

Modelagem e Previsão

Programa e informações gerais do curso

AUTOR

Fernando da Silva

AFILIAÇÃO

Análise Macro

Objetivo

O objetivo do curso de Modelagem e Previsão é o de capacitar estudantes e profissionais de mercado que estejam interessados em desenvolver modelos preditivos usando estatística, econometria e aprendizado de máquina, contemplando problemas de regressão e classificação, com aplicações nas linguagens R e Python.

A abordagem do curso é teórica e prática: introduzindo e formalizando, quando necessário, os modelos propostos e mostrando como aplicar com conjuntos de dados reais.

Ementa

O curso de Modelagem e Previsão percorrerá todo o ciclo de ciência de dados, desde a identificação do problema, coleta de dados, análise exploratória/visualização, modelagem, validação e tomada de decisão.

O curso é dividido em 6 partes e conta com mais de 40 aulas.

Introdução à Modelagem Preditiva

1. Boas vindas e Apresentação do Curso
2. Ciclo de Ciência de Dados
 - a. Entendendo o Ciclo
 - b. Análise Descritiva vs. Preditiva
3. Estruturas de Dados
 - a. Série temporal
 - b. Corte transversal
4. Na Prática: Construindo uma Análise de Dados

Fontes de Dados

1. Formatos de Dados
2. Fontes Nacionais: BCB, IBGE e IPEADATA
3. Fontes Internacionais: Banco Mundial, FMI, FRED e OCDE

Análise Exploratória de Dados e Visualização

1. Variáveis Quantitativas vs. Qualitativas
2. Processo Gerador de Dados
3. Sumário de Estatísticas
4. Gráfico de Dispersão
5. Gráfico de Histograma
6. Gráfico de Boxplot
7. Gráfico de Linha
8. Análise de Correlação
9. Análise de Autocorrelação
10. Estacionariedade
11. Análise de Séries Temporais

Problemas Práticos de Modelagem Preditiva

1. Fluxo de Modelagem e Previsão
2. Separação de amostras: treino e teste
3. Pré Processamento de Dados
 - a. Valores extremos
 - b. Valores ausentes
4. Viés e Variância

Modelos e Aplicações

1. Regressão Linear
2. Suavização exponencial
3. ARIMA
4. Prophet
5. Regressão Logística
6. Naive Bayes
7. Algoritmos k-NN
8. Support Vector Machine (SVM)
9. Árvores de Decisão
10. Redes Neurais

Validação

1. Avaliando o desempenho do modelo
2. Validação Cruzada
3. Aumentando o desempenho do modelo
 - a. Bootstrap
 - b. Bagging
 - c. Random Forests
 - d. Boosting
4. Interpretação de resultados e tomada de decisão

Requisitos

Para o melhor aproveitamento é esperado que o interessado em adquirir o curso cumpra os seguintes requisitos:

- Computador com acesso à internet;
- Cadastro de conta gratuita no [Posit Cloud](#);
- Cadastro de conta gratuita no [Google](#);
- Matemática a nível de cálculo diferencial/integral;
- Estatística básica;
- Nível introdutório/intermediário de R;
- Nível introdutório/intermediário de Python;
- Domínio de inglês para leitura.

Materiais

O curso oferece como material:

1. Notas de aula em arquivo HTML;
2. Vídeo aula gravada;
3. Códigos reproduzíveis de R e/ou Python em arquivo `.R` e/ou `.ipynb`;
4. Dados utilizados em formato `.csv` *offline* e/ou *online*.
5. Referências bibliográficas e materiais de apoio.