

PREDICTIVE ANALYTICS Assignment Project Exam Help

<https://powcoder.com>

VIDEO 4: OPTIMISATION Add WeChat powcoder

A P MOORE

Assignment Project Exam Help

<https://powcoder.com>
Data + model → prediction
Add WeChat powcoder

Assume there is enough data to find statistical associations to solve specific tasks

Assignment Project Exam Help

<https://powcoder.com>
Data + model → prediction

Add WeChat powcoder

Define how well the model solves the task and adapt the parameters to maximize performance

A. CIAssification

Week 2 – Classification and Regression

Assignment Project Exam Help

Week 3 – Trees and Ensembles

<https://powcoder.com>

Add WeChat powcoder

C. Clustering

Week 5 – Clustering

D. Decomposition

Week 4 – Kernel spaces and Decomposition

- Discover
- Explore
- Visualize
- Clean
- Sample
- Impute
- Encode
- Transform
- Scale
- Features
- Pipelines
- Training/Validation splits
- Modeling
- Tuning
- Error Analysis
- Documentation
- Presentation
- Launch
- Monitor
- Maintain

Assignment Project Exam Help

<https://powcoder.com>

Add WeChat powcoder

- Pipelines
- Training/Validation splits
- Modeling
- Tuning
- Error Analysis

- Discover
- Explore
- Visualize

- Clean
- Sample
- Impute

- Documentation
- Presentation

Assignment Project Exam Help

- Encode

<https://powcoder.com>

- Transform
- Launch

Add WeChat powcoder

- Scale
- Modeling

- Monitor
- Maintain

- Overfitting
- Optimization
- Model Selection
- Regularization
- Generalization

$$x \rightarrow y$$

Assignment Project Exam Help

$$x \rightarrow f(x) \rightarrow y$$

<https://powcoder.com>

Add WeChat powcoder

LEARNING A FUNCTION

$$x \rightarrow y$$

Assignment Project Exam Help

$$x \rightarrow f(x) \rightarrow y$$

<https://powcoder.com>

Measured data

$$x \xrightarrow{\text{Features}} \text{Add WeChat powcoder} \xrightarrow{\text{Inferred/Predicted/Estimated value}} \hat{x} \rightarrow f(x) = \hat{y} \rightarrow y$$

True initial value
(world state)

Learned/Fitted function
From n observations

True target value
(world state)

COMPONENTS OF A MACHINE LEARNING SOLUTION

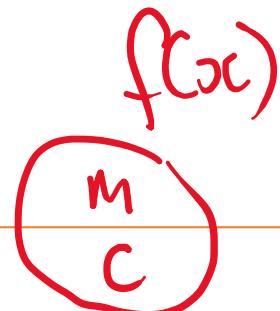
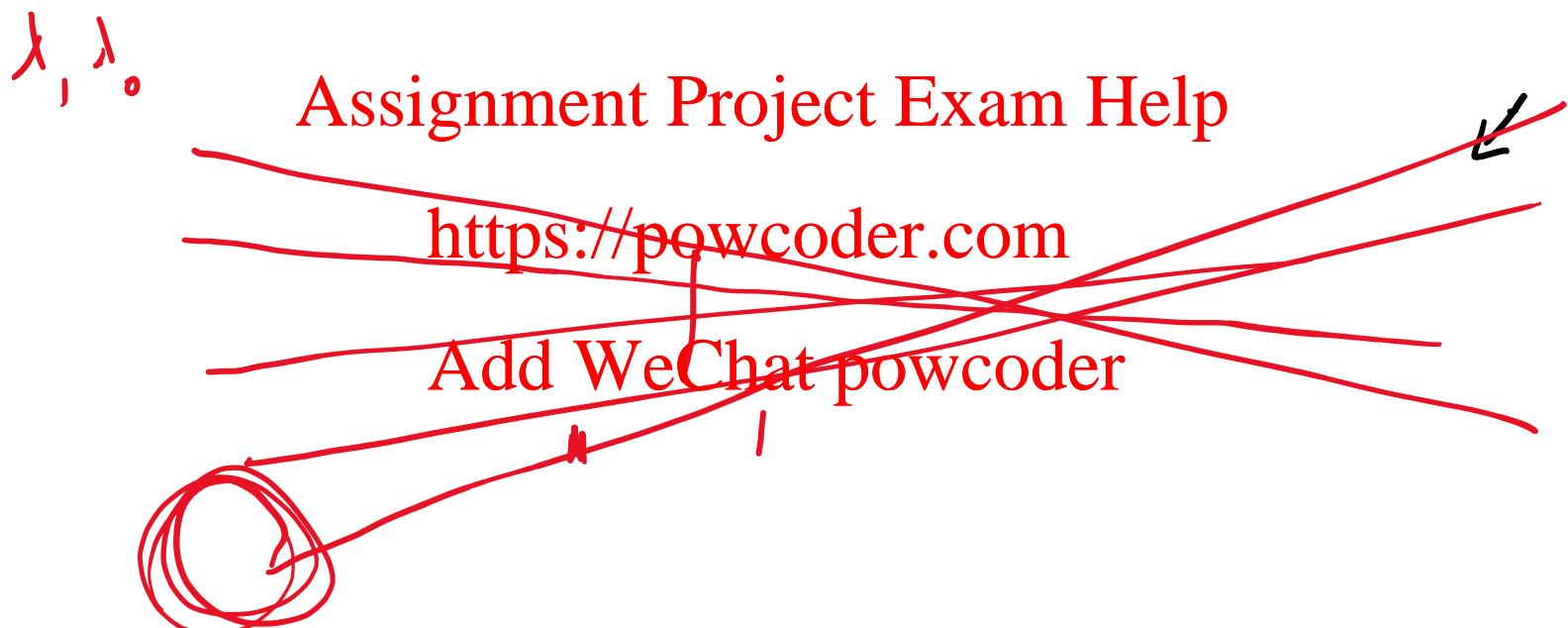
To solve a machine learning problem we need three components:

Assignment Project Exam Help

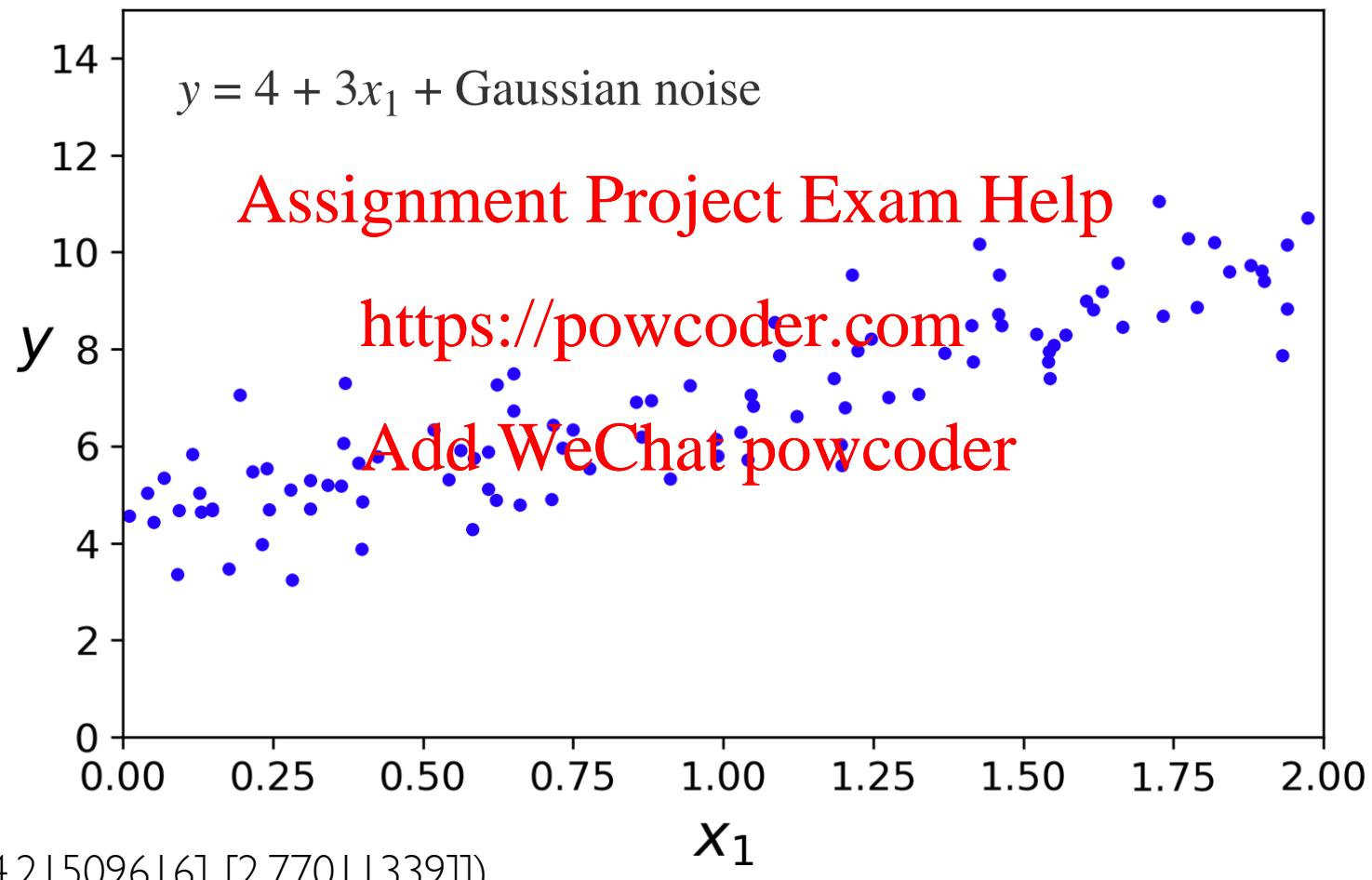
- We need a **model** that mathematically relates the data x and the world state y .
The model specifies a family of possible relationships between x and y and the particular relationship is determined by the model parameters θ
- We need a **learning algorithm** that allows us to fit the parameters θ using paired training examples $\{x_i, y_i\}$ where we know both the measurements and the underlying state. (*Stated like this this is a **supervised learning problem***)
- We need an **inference algorithm** that takes a new observation x and uses the model to return the posterior $Pr(y|x, \theta)$ over the world state y .

B. REGRESSION

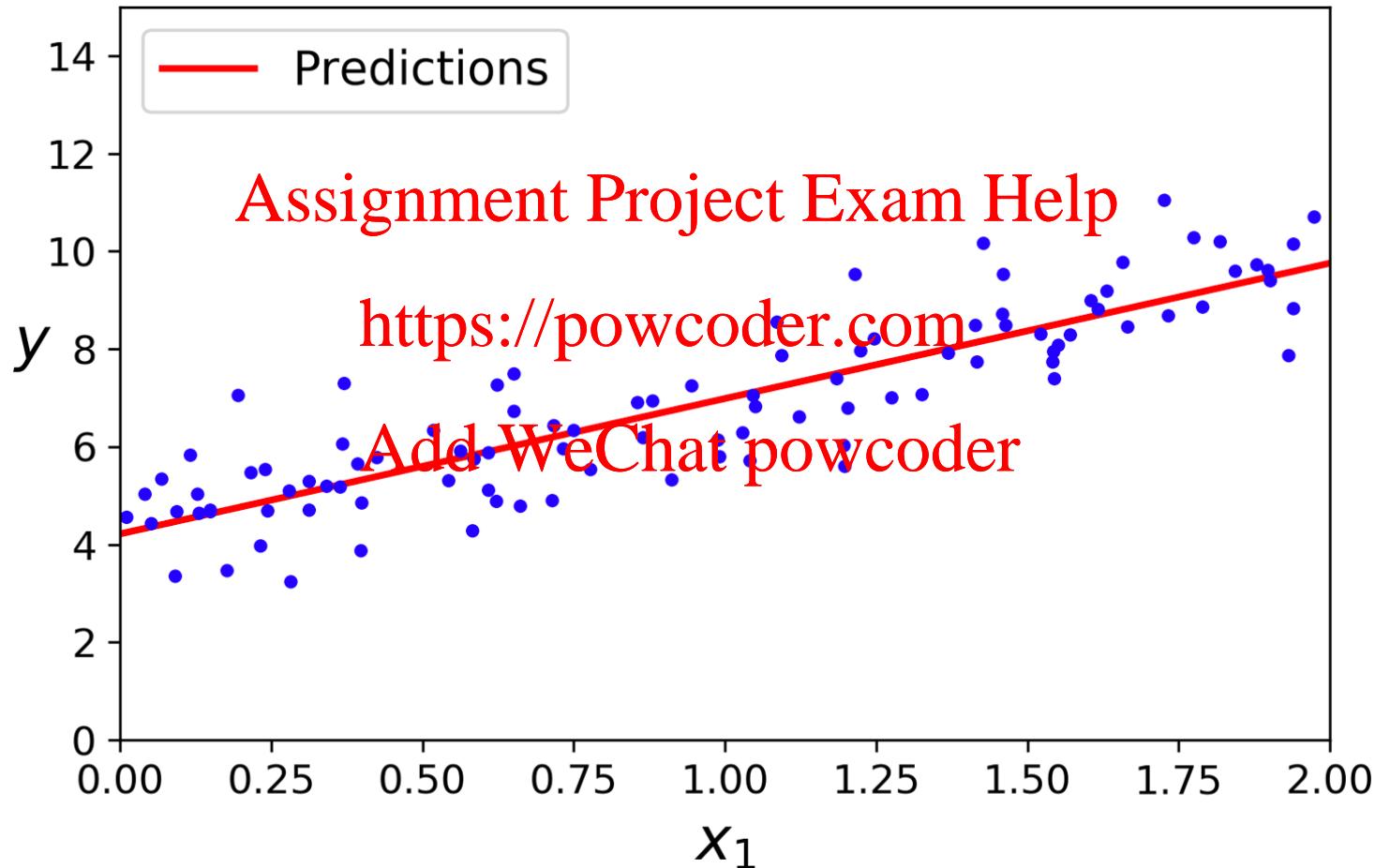
REAL VALUED VARIABLE



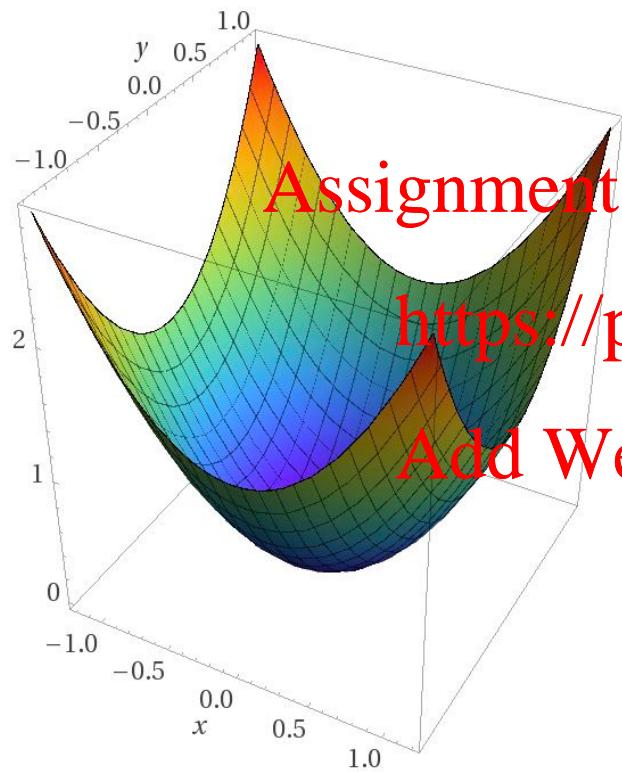
```
import numpy as np  
  
X = 2 * np.random.rand(100, 1)  
y = 4 + 3 * X + np.random.randn(100, 1)
```



MAKING PREDICTIONS



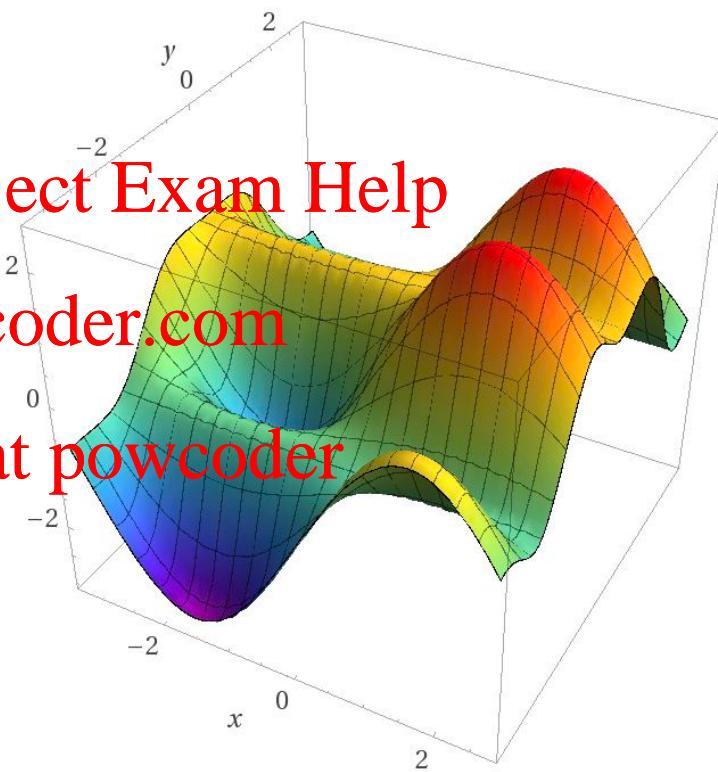
ERROR SURFACE



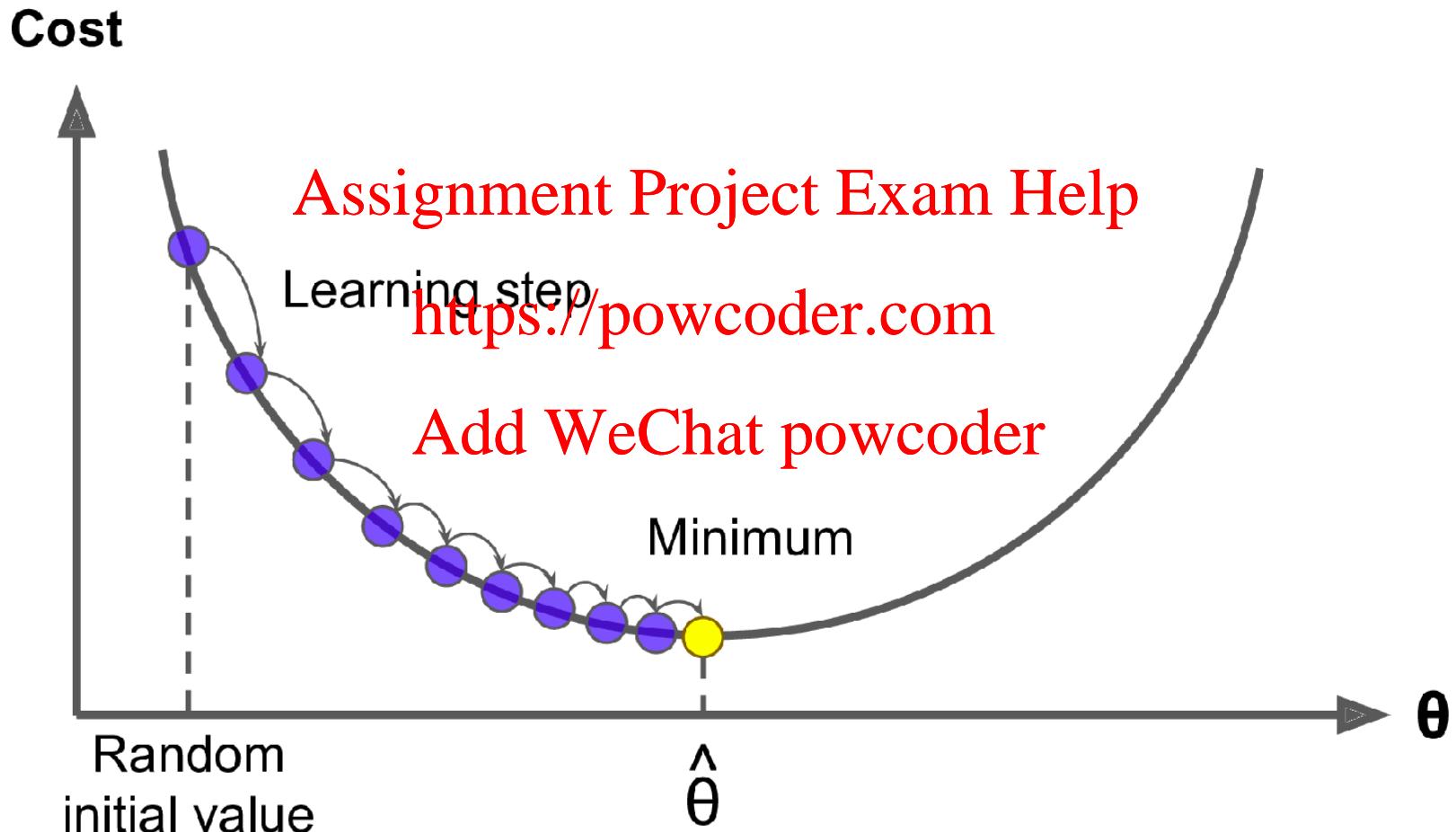
Assignment Project Exam Help

<https://powcoder.com>

Add WeChat powcoder

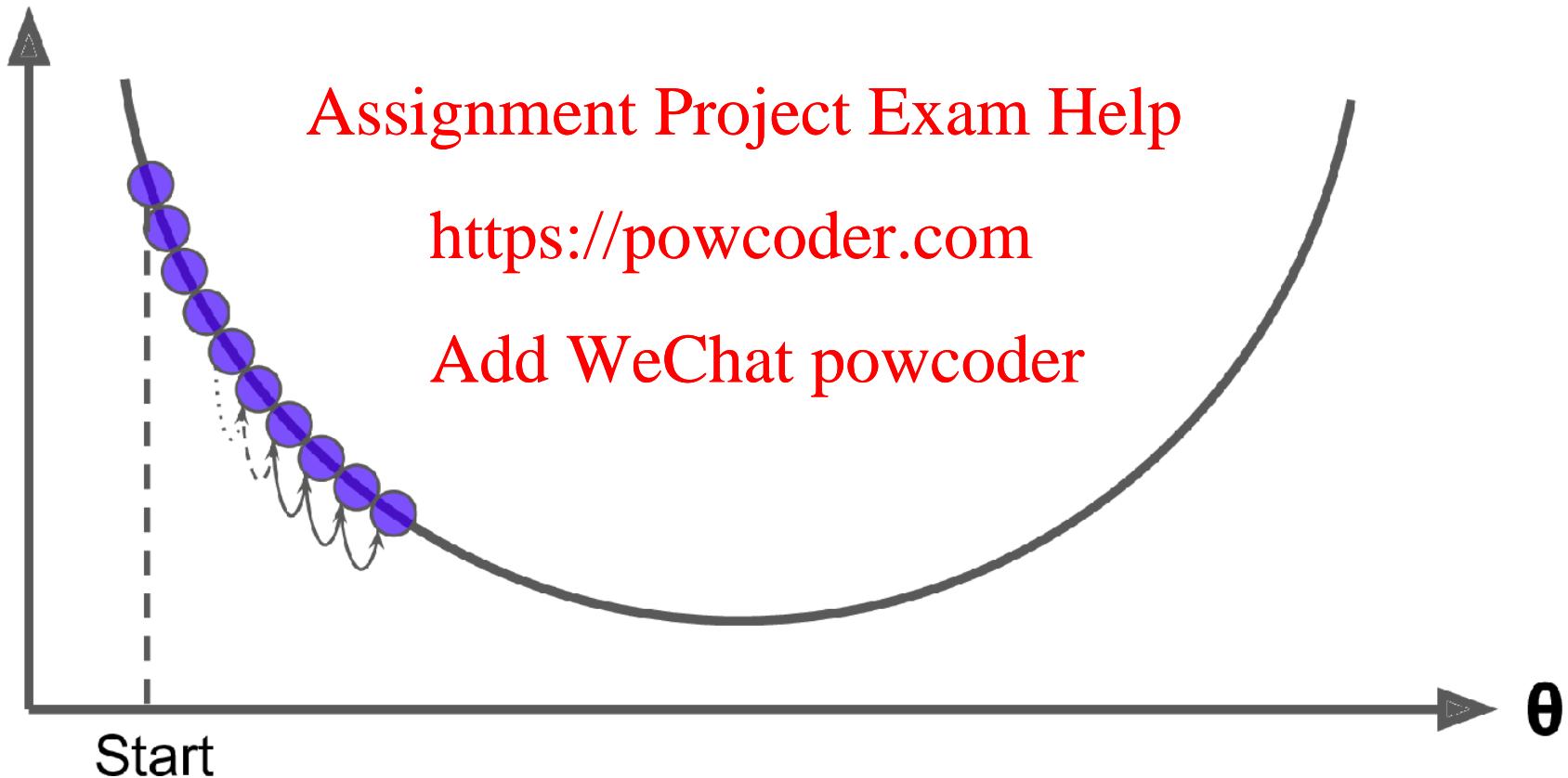


GRADIENT DESCENT



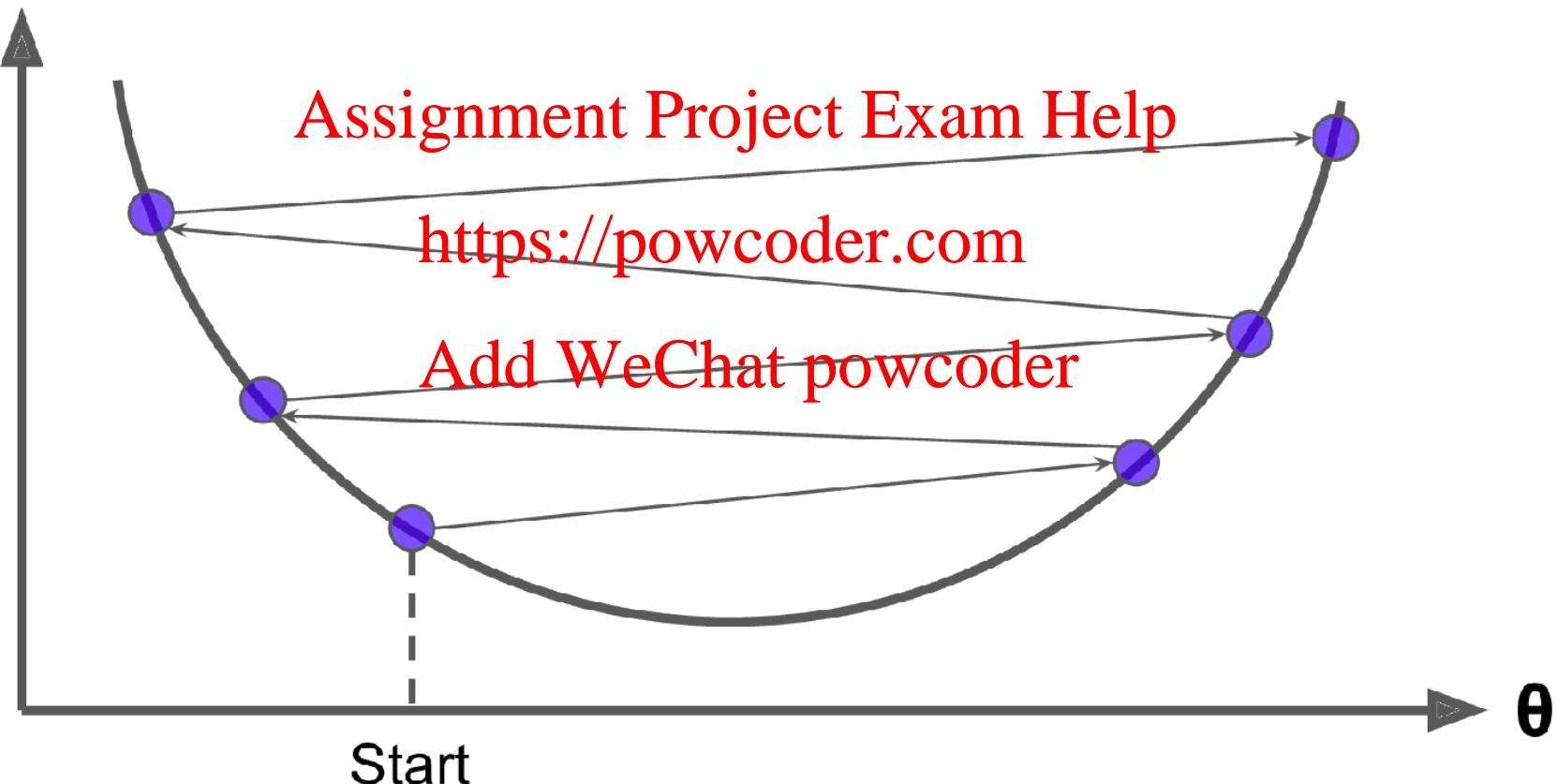
LEARNING RATE TOO SMALL

Cost

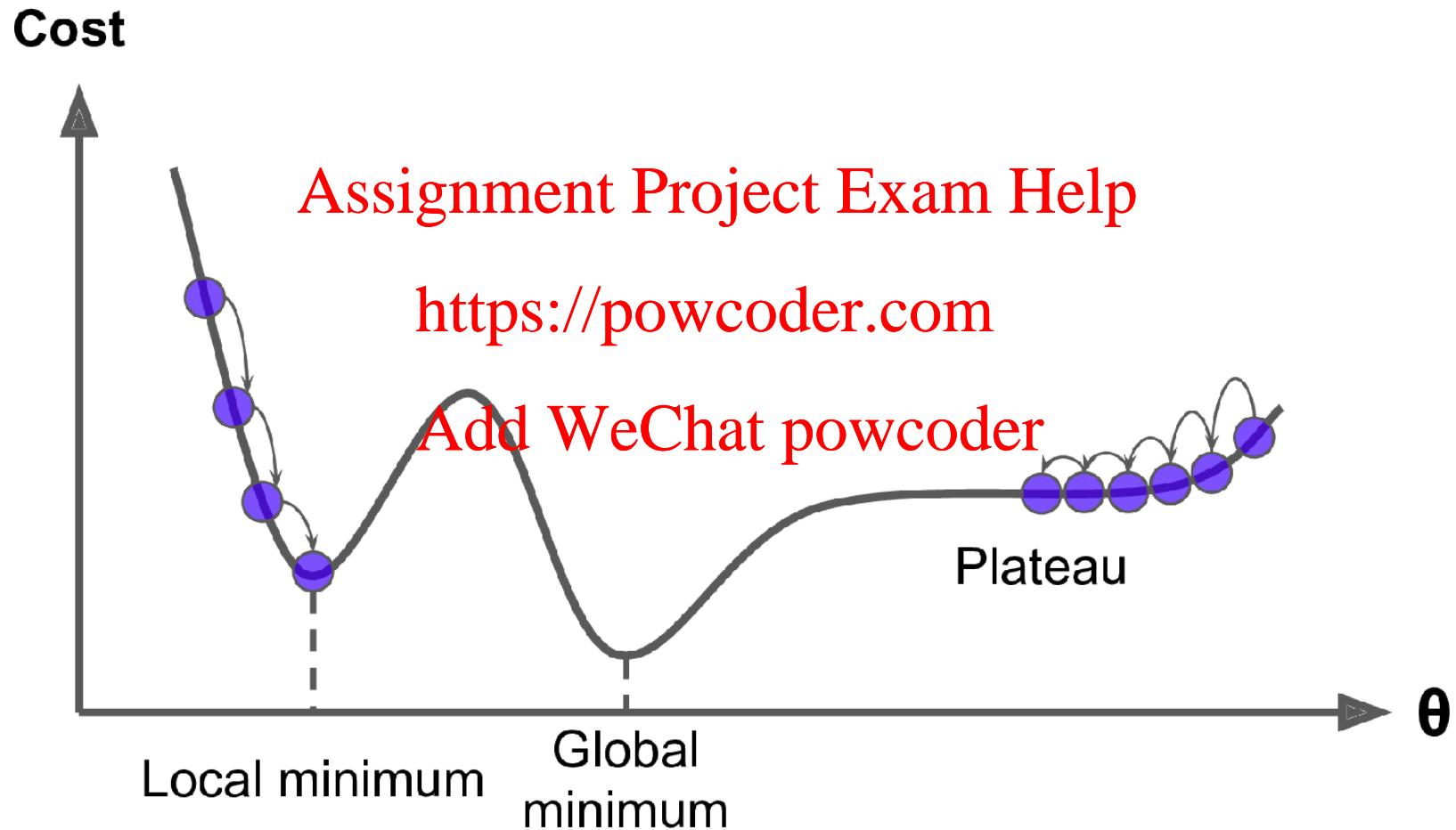


LEARNING RATE TOO LARGE

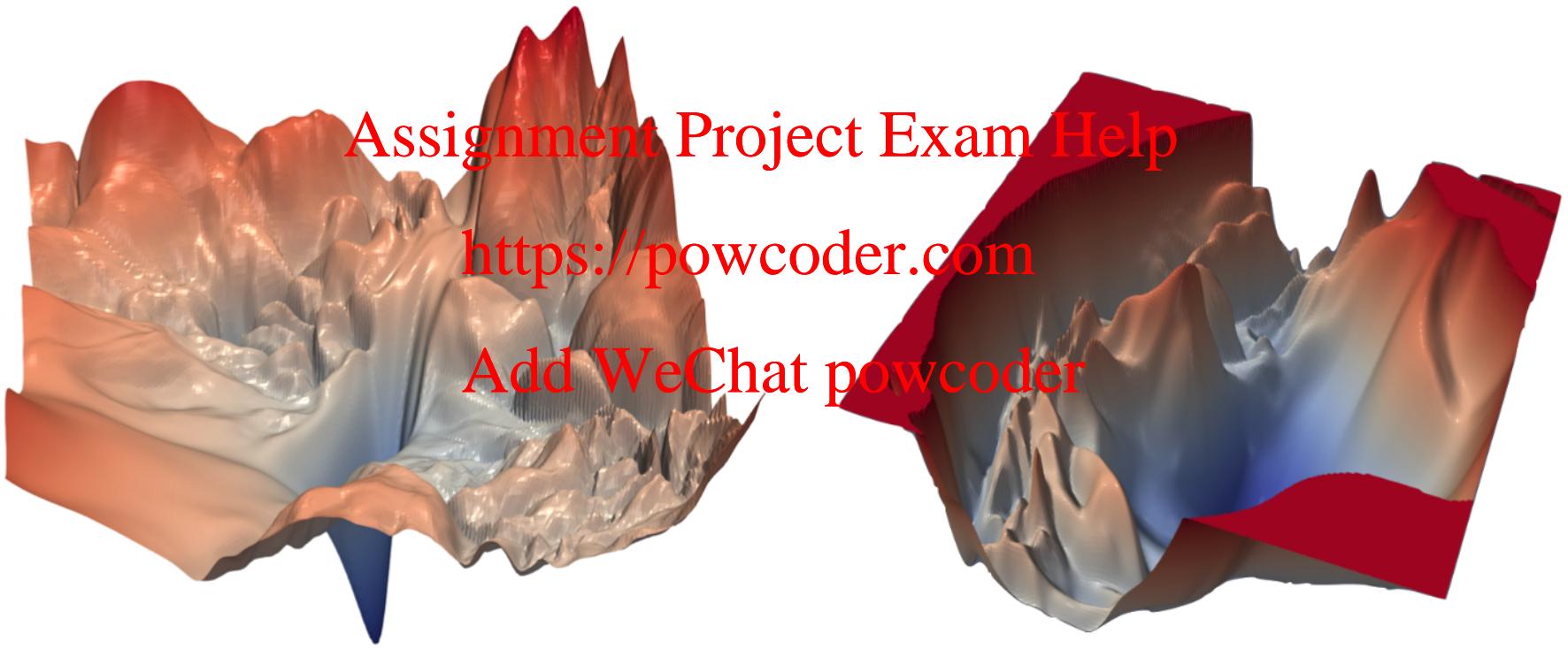
Cost



MINIMA AND PLATEAUS



ERROR SURFACES



Assignment Project Exam Help

<https://powcoder.com>

Add WeChat powcoder

Assignment Project Exam Help

<https://powcoder.com>

Add WeChat powcoder

$$\text{MSE}(\mathbf{X}, h_{\boldsymbol{\theta}}) = \frac{1}{m} \sum_{i=1}^m (\boldsymbol{\theta}^\top \mathbf{x}^{(i)} - y^{(i)})^2$$

PARTIAL DERIVATIVES IN PARAMETER SPACE

$$\hat{y} = h_{\theta}(\mathbf{x}) = \boldsymbol{\theta} \cdot \mathbf{x}$$

$$\text{MSE}(\mathbf{X}, h_{\theta}) = \frac{1}{m} \sum_{i=1}^m (\boldsymbol{\theta}^T \mathbf{x}^{(i)} - y^{(i)})^2$$

$$\nabla_{\theta} \text{MSE}(\boldsymbol{\theta}) = \begin{pmatrix} \frac{\partial}{\partial \theta_0} \text{MSE}(\boldsymbol{\theta}) \\ \frac{\partial}{\partial \theta_1} \text{MSE}(\boldsymbol{\theta}) \\ \vdots \\ \frac{\partial}{\partial \theta_n} \text{MSE}(\boldsymbol{\theta}) \end{pmatrix} = \frac{2}{m} \mathbf{X}^T (\mathbf{X}\boldsymbol{\theta} - \mathbf{y})$$

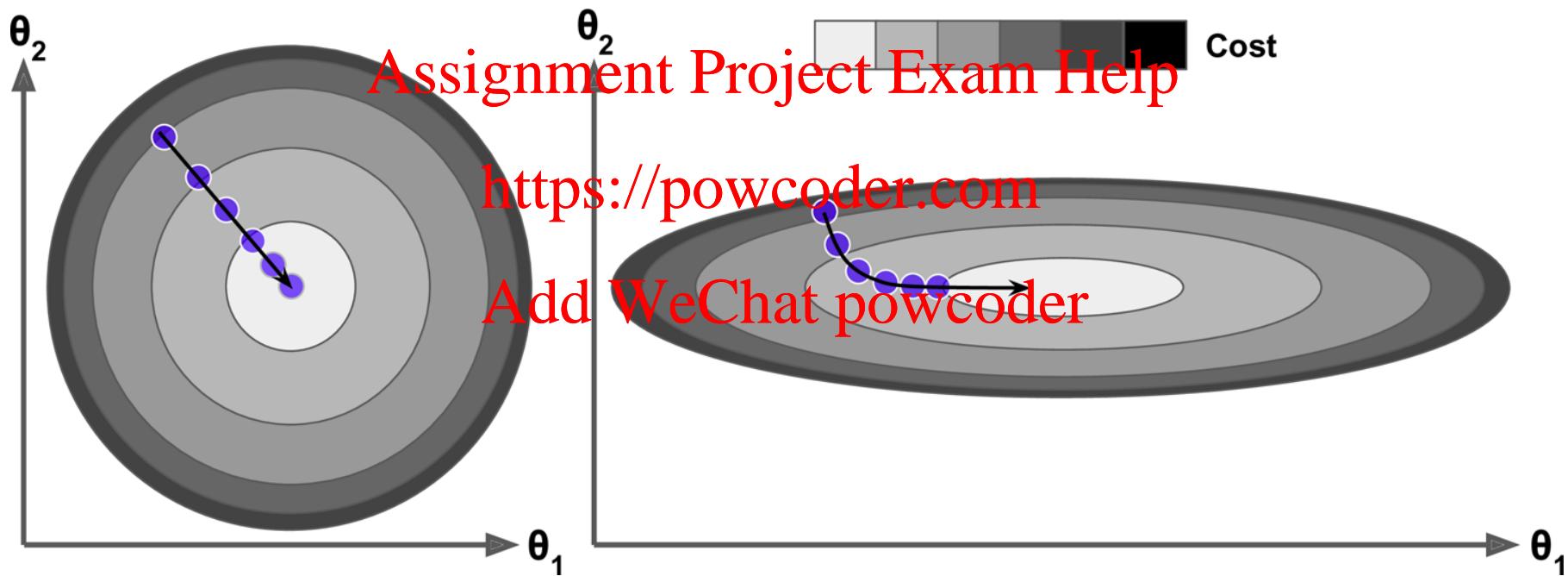
$$\boldsymbol{\theta}^{(\text{next step})} = \boldsymbol{\theta} - \eta \nabla_{\theta} \text{MSE}(\boldsymbol{\theta})$$

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

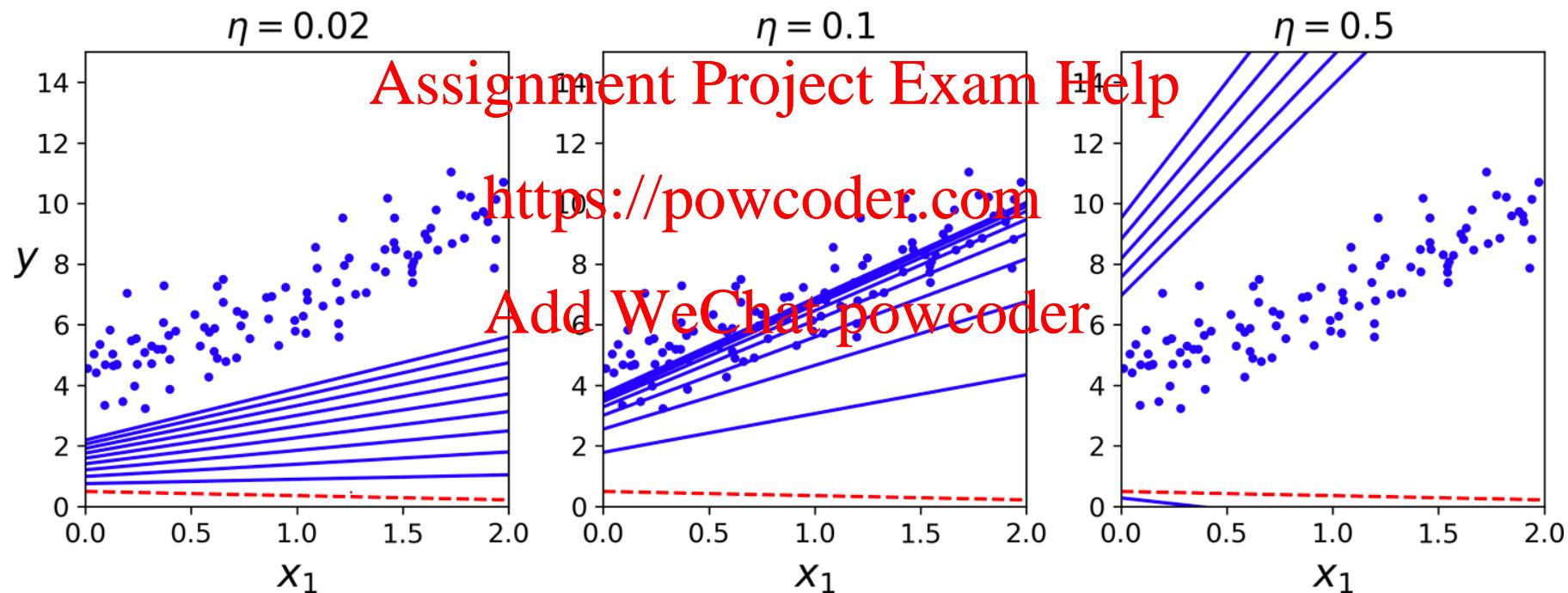
```
eta = 0.1 # learning rate
n_iterations = 1000
m = 100
theta = np.random.randn(2,1) # random initialization

for iteration in range(n_iterations):
    gradients = 2/m * X_b.T.dot(X_b.dot(theta) - y)
    theta = theta - eta * gradients
```

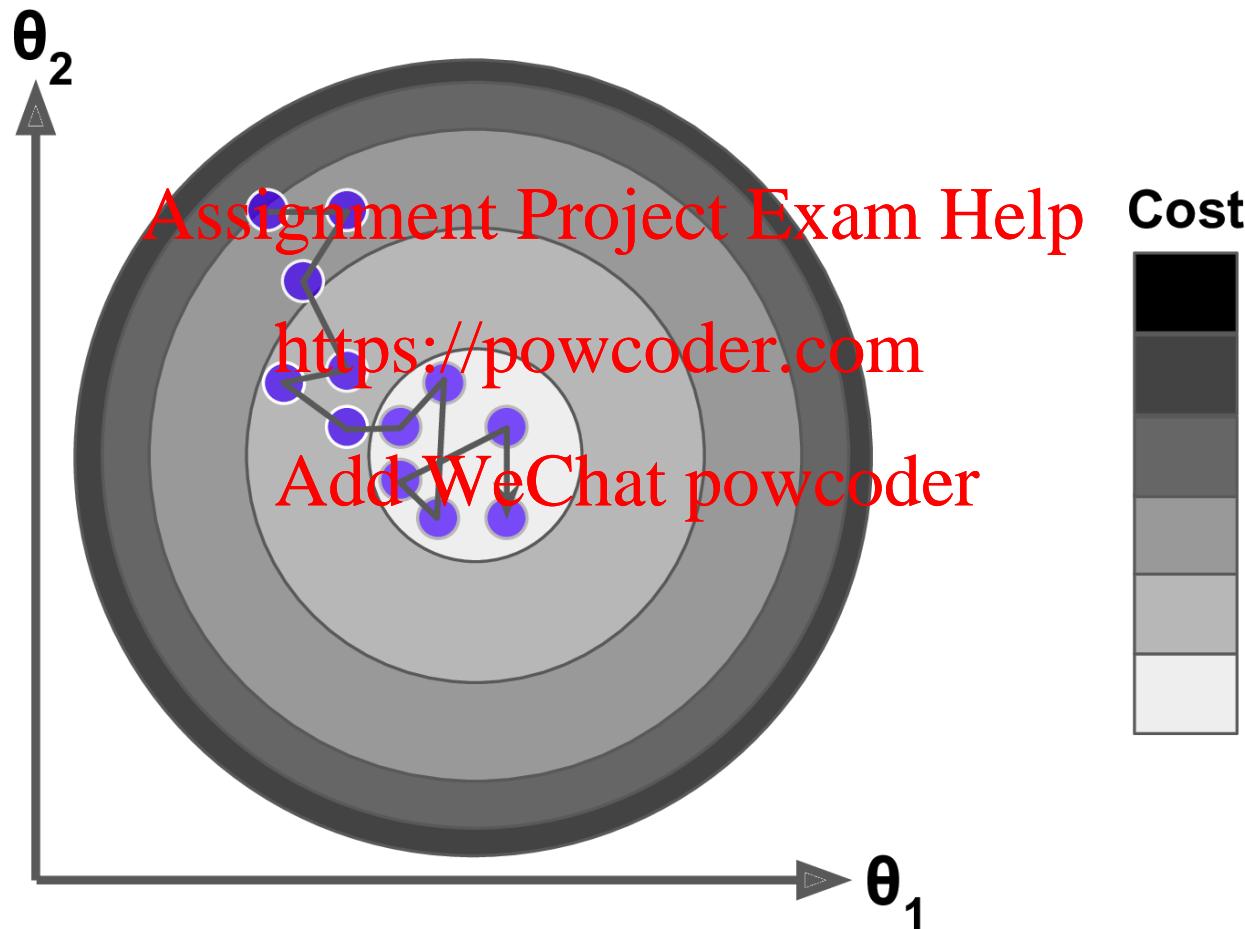
FEATURE SCALING



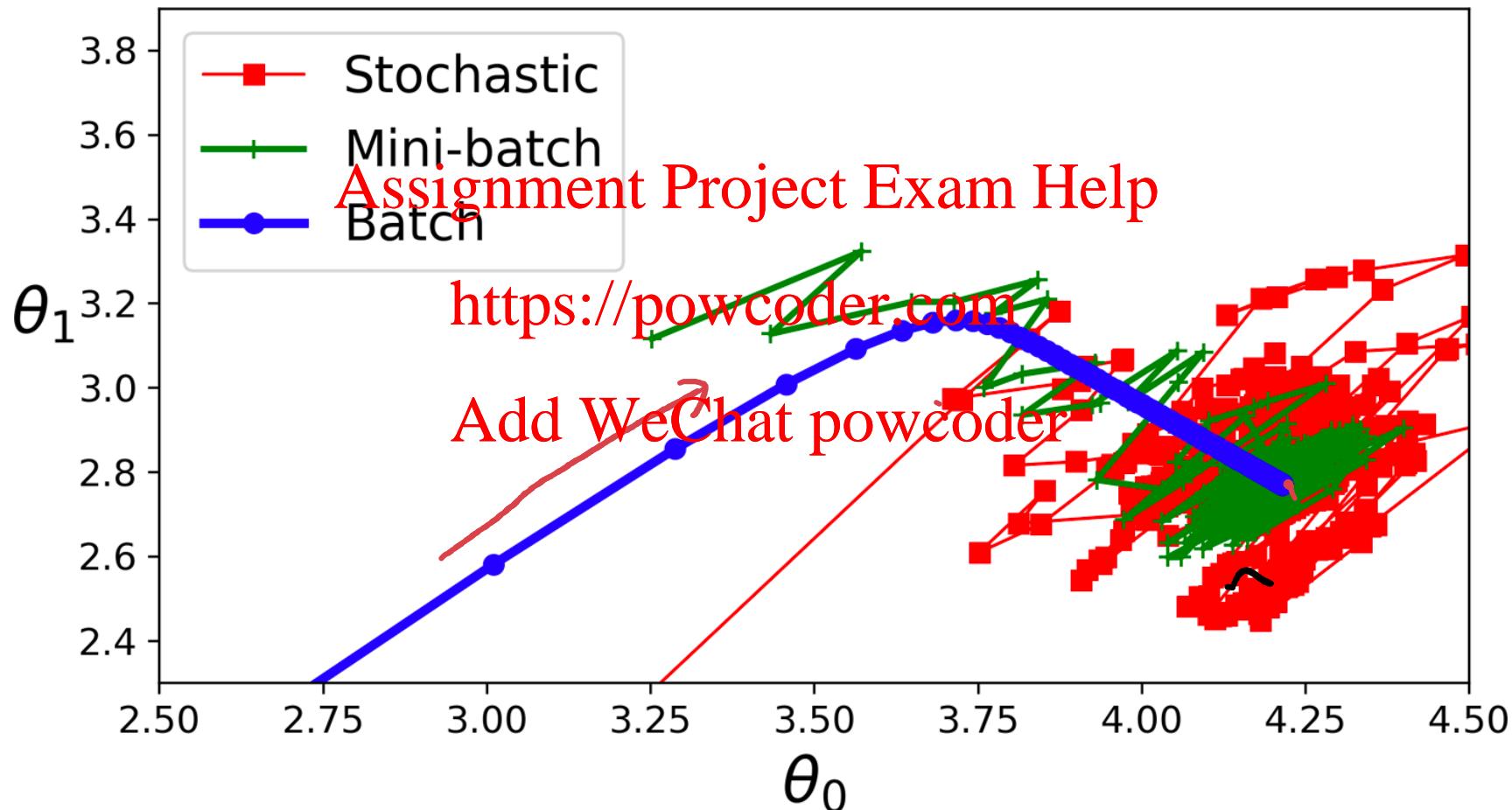
LEARNING RATES



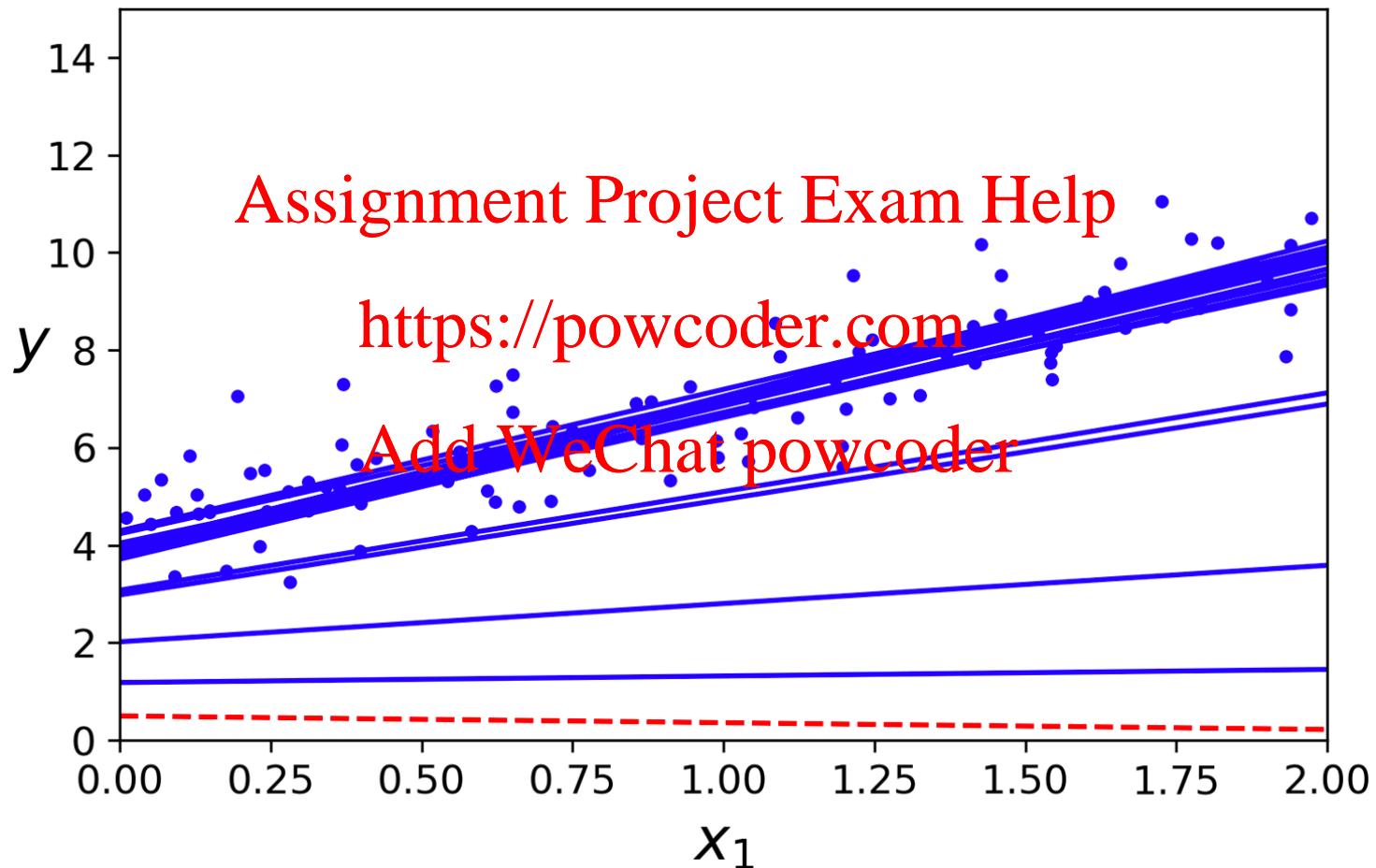
STOCHASTIC GRADIENT DESCENT

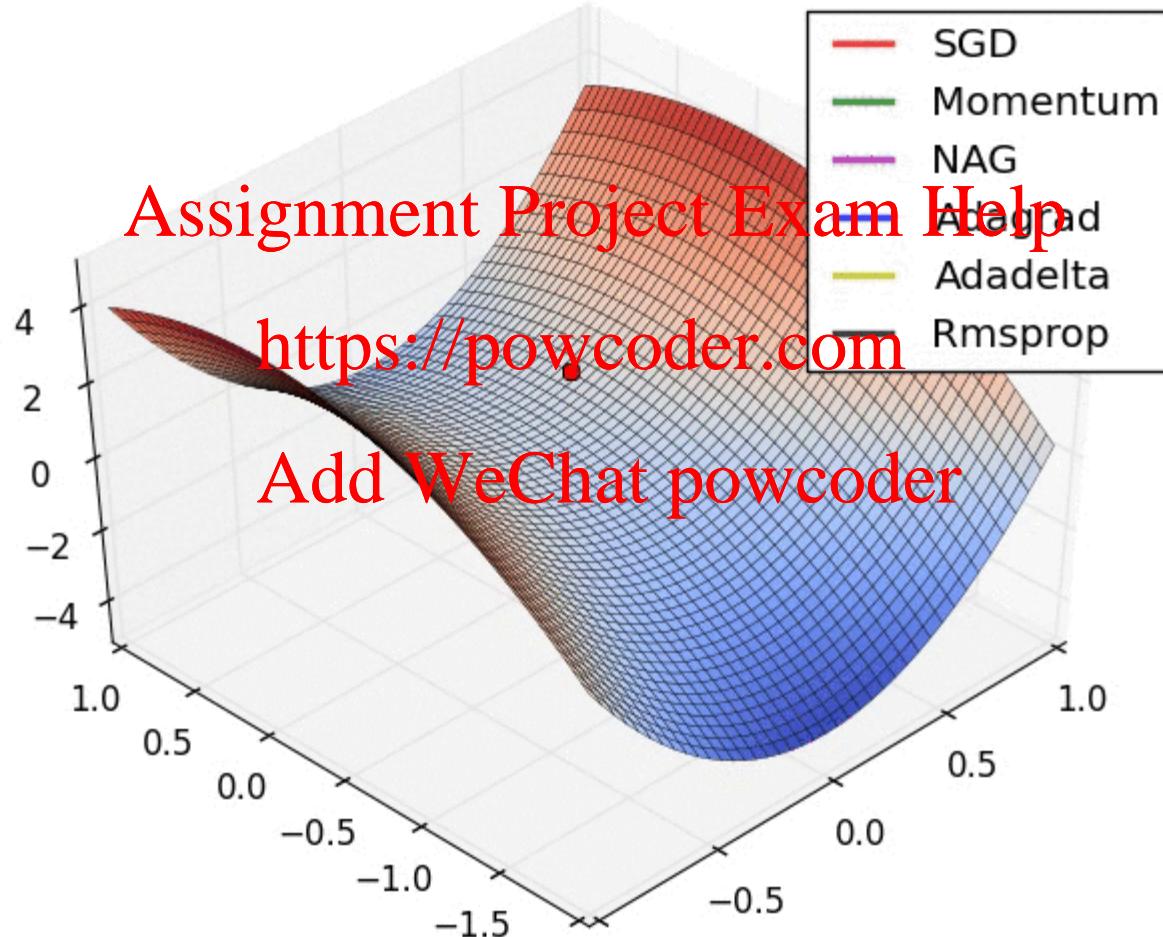


PATHS IN THE SEARCH SPACE



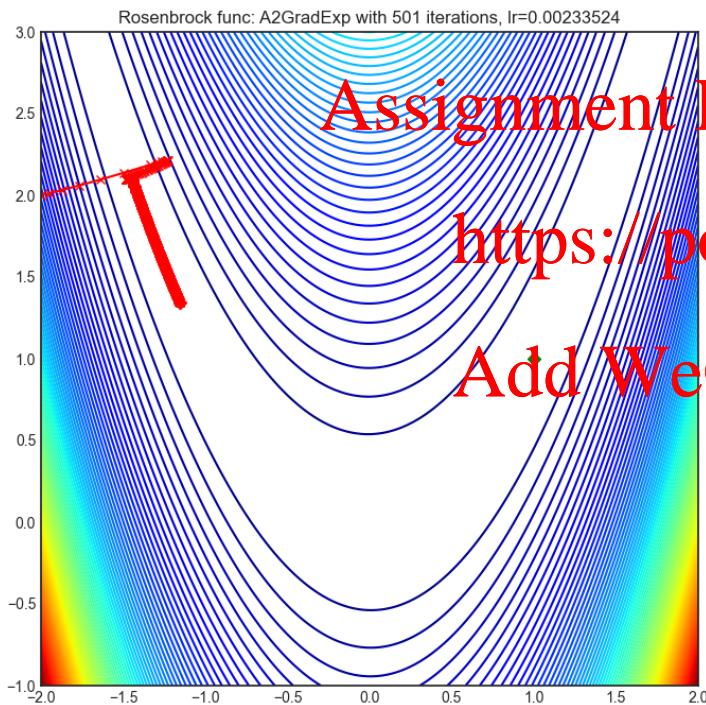
LEARNING SCHEDULE



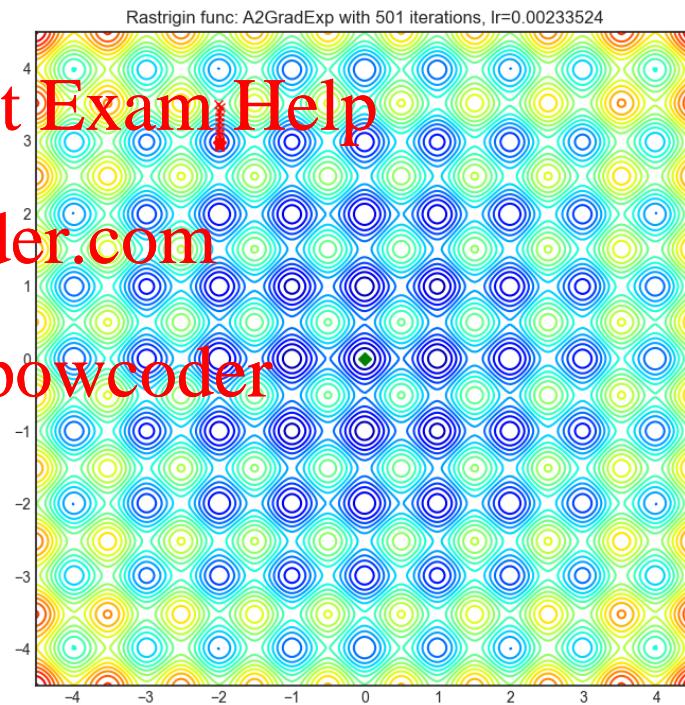


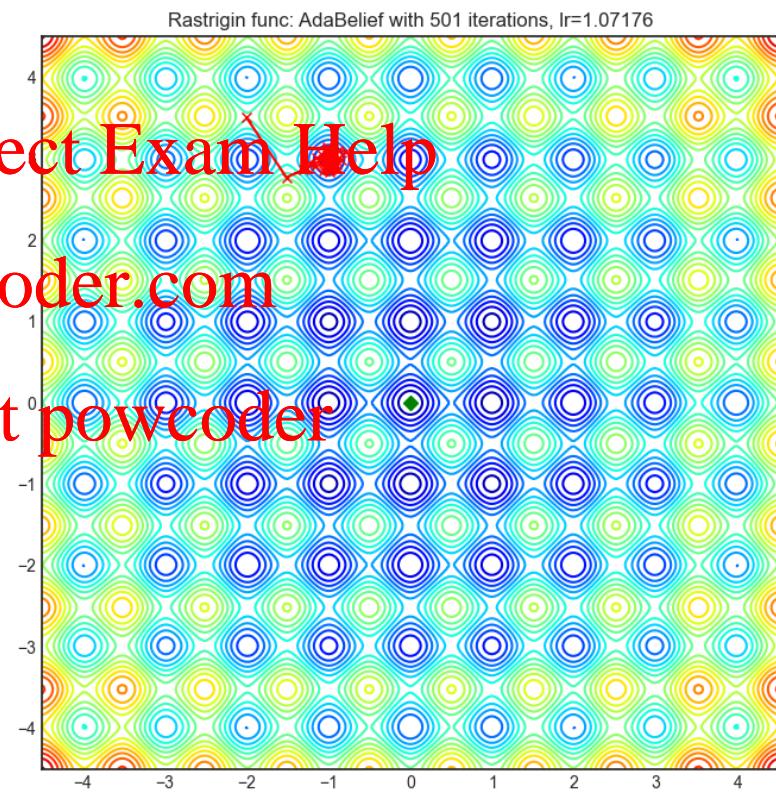
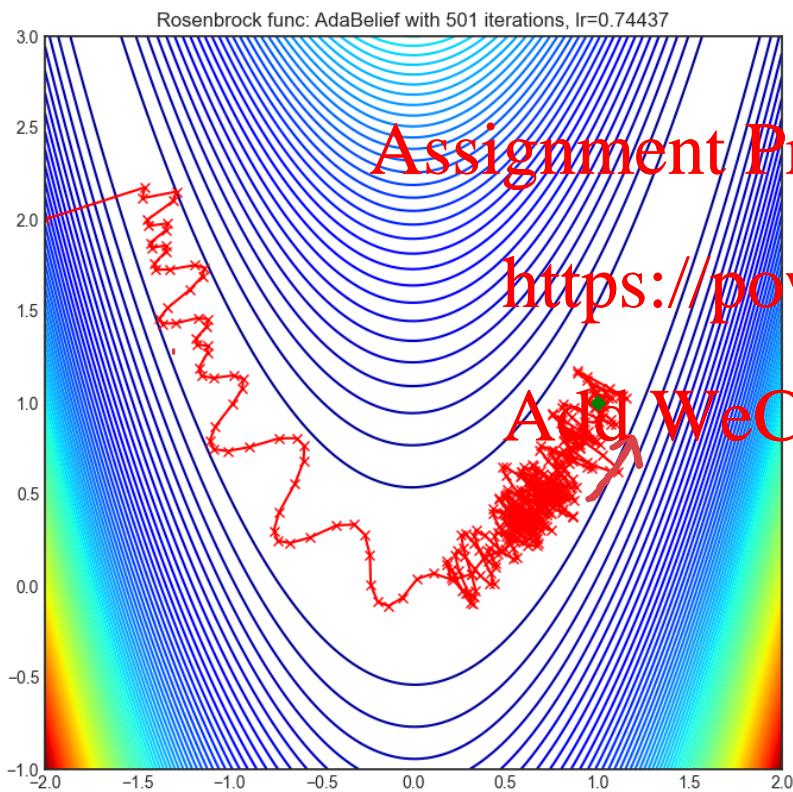
OPTIMIZERS

A2GRADEXP -



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





Assignment Project Exam Help

<https://powcoder.com>

Add WeChat powcoder

HYPER-PARAMETERS

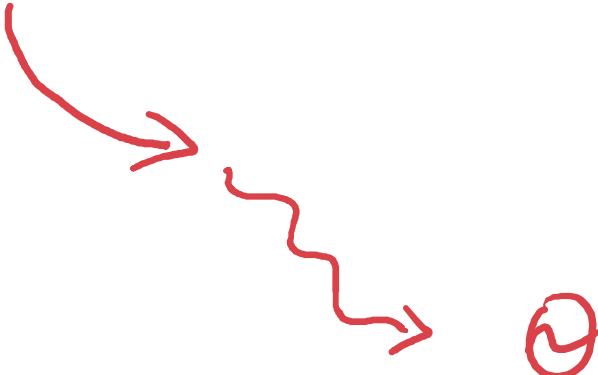
η μ - learning rate

Assignment Project Exam Help

b - batch size

<https://powcoder.com>

Add WeChat powcoder



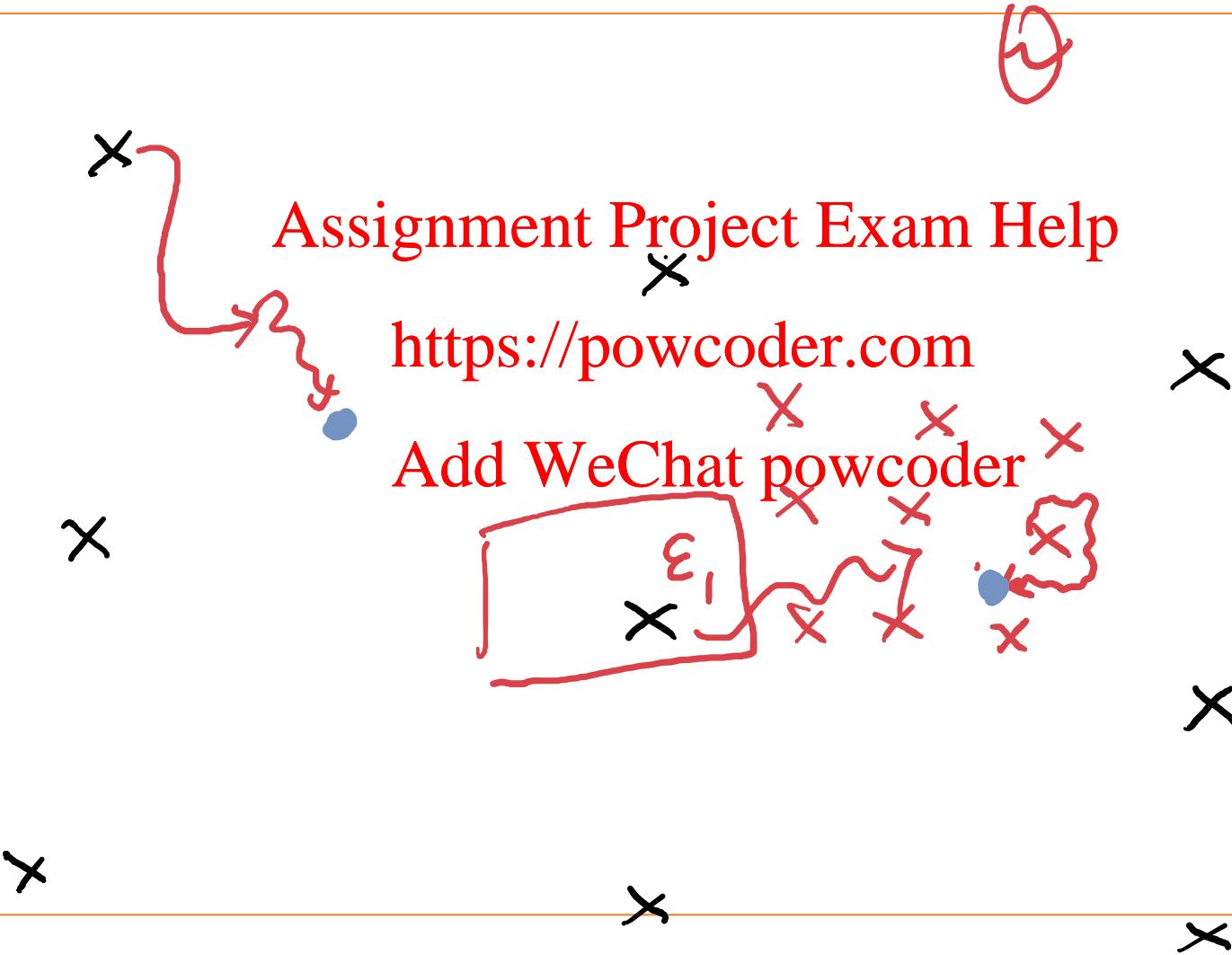
Assignment Project Exam Help

<https://powcoder.com>

Add WeChat powcoder



GRID SEARCH



RANDOM SEARCH



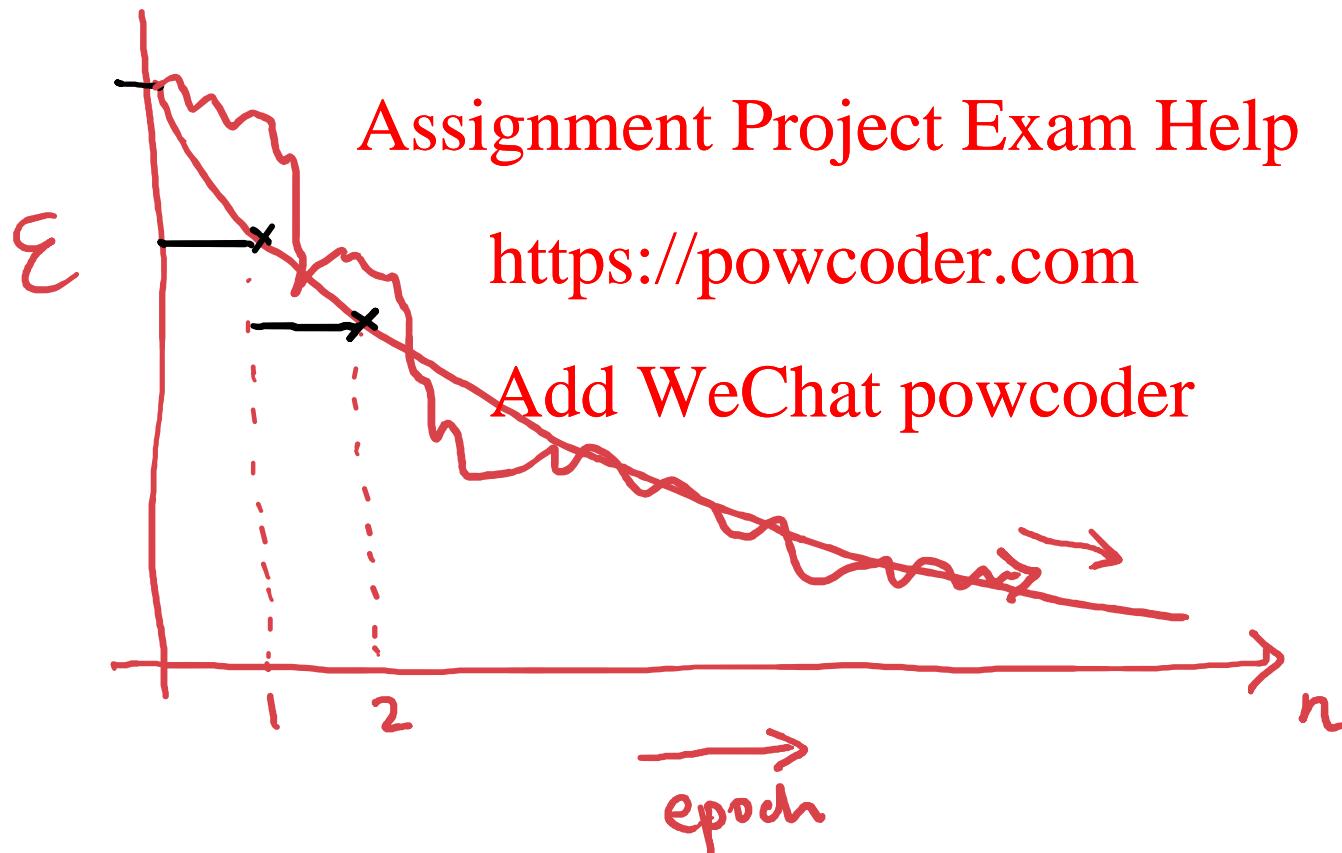
Assignment Project Exam Help

<https://powcoder.com>

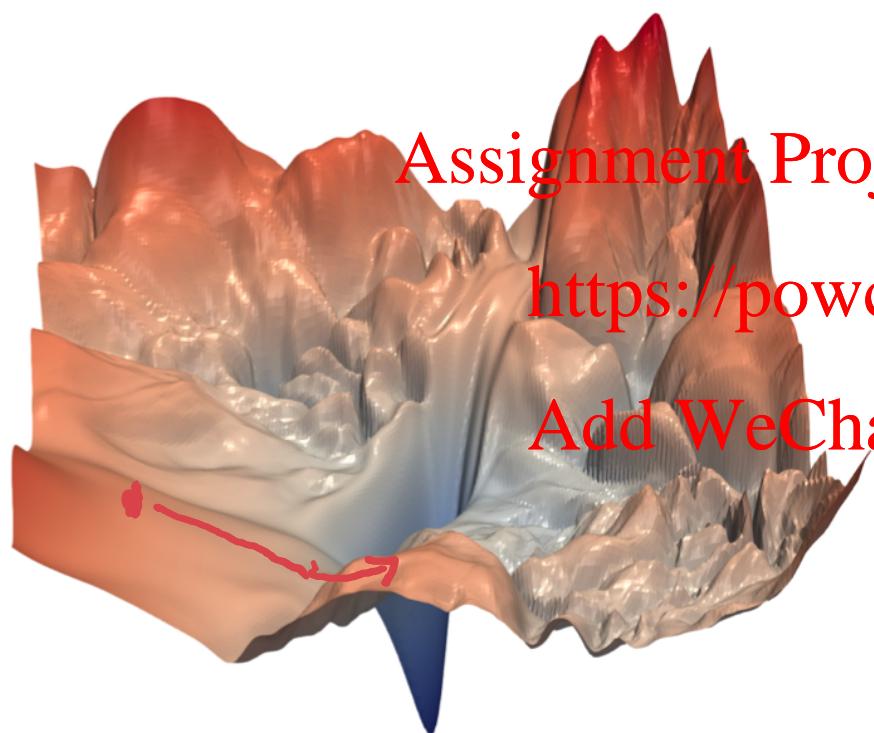
Add WeChat powcoder



TRAINING EPOCHS



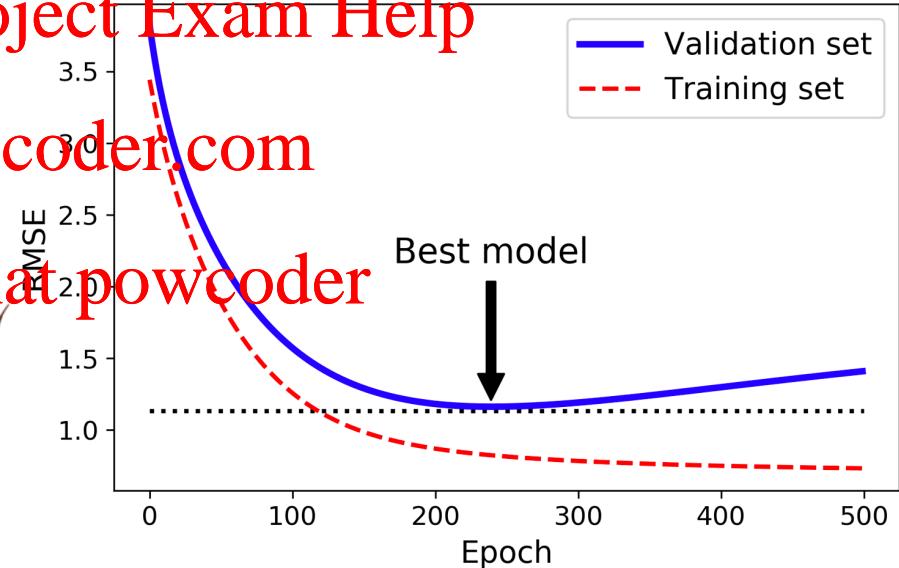
ERROR SURFACES



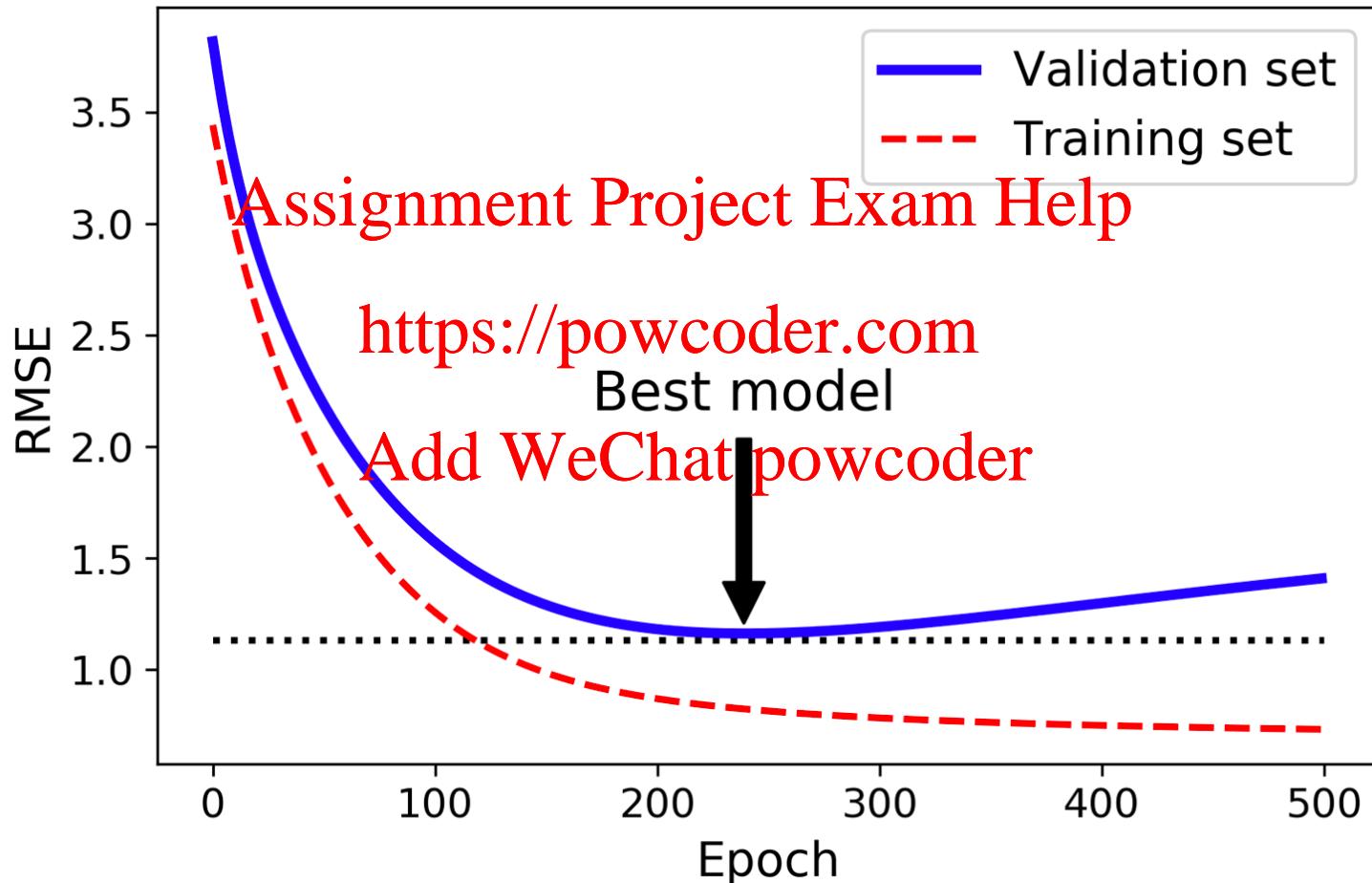
Assignment Project Exam Help

<https://powcoder.com>

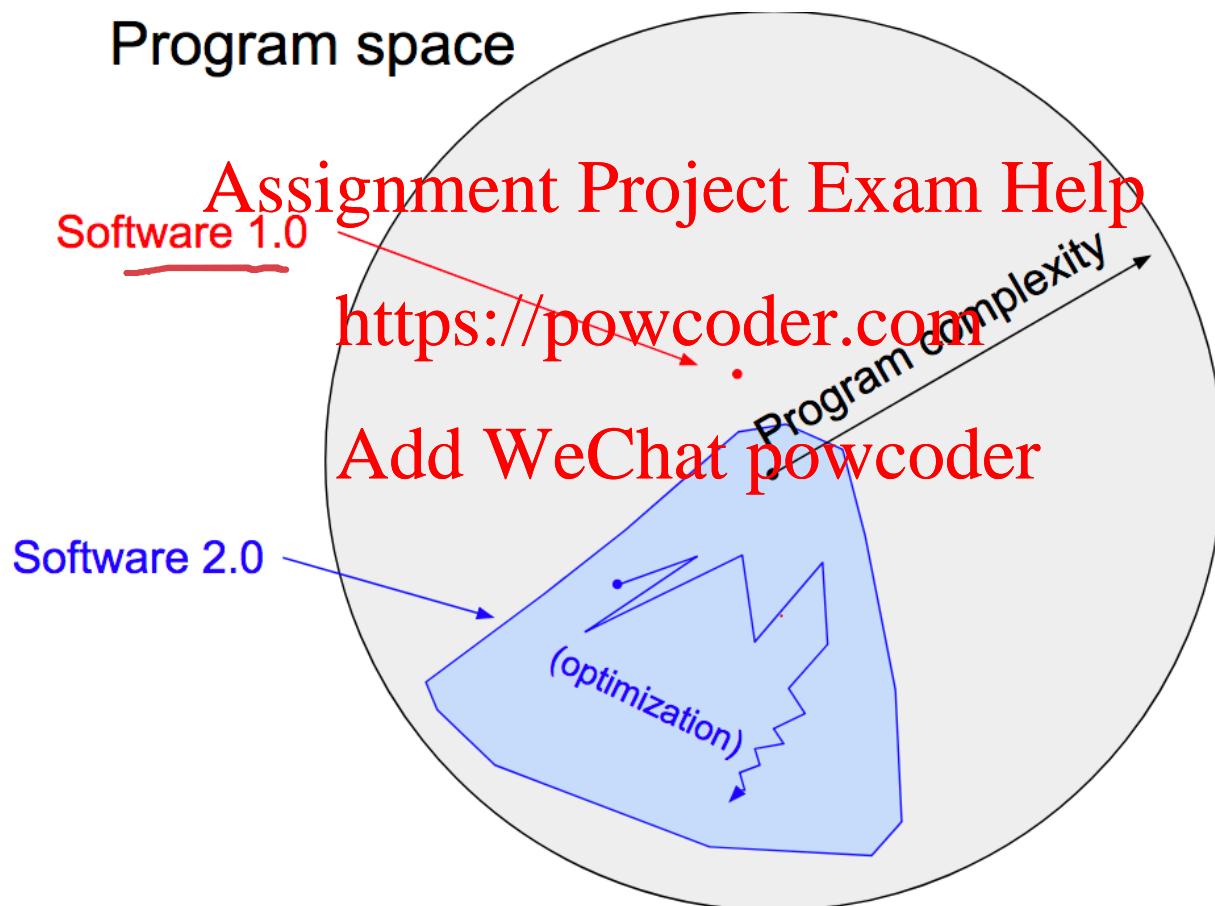
Add WeChat powcoder

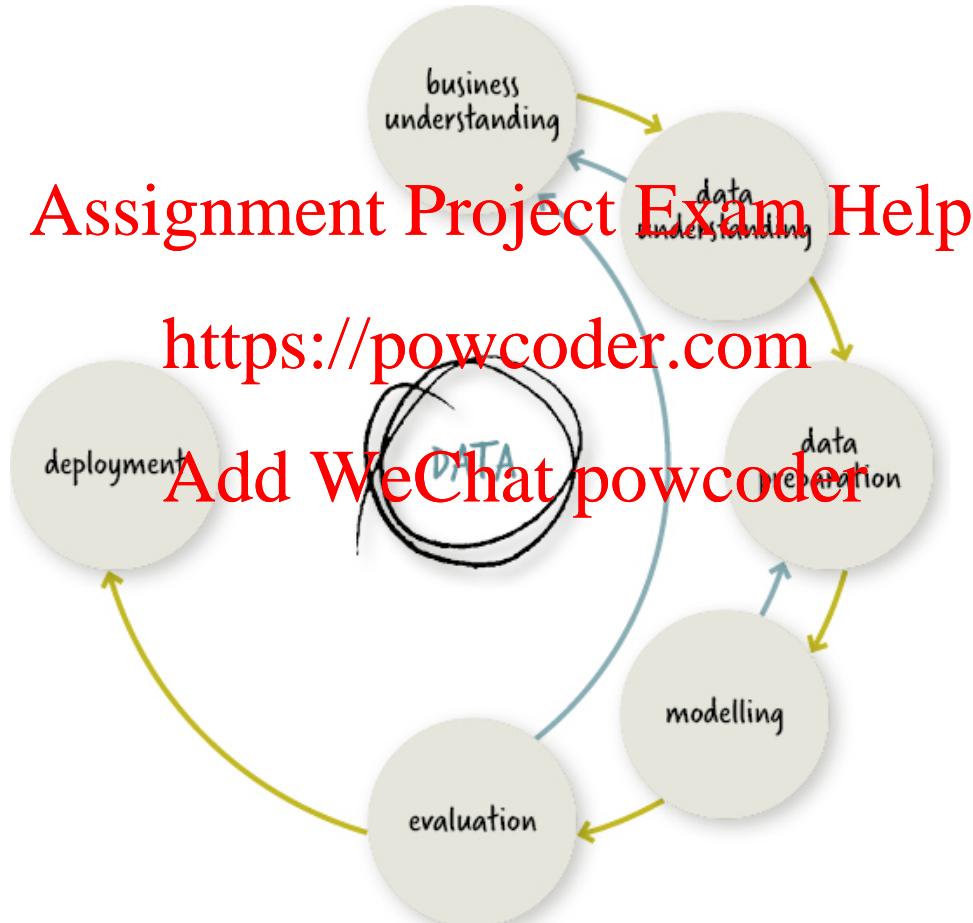


EARLY STOPPING



PROGRAM SPACE





PREDICTIVE ANALYTICS Assignment Project Exam Help

<https://powcoder.com>

Add WeChat powcoder

A P MOORE