Tree-Based Machine Learning Methods in SAS® Viya®

Course code: VBBF35

Decision trees and tree-based ensembles are supervised learning models used for problems involving classification and regression. This course covers everything from using a single tree to more advanced bagging and boosting ensemble methods in SAS Viya. The course includes discussions of tree-structured predictive models and the methodology for growing, pruning, and assessing decision trees, forest models, and gradient boosting models. The course also explains isolation forest (an unsupervised learning algorithm for anomaly detection), deep forest (an alternative for neural network deep learning), and Poisson and Tweedie gradient boosted regression trees. In addition, many of the auxiliary uses of trees, such as exploratory data analysis, dimension reduction, and missing value imputation, are examined, and running open source in SAS and running SAS in open source are demonstrated for tree-based ensemble models.

Who is the course for

Predictive modelers and data analysts who want to build decision trees and ensembles of decision trees using SAS Visual Data Mining and Machine Learning in SAS Viya

What we teach you

- Build tree-structured models, including classification trees and regression trees
- Use the methodology for growing, pruning, and assessing decision trees
- Build tree-based ensemble models, including forest and gradient boosting
- Run isolation forest and Poisson and Tweedie gradient boosted regression tree mogels
- Provide an introduction to deep forest models
- Implement open source in SAS and SAS in open source
- Use decision trees for exploratory data analysis, dimension reduction, and missing value imputation

Required skills

Before attending this course, you should have the following:

- An understanding of basic statistical concepts. You can gain this knowledge from the SAS® Visual Statistics in SAS® Viya®: Interactive Model Building course
- Familiarity with SAS Visual Data Mining and Machine Learning software. You can gain this knowledge from the Machine Learning Using SAS Viya course

Course outline

Introduction to Decision Trees

- Tree-structured models
- Tree-based models in SAS Viya
- Regression trees

Growing a Decision Tree

- Recursive partitioning
- Split search
- Splitting criteria
- Missing values and variable importance

Preventing Overfitting in Decision Trees

- Pruning
- Cross validation
- Autotuning
- Secondary uses of decision trees

Ensembles of Trees: Bagging, Boosting, and Forest

- Bagging and boosting

GOPAS Praha
Kodaňská 1441/46
101 00 Praha 10
Tel.: +420 234 064 900-3
info@gopas.cz

GOPAS Brno
Nové sady 996/25
602 00 Brno
Tel.: +420 542 422 111
info@gopas.cz

GOPAS Bratislava
Dr. Vladimíra Clementisa 10
Bratislava, 821 02
Tel.: +421 248 282 701-2
info@gopas.sk

- Forest models
- Open-source random forest models
- Isolation forest models
- Introducing deep forest models

Tree-Based Gradient Boosting Machines

- Overview of gradient boosting
- Tuning a gradient boosting model
- Gradient boosting for transfer learning
- Gradient boosted Poisson and Tweedie regression trees
- SAS gradient boosting using open source

A Practice Case Study

- Data exploration
- Creating decision trees
- Out-of-time testing
- Creating tree-based ensemble models
- Hyperparameter tuning
- Model comparison

GOPAS Praha
Kodaňská 1441/46
101 00 Praha 10
Tel.: +420 234 064 900-3
info@gopas.cz

GOPAS Brno
Nové sady 996/25
602 00 Brno
Tel.: +420 542 422 111
info@gopas.cz

GOPAS Bratislava
Dr. Vladimíra Clementisa 10
Bratislava, 821 02
Tel.: +421 248 282 701-2
info@gopas.sk