



BITTIGER

DS 501 Data scientist express bootcamp

Week 3 [Ella]

版权声明

所有太阁官方网站以及在第三方平台课程中所产生的课程内容，如文本，图形，徽标，按钮图标，图像，音频剪辑，视频剪辑，直播流，数字下载，数据编辑和软件均属于太阁所有并受版权法保护。

对于任何尝试散播或转售BitTiger的所属资料的行为，太阁将采取适当的法律行动。



有关详情，请参阅

<https://www.bittiger.io/termsfuse> <https://www.bittiger.io/termservice>

Copyright Policy

All content included on the Site or third-party platforms as part of the class, such as text, graphics, logos, button icons, images, audio clips, video clips, live streams, digital downloads, data compilations, and software, is the property of BitTiger or its content suppliers and protected by copyright laws.

Any attempt to redistribute or resell BitTiger content will result in the appropriate legal action being taken.



We thank you in advance for respecting our copyrighted content.

For more info:

see <https://www.bittiger.io/termsfuse>

and <https://www.bittiger.io/termservice>



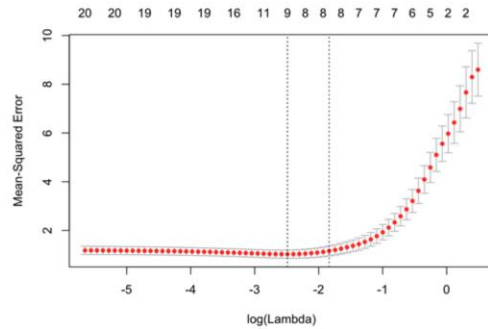
Summary

- Regularization in R
 - Ridge and lasso
 - Cross validation to choose best parameter
 - Linear regression with regularization
- Build logistic regression model
 - Coefficient estimation
 - Model performance
 - Add regularization



Regularization

- Recap the definition
 - Ridge (L2)
 - Lasso (L1)
 - Elastic net (mixed)
- Choose best penalty
 - Cross validation
- Library(glmnet)





glmnet() function

- glmnet(x, y)
 - alpha
 - weights
 - nlambda
 - lambda
 - **Standardize**, default is T
- Exercise



logistic regression model

- Build logistic regression model
 - Data exploration
 - Eye-balling the decision boundary
 - Hypothesis Function and likelihood
 - Gradient descent/ascent
 - Evaluate model performance
 - Exercise