

LECTURE 3 TERM 2:

MSIN0097

UCL
SCHOOL OF
MANAGEMENT

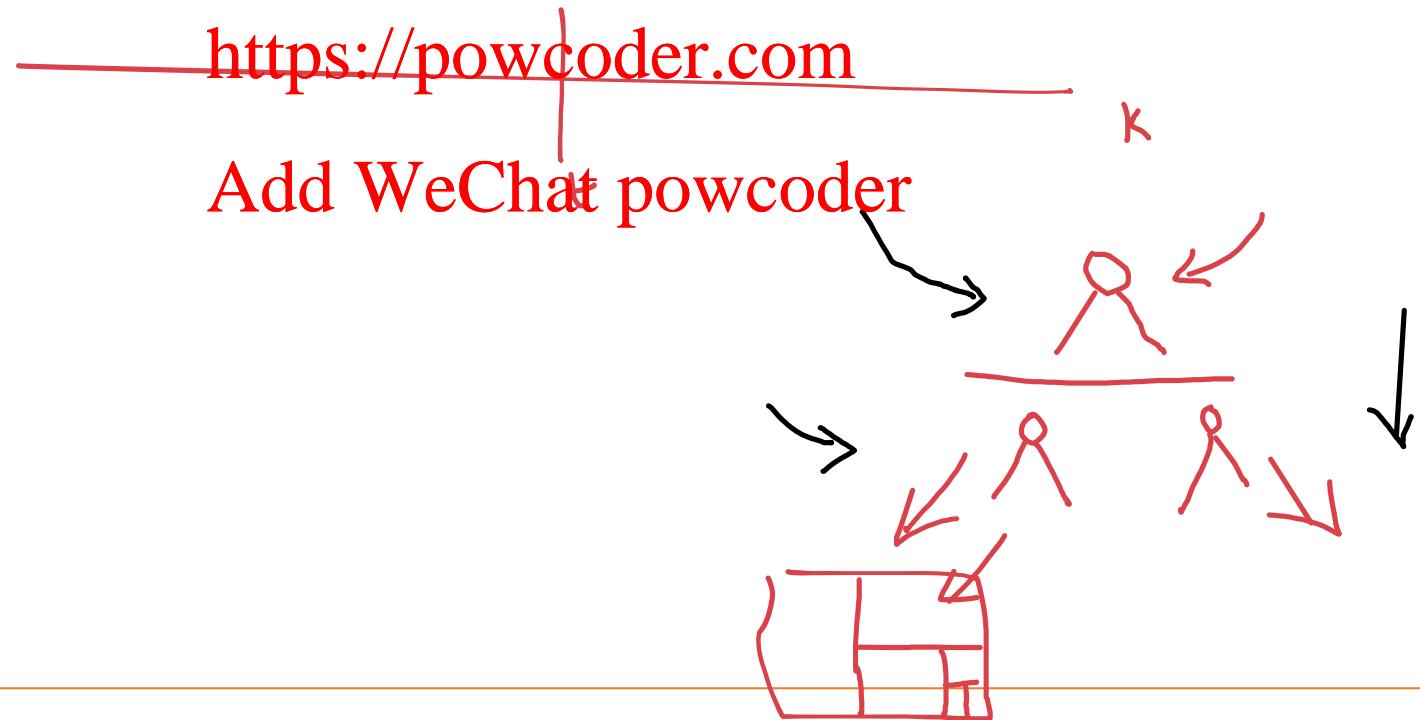
PREDICTIVE ANALYTICS Assignment Project Exam Help

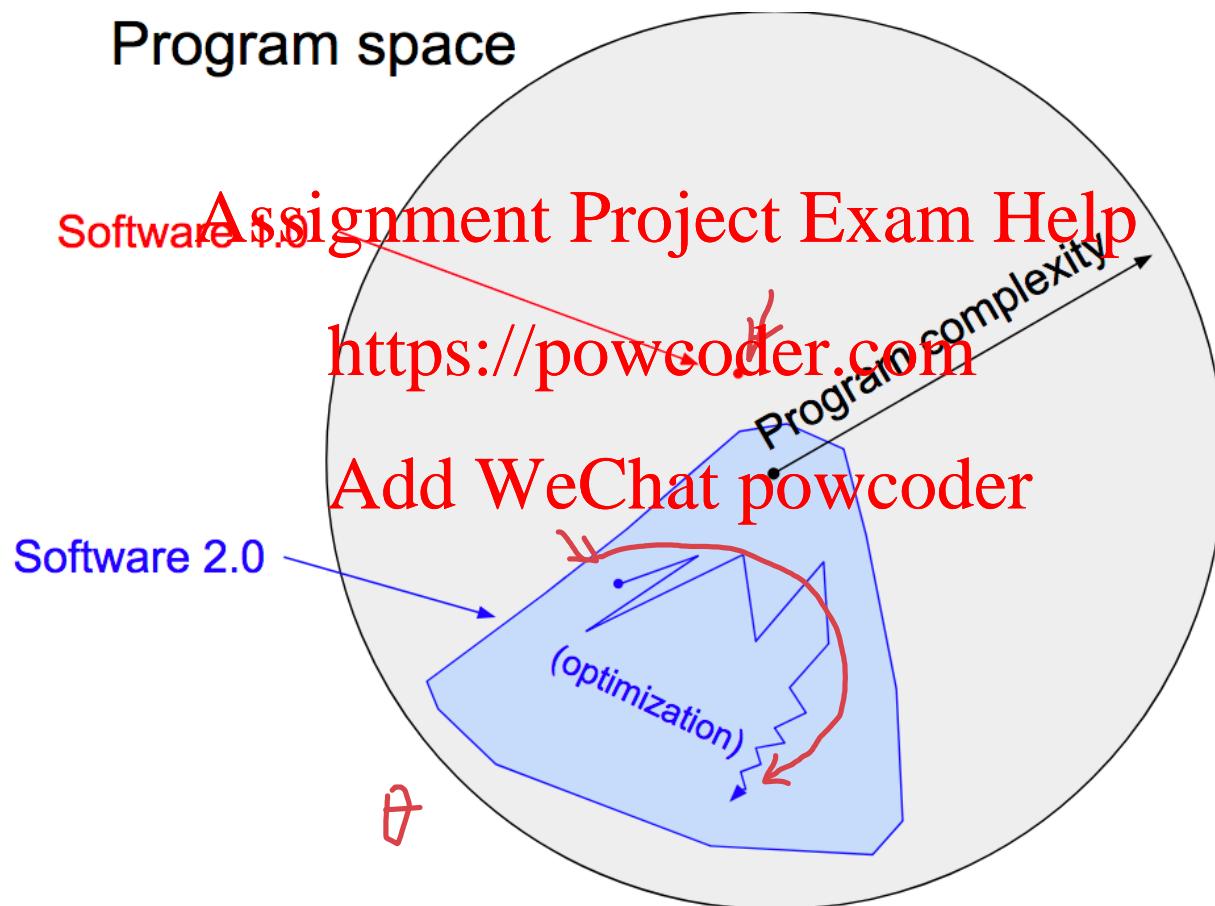
<https://powcoder.com>

Add WeChat powcoder

A P MOORE

Assignment Project Exam Help



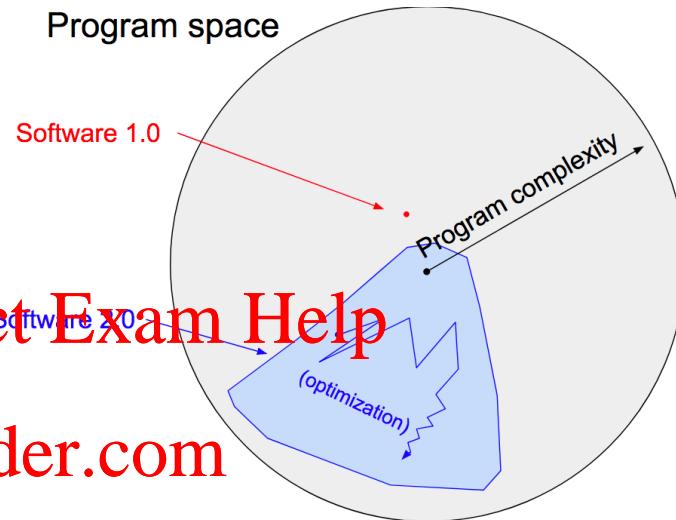
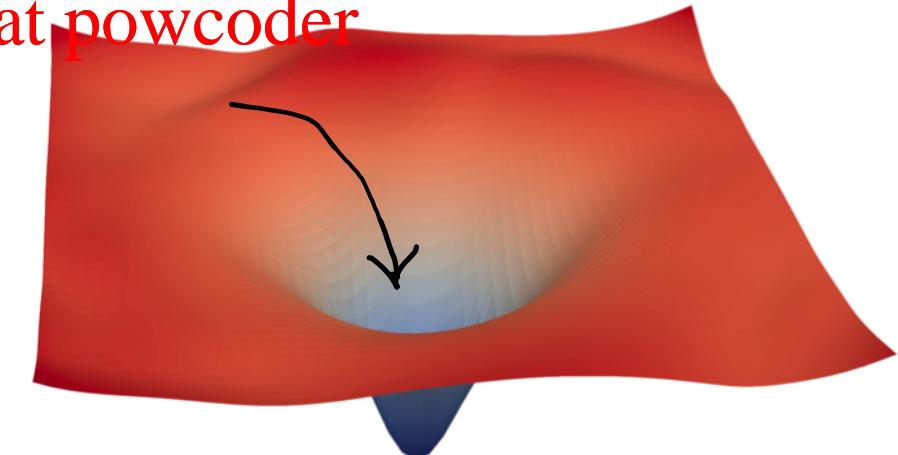
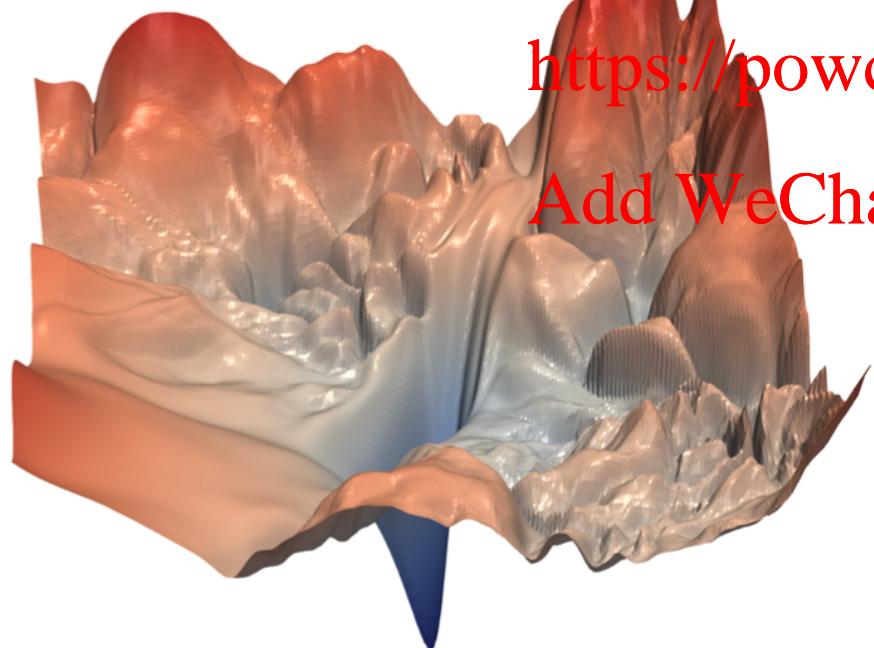


LOSS SURFACES

Assignment Project Exam Help

<https://powcoder.com>

Add WeChat powcoder



— Improving bad solutions

- Start with a bad Solution (weak learner) and improve it
- Build up a better solution by thinking about how partial solutions can support/correct each others mistakes



Assignment Project Exam Help

<https://powcoder.com>

Add WeChat powcoder

- Improving bad solutions
 - Start with a bad Solution (weak learner) and improve it
 - Build up a better solution by thinking about how partial solutions can support/correct each others mistakes

Assignment Project Exam Help
<https://powcoder.com>

- Make the problem simpler
 - Divide and conquer
 - Problem decomposition

I. Improving bad solutions

- Start with a bad Solution (weak learner) and improve it
- Build up a better solution by thinking about how partial solutions can support/correct each others mistakes

Assignment Project Exam Help
<https://powcoder.com>

2. Make the problem simpler

- Divide and conquer
- Problem decomposition

Add WeChat powcoder

3. Building much better solutions

- Deep models



HOW DO WE FIND THE RIGHT ANSWER?

Assignment Project Exam Help

<https://powcoder.com>

Add WeChat powcoder

HOW DO WE FIND THE RIGHT ANSWER?

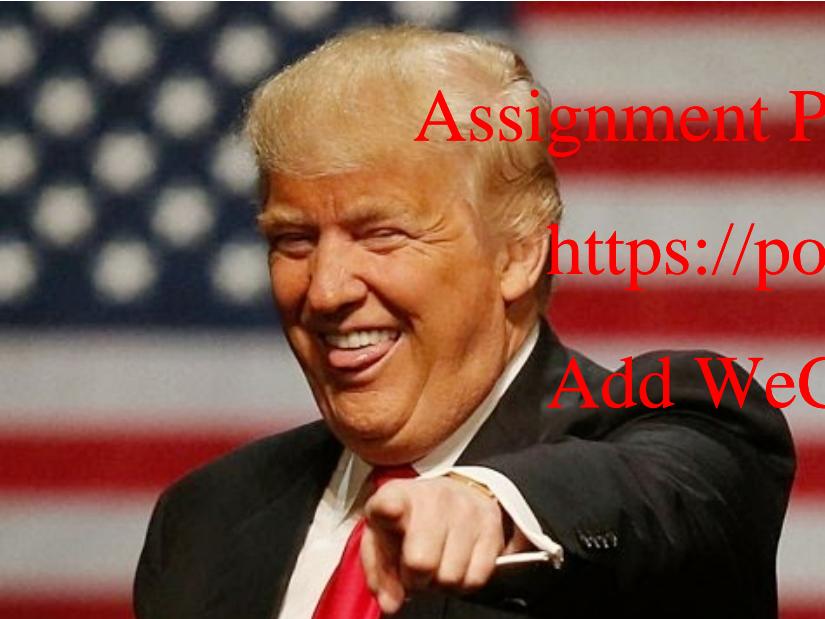


Assignment Project Exam Help

<https://powcoder.com>

Add WeChat powcoder

CORRECT?



Assignment Project Exam Help

<https://powcoder.com>

Add WeChat powcoder



- Discover
- Explore
- Visualize

- Clean
- Sample
- Impute
- Encode
- Transform

- Documentation
- Presentation

Assignment Project Exam Help

<https://powcoder.com>

Add WeChat powcoder
Modeling ^{Scale}

- Overfitting
- Optimization
- Model Selection
- Regularization
- Generalization



- Discover
- Explore
- Visualize

- Clean
- Sample
- Impute
- Encode
- Transform

- Documentation
- Presentation

Assignment Project Exam Help

<https://powcoder.com>

Add WeChat powcoder

- Voting
- Bagging
- Boosting
- Pasting
- Stacking



- Discover
- Explore
- Visualize
- Clean
- Sample
- Import
- Encode
- Transform
- Scale
- Modeling
- Overfitting
- Optimization
- Model Selection
- Regularization
- Generalization

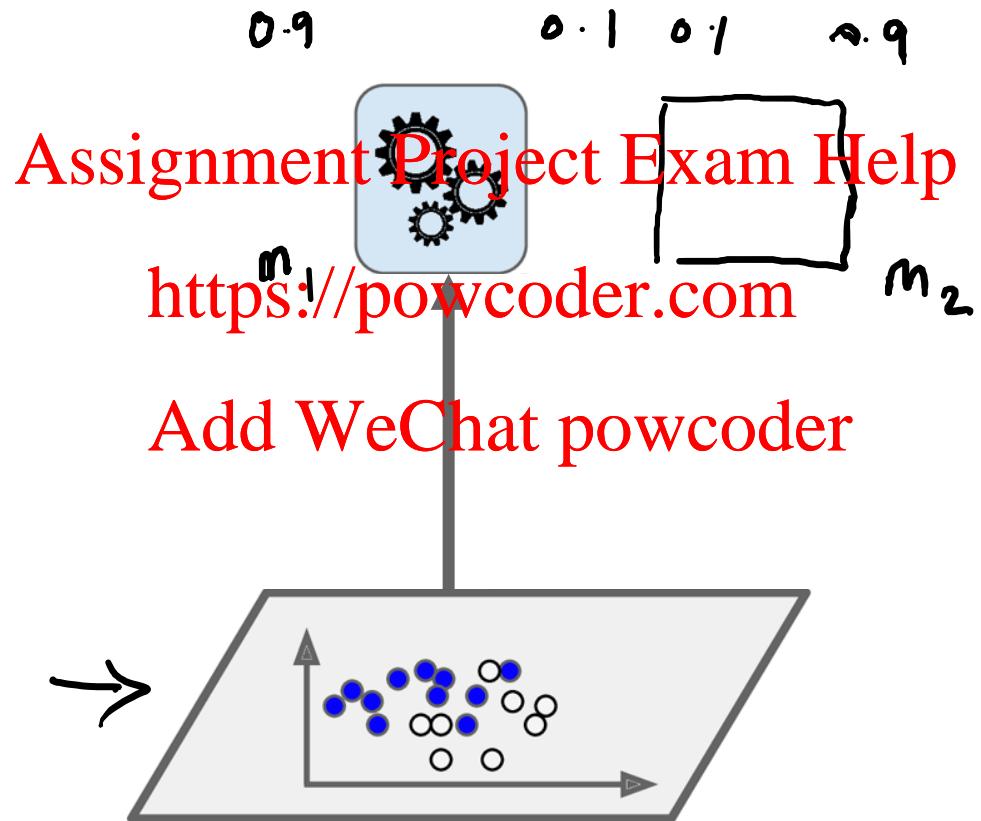
Assignment Project Exam Help

<https://powcoder.com>

Add WeChat powcoder

- Voting
 - Majority voting
 - Bagging and Pasting
- Out-of-bag evaluation
- Boosting
 - Adaptive Boosting (Adaboost)
 - Gradient Boosting
 - XGBoost
- Stacking

MULTIPLE MODELS



ENSEMBLES

IMPROVING BAD SOLUTIONS

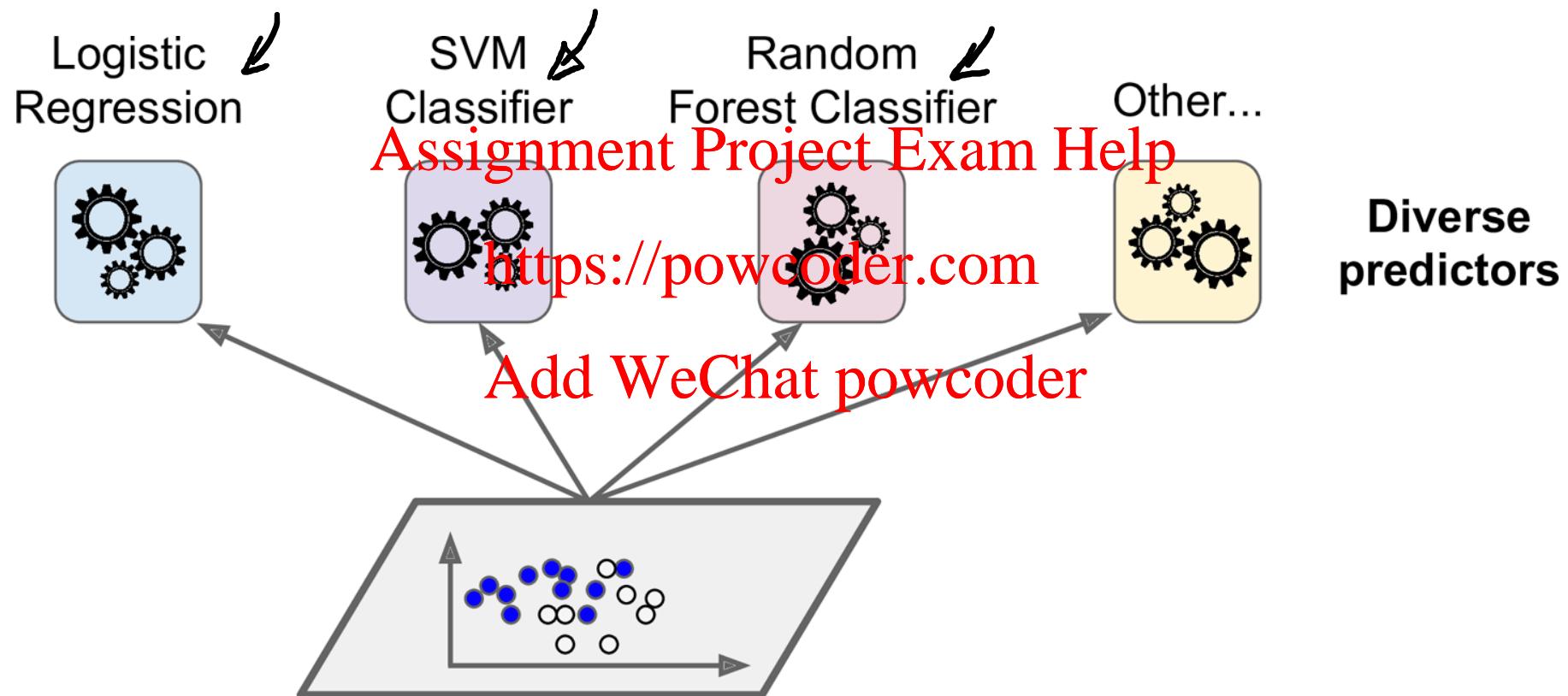
- Voting
- Bagging (bootstrap aggregating) and Pasting (non-replacement)
- Boosting
- Stacking

Assignment Project Exam Help

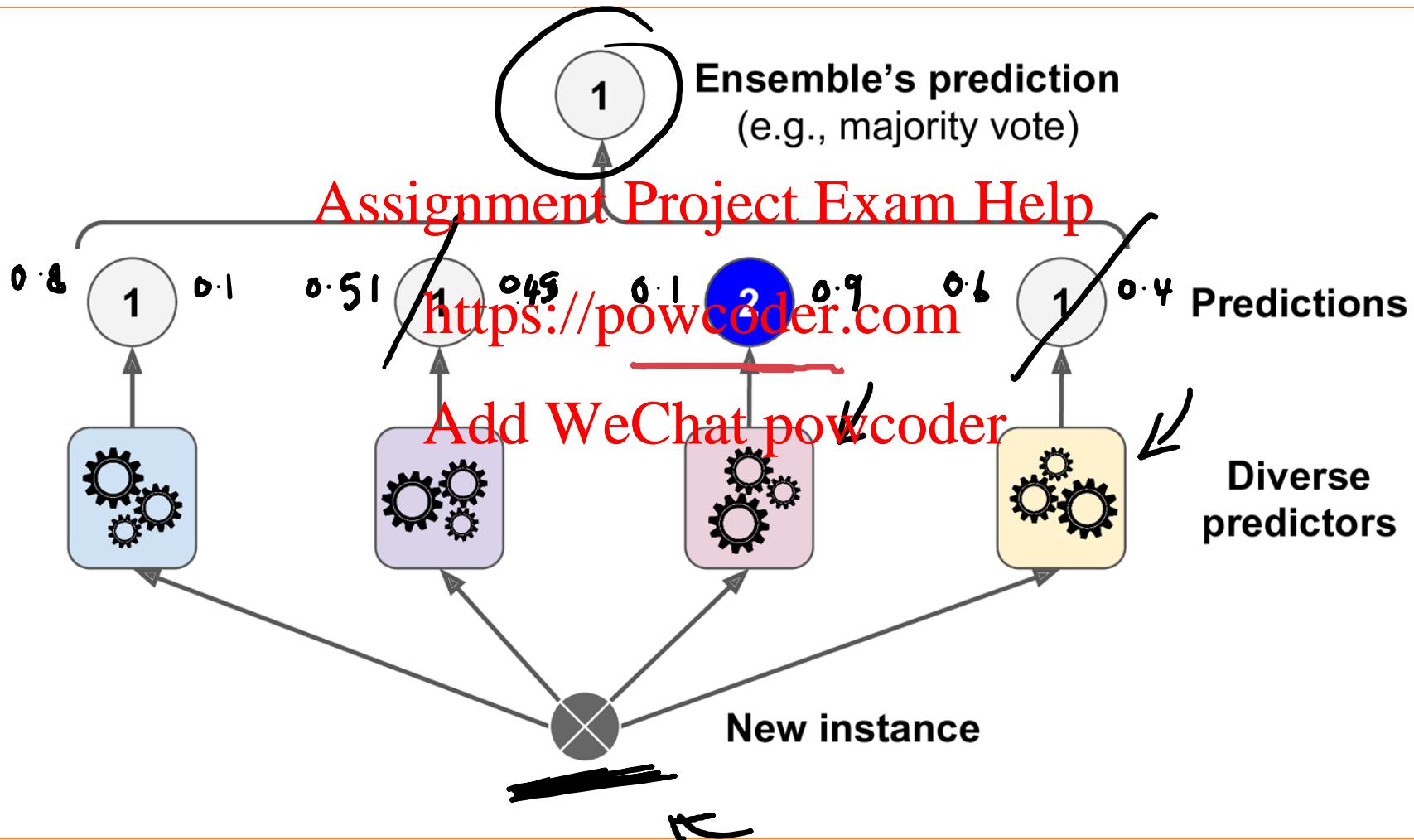
<https://powcoder.com>

Add WeChat powcoder

MULTIPLE MODELS



MAJORITY VOTING

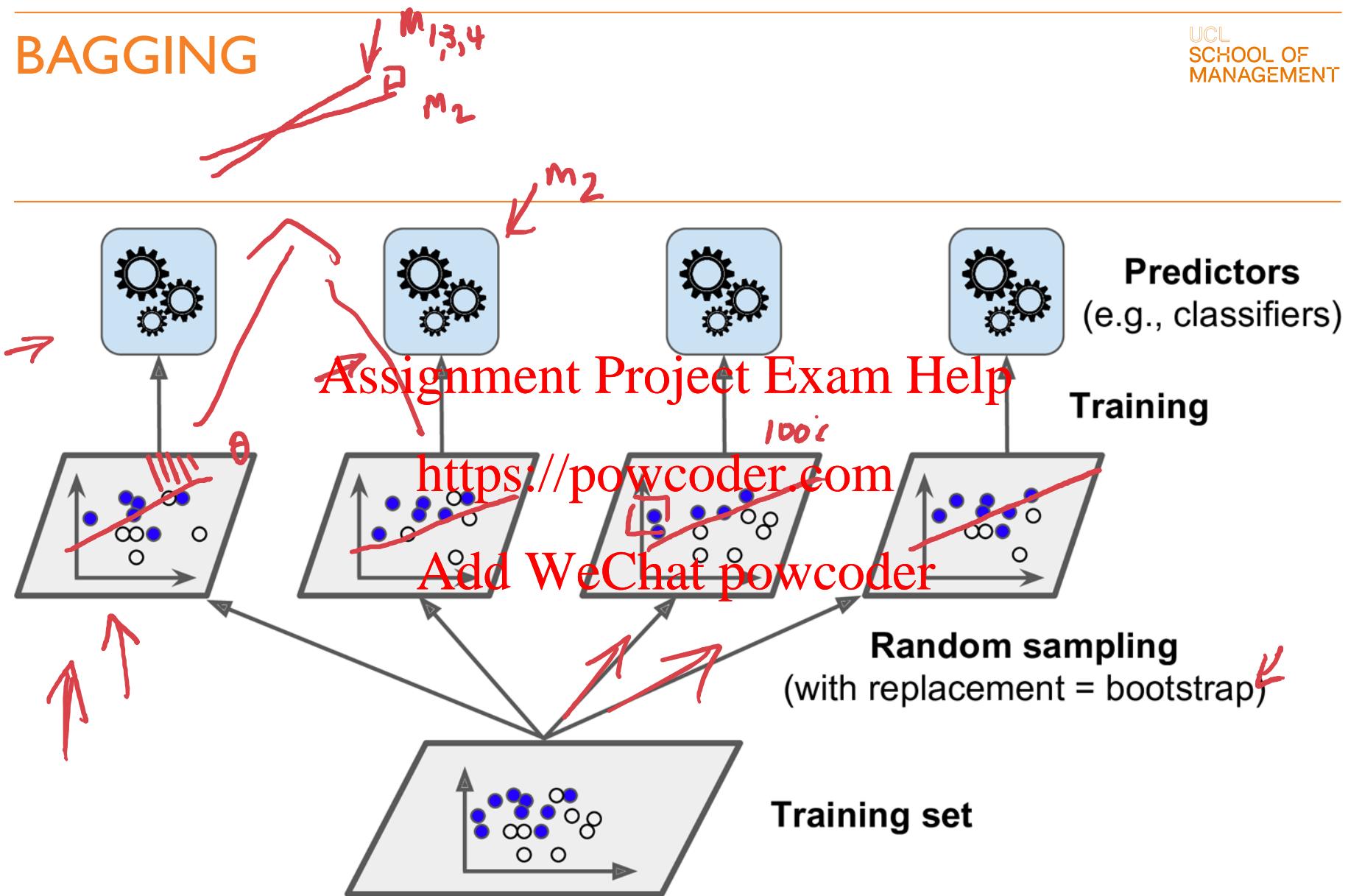


Assignment Project Exam Help

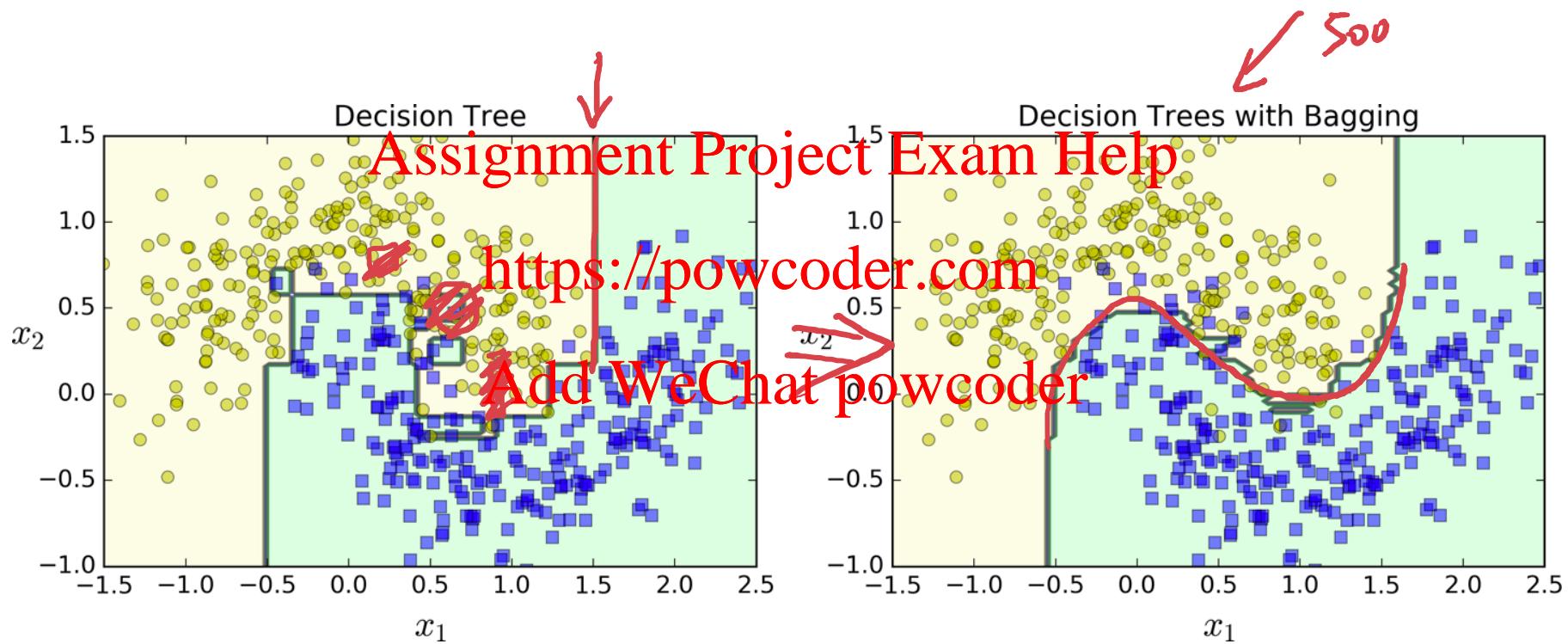
<https://powcoder.com>
Partitioning data

Add WeChat powcoder

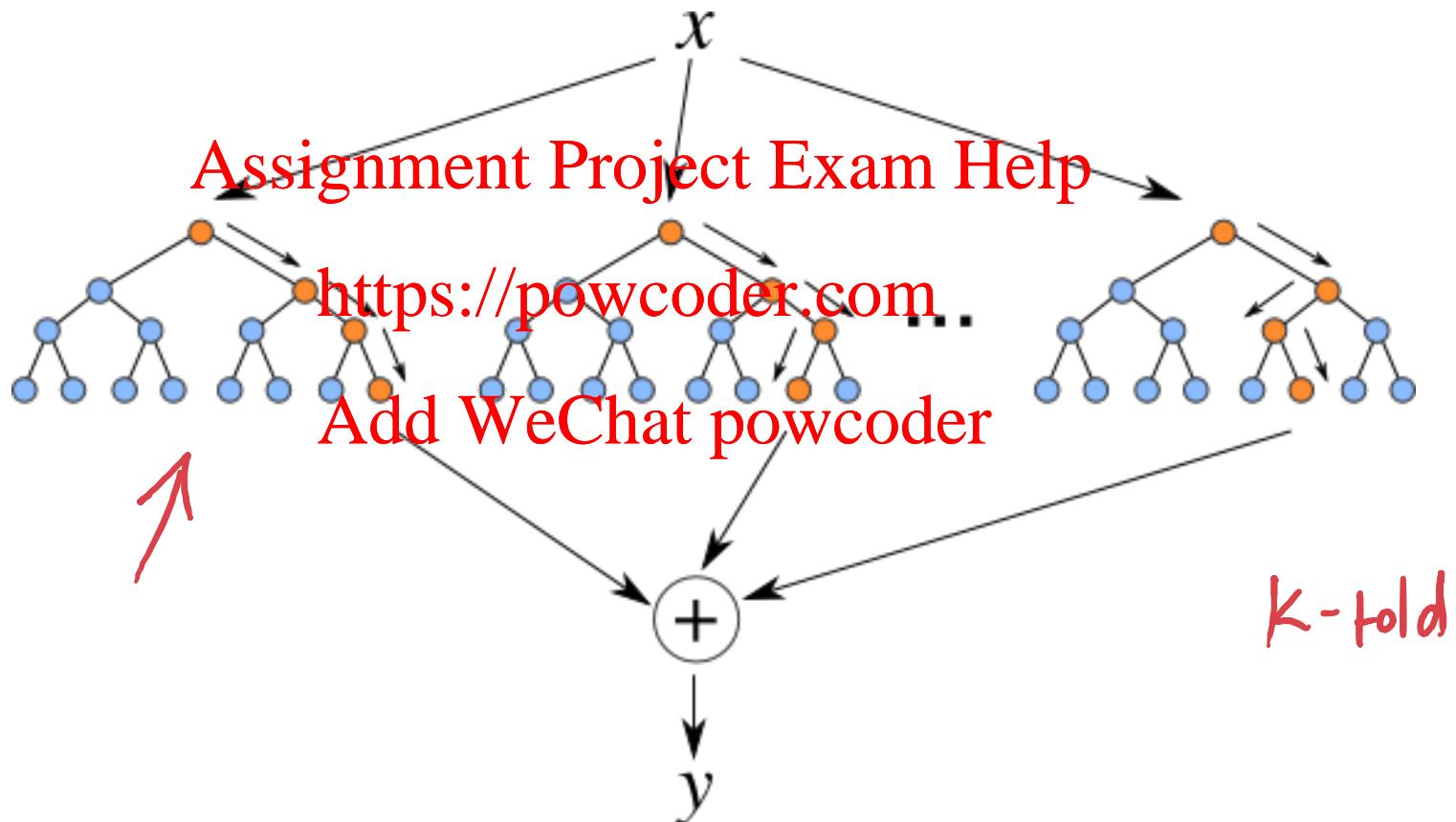
BAGGING



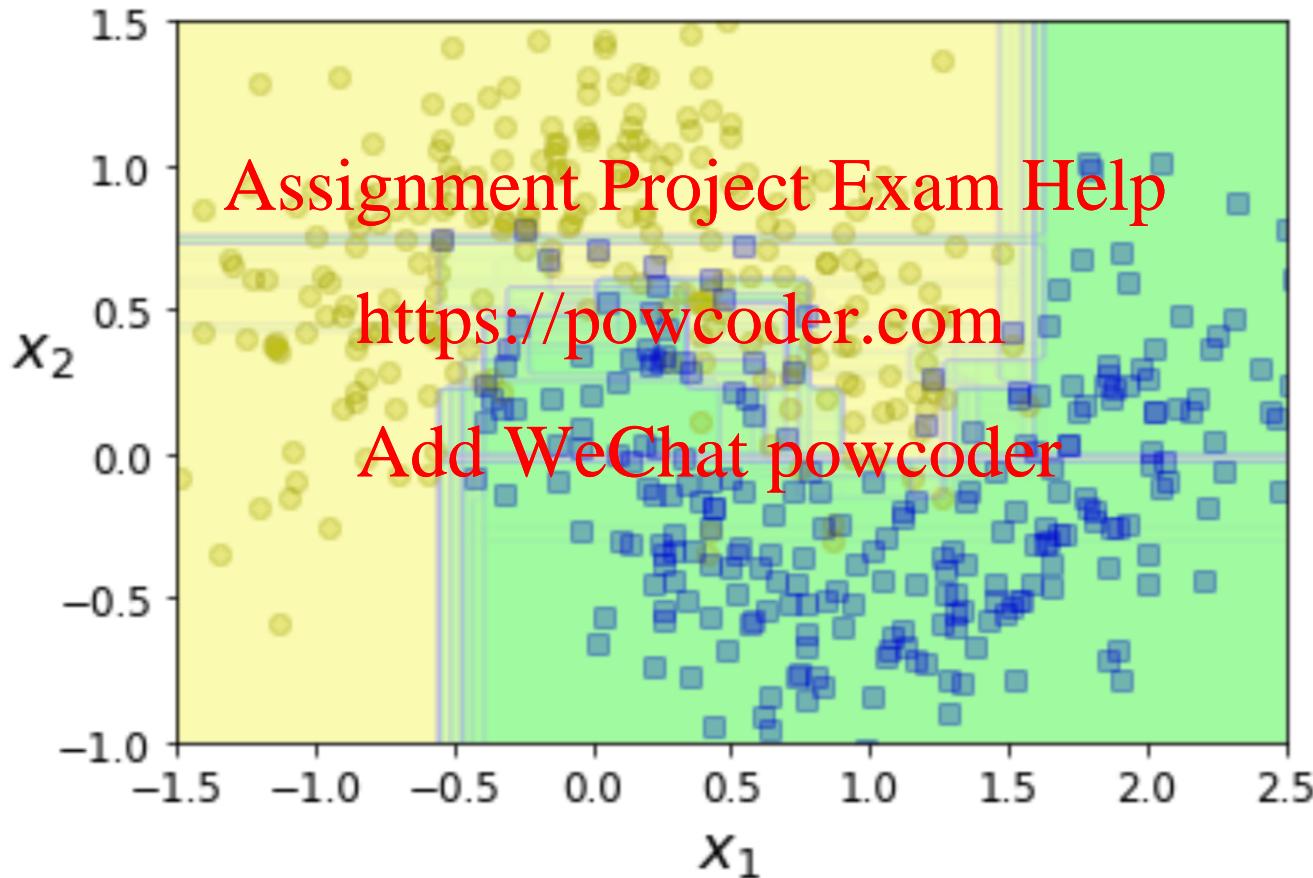
BAGGING



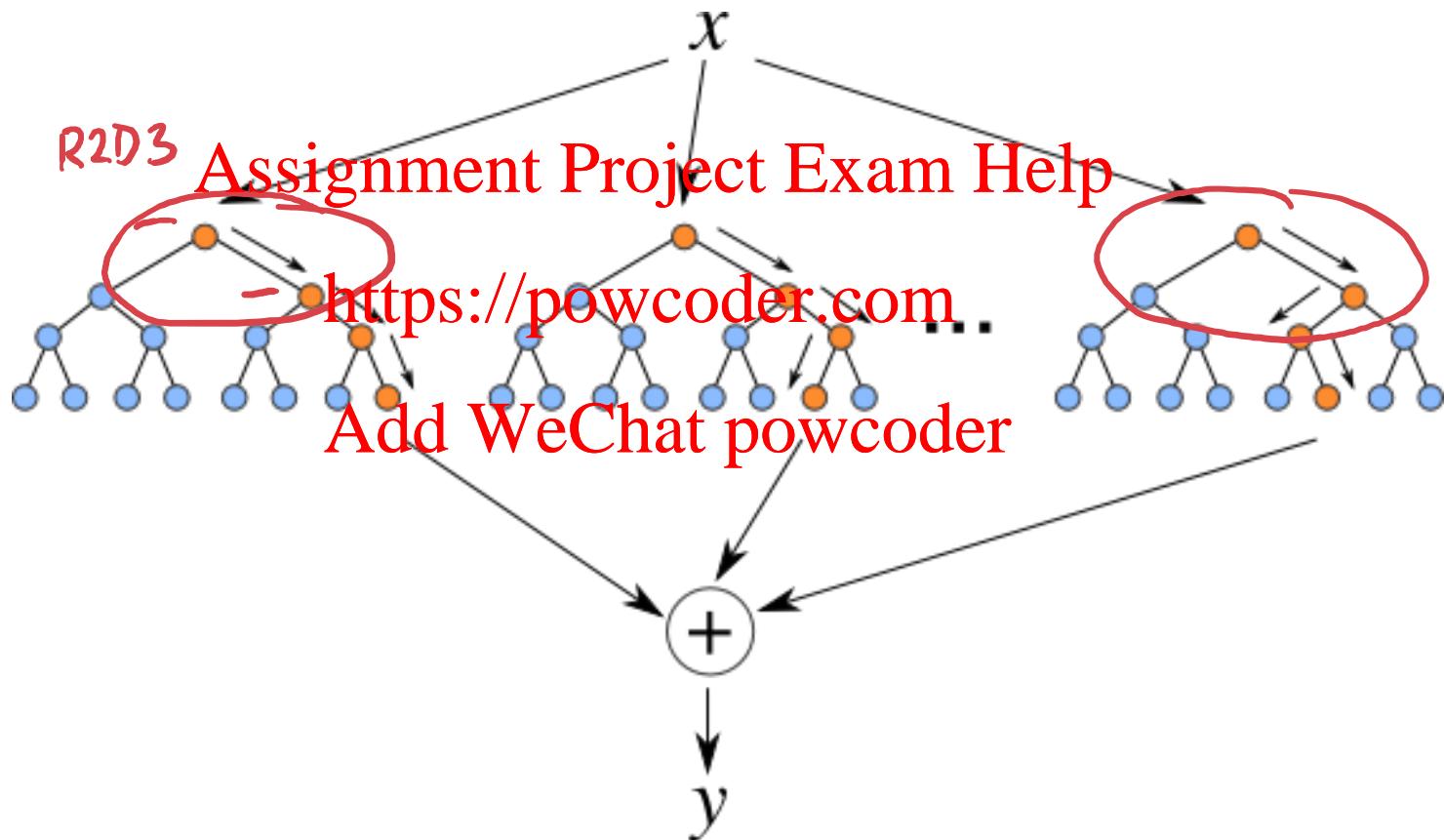
RANDOM FORESTS



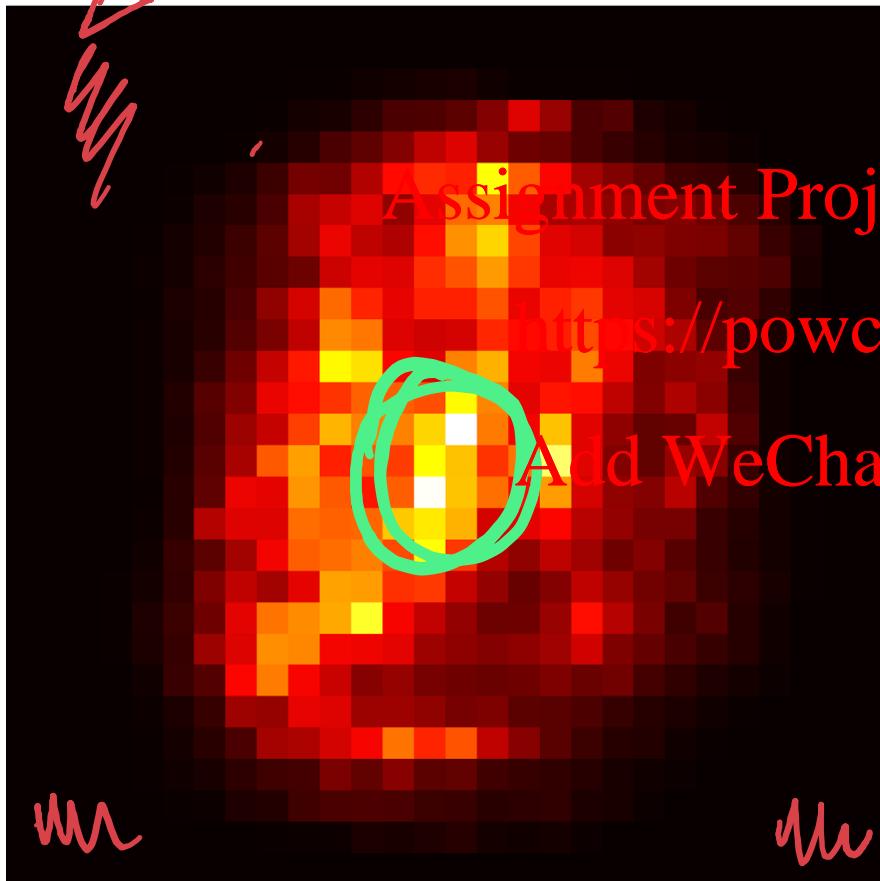
DECISION BOUNDARIES



RANDOM FORESTS



FEATURE IMPORTANCE



Very important

Not important



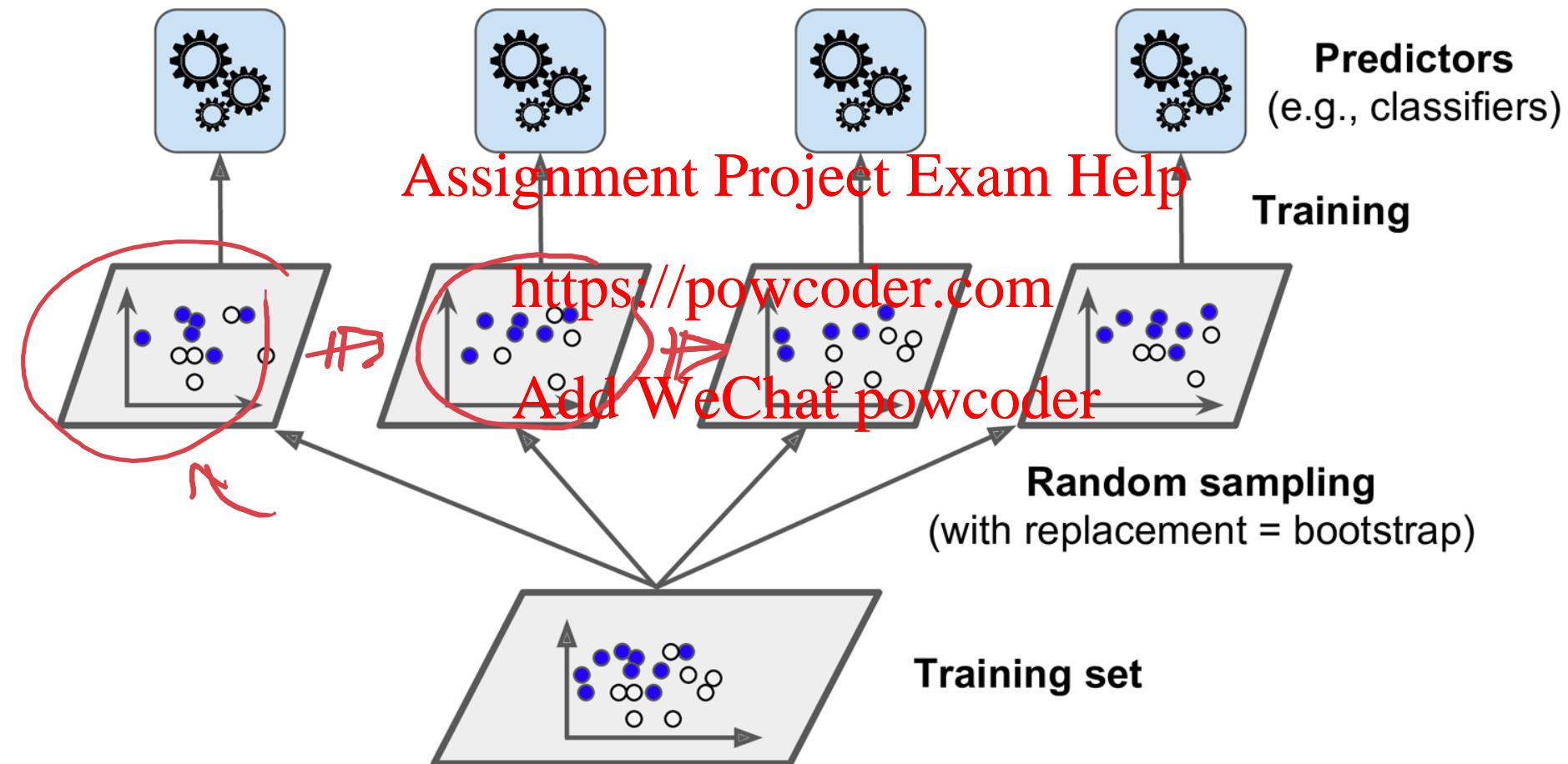
5	0	4	1	9	2	1	3	1	4
3	5	3	6	1	7	2	8	6	9
4	0	9	1	1	2	4	3	2	7
3	8	6	9	0	5	6	0	7	6
1	8	7	9	3	9	8	5	9	3
3	0	7	4	9	8	0	9	4	1
4	4	6	0	4	5	6	1	0	0
1	7	1	6	3	0	2	1	1	7
9	0	2	6	7	8	3	9	0	4
6	7	4	6	8	0	7	8	3	1



Assignment Project Exam Help

<https://powcoder.com>
Sequential data
Add WeChat powcoder

SEQUENTIAL MODELS



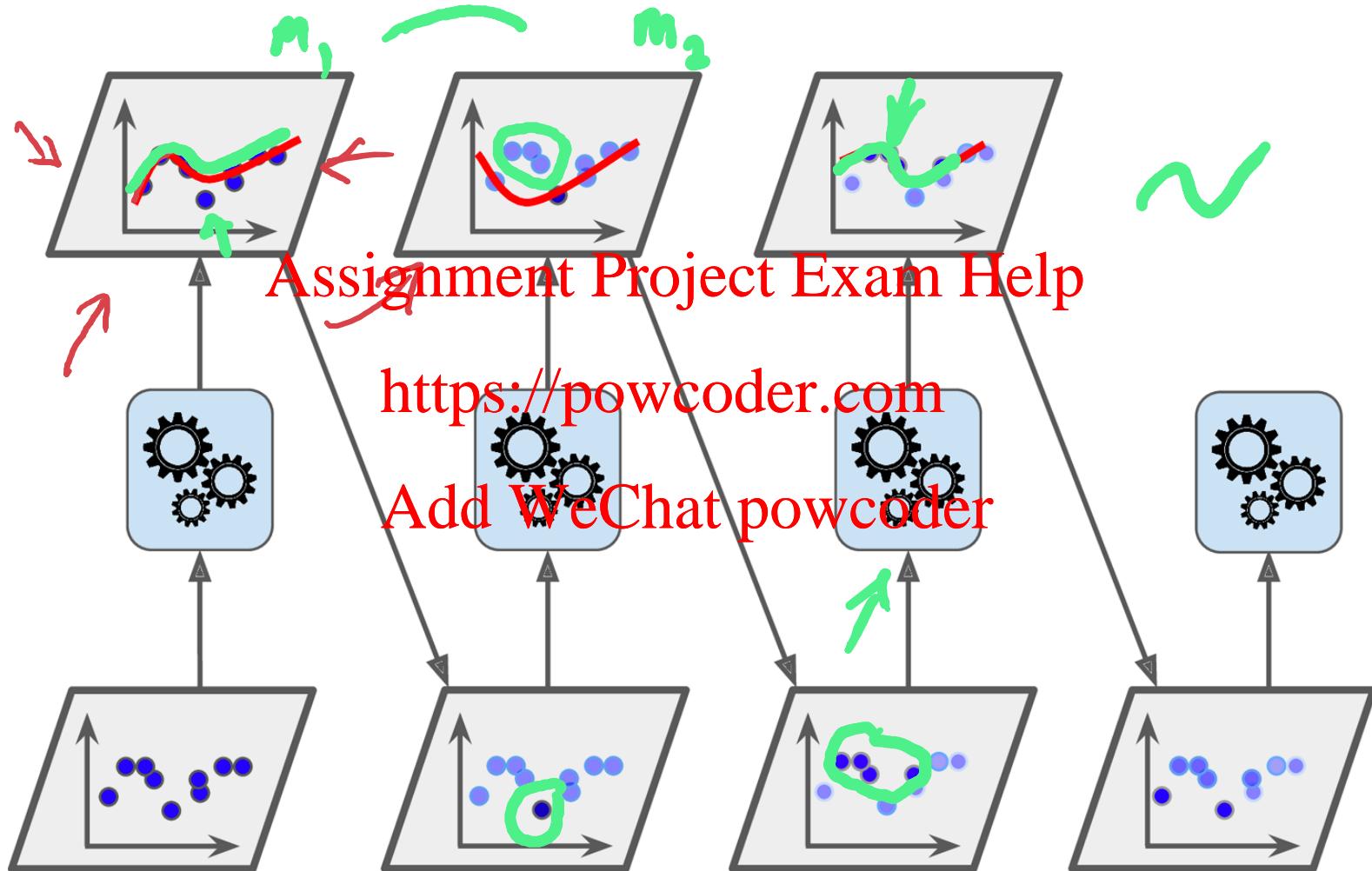
Assignment Project Exam Help

<https://powcoder.com>
Boosting data

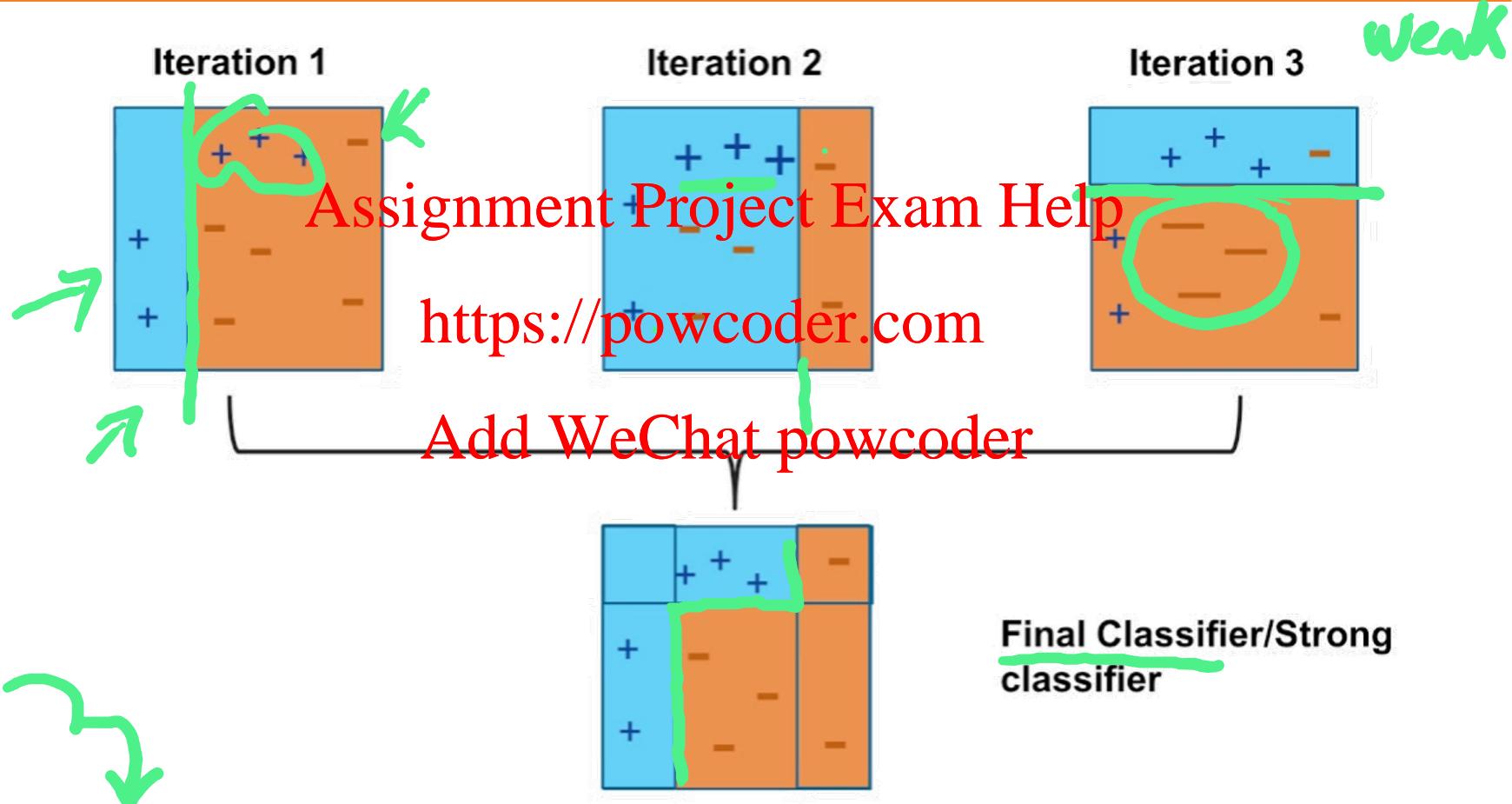
Add WeChat powcoder

~ 2001

BOOSTING

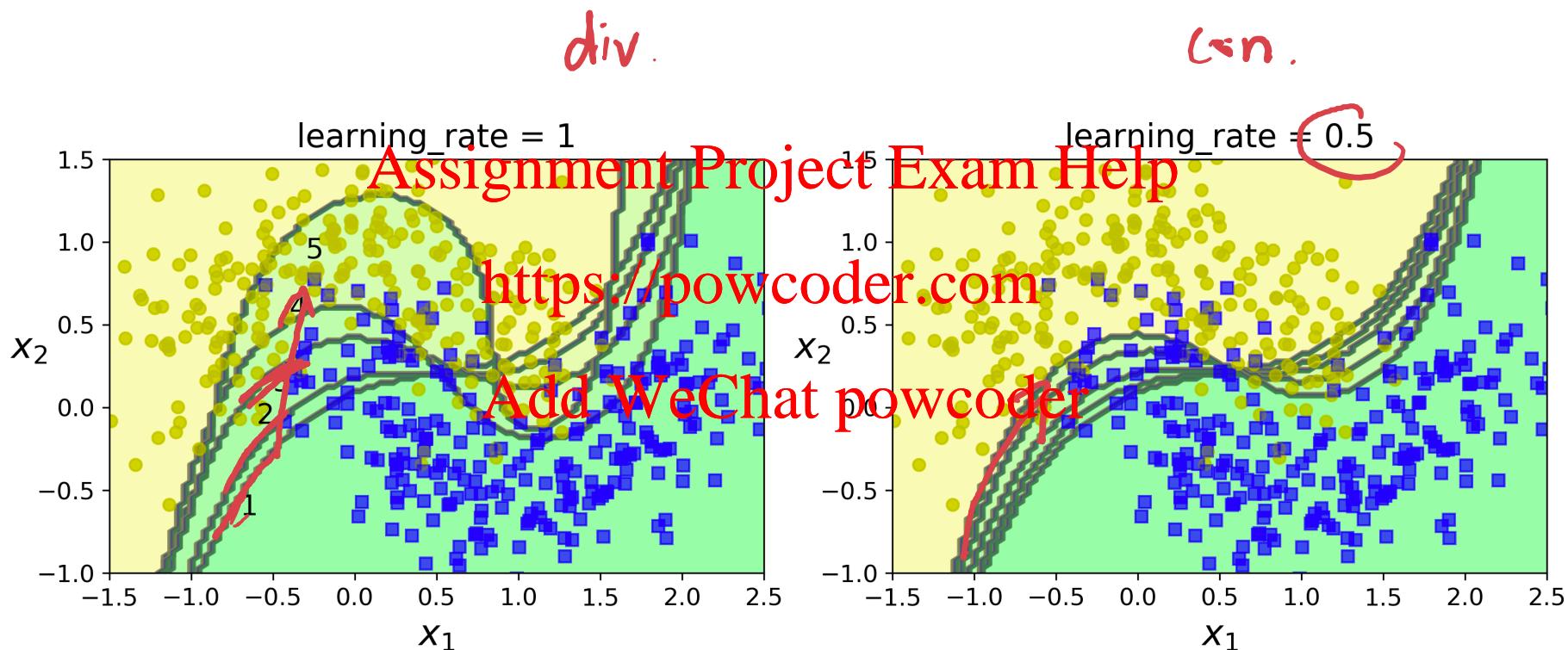


ADABOOST

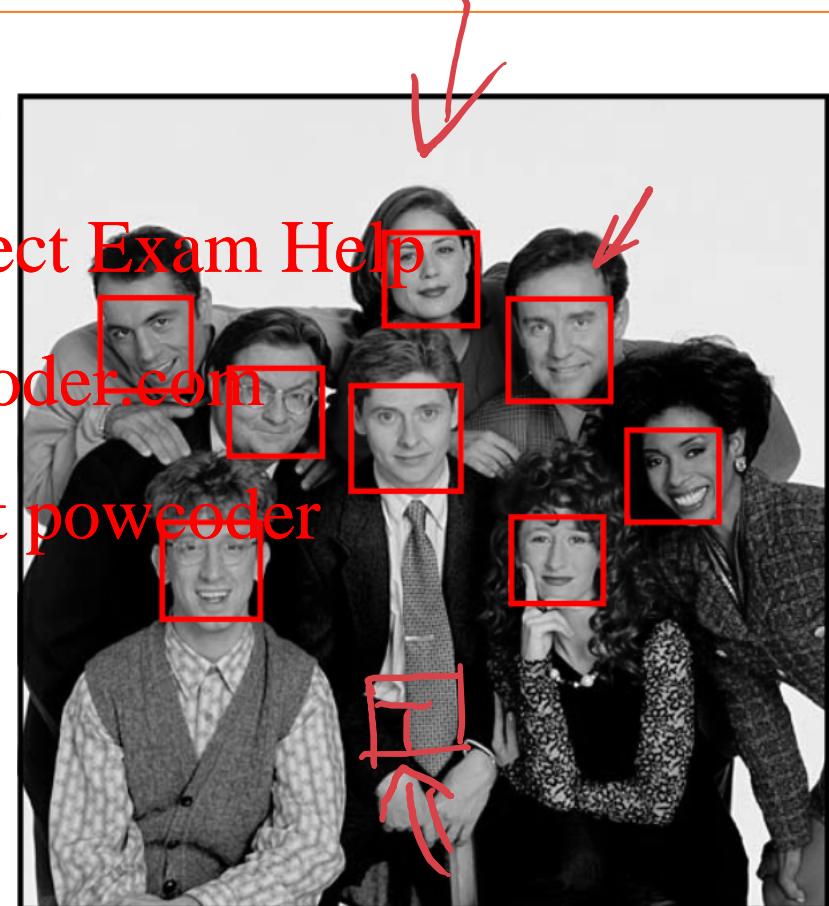
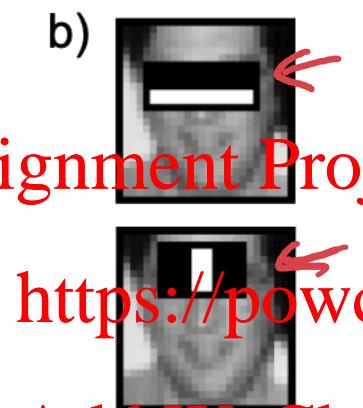
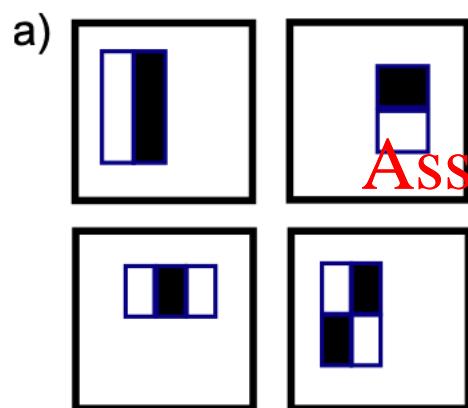


DECISION BOUNDARIES

CONSECUTIVE PREDICTORS



100



Assignment Project Exam Help

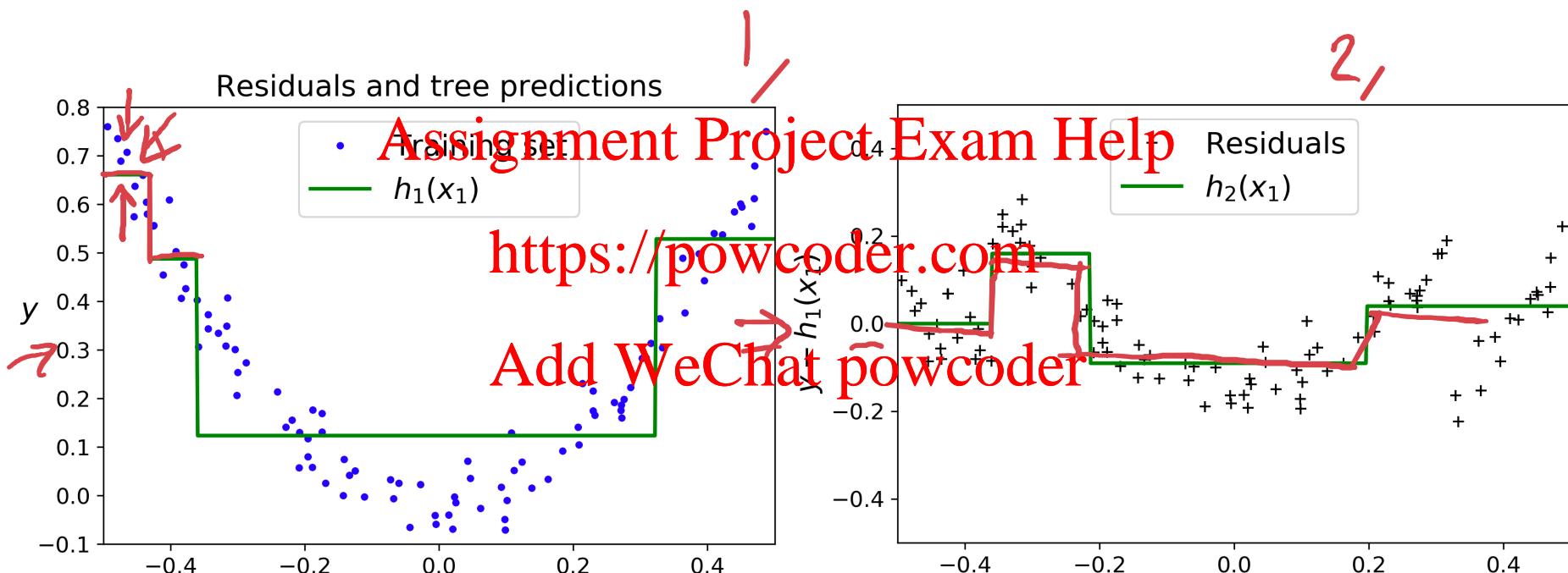
<https://powcoder.com>

Gradient Boosting

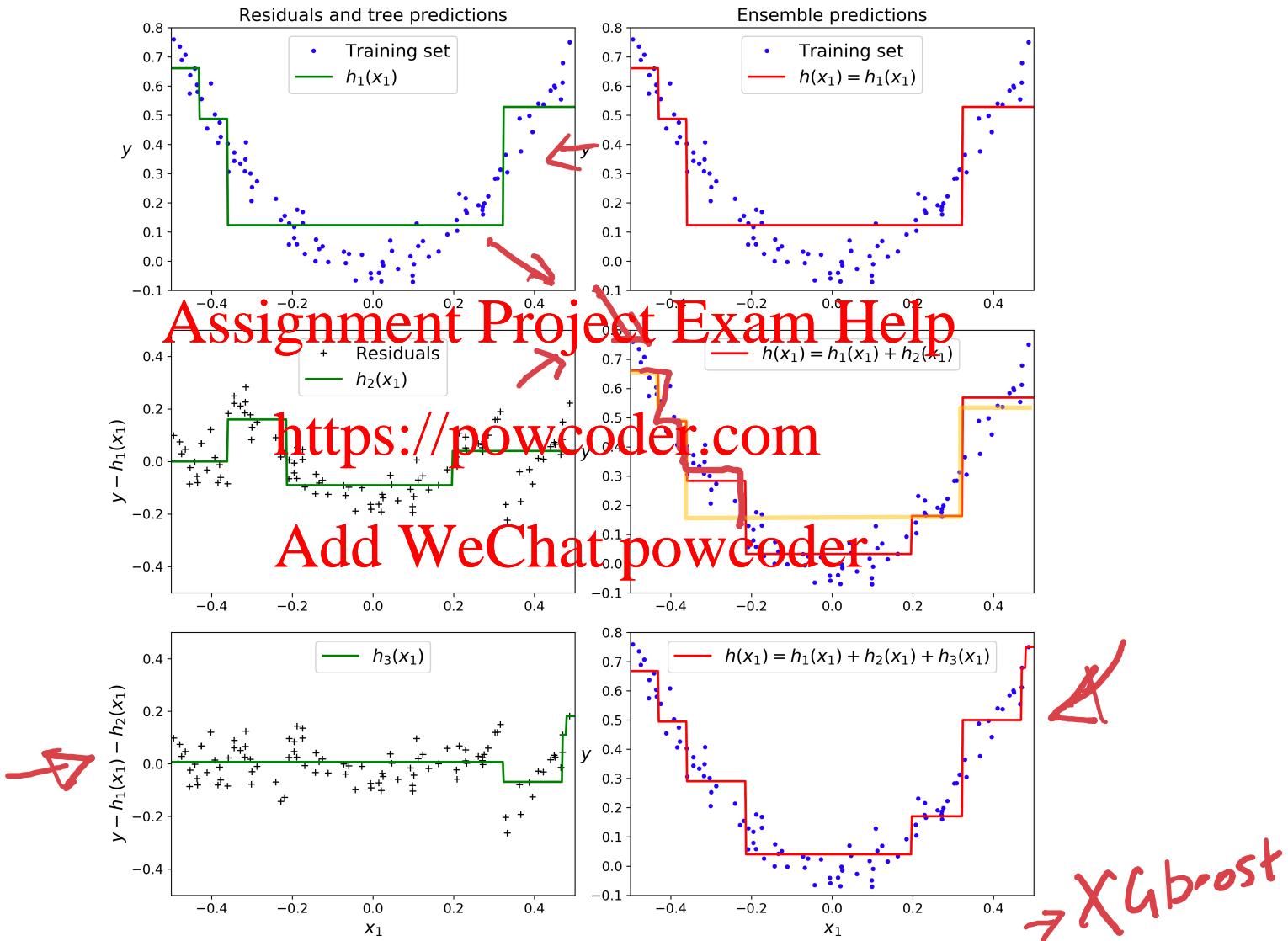
Add WeChat powcoder

GRADIENT BOOSTING

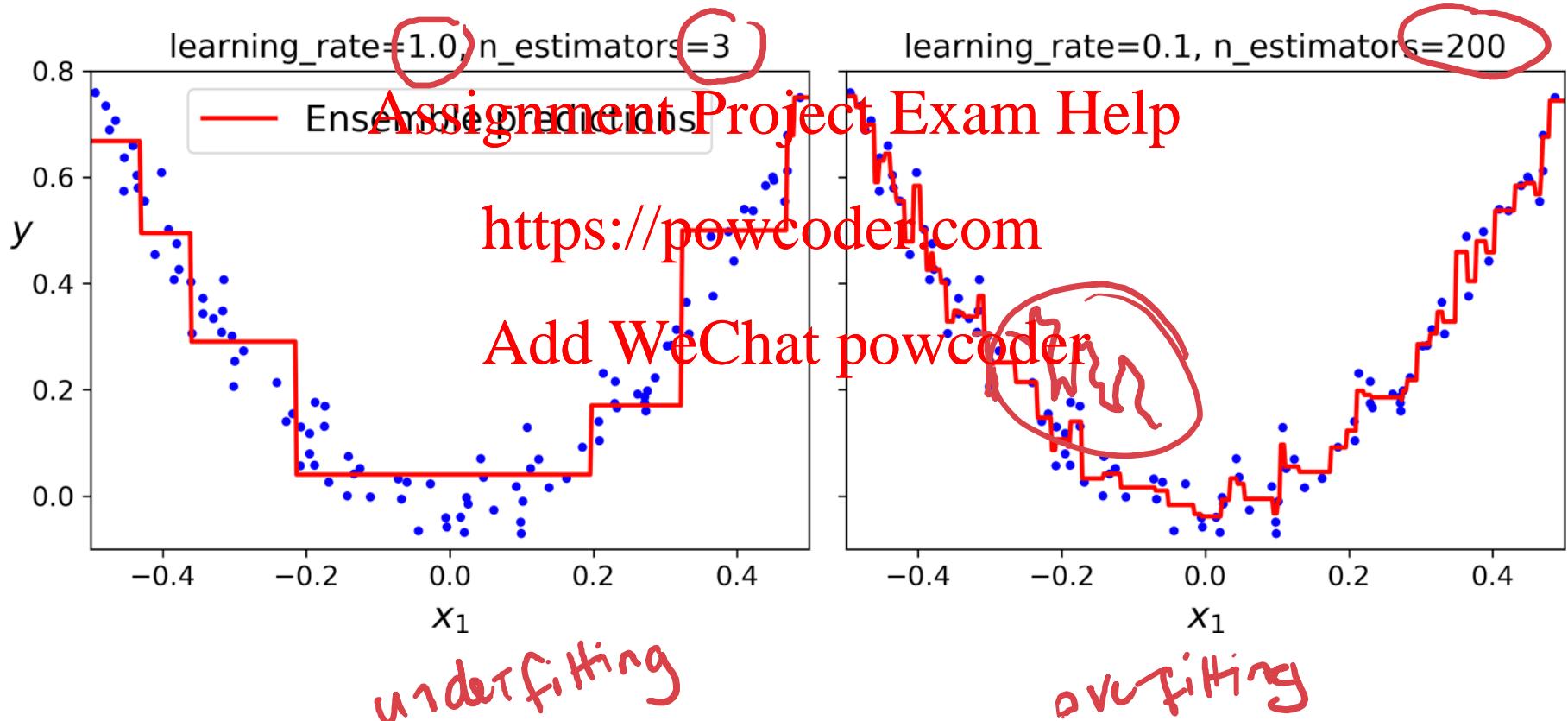
FITTING RESIDUAL ERRORS



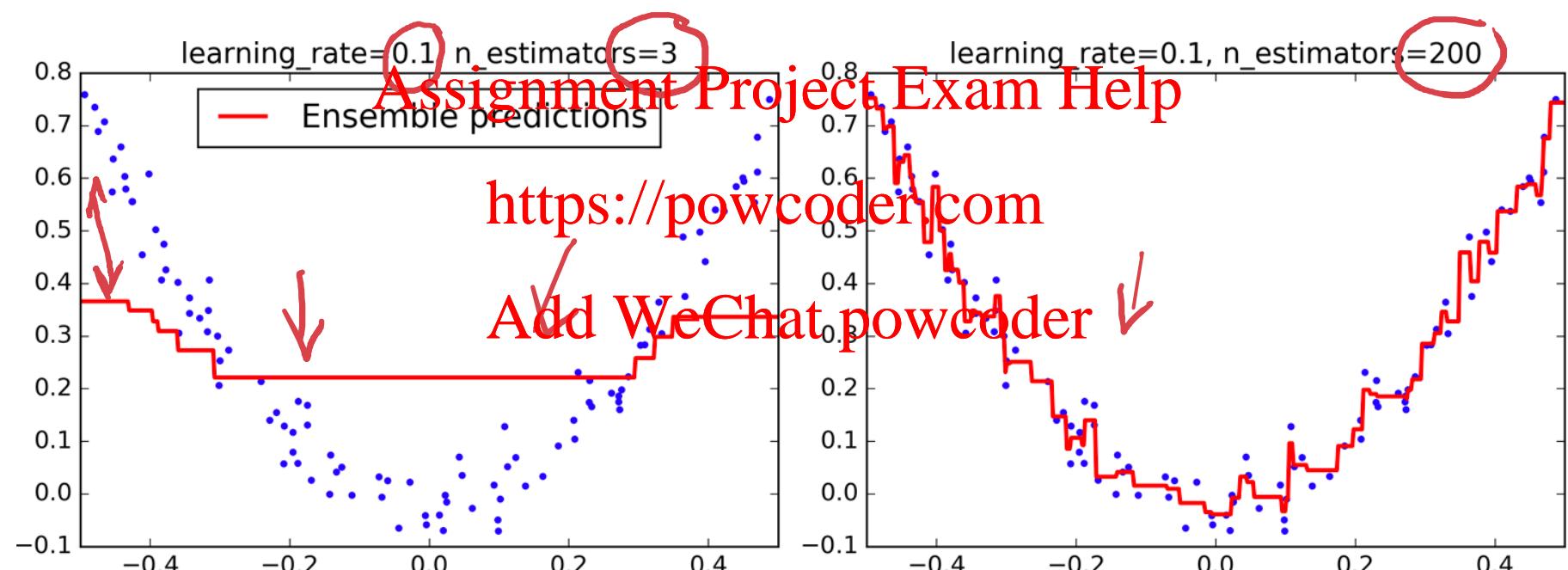
GRADIENT BOOSTING



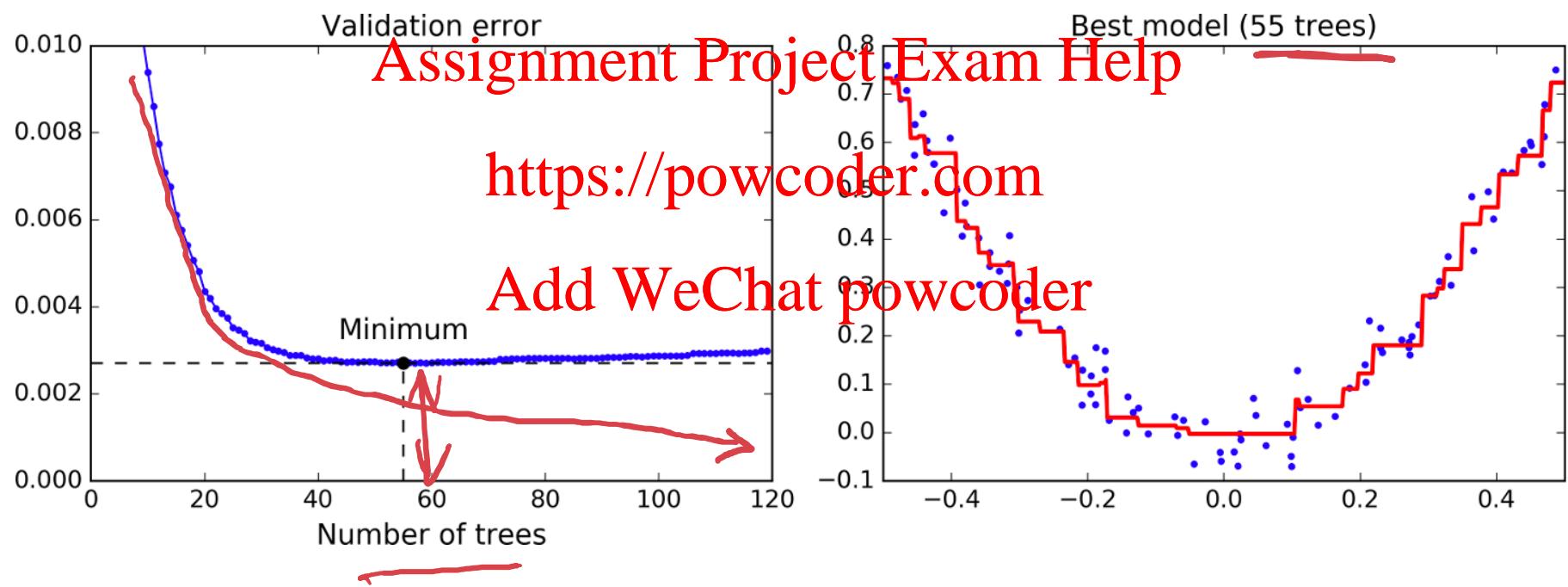
GBRT ENSEMBLES



OVERFITTING



VALIDATION ERROR

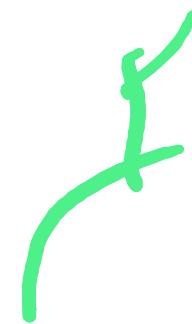


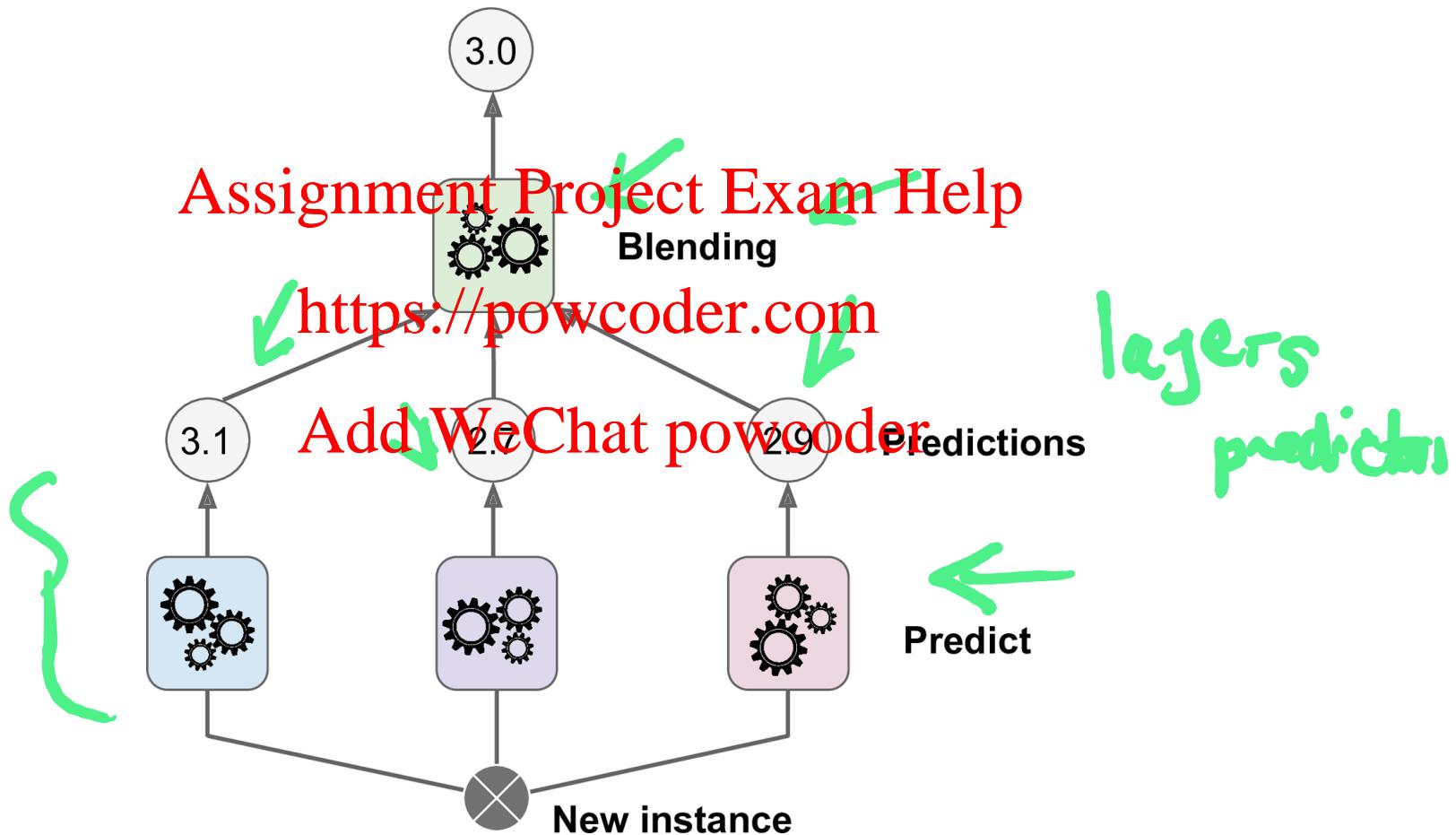
Assignment Project Exam Help

<https://powcoder.com>

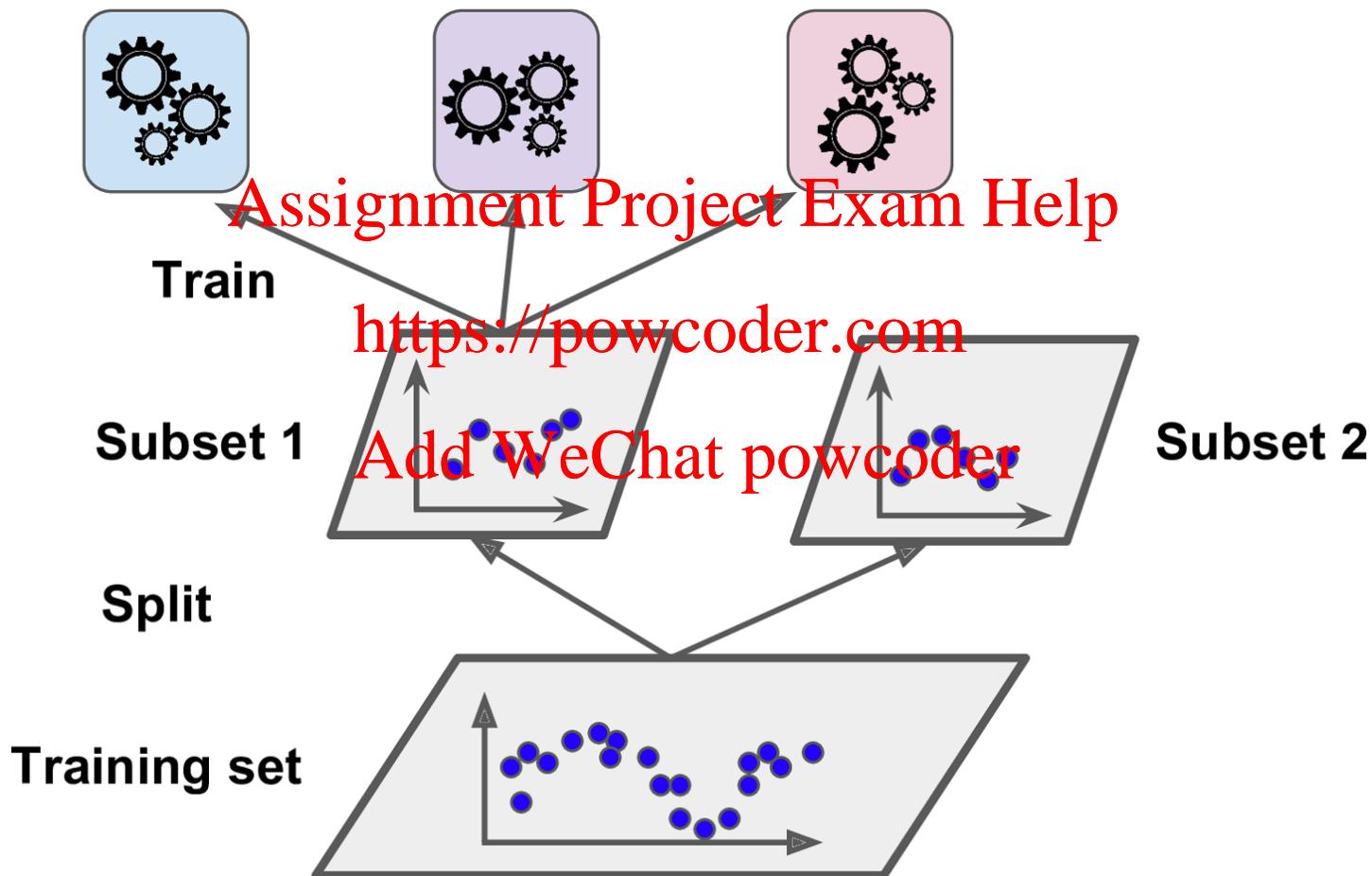
Add WeChat powcoder

⋮



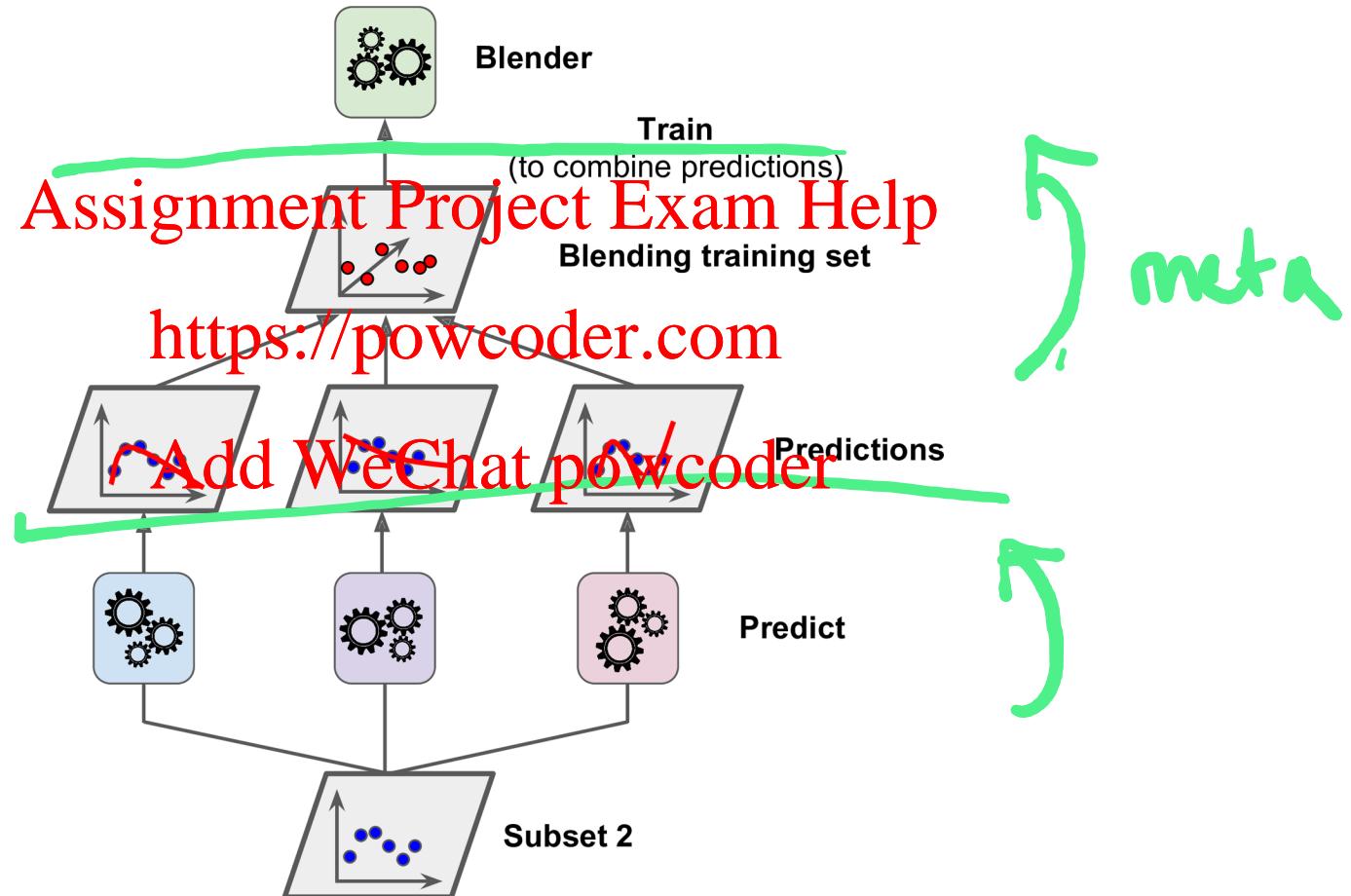


SUB-PROBLEMS



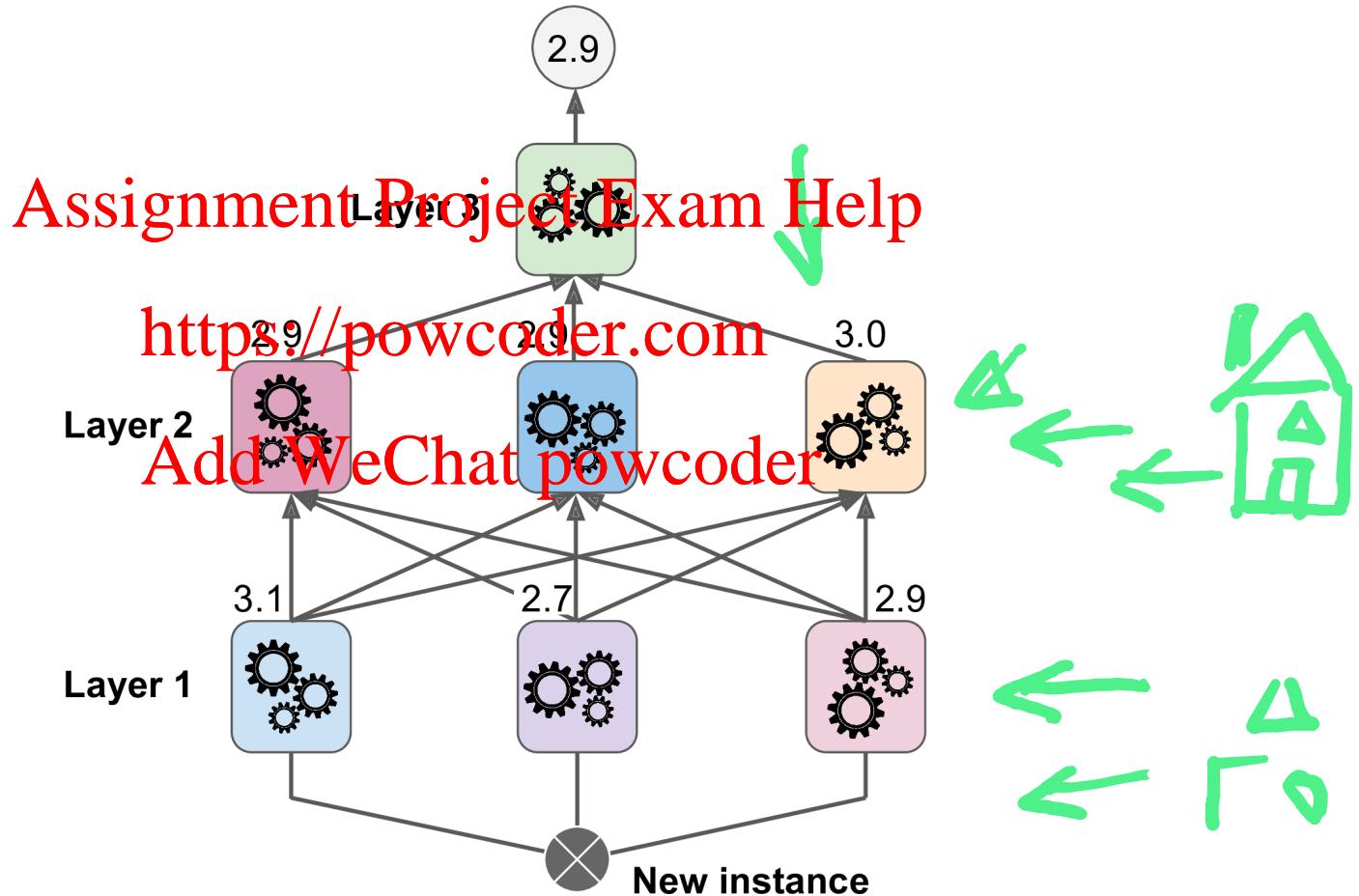
STACKING

BLENDED OR META LEARNER



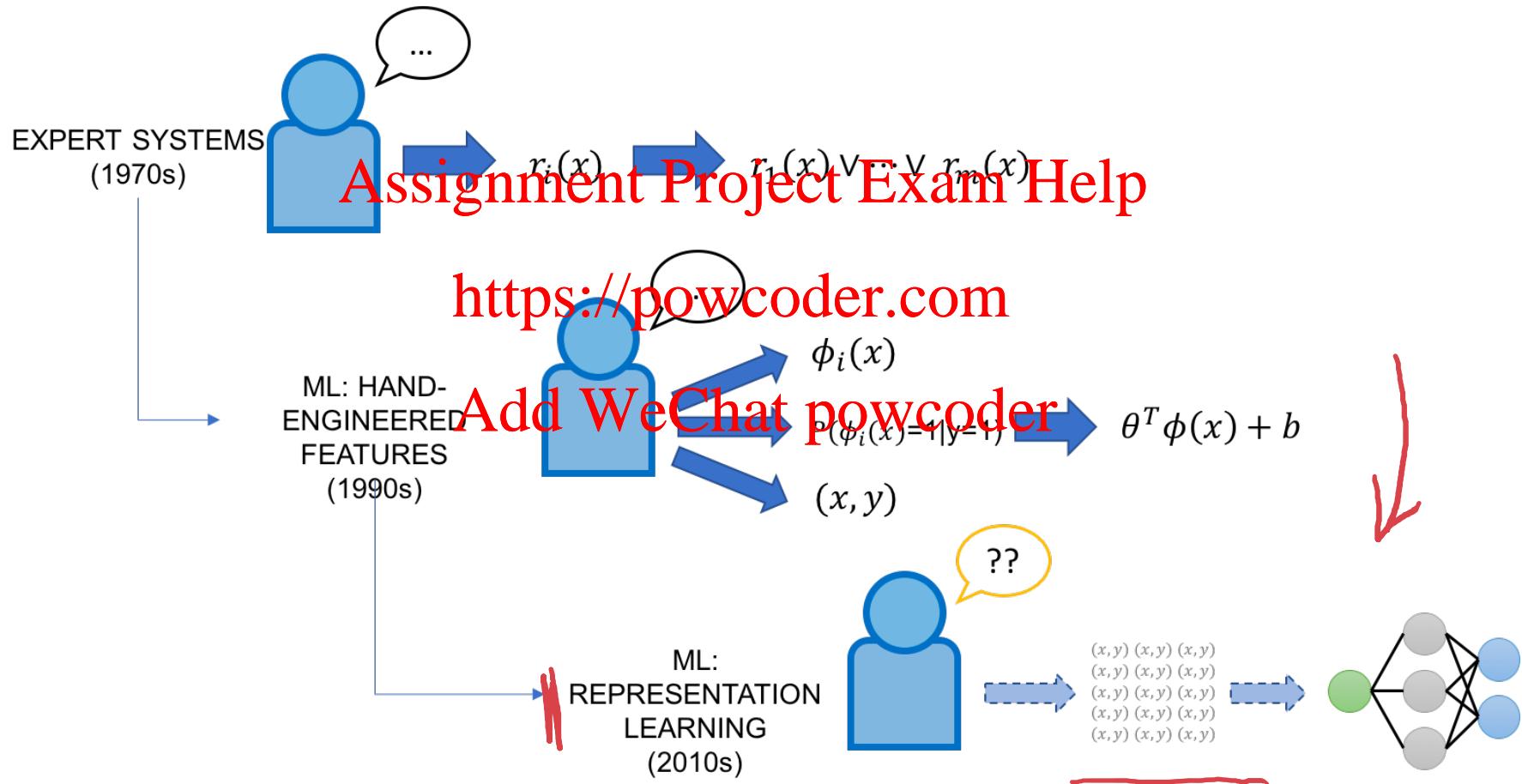
STACKING

BLENDED OR META LEARNER



DEEP MODELS

ANNs 2 GANs





Chip Huyen
@chipro

...

→ XGBoost was popular among winning solutions on Kaggle in 2016. Not anymore.

Assignment Project Exam Help

On recent competitions

- * Most winning solutions use deep learning.
- * Ensemble is still popular.
- * Data augmentation has become a standard.

<https://powcoder.com>

Maybe classical ML algorithms aren't really that hard to beat.

Add WeChat powcoder! 

Year	Competition	Data-type	Model	Architecture	Data augmentation
2020	Plant Pathology - FGVC7	Image	0	SE-ResNeXt-50	Blur, crop, shift, rotate, contrast
2020	NFL Big Data Bowl	Image + Time	75k	CNN	Flip
2020	Herbarium - FGVC7	Image	0	EfficientNet, ensemble	Crop, erase
2020	Ion Switching	Tabular	25k	Wavenet	Flip, random shift
2020	Abstraction and Reasoning	Image	20k	DSL, ensemble	Flip
				EfficientNet, CycleGAN, ensemble	
2020	Handwritten Grapheme	Image	10k		AutoAugment
				MTCNN, EfficientNet, ensemble	image compression, noise, blur, resize with different interpolations, color jittering, scaling and rotations, domain specific augmentations
2020	Deepfake Detection	Videos	1M		
2020	Santa's Workshop	Tabular	10k	MIP	
				BERT, BART, ensemble	pseudo-labels
2020	Google QUEST Q&A Label Text		25k		

LECTURE 3 TERM 2:

MSIN0097

UCL
SCHOOL OF
MANAGEMENT

PREDICTIVE ANALYTICS Assignment Project Exam Help

<https://powcoder.com>

Add WeChat powcoder

A P MOORE