




Predição Nota Final

Escola Todos Unidos

Workshop ML para mulheres - AI Girls
Equipe: Ana Lícia, Ariadne e Mirella



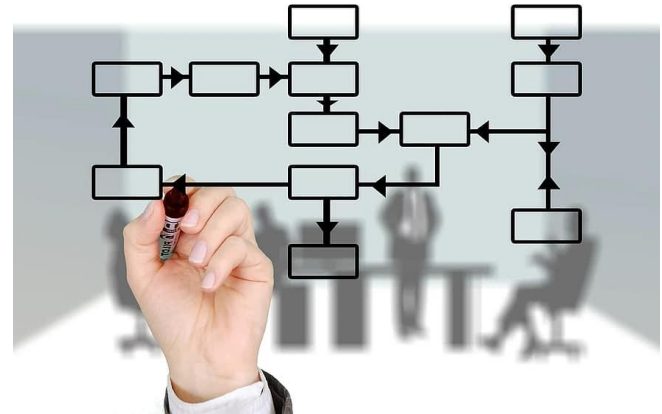
Descrição do Objetivo

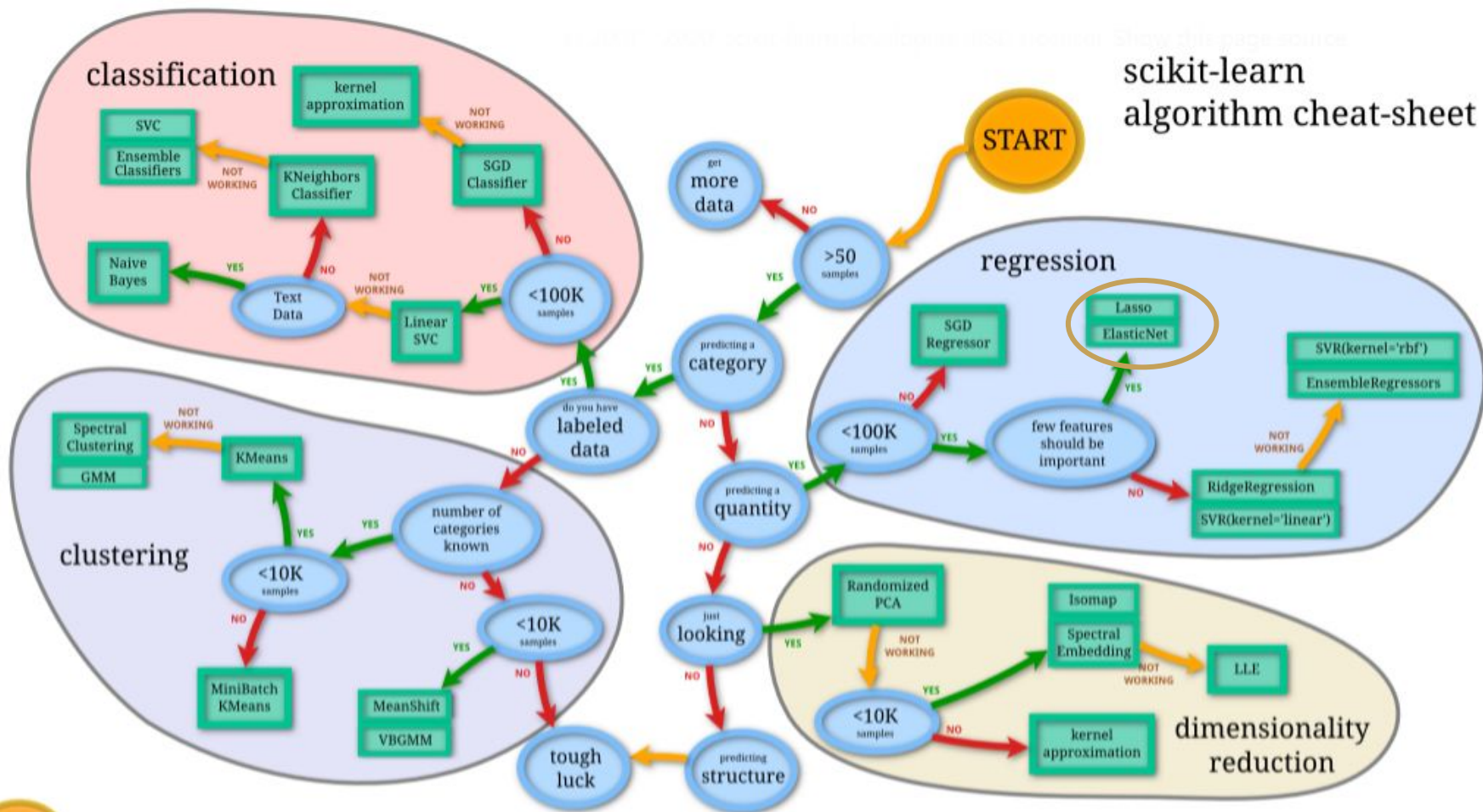
Criar um código que seja capaz de prever a nota final dos alunos de um grupo de escolas chamado Todos Unidos.

- Planilha: [Kaggle](#)
- Quantidade de dados: 395
- Total de variáveis: 33
- Variável-alvo: G3

Escolha do Método

- Algoritmo para dados não binários e supervisionados:
 - Variável alvo definida
 - Dados do tipo contínuo
- Requisitos:
 - Algoritmos baseados em árvores de decisão
 - Regressão
- Opção escolhida:
 - Random Forest Regressor

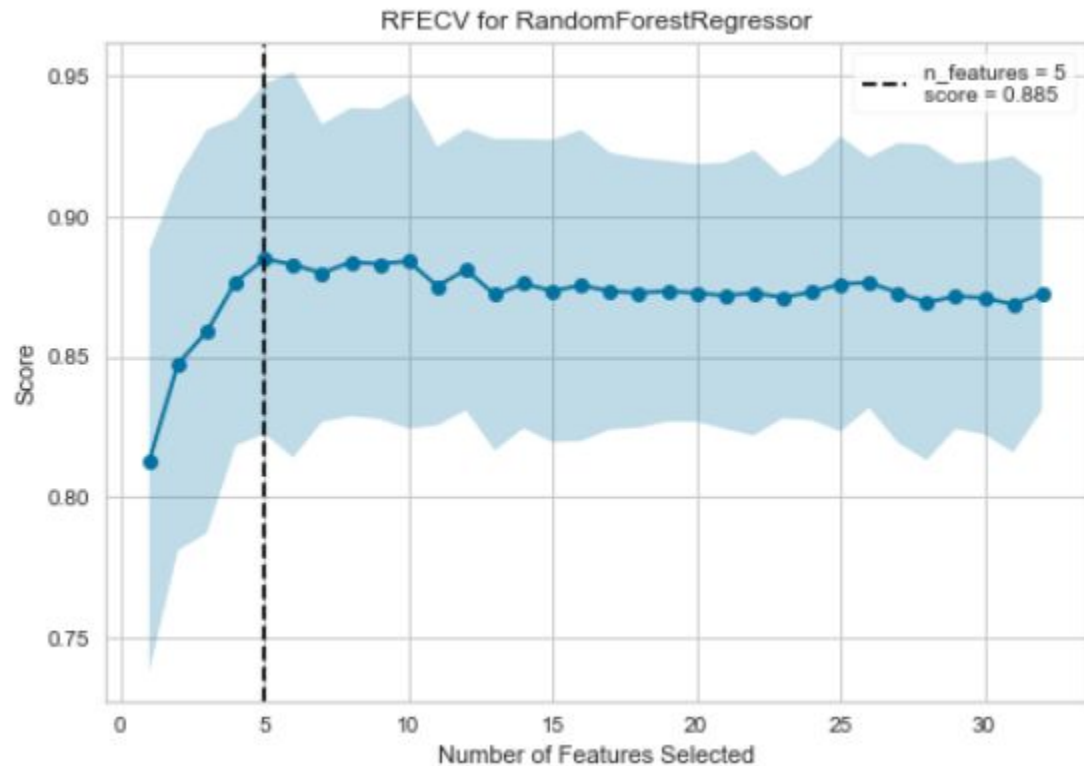




Escolha das variáveis

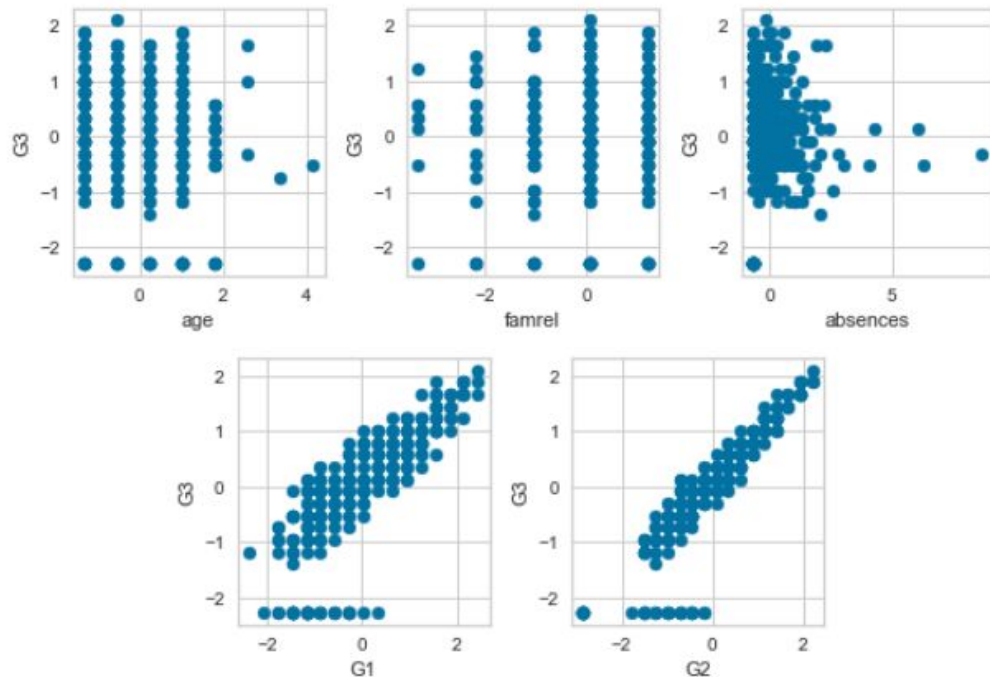
- Total de variáveis: 33
- Método escolhido:
 - RFECV: eliminação recursiva de variáveis com validação cruzada
 - Apresentação gráfica com Yellowbrick
 - Parâmetros:
 - Função Random Forest Regressor
 - cv: 10





Variáveis apresentadas como ideais para o modelo: 5 ('age', 'famrel', 'absences', 'G1', 'G2')

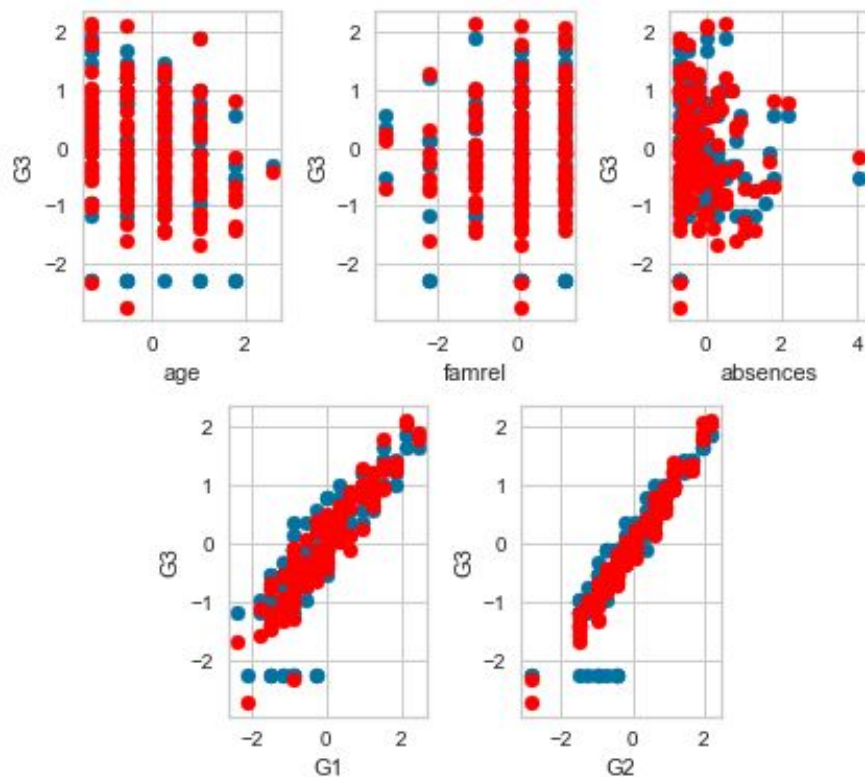
Relações das variáveis escolhidas com a variável-alvo



Aplicação do modelo

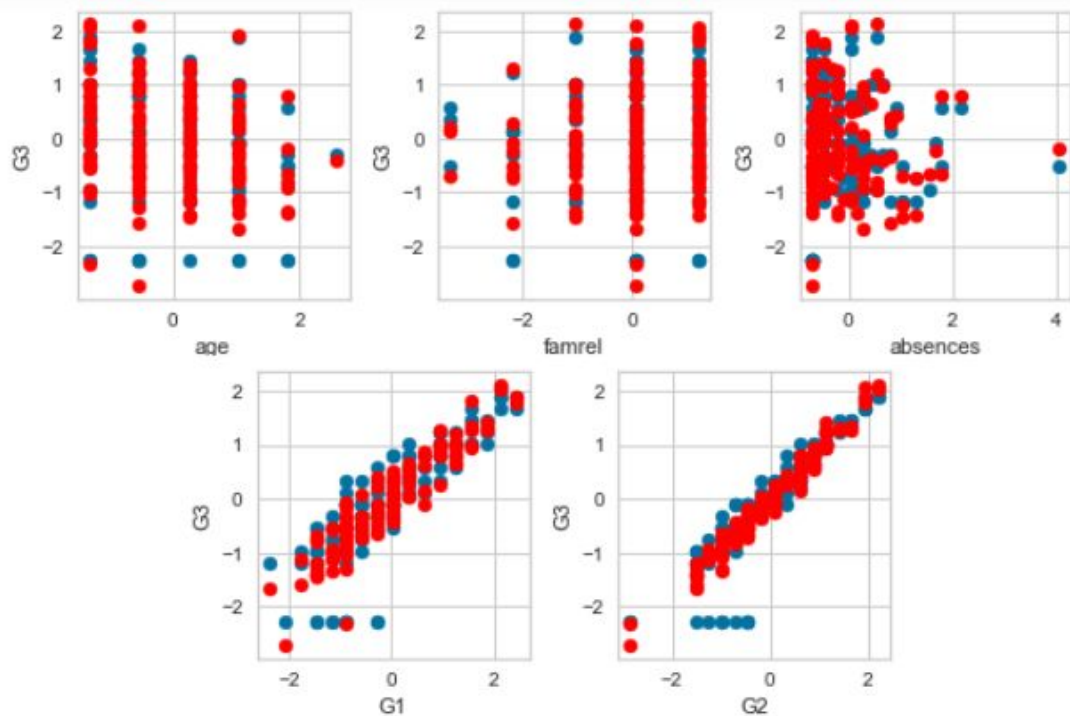
- Separação dos dados para treino e os dados para testes;
- Separação dos dados de target para treino e para teste
 - As métricas utilizadas para avaliar o desempenho do modelo e o desempenho final do modelo (30% - 70%)

Resultado - 33 variáveis



- Score: 0.7947778147438245
- MAE: 0.3058650060022317
- MSE: 0.21547115338548534
- RMSE: 0.4641887045001045

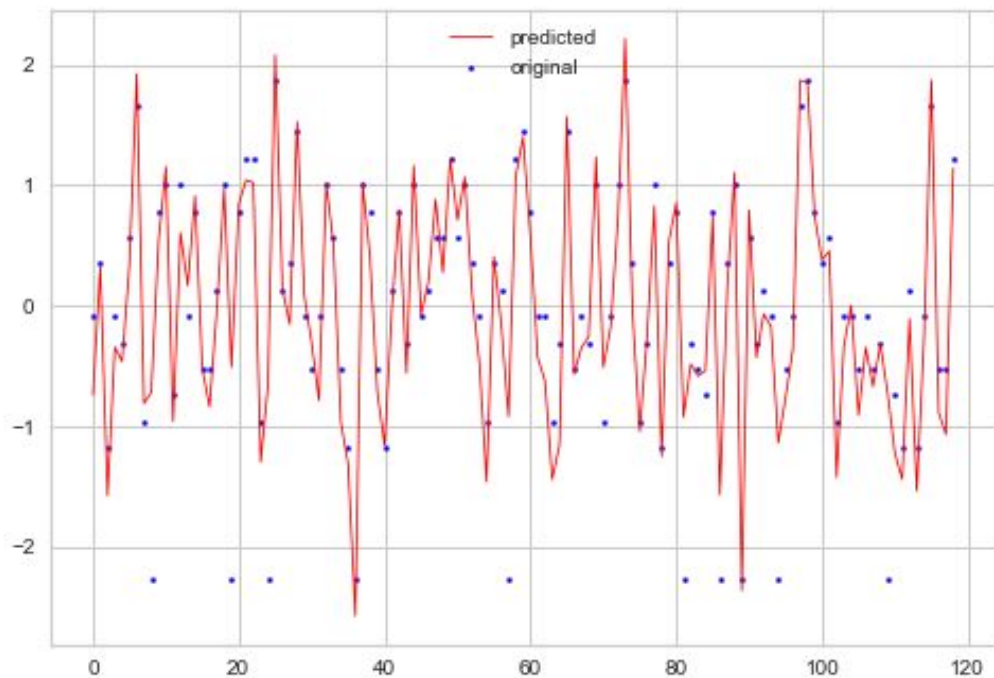
Resultado - 5 variáveis



- Score: 0.8091841086363154
- MAE: 0.2893939136456176
- MSE: 0.20034539708798535
- RMSE: 0.44759959460212356

Avaliação do Modelo

Comparação entre as notas obtidas pelo modelo e as notas reais



Comparando métricas

32 variáveis

- Score: 0.7947778147438245
- MAE: 0.3058650060022317
- MSE: 0.21547115338548534
- RMSE: 0.4641887045001045

5 variáveis

- Score: 0.8091841086363154
- MAE: 0.2893939136456176
- MSE: 0.20034539708798535
- RMSE: 0.44759959460212356

Conclusões

- O modelo consegue prever a nota final?
 - Inconsistência
- Quais são os pontos que mais influenciam a nota dos alunos?
 - G2 e G1
- O que a escola pode propor para melhorar a nota dos alunos?