

Econometria Tradicional e Machine Learning

Classificando cursos no ProUni

Pedro Cavalcante Oliveira

Faculdade de Economia, UFF

28 de Novembro, 2018

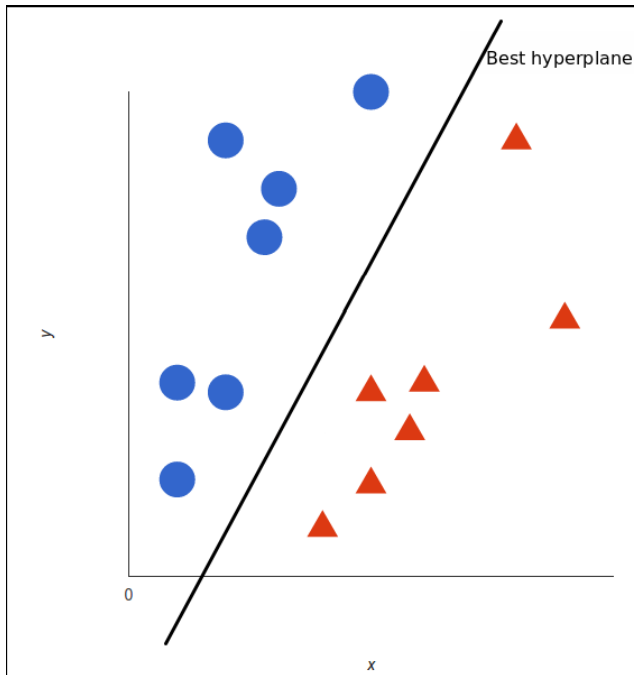
Existem 3 principais técnicas de classificação na econometria tradicional:

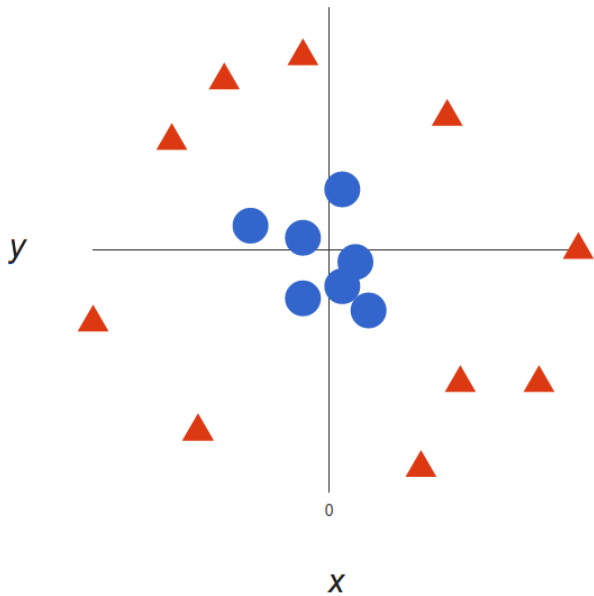
- Probit
- Logit
- Modelo Linear de Classificação

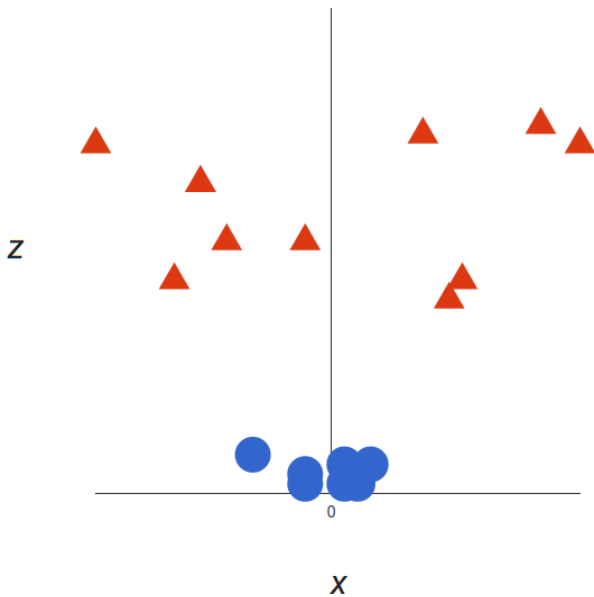
Existem várias em Machine Learning:

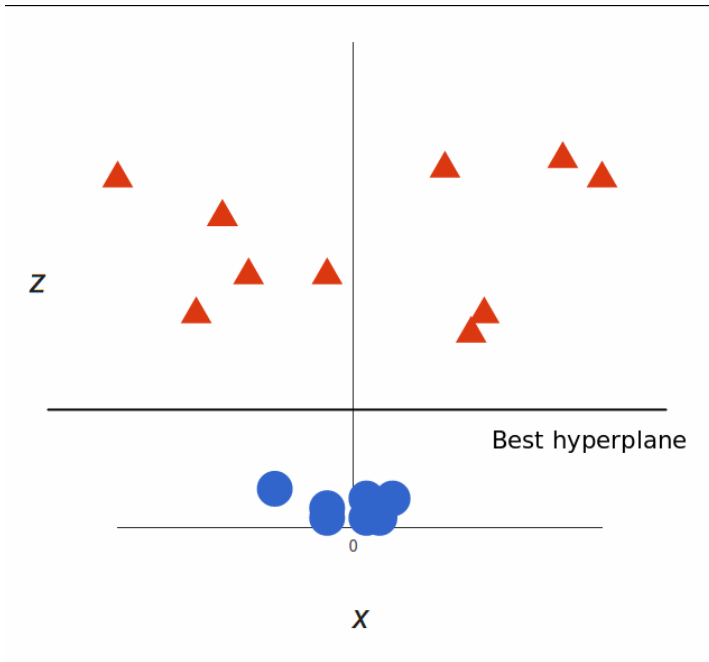
- Support Vector Machines (SVM)
- Clustering k -means
- Random Forests
- Hierarchical Clustering
- Mais uma infinidade...

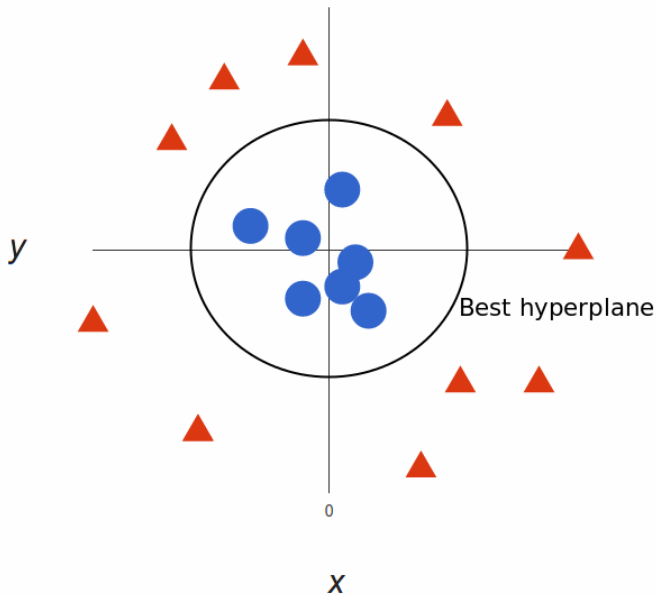
O que é Support Vector Machine?









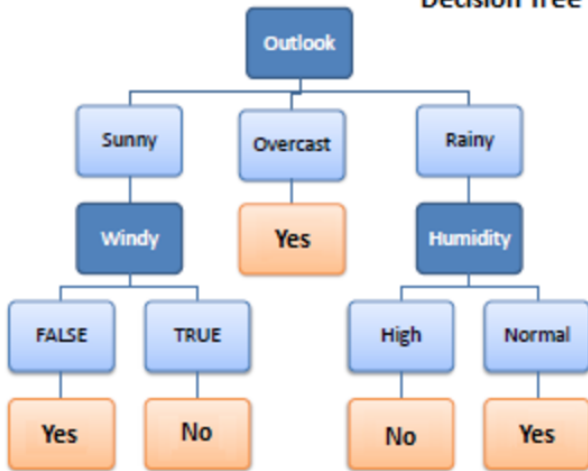


A função que define o hiperplano separador se chama *kernel*. Ela pode ser de 3 tipos:

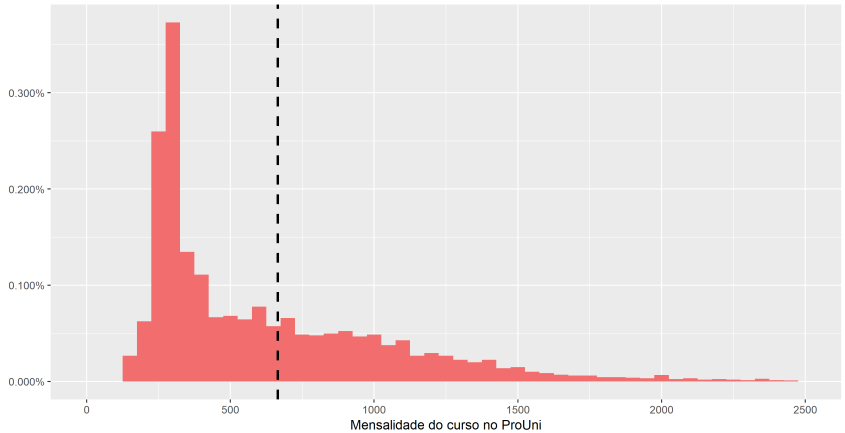
- Linear, em que não se cria dimensões extras
- Radial, em que o hiperplano é circular
- Polinomial, em que não há forma específica do hiperplano

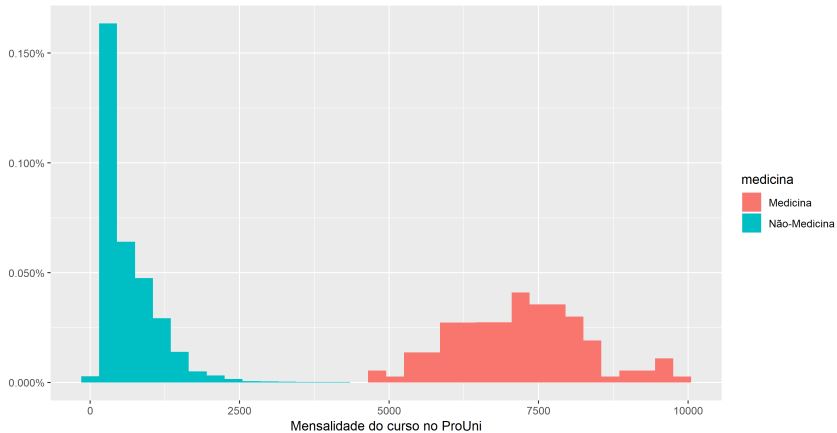
O que é Random Forest?

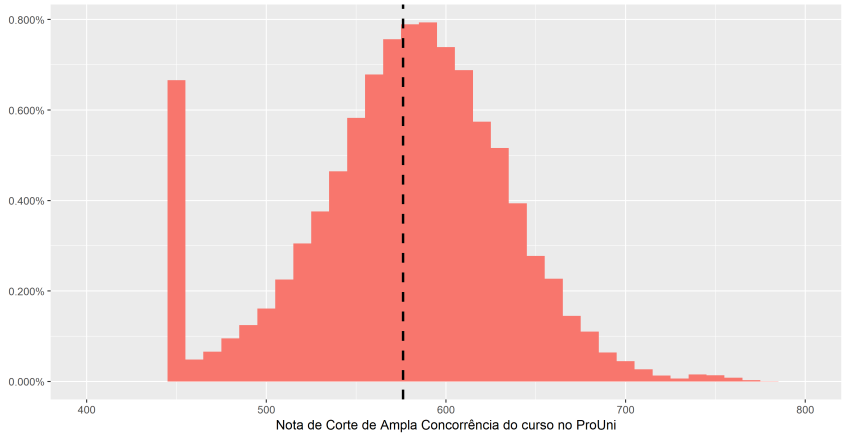
Decision Tree

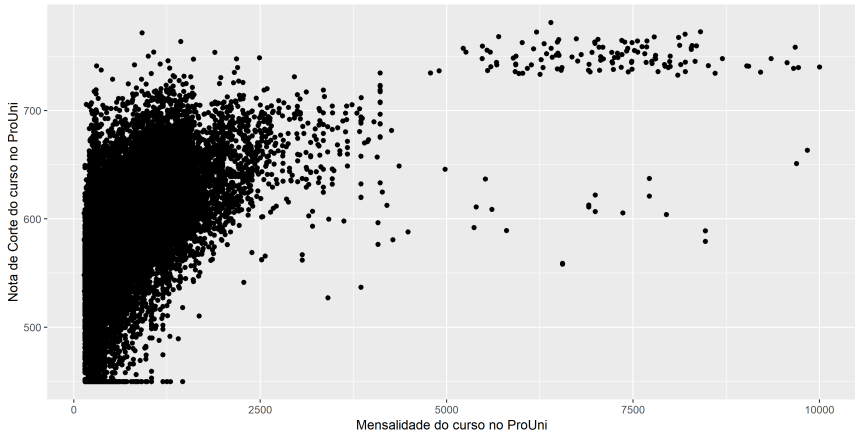


Explorando a amostra

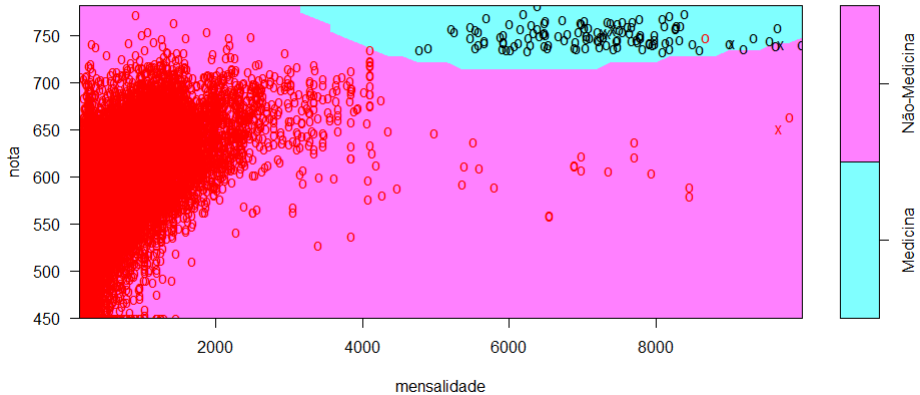








SVM classification plot



Para estimar a classificação o vamos rodar 3 modelos:

- SVM com kernel polinomial
- Random Forest
- Uma média ponderada de classificação com probit e com OLS

