

$R^2$  for the test and train datasets vs number of epochs

