Tópicos Gerais

Viés x Variância:

- https://youtu.be/EuBBz3bI-aA (~ 6 minutos) (STATQUEST)
- https://towardsdatascience.com/understanding-the-bias-variance-tradeoff-165e6942b229 (~4 minutos) (Towards Data Science)

Validação Cruzada:

- https://scikit-learn.org/stable/modules/cross_validation.html (Sklearn)
- https://youtu.be/fSytzGwwBVw (~ 6 minutos) (STATQUEST)
- https://medium.com/@eijaz/holdout-vs-cross-validation-in-machine-learning-7637112d3f8f (~ 2 minutos) (Medium)
- https://towardsdatascience.com/why-and-how-to-cross-validate-a-model-d6424b45261f (~ 4 minutos) (Towards Data Science)
- https://towardsdatascience.com/why-isnt-out-of-time-validation-more-ubiquitous-7397098c4ab6 (~ 7 minutos) (Towards Data Science)

Gradiente Descendente:

- https://youtu.be/sDv4f4s2SB8 (~ 24 minutos) (STATQUEST)
- https://youtu.be/htfh2xrnlaE (~ 51 minutos) (Didática Tech)
- https://arshren.medium.com/gradient-descent-5a13f385d403 (~ 5 minutos) (Medium)

Likelihood x Odds x Probability:

- https://youtu.be/pYxNSUDSFH4 (~ 5 minutos) (STATQUEST)
- https://youtu.be/XepXtl9YKwc (~ 6 minutos) (STATQUEST)
- https://youtu.be/ARfXDSkQf1Y (~ 11 minutos) (STATQUEST)

Distâncias:

- https://towardsdatascience.com/importance-of-distance-metrics-in-machine-learning-modelling-e51395ffe60d (~ 11 minutos) (Towards Data Science)

Regressão

Essenciais (~ 66 minutos):

(1) **Regressão Linear Simples**: https://youtu.be/PaFPbb66DxQ (~ 9 minutos) (STATQUEST)

(2) **Regressão Linear Múltipla**: https://youtu.be/yscO3epJTyQ (~ 2 minutos) (Estudar Com Você: Econometria)

(3) **Métricas de avaliação de regressão**: https://youtu.be/PjnKeAv5WmE (assistir do 3:45 até 14:10) (~ 10 minutos) (StatiR)

(4) **Regularização parte 1 - Ridge**: https://youtu.be/Q81RR3yKn30 (~ assistir até 13:22) (~ 13 minutos) (STATQUEST)

(5) **Regularização parte 2 - Lasso**: https://youtu.be/NGf0voTMlcs (~ 8 minutos) (STATQUEST)

(6) **Regularização parte 3 - Elastic Net**: https://youtu.be/1dKRdX9bfIo (~ 5 minutos) (STATQUEST)

(7) **Regularização parte 4 – Visualização Ridge x Lasso**: https://youtu.be/Xm2C_gTAl8c (~ 9 minutos) (STATQUEST)

(8) **Suposições parte 1**: https://youtu.be/ui0Hdd0U_qc (~ 5 minutos) (Estudar com você)

(9) **Suposições parte 2**: https://youtu.be/let-pzq5rp8 (~ 5 minutos) (Estudar com você)

Complementares (~ 43 minutos):

(1) **Matemática das regressões lineares** (beeeeem opcional...caso queria ver as contas, não exatamente conceitos): https://youtu.be/K_EH2abOp00 (~ 13 minutos) (CodeEmporium)

(2) **MSE x MAE**: https://www.coursera.org/lecture/competitive-data-science/regression-metrics-review-i-UWhYf (~ 14 minutos) (Coursera)

(3) **Métricas de avaliação**: https://www.dataquest.io/blog/understanding-regression-error-metrics/ (~ 10 minutos) (Dataquest)

(4) **Regressão Polinomial**: https://towardsdatascience.com/polynomial-regression-bbe8b9d97491 (~ 6 minutos) (Towards Data Science)

Se gostar muito do assunto:

Regularização (Lasso x Ridge) (complicado, tem que sentar e estudar, não é leitura de ônibus)**:** https://explained.ai/regularization/index.html

Curso de Econometria IE/Unicamp (livro-texto, linguagem acadêmica)**:** https://www.youtube.com/channel/UCEIgLZMzF76ifRnt2wta40A/videos

GLM

Essenciais (~ 16 minutos):

(1) **Generalized Linear Models 1**: https://towardsdatascience.com/generalized-linear-models-9cbf848bb8ab (~ 6 minutos) (Towards Data Science)

(2) **Generalized Linear Models 2**: http://www.est.ufmg.br/~enricoc/pdf/categoricos/mlg.pdf (até o slide 14) (~ 10 minutos) (UFMG)

Complementares (~ 7 minutos):

(1) **Detalhamento GLM's** (um pouco mais técnico, só pra quem quer ver como as funções se relacionam)**:** https://statmath.wu.ac.at/courses/heather_turner/glmCourse_001.pdf (até o slide 29) (~ 7 minutos) (University of Warwick)

Se gostar muito do assunto:

Aula do MIT (a matemática não é simples e a letra do professor é horrível, mas se entendeu bem os conceitos antes de ver a aula dá pra acompanhar)**:** https://youtu.be/X-ix97pw0xY (MIT)

Regressão Logística

Essenciais (~ 85 minutos):

(1) **Regressão Logística**: https://youtu.be/yIYKR4sgzI8 (~ 9 minutos) (STATQUEST)

(2) **Regressão Logística 2**: https://towardsdatascience.com/understanding-logistic-regression-step-by-step-704a78be7e0a (~ 6 minutos) (Towards Data Science)

(3) **Coeficientes**: https://youtu.be/vN5cNN2-HWE (~ 19 minutos) (STATQUEST)

(4) **Maximum Likelihood**: https://youtu.be/BfKanl1aSG0 (~ 10 minutos) (STATQUEST)

(5) **R^2 e p-value**: https://youtu.be/xxFYro8QuXA (~ 15 minutos) (STATQUEST)

(6) **Métricas de classificação:** https://medium.com/@MohammedS/performance-metrics-for-classification-problems-in-machine-learning-part-i-b085d432082b (~ 10 minutos) (Medium)

(7) **ROC-AUC:** https://youtu.be/4jRBRDbJemM (~ 16 minutos) (STATQUEST)

Complementares (~ 36 minutos):

(1) **Regressão Logística – Passo a passo**: https://towardsdatascience.com/logistic-regression-explained-9ee73cede081 (~ 6 minutos) (Towards Data Science)

(2) **Guia de Métricas**: https://towardsdatascience.com/the-ultimate-guide-to-binary-classification-metrics-c25c3627dd0a (30 minutos) (Towards Data Science)

Se gostar muito do assunto:

Regressão Logística e Perceptron (é bom ter pelo menos uma noção de redes neurais antes de ver): https://youtu.be/jbluHIgBmBo (Serrano Academy)

Classificadores Bayesianos

Essenciais (~ 38 minutos):

(1) **Introdução Naive Bayes:** https://medium.com/@srishtisawla/introduction-to-naive-bayes-for-classification-baefefb43a2d (~ 4 minutos) (Medium)

(2) **Naive Bayes**: https://youtu.be/O2L2Uv9pdDA (~ 15 minutos) (STATQUEST)

(3) **Gaussian Naive Bayes**: https://youtu.be/H3EjCKtlVog (~ 9 minutos) (STATQUEST)

(4) **LDA e QDA:** https://scikit-learn.org/stable/modules/lda_qda.html#lda-qda-math (~ 10 minutos) (Sklearn)

Complementares (~ 5 minutos):

(1) **Correção Laplaciana**: https://courses.cs.washington.edu/courses/cse446/20wi/Section7/naive-bayes.pdf (~ 5 minutos) (University of Washington)

SVM

Essenciais (~ 43 minutos):

(1) **Support Vector Machines 1:** https://youtu.be/efR1C6CvhmE (~ 20 minutos) (STATQUEST)

(2) **Kernel:** https://towardsdatascience.com/understanding-the-kernel-trick-e0bc6112ef78 (~ 4 minutos) (Towards Data Science)

(3) **Support Vector Machines 2:** https://youtu.be/Toet3EiSFcM (~ 7 minutos) (STATQUEST)

(4) **Support Vector Machines 3:** https://youtu.be/Qc5IyLW_hns (~ 16 minutos) (STATQUEST)

Complementares (~ 5 minutos):

(1) **Support Vector Regressor:** https://towardsdatascience.com/an-introduction-to-support-vector-regression-svr-a3ebc1672c2 (~ 5 minutos) (Towards Data Science)

Se gostar muito do assunto:

Aula SVM (passa por pré-processamento, kerneis, otimização de hiper parâmetros, validação cruzada etc tudo em python, bom demais!)**:** https://youtu.be/8A7L0GsBiLQ (STATQUEST)

Aula do MIT (parte matemática do algoritmo)**:** https://youtu.be/_PwhiWxHK8o (MIT)

KNN

Essenciais (~ 62 minutos):

(1) **KNN:** https://youtu.be/HVXime0nQeI (~ 5 minutos) (STATQUEST)

(2) **KNN 2 + Exemplo em Python**: https://youtu.be/4HKqjENq9OU (~ 28 minutos) (Simplilearn)

(3) **Geral + Python:** https://towardsdatascience.com/k-nearest-neighbors-knn-explained-cbc31849a7e3 (~ 6 minutos) (Towards Data Science)

(4) **Geral + Distâncias:** https://medium.datadriveninvestor.com/k-nearest-neighbors-knn-7b4bd0128da7 (~ 6 minutos) (Medium)

(5) **Geral + Regressor:** https://medium.com/roottech/knn-understanding-k-nearest-neighbor-algorithm-in-python-71488b8802f0 (~ 9 minutos) (Medium)

(6) **KD Tree x Ball Tree x Brute Force:** https://towardsdatascience.com/tree-algorithms-explained-ball-tree-algorithm-vs-kd-tree-vs-brute-force-9746debcd940 (~ 8 minutos) (Towards Data Science)

Se gostar muito do assunto:

Aula KD Tree e Ball Tree 1: https://youtu.be/BzHJ57QCdVo (assistir a partir de 30 minutos) (Cornell's Machine Learning Course)
Aula KD Tree e Ball Tree 2: https://youtu.be/_PwhiWxHK8o (assistir até 33 minutos) (Cornell's Machine Learning Course)

Árvores de decisão

Essenciais (~ 86 minutos):

(1) **Decision Tree:** https://youtu.be/7VeUPuFGJHk (~ 17 minutos) (STATQUEST)

(2) **Regression Tree**: https://youtu.be/g9c66TUylZ4 (~ 22 minutos) (STATQUEST)

(3) **Árvores de Classificação e Regressão + Python:** https://towardsdatascience.com/https-medium-com-lorrli-classification-and-regression-analysis-with-decision-trees-c43cdbc58054 (~ 8 minutos) (Towards Data Science)

(4) **Geral + Distâncias:** https://medium.datadriveninvestor.com/k-nearest-neighbors-knn-7b4bd0128da7 (~ 6 minutos) (Medium)

(5) **Regression Tree Prunning:** https://youtu.be/D0efHEJsfHo (~ 16 minutos) (STATQUEST)

(6) **Decision Tree Prunning:** https://youtu.be/u4kbPtiVVB8 (~ 17 minutos) (Sebastian Mantey)

Esembles

Essenciais (~ 66 minutos):

(1) **Random Forest 1:** https://youtu.be/J4Wdy0Wc_xQ (~ 10 minutos) (STATQUEST)
(2) **Random Forest 2:** https://youtu.be/sQ870aTKqiM (~ 12 minutos) (STATQUEST)
(3) **Esembles:** https://towardsdatascience.com/basic-ensemble-learning-random-forest-adaboost-gradient-boosting-step-by-step-explained-95d49d1e2725 (~ 6 minutos) (Towards Data Science)

(4) **Bagging for dummies:** https://medium.com/machine-learning-through-visuals/machine-learning-through-visuals-part-1-what-is-bagging-ensemble-learning-432059568cc8 (~ 2 minutos) (Medium)

(5) **Bagging, Boosting e Stacking:** https://towardsdatascience.com/ensemble-methods-bagging-boosting-and-stacking-c9214a10a205 (~ 20 minutos) (Towards Data Science)

(6) **AdaBoost:** https://medium.com/analytics-vidhya/implementing-an-adaboost-classifier-from-scratch-e30ef86e9f1b (~ 8 minutos) (Medium)

(7) **Gradient Boosting:** https://blog.mlreview.com/gradient-boosting-from-scratch-1e317ae4587d (~ 8 minutos) (Medium)

Agrupamento

Essenciais (~ 74 minutos):

(1) **Introdução Algoritmos de Clusterização:** https://towardsdatascience.com/overview-of-clustering-algorithms-27e979e3724d (~ 6 minutos) (Towards Data Science)

(2) **KMeans 1:** https://towardsdatascience.com/k-means-clustering-algorithm-applications-evaluation-methods-and-drawbacks-aa03e644b48a (~ 13 minutos) (Towards Data Science)

(3) **KMeans 2:** https://youtu.be/4b5d3muPQmA (~ 8 minutos) (STATQUEST)

(4) **KMedoids:** https://towardsdatascience.com/k-medoids-clustering-on-iris-data-set-1931bf781e05 (~ 7 minutos) (Towards Data Science)

(5) **Método do Cotovelo:** https://medium.com/analytics-vidhya/elbow-method-of-k-means-clustering-algorithm-a0c916adc540 (~ 3 minutos) (Medium)

(6) **Silhueta:** https://towardsdatascience.com/silhouette-coefficient-validating-clustering-techniques-e976bb81d10c (~ 3 minutos) (Towards Data Science)

(7) **KMeans x Kmedian** (quando usar cada um)**:** https://stats.stackexchange.com/questions/109547/k-means-vs-k-median (~ 2 minutos) (StackExchange)

(8) **Clusterização Hierárquica:** https://youtu.be/7xHsRkOdVwo (~ 11 minutos) (STATQUEST)

(9) **Clusterização Hierárquica 2:** https://towardsdatascience.com/understanding-the-concept-of-hierarchical-clustering-technique-c6e8243758ec (~ 7 minutos) (Towards Data Science)

(10) **Linkage:** https://towardsdatascience.com/introduction-to-hierarchical-clustering-part-1-theory-linkage-and-affinity-e3b6a4817702 (~ 7 minutos) (Towards Data Science)

(11) **Introdução DBSCAN:** https://towardsdatascience.com/understanding-the-concept-of-hierarchical-clustering-technique-c6e8243758ec (~ 3 minutos) (Great Learning)

(12) **DBSCAN:** https://towardsdatascience.com/machine-learning-clustering-dbscan-determine-the-optimal-value-for-epsilon-eps-python-example-3100091cfbc (~ 4 minutos) (Towards Data Science)

Complementares (~ 8 minutos):

(1) **Bisecting KMeans:** https://youtu.be/ZvXK1HH16vM (~ 8 minutos) (Ranji Raj)

GMM

Essenciais (~ 45 minutos):

(1) **GMM 1:** https://towardsdatascience.com/gaussian-mixture-models-explained-6986aaf5a95 (~ 12 minutos) (Towards Data Science)

(2) **GMM 2:** https://youtu.be/q71Niz856KE (~ 17 minutos) (Serrano Academy)

(3) **GMM + Python:** https://jakevdp.github.io/PythonDataScienceHandbook/05.12-gaussian-mixtures.html (~ 10 minutos) (Python Data Science Handbook)

(5) **AIC x BIC:** https://medium.com/analytics-vidhya/probabilistic-model-selection-with-aic-bic-in-python-f8471d6add32 (~ 6 minutos) (Medium)

Complementares (~ 25 minutos):

(1) **Guia GMM** (meio complicado, mas o melhor material)**:** https://brilliant.org/wiki/gaussian-mixture-model/ (~ 20 minutos) (Brilliant)

(2) **GMM + Python** (aqui o Python é do zero)**:** https://towardsdatascience.com/gaussian-mixture-models-implemented-from-scratch-1857e40ea566 (~ 5 minutos) (Towards Data Science)

Redução de Dimensionalidade

Essenciais (~ 34 minutos):

(1) **PCA 1:** https://youtu.be/HMOI_lkzW08 (~ 6 minutos) (STATQUEST)

(2) **PCA 2:** https://youtu.be/FgakZw6K1QQ (~ 22 minutos) (STATQUEST)

(3) **PCA + Python:** https://towardsdatascience.com/principal-component-analysis-pca-from-scratch-in-python-7f3e2a540c51 (~ 6 minutos) (Towards Data Science)

Se gostar muito do assunto:

Outros métodos: https://www.analyticsvidhya.com/blog/2018/08/dimensionality-reduction-techniques-python/ (Analytics Vidhya)

Estatística

Essenciais (~ 76 minutos):

(1) **Fundamentos 1:** https://www.cienciaedados.com/probabilidade-e-estatistica-os-fundamentos-para-cientistas-de-dados-parte-1/ (~ 5 minutos) (Ciência e Dados)

(2) **Fundamentos 2:** https://www.cienciaedados.com/probabilidade-e-estatistica-os-fundamentos-para-cientistas-de-dados-parte-2/ (~ 3 minutos) (Ciência e Dados)

(3) **Descritiva + Python:** https://realpython.com/python-statistics/ (~ 20 minutos) (Real Python)

(4) **Distribuições + Python:** https://www.datacamp.com/community/tutorials/probability-distributions-python (~ 15 minutos) (Datacamp)

(5) **Teste de Hipóteses:** https://medium.com/data-hackers/o-que-realmente-um-teste-de-hip%C3%B3teses-quer-nos-dizer-b82801b03529 (~ 18 minutos) (Medium)

(6) **Teste de Hipóteses 2:** https://medium.com/rodrigo-lampier/usando-teste-de-hip%C3%B3teses-para-responder-quest%C3%B5es-de-neg%C3%B3cio-8a3d5ae9ebc0 (~ 15 minutos) (Medium)

Se gostar muito do assunto:

Fundamentos (aqui tem um guia com links que redirecionam para qualquer assunto de fundamentos de estatística que interessar)**:** https://www.statlect.com (StatLect)

Redes Neurais
