

Module 2 – Regression and Prediction

CASE STUDY ACTIVITY TUTORIAL

Case Study 4: Predicting Wages 2

Contents of Case Study 4:

The video that covers this case study is given in Module 2, Segment 3.3.

1. Codebook.txt contains the name of the variables and a brief description.
2. wage2015.Rdata: The dataset contains the variables used in the prediction analysis.
3. Regression3.3.CaseStudy.R performs a prediction analysis where weekly wage is predicted using several demographic and job related characteristics. Several machine learning methods are used for prediction and their performance are compared based on out-of-sample mean squared error and R-squared. The final part of the program aggregates the predictions obtained by each machine learning method.
4. Regression.3.3.pdf is the set of slides that describes the estimation technique and present the results.