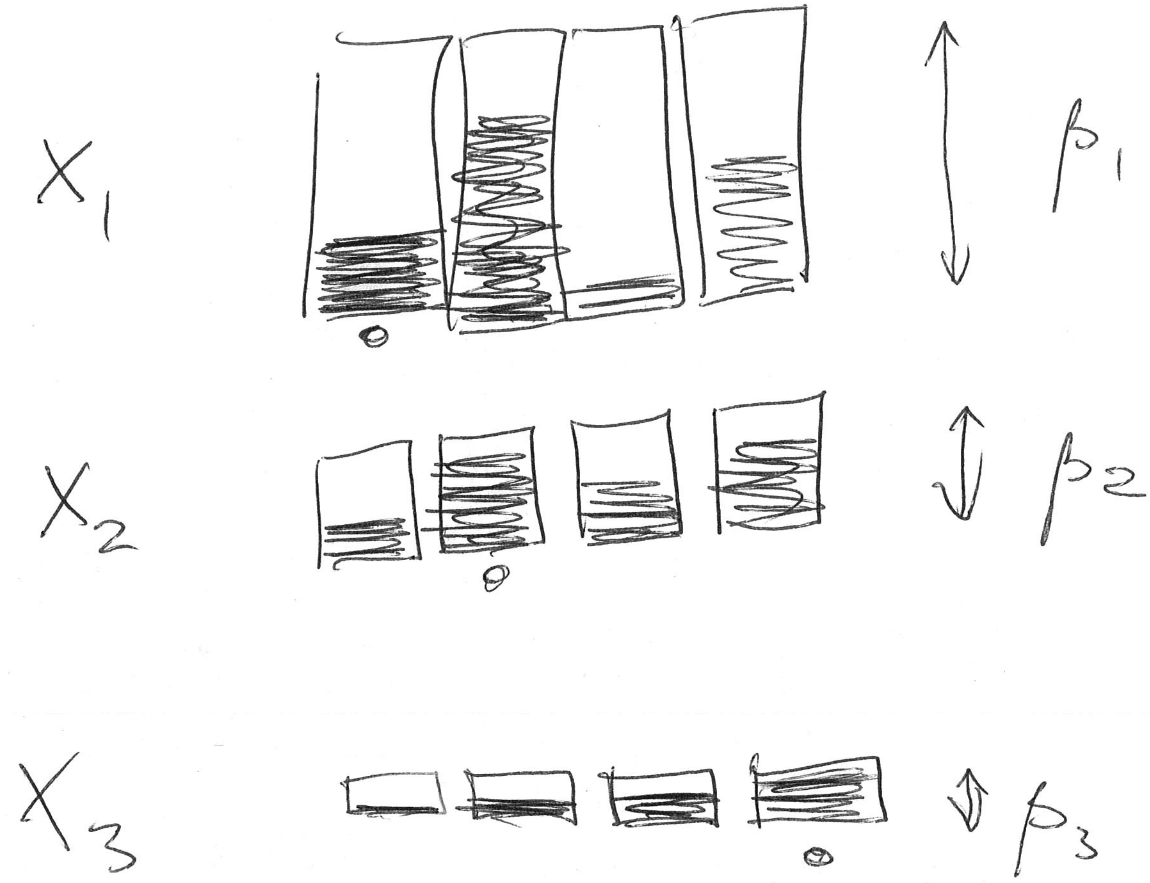
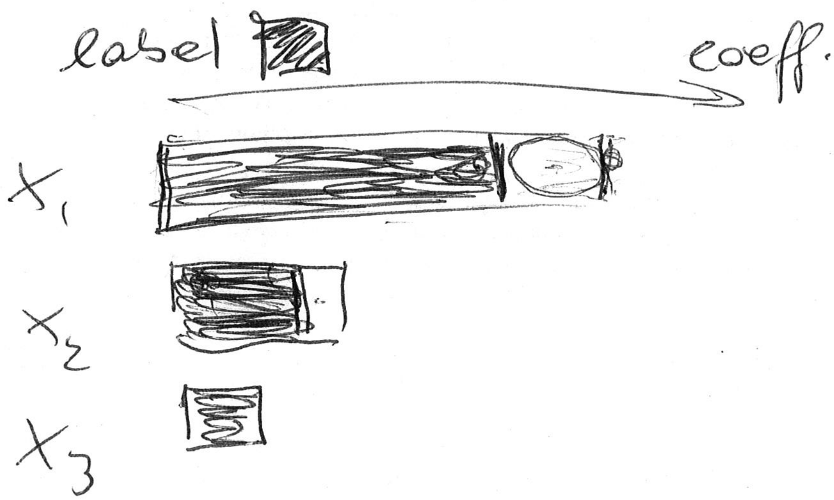
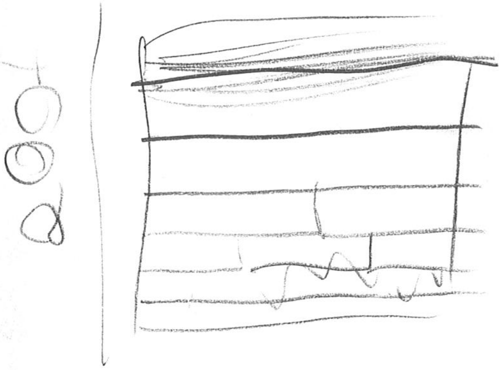




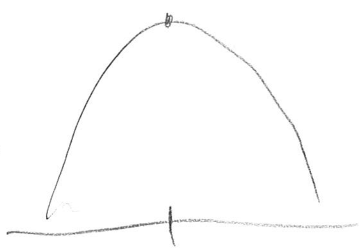
$$|X_1 \cap X_2 \cap X_3| = \frac{1}{64} \cdot n$$

Pred	Obs	dop
1	1	rough
	⋮	ccf
	⋮	ccf
⋮	⋮	
10	1	rough_cor
1	2	ccf
⋮	⋮	⋮
10	2	ccf

Obs	rough_cor	ccf	nd
1	0.3	0.7	0



$$p_i = (p_{i1}, \text{—} \, p_{iq}) \sim \text{Mult}$$



$p(1-p)$

$$\exp\left(-\sum_{k=1}^{q} p_{ik} \log p_{ik}\right)$$

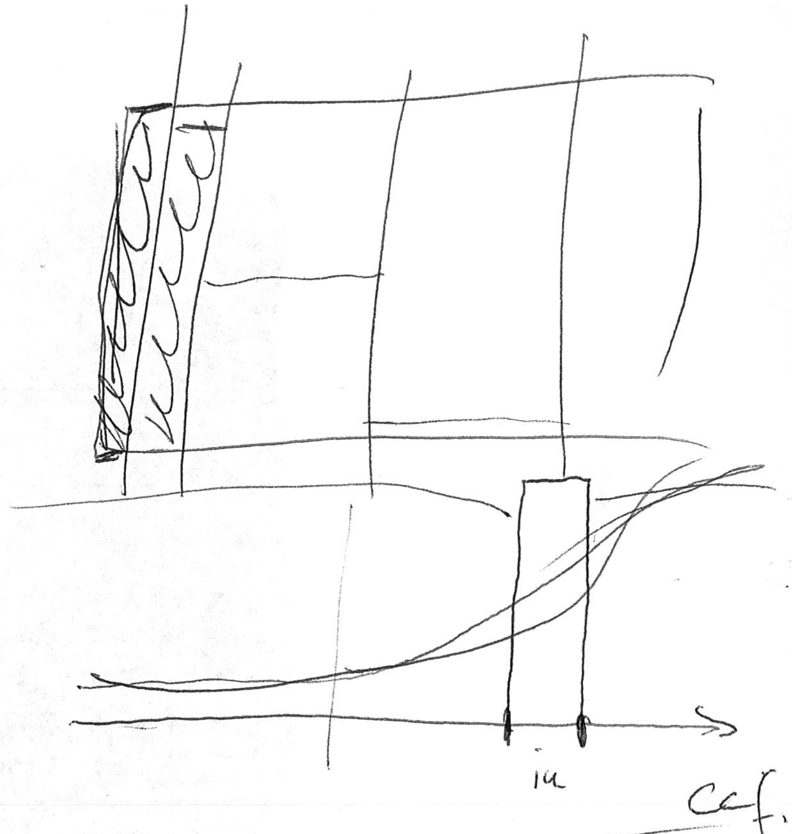
Shannon entropy

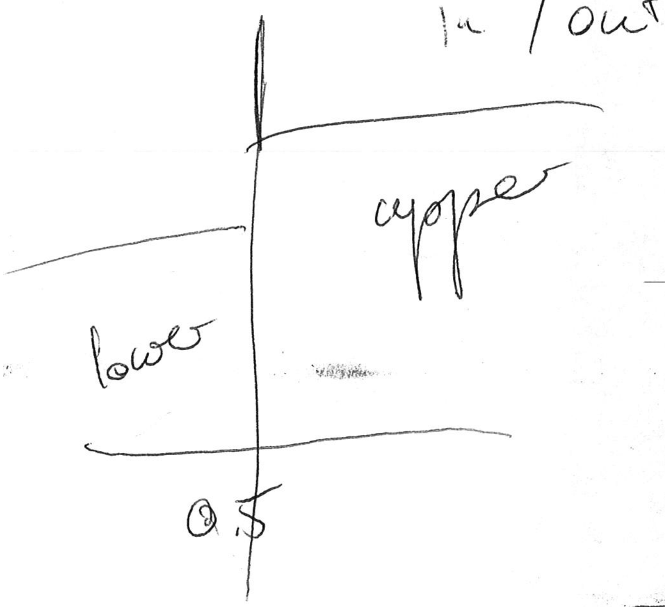
$$\left(1 - \sum_{k=1}^{q} p_{ik}^2\right) \in [0,1]$$

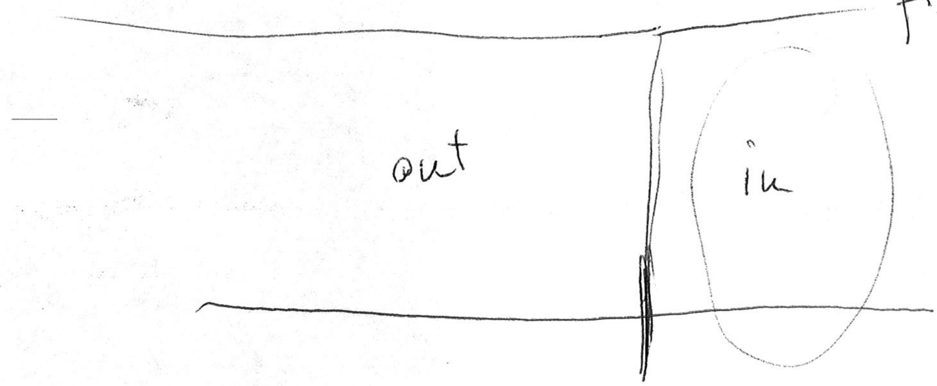
$$\sum_{k=1}^{q} p_{ik} = 1$$

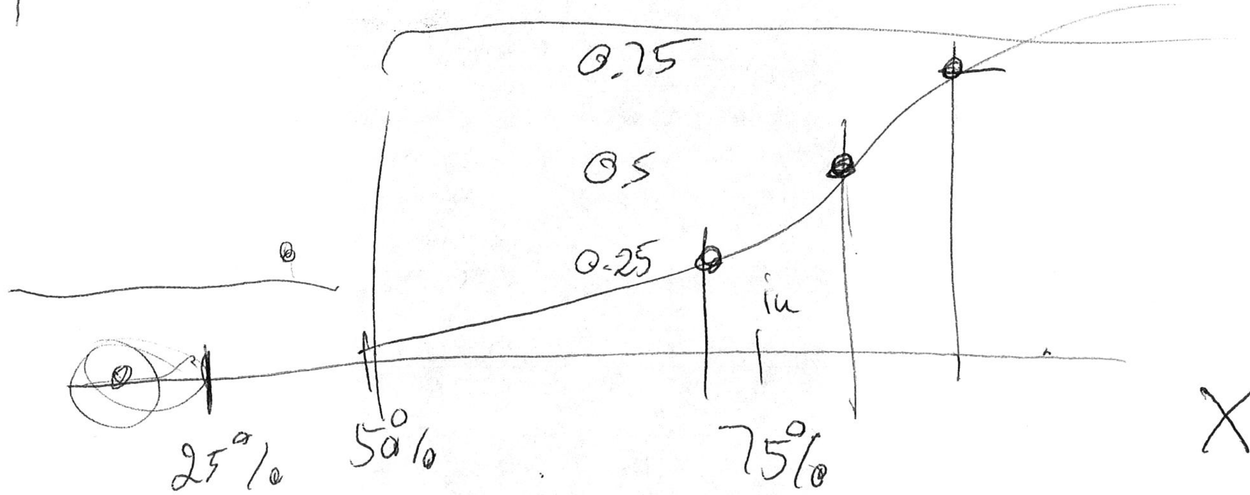
ccf ≧ 0.5

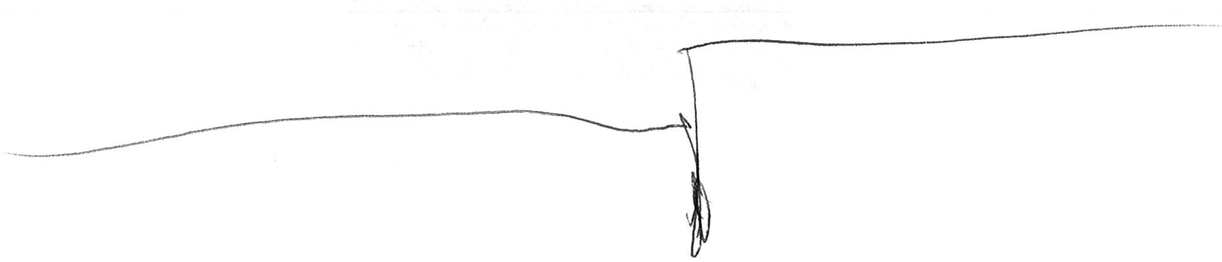
ccf ≥ 0.75





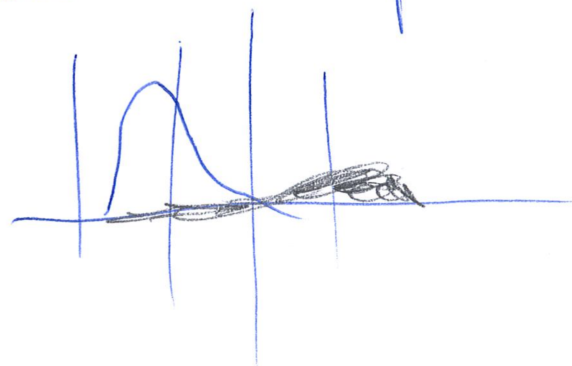






~~Same~~ Equalbin method

→ skew data will be problematic



Quantile method is better

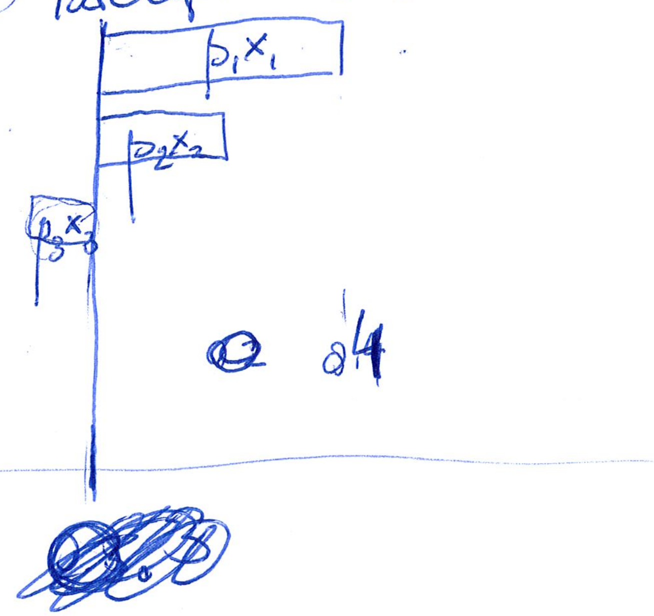
binning + ~~lasso~~ ridge regression

$$y_{ij} = \sum \beta_j x_{ij} + \text{penalty} + \varepsilon$$

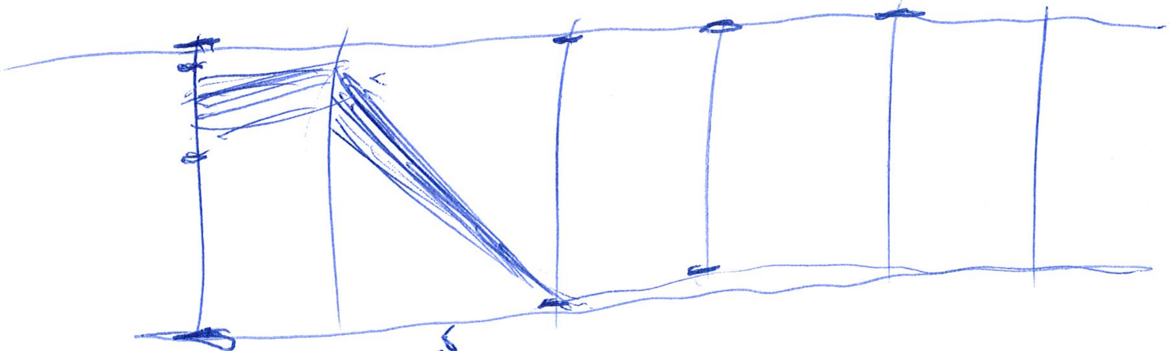
case i:

$\beta_j x_{ij}$ for each $j = 1, \to p$

intercept is 0.3

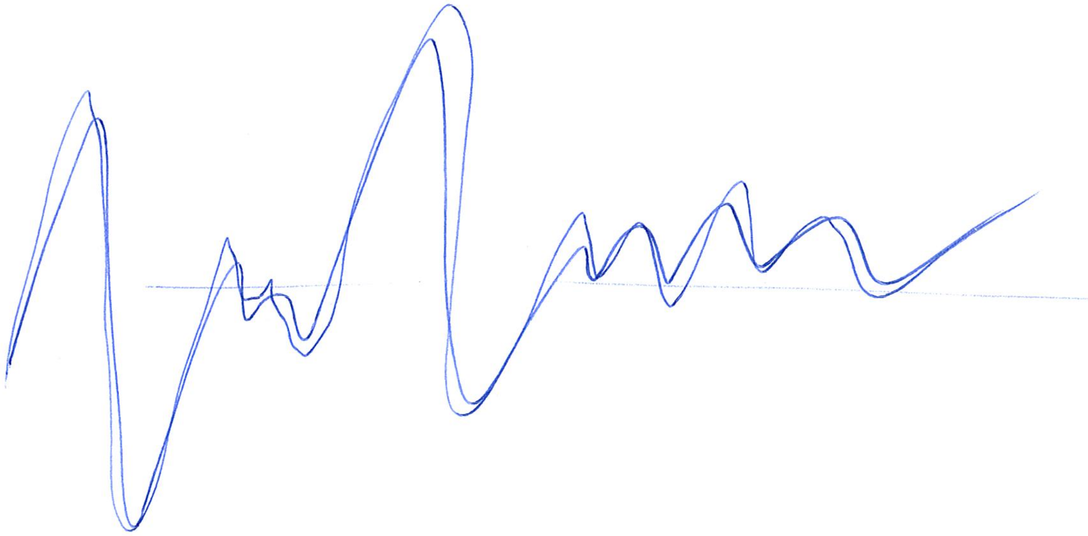


col = rgb(1, 1, 1, alpha = 0.1)



cor

doegh-cor



ASA Data ~~Science~~ Journal

Visual Diagnostics of a model explainer

at the example of LIME

main objective of model explainer:

- understand and explain model performance

LIME does....

Conceptually: models at two levels:

explainer model — "simple"

original "black box" model — "complicated"

Usually: model predictions, maybe with ground truth

Type I error

model is wrong

Type II error

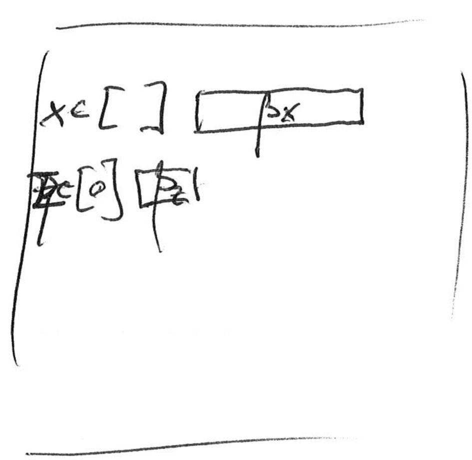
explanation is also a prediction — how reliable is that explanation?

Explained model has very low R^2 generally
- probably due to binning

"Local" explanations are not local, but are driven by the (global) marginal distribution of covariates

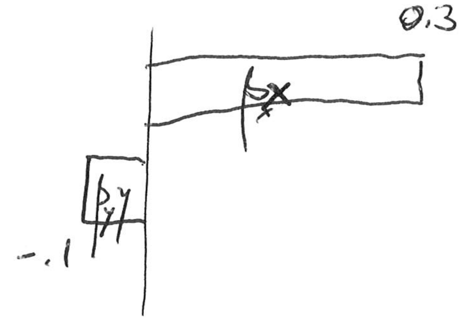
Describe LIME, including details on binning and linear regression in binned features

Motivation



$X = 8$

$y = 0.5$



$\mu = 0.5$

.sty
Style file

Rmarkdown:

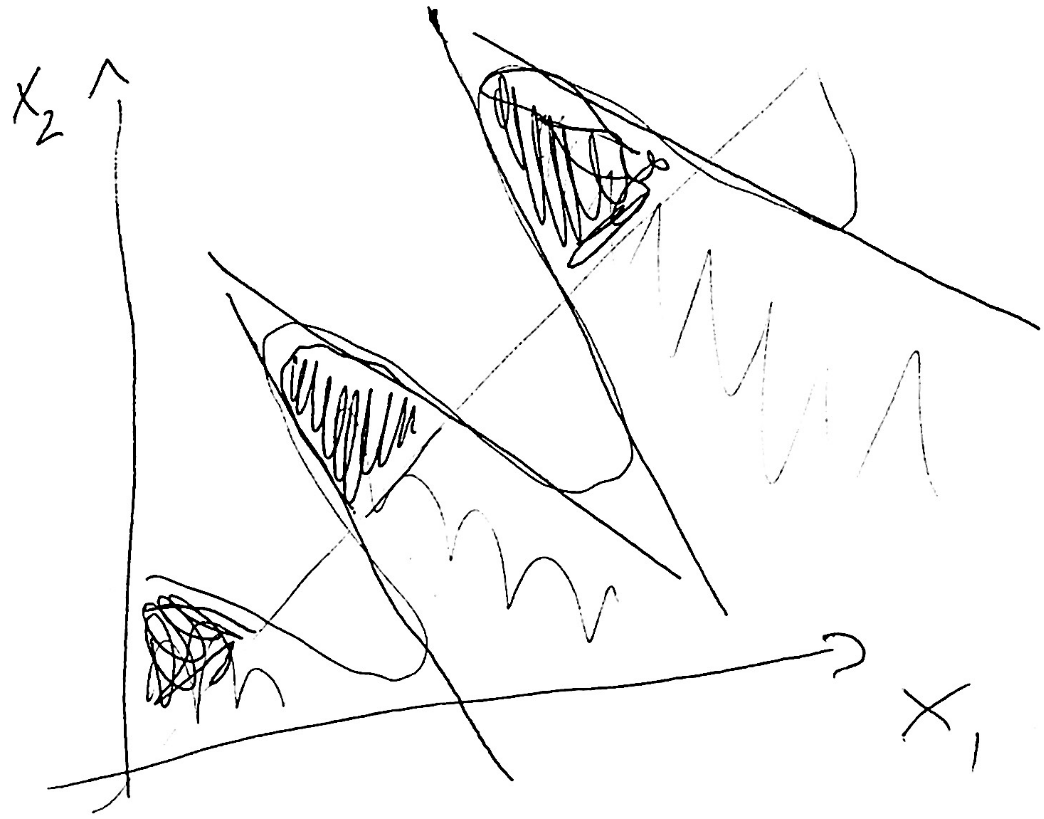
@bibtextag Author (2010)

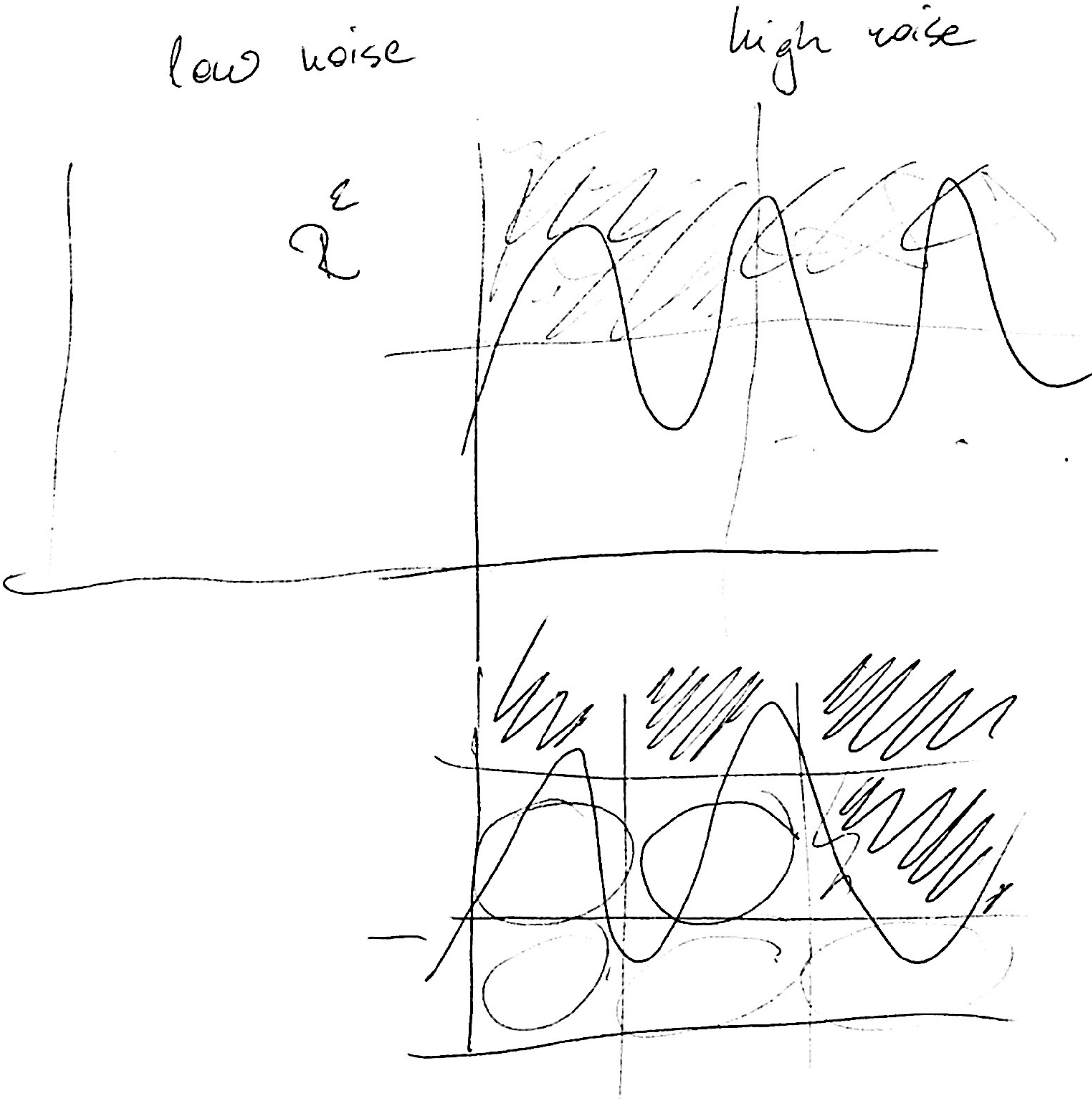
[@bibtextag] (Author 2010)

[@ref1; @ref2] (Author1 2010, Author2 et al 2013)

yaml:

bibliography: name.bib


X_2
X_1


low noise
high noise

Local (interpretable) model agnostic explanations

4^2 is low number ~~simple~~

4^p might not be very low

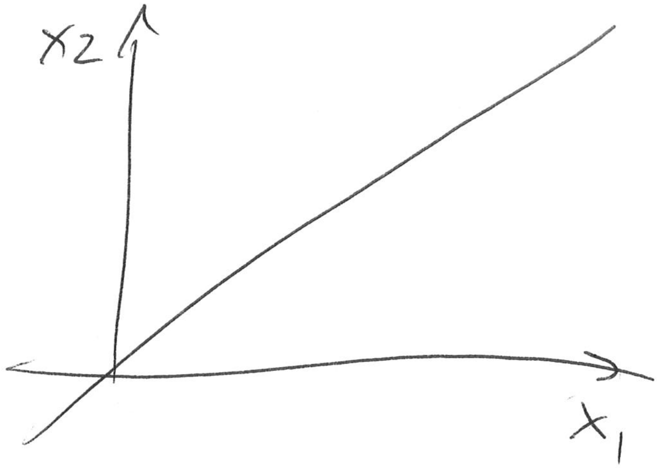
Expectations for explanations

- data driven by relevant features
- "explain": difference in coefficients

deterministic/ non-deterministic

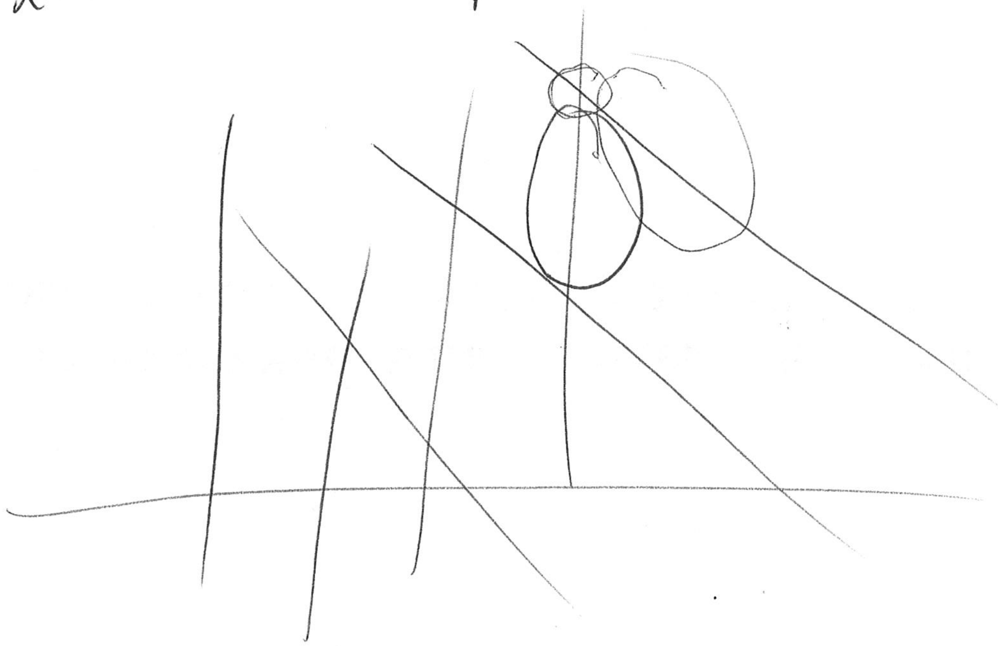
Random forest variable importance for identify relevant features

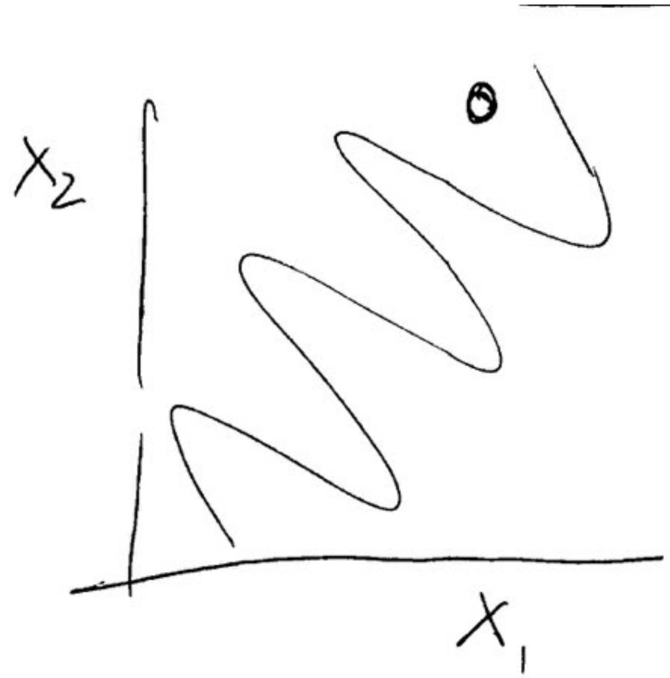
Score based Likelihood Ratio/Bayes factor global summary

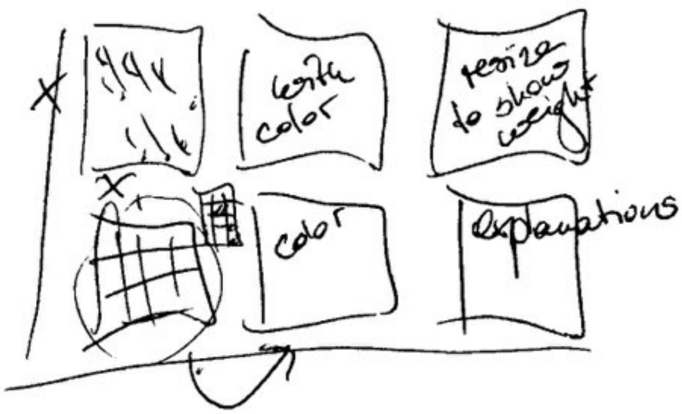


$$z = ax_1 + bx_2$$

local summaries







Create permutations
obtain RF predictions
weight (gower)
bin
weight (not gower)
fit ridge regression
~~get expla~~
Select features
re-fit ridge
get explanations

Suggestion

Discussion

random Forest

importance features

x_1, x_2 important

E ok

Then a tree on
x_1, x_2 < on weighted
simulated data of fig 7
using tfscore as y

binned versions
of x_1 and x_2

Complexity of tree
directly related to
complexity of explanation
tradeoff between accuracy
and simplicity of explanation

$$\left[\begin{array}{c} < X_1 < \\ < X_2 < \end{array}\right]$$

simple versus useful