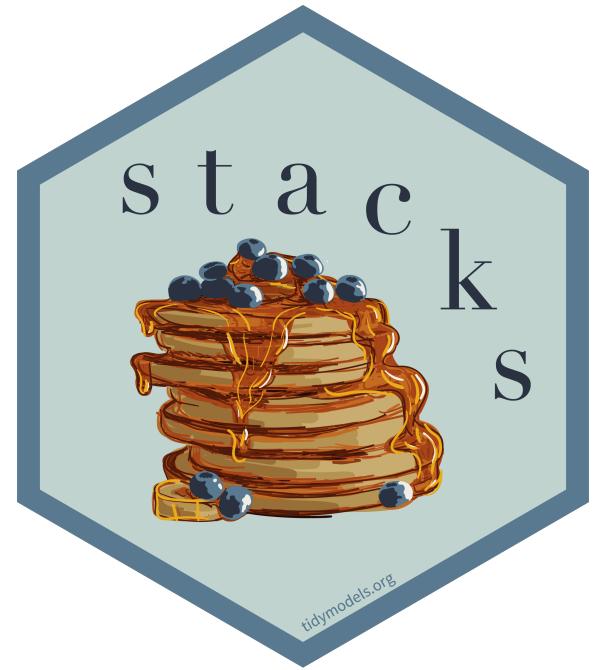


Stacked Ensemble Modeling with Tidy Data Principles

Simon Couch  @simonpcouch

JSM 2021 - John M. Chambers Award Recipient



Stacked Ensemble Modeling with Tidy Data Principles



@simonpcouch

@simonpcouch

R development "ecosystems"

@simonpcouch

R development "ecosystems"

"base"

@simonpcouch

R development "ecosystems"

"base"

"tidyverse"

@simonpcouch

R development "ecosystems"

"base"

"tidyverse"



...

@simonpcouch

R development "ecosystems"

"base"

"tidyverse"



...

"tidymodels"

@simonpcouch

R development "ecosystems"

"base"

"tidyverse"



...

"tidymodels"



...

@simonpcouch

R development "ecosystems"

"base"

"tidyverse"



...

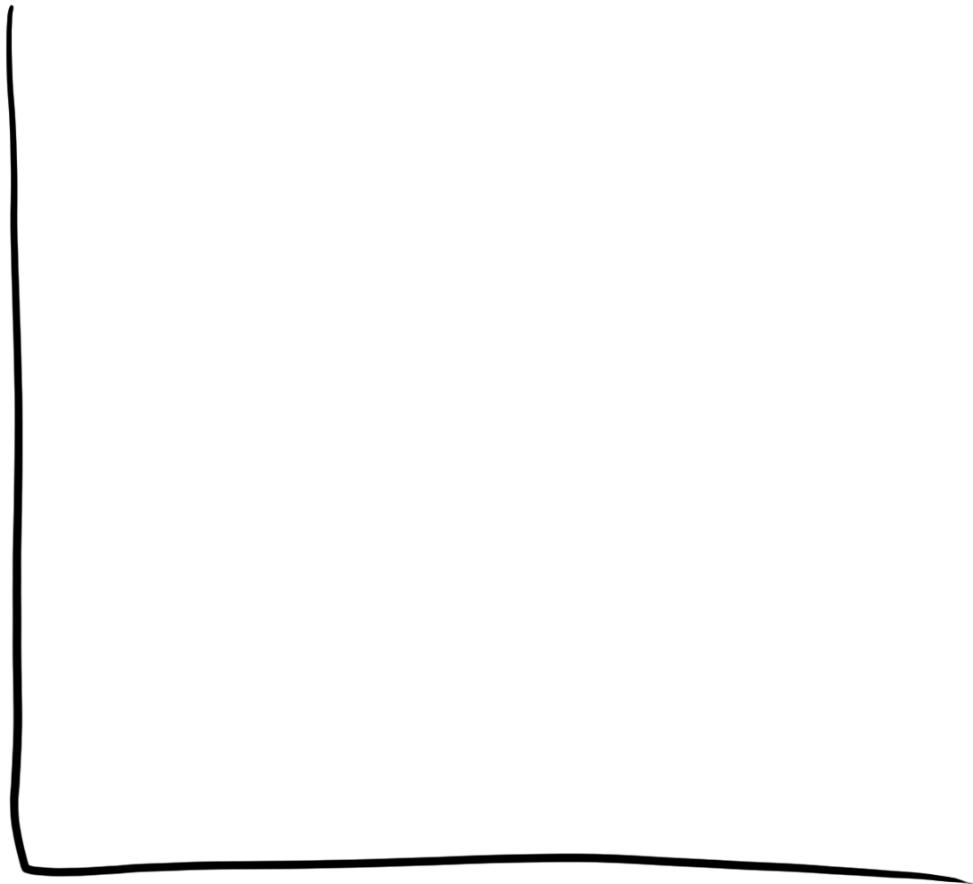
"tidymodels"



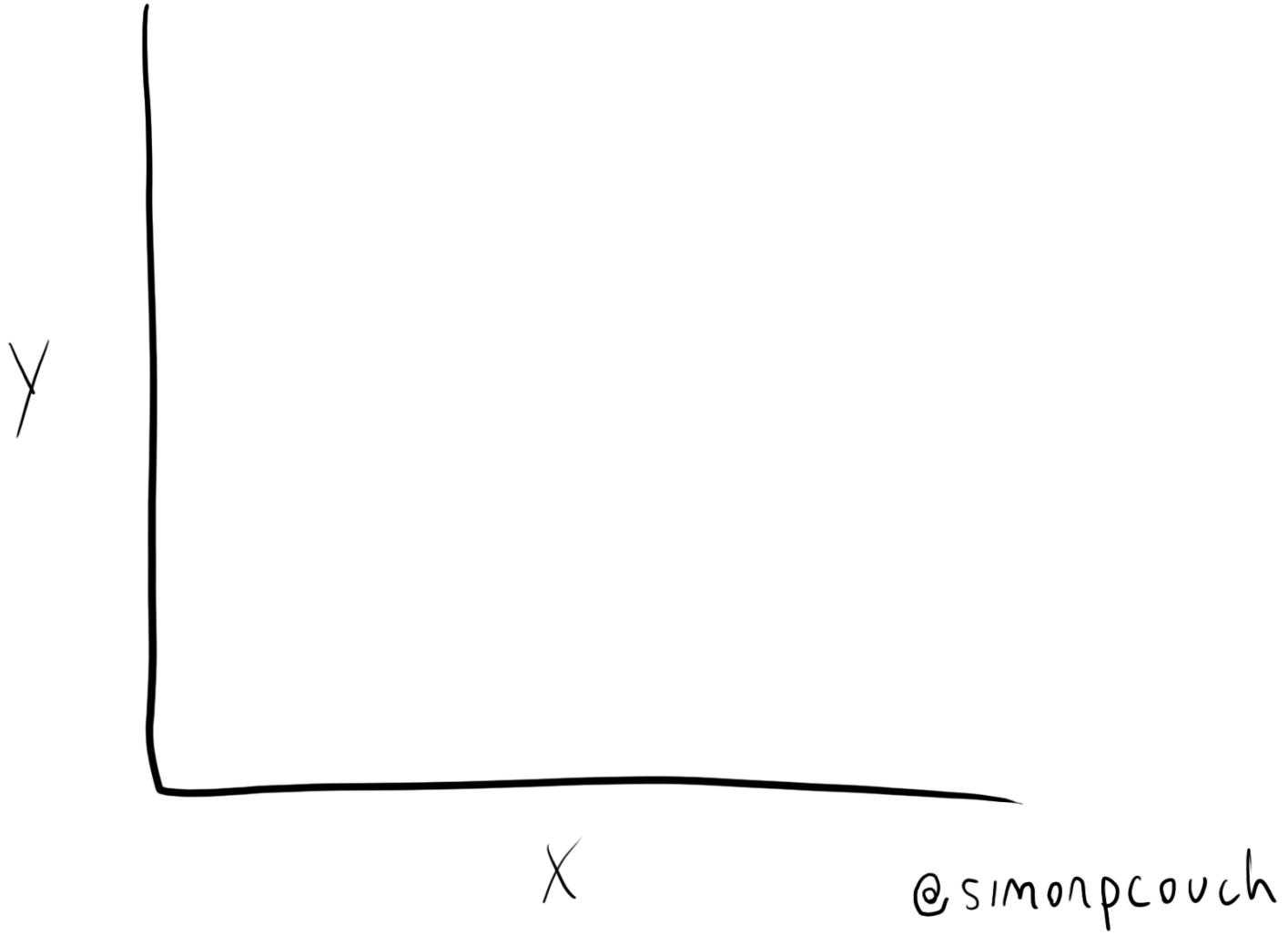
...

@simonpcouch

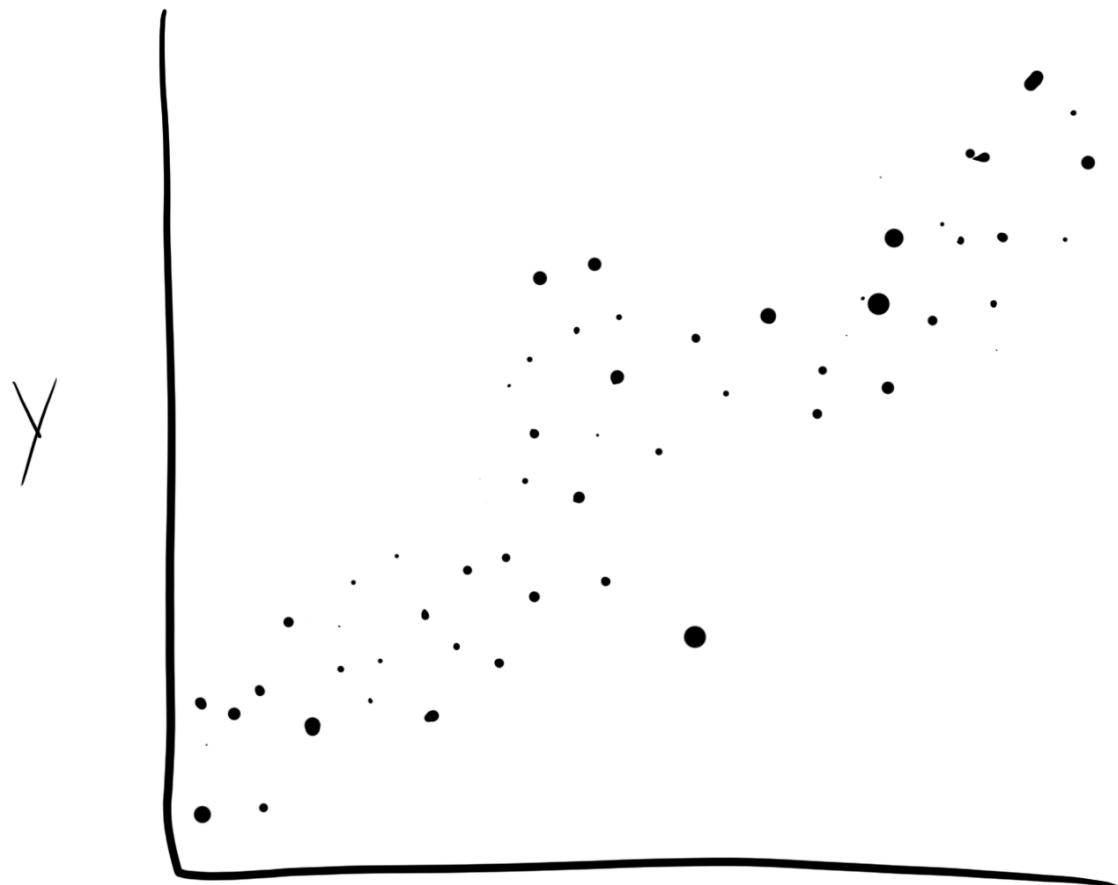
@simonpcouch



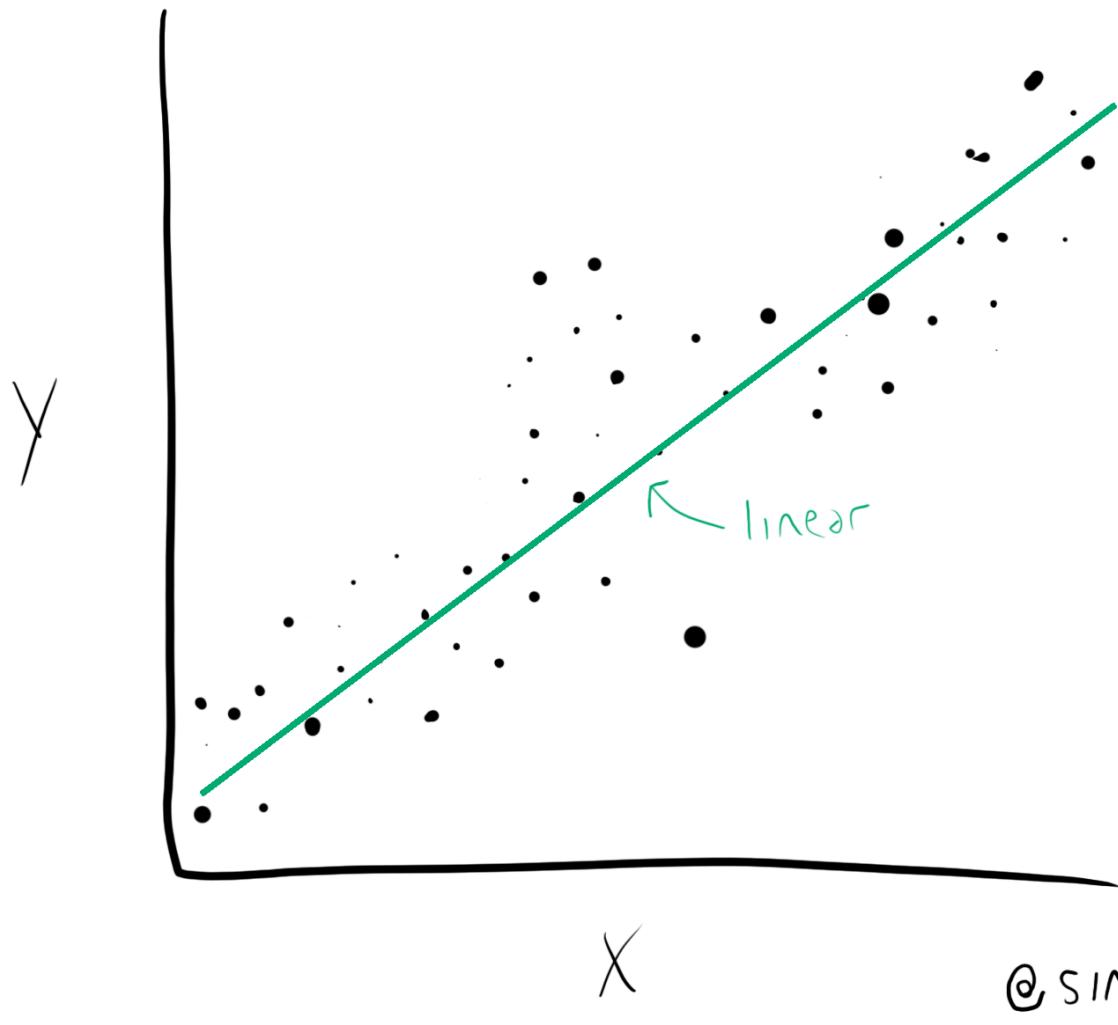
@simonpcouch



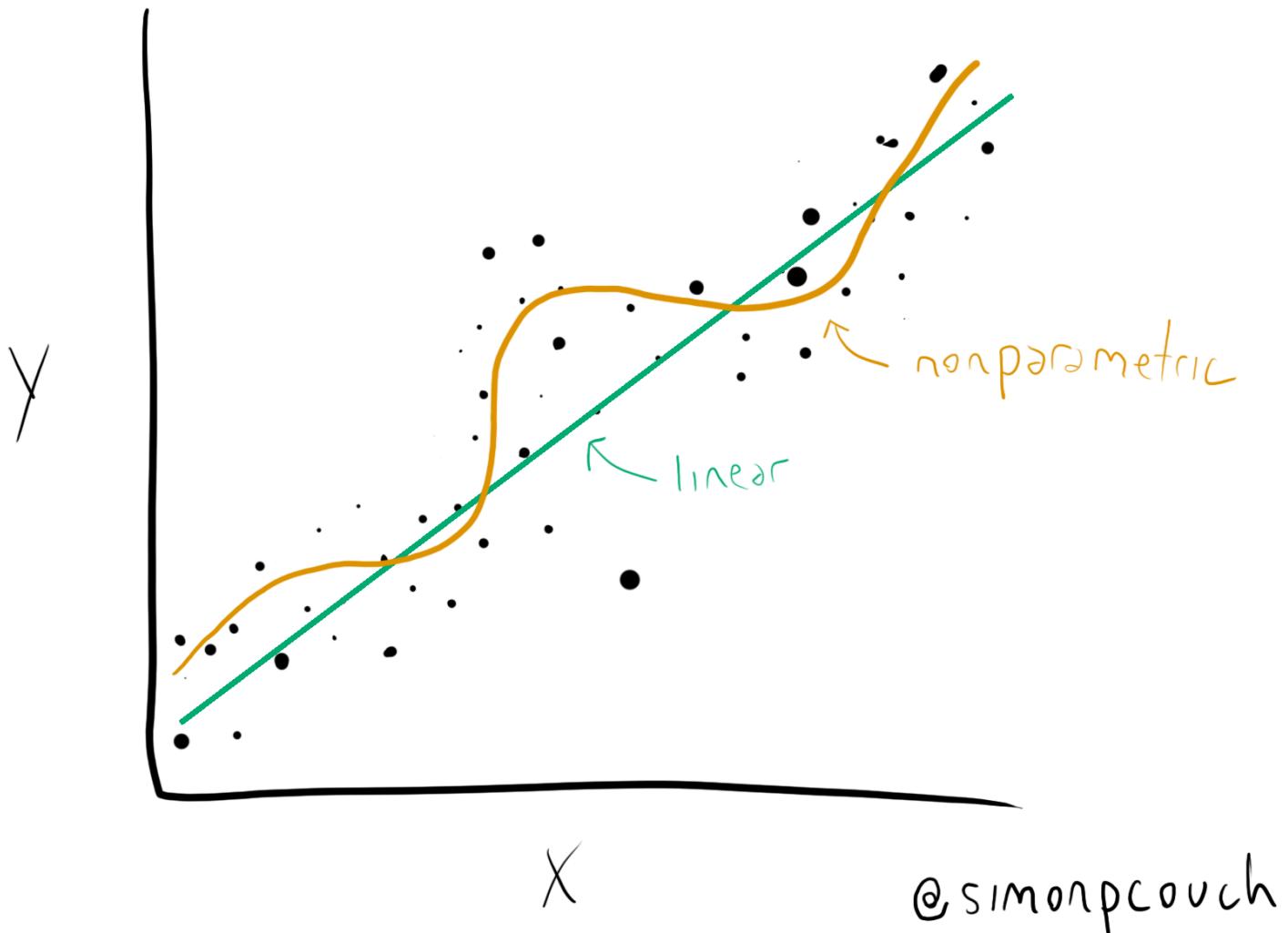
@simonpcouch



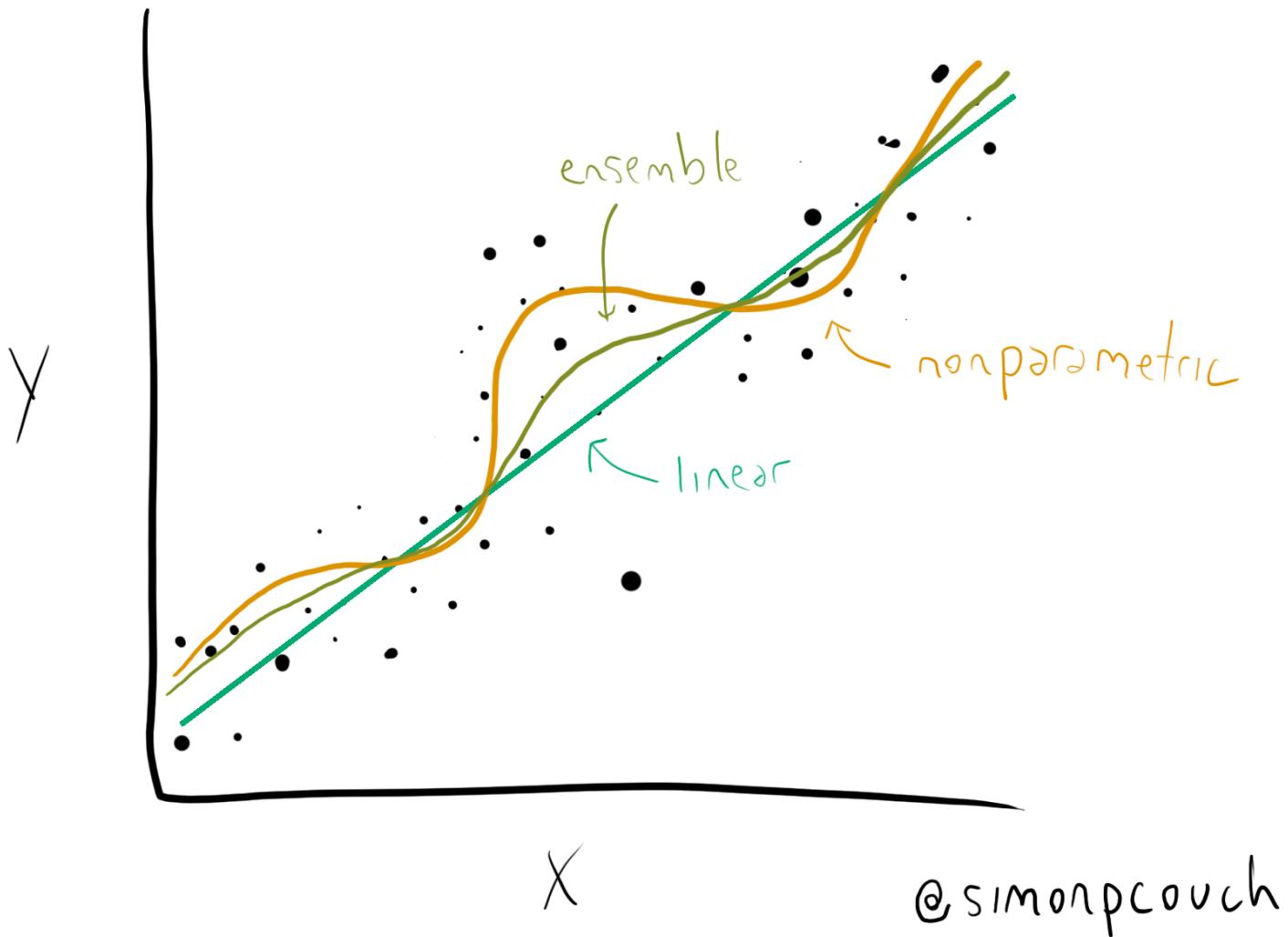
@simonpcouch



@simonpcouch



@simonpcouch



 @simonpcouch

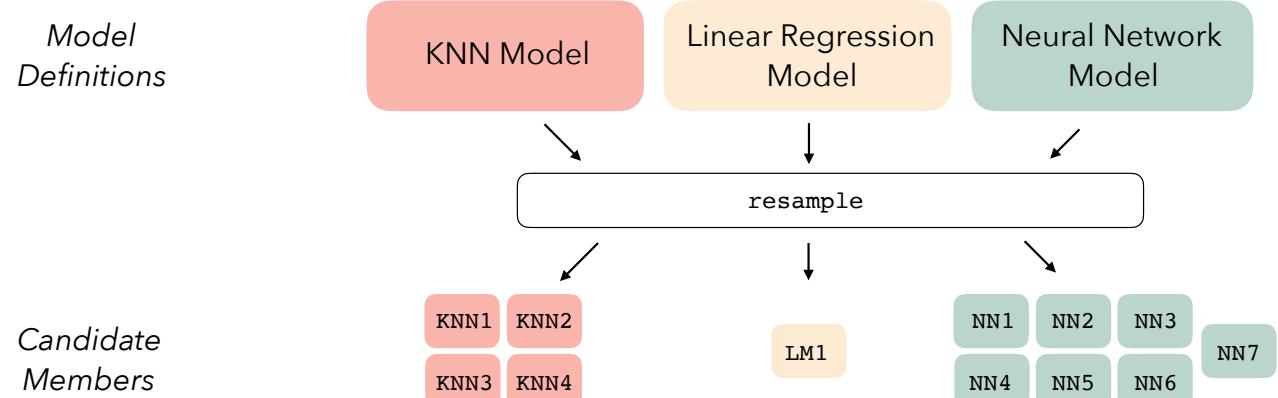
*Model
Definitions*

KNN Model

Linear Regression
Model

Neural Network
Model

1: Resample candidate ensemble members from model definitions



1: Resample candidate ensemble members from model definitions

Model Definitions

KNN Model

Linear Regression Model

Neural Network Model

2: Collect the validation set predictions for each candidate

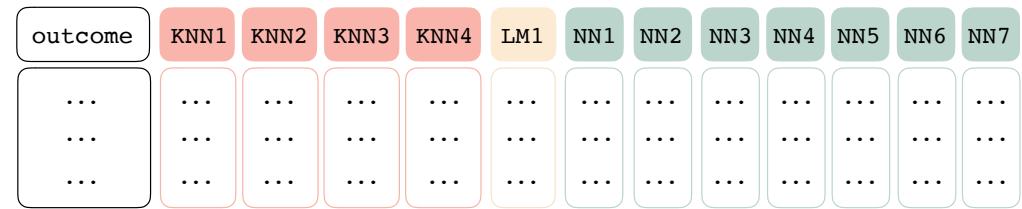
Candidate Members

KNN1 KNN2
KNN3 KNN4

LM1

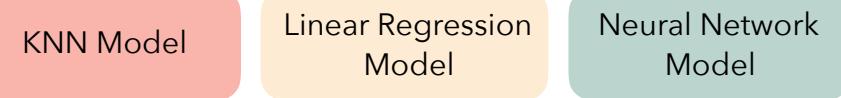
NN1 NN2 NN3
NN4 NN5 NN6
NN7

Data Stack



1: Resample candidate ensemble members from model definitions

Model Definitions



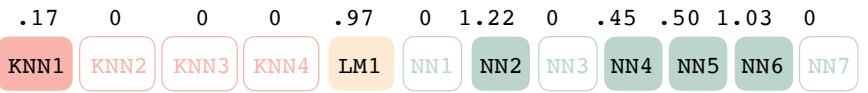
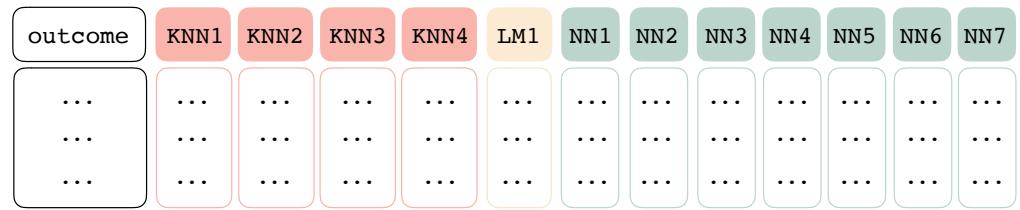
2: Collect the validation set predictions for each candidate

Candidate Members



3: Estimate the stacking coefficients for each candidate

Data Stack



1: Resample candidate ensemble members from model definitions

Model Definitions



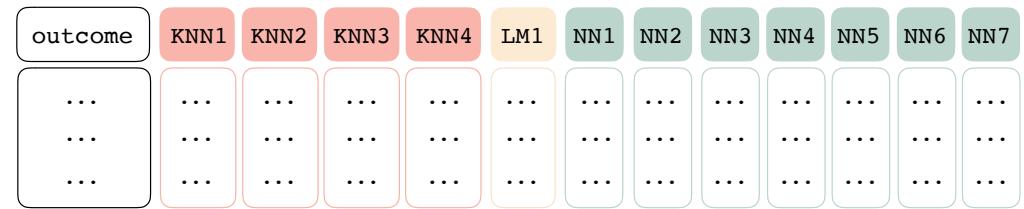
2: Collect the validation set predictions for each candidate

Candidate Members



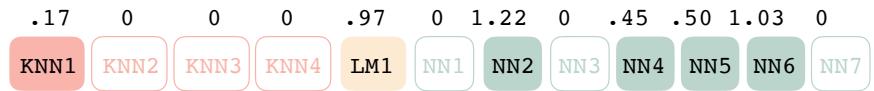
3: Estimate the stacking coefficients for each candidate

Data Stack



4: Train each candidate with nonzero stacking coefficients

Stacking Coefficients



Members



Stacked Ensemble Modeling with Tidy Data Principles

Simon Couch



@simonpcouch

JSM 2021 - John M. Chambers Award Recipient

