



Lab Guide

Machine Learning with Pentaho Data Integration

Credit Card Fraud – randomForest in R

Courseware Version: 1.0

Table of Contents

Lab 1	AutoML	1
Lab 2	randomForest	14
Appendix A	Python script with PDI	22
Appendix B	TPOT Parameters	23
Appendix C	R script for Train	24
Appendix D	R script for Predict	25

Hitachi Vantara



Corporate Headquarters

2535 Augustine Drive
Santa Clara, CA 95054 USA
www.HitachiVantara.com | community.HitachiVantara.com

Regional Contact Information

Americas: +1 866 374 5822 or info@hitachivantara.com
Europe, Middle East and Africa: +44 (0) 1753 618000 or info.emea@hitachivantara.com
Asia Pacific: +852 3189 7900 or info.marketing.apac@hitachivantara.com