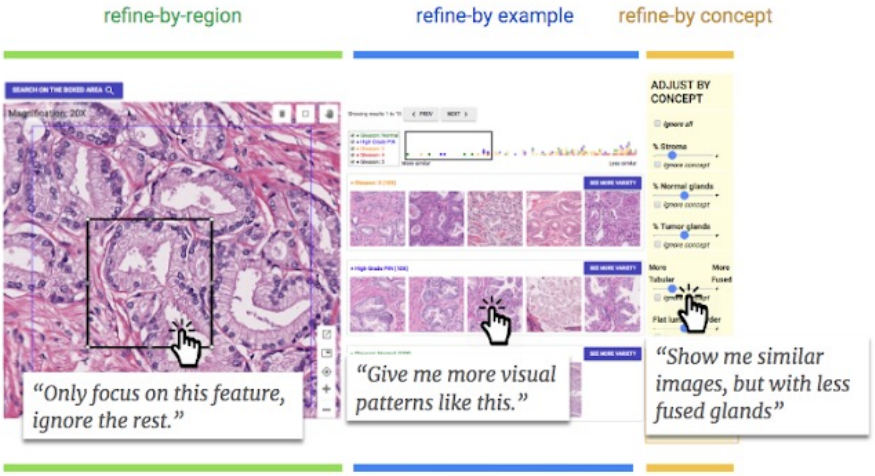
37

Use Case: Understanding and Improving Diabetes Classification

Ming, Qu, & Bertini (2019)

38

SMILY (Similar Medical Images Like Yours)



Cai et al. (2019)

39

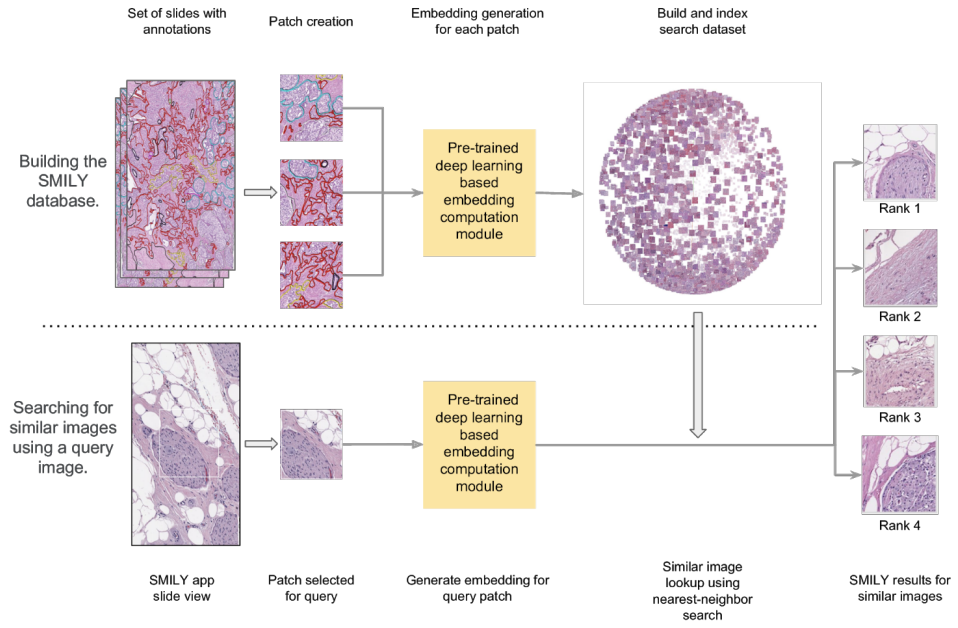
Domain Task

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

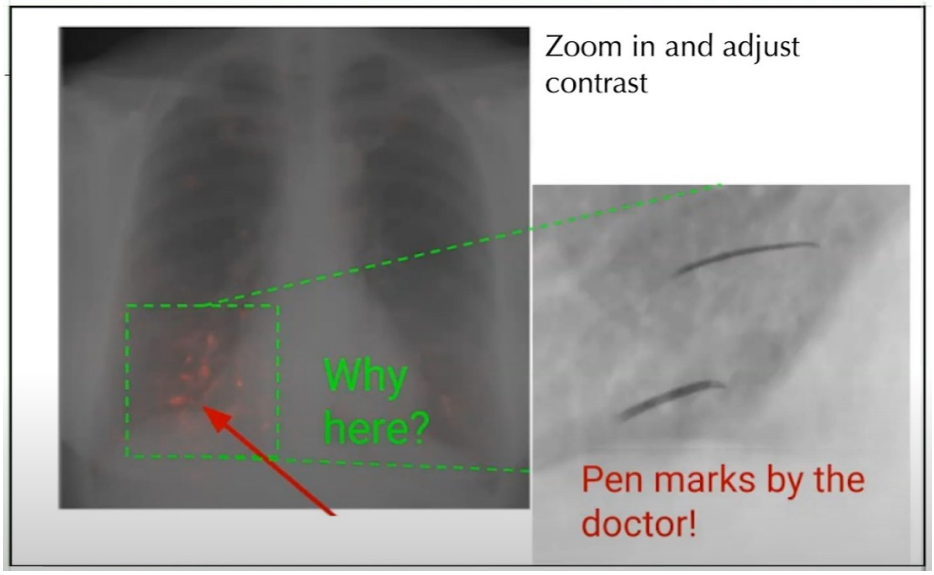


Cai et al. (2019)

41

Cai et al. (2019)

42



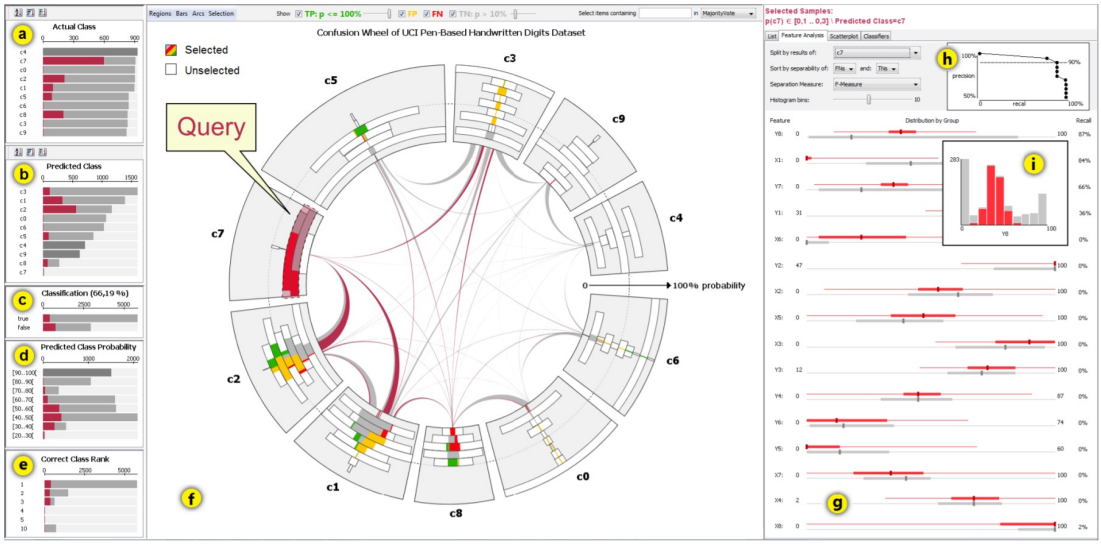
https://www.youtube.com/watch?v=AlK4_MznKJY

43

Visualizations for Model Evaluation

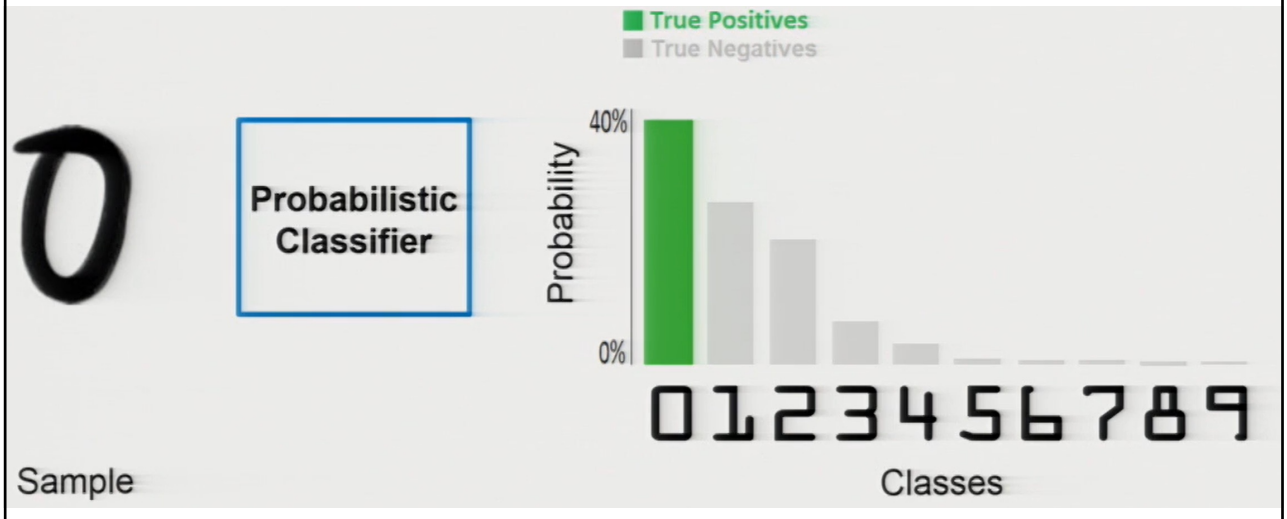
44

Confusion Wheel



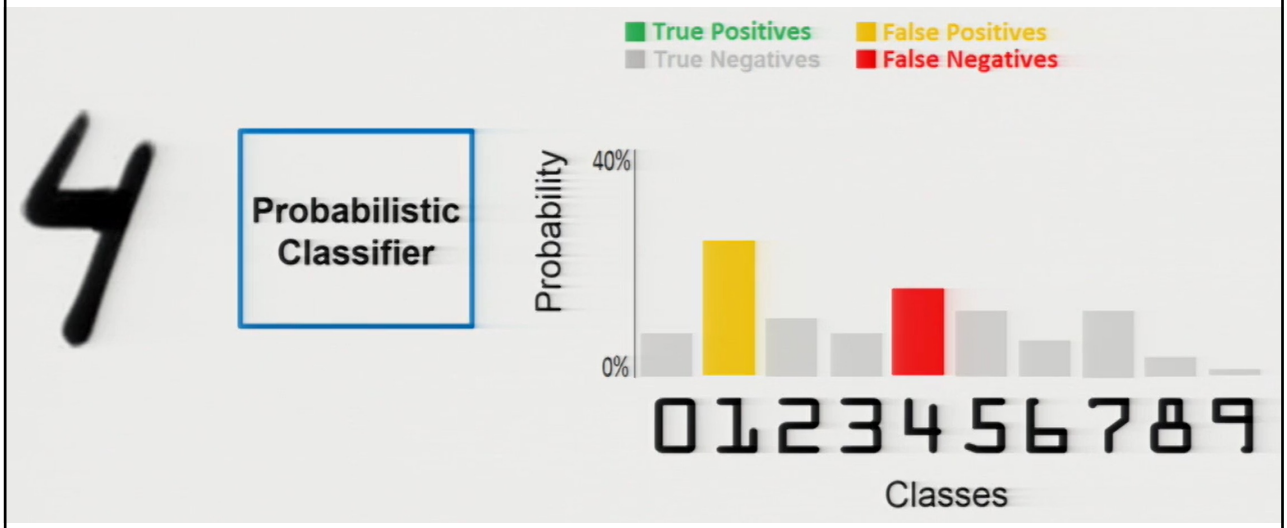
Alsallakh et al. (2014)

45



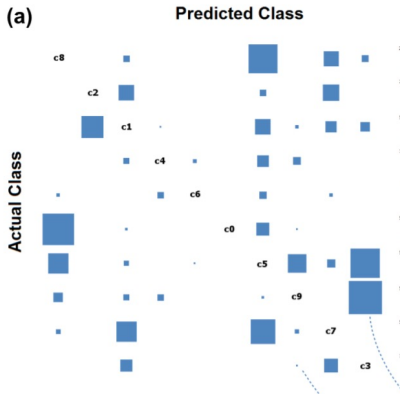
Alsallakh et al. (2014)

46



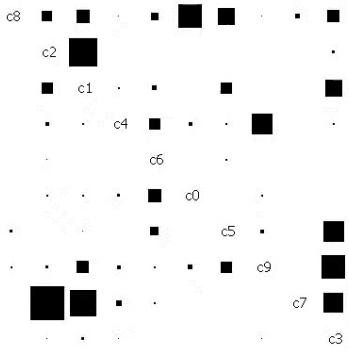
Alsallakh et al. (2014)

47



Alsallakh et al. (2014)

48



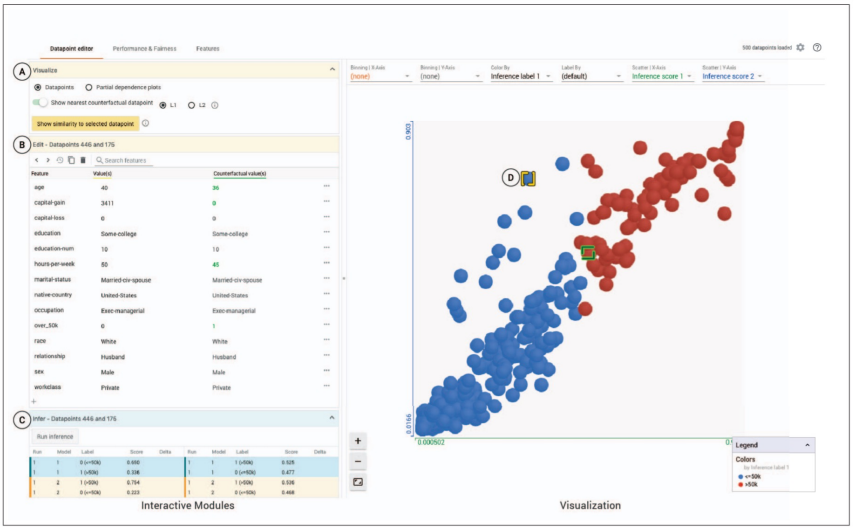
Alsallakh et al. (2014)

49

In what other ways can we visualize this data?

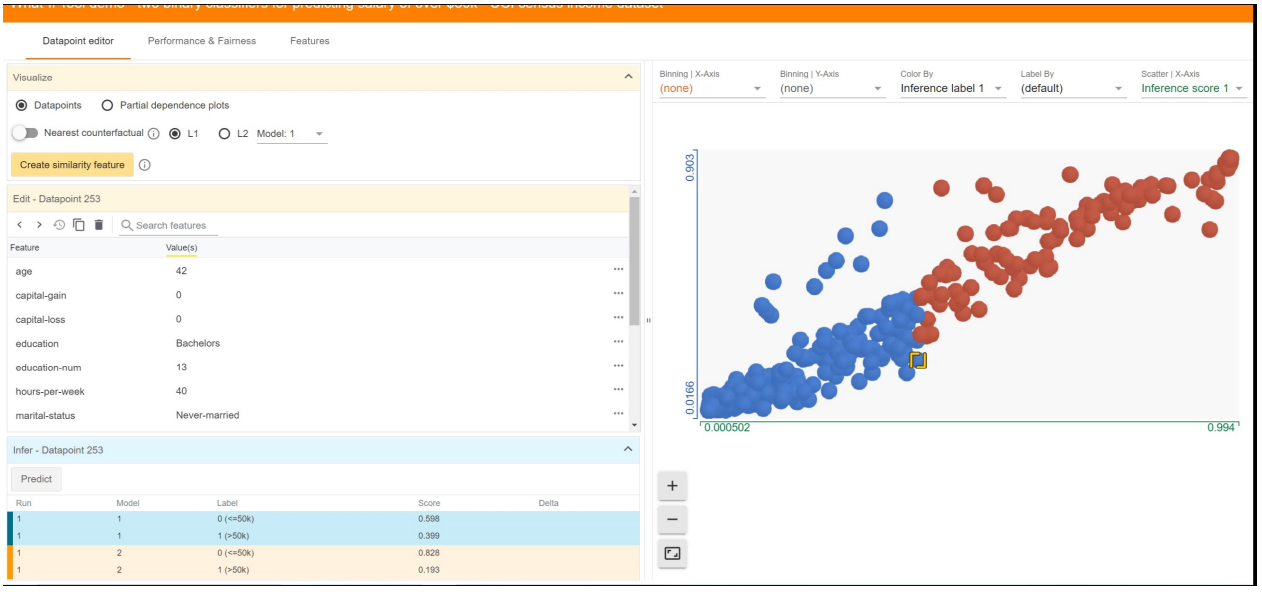
50

The What-If Tool



Wexler et al. (2019)

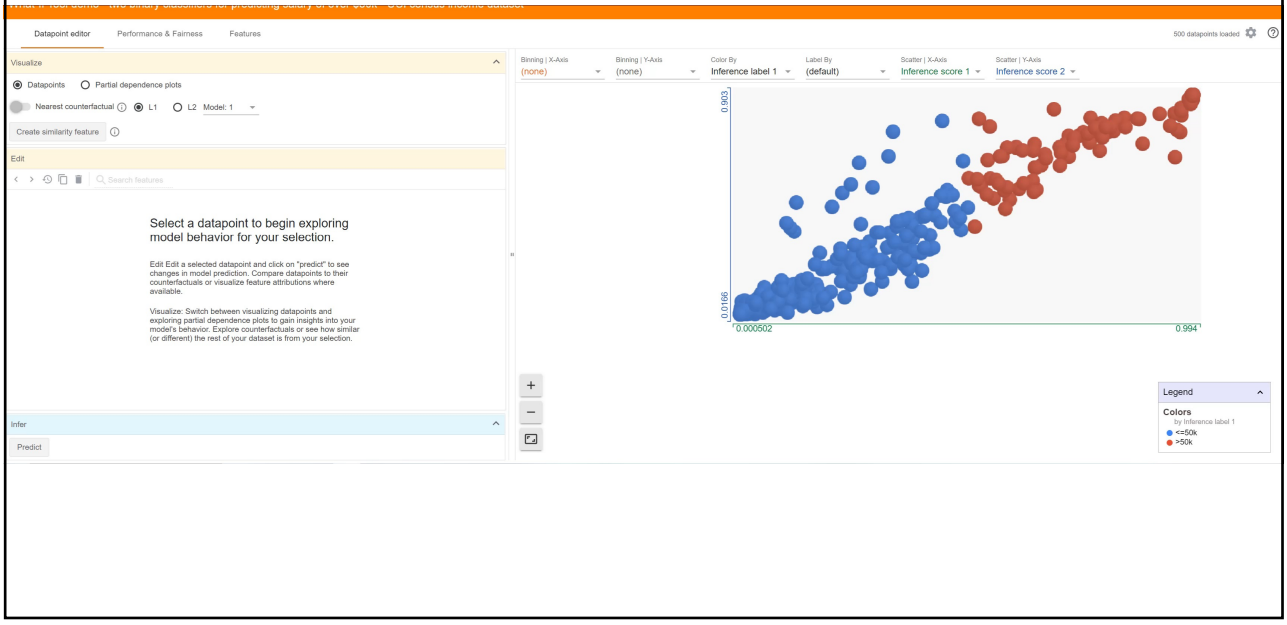
51



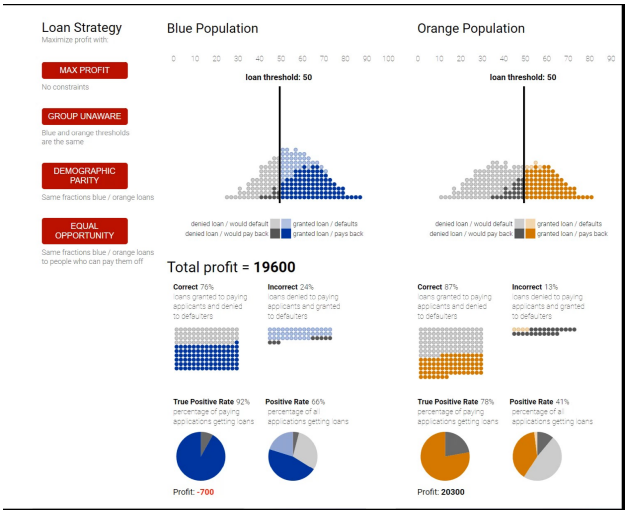
52


Datapoint editor
Performance & Fairness
Features
Visualize
Datapoints
Partial dependence plots
Nearest counterfactual
L1
L2
Model: 1
Create similarity feature
Edit - Datapoint 332
Search features
Feature
Value(s)
age
57
capital-gain
0
capital-loss
0
education
HS-grad
education-num
9
hours-per-week
40
marital-status
Married-civ-spouse
Infer - Datapoint 332
Predict
Run
Model
Label
Score
Delta
1
1
0 (<=50k)
0.656
1
1
1 (>50k)
0.347
1
2
0 (<=50k)
0.766
1
2
1 (>50k)
0.247
Binning | X-Axis
(none)
Binning | Y-Axis
(none)
Color By
Inference label 1
Label By
(default)
Scatter | X-Axis
Inference score 1
Scatter | Y-Axis
Inference score 2
0.903
0.0166
0.000502
0.994
Legend
Colors
<=50k
>50k

53


Datapoint editor
Performance & Fairness
Features
500 datapoints loaded
Visualize
Datapoints
Partial dependence plots
Nearest counterfactual
L1
L2
Model: 1
Create similarity feature
Edit
Search features
Select a datapoint to begin exploring model behavior for your selection.
Edit Edit a selected datapoint and click on "predict" to see changes in model prediction. Compare datapoints to their counterfactuals or visualize feature attributions where available.
Visualize: Switch between visualizing datapoints and exploring partial dependence plots to gain insights into your model's behavior. Explore counterfactuals or see how similar (or different) the rest of your dataset is from your selection.
Infer
Predict
Binning | X-Axis
(none)
Binning | Y-Axis
(none)
Color By
Inference label 1
Label By
(default)
Scatter | X-Axis
Inference score 1
Scatter | Y-Axis
Inference score 2
0.903
0.0166
0.000502
0.994
Legend
Colors
by Inference label 1
<=50k
>50k

54



https://research.google.com/bigpicture/attacking-discrimination-in-ml/

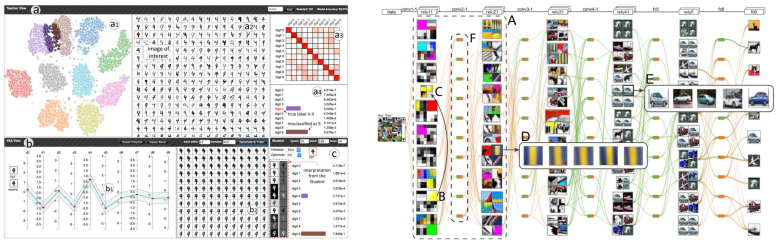
FairSight

Ahn & Lin (2019).

55

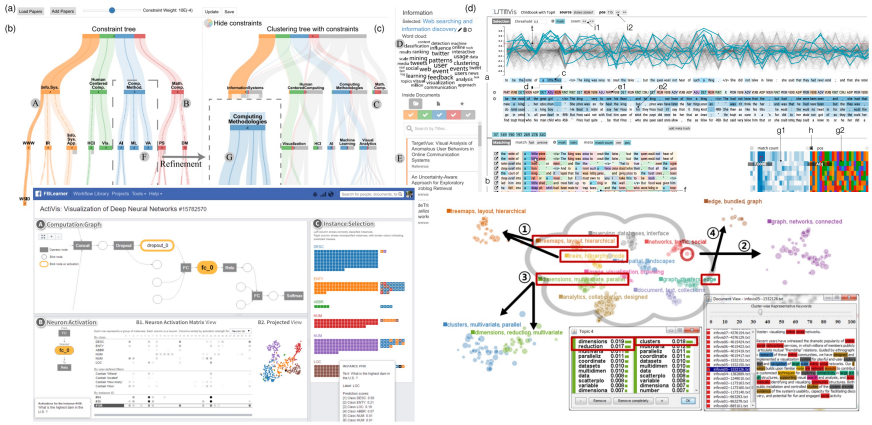
What is different/unique about visualizing ML data?

56



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.



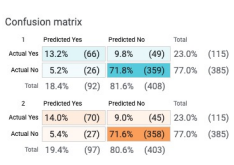
57

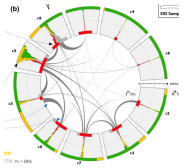
Some Guidelines for ML Visualizations

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



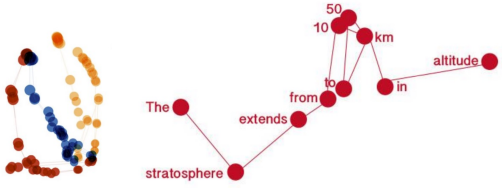


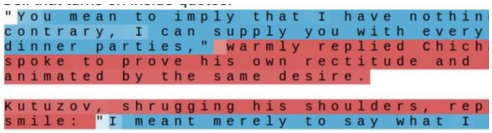
Dudley & Kristensson (2018)

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



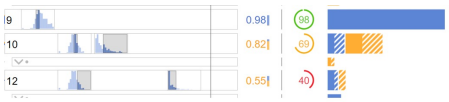


Dudley & Kristensson (2018)

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.

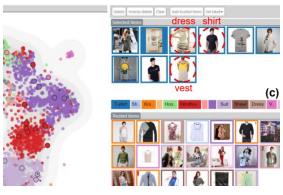


Dudley & Kristensson (2018)

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
5. Support expressive inputs

- Editing data points
- Evaluating Hypotheses
- Constructing Explanations



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

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.

64