Statistical learning – hastie and tibshirani

Statistical LearningModel: Y= f(X) + epsilonWhat can a good f do- Predict

- Help understand which variables are relevant
- How each feature X_i affects target Y ONSTATISTICAL LEARNING – HASTIE & TIBSHIRANI SPECIFICALLY FOR YOUFOR ONLY$13. 90/PAGEOrder NowRegression Function- Ideal function: one that minimizes some loss func, e. g. MSE
- Turns out to be f(x) = E(Y| X) or average
- optimizes MSE (mean squared error)Nearest Nbr AveragingTo account for x without any observations, we can relax f(x) = E(Y| X) to f(x) = E[Y| X in N(x)] where N denoted neighborhoodCurse of dimenisnalityReducible vs Irreducible ErrorE[(Y - f''(X))^2| X= x] = [f''(x) - f(x)]^2 + Var(epsilon)Model Tradeoffs- Prediction accuracy vs interpretability
- under-fit vs over-fit
- Simple Model vs Black BoxBias vs Variance tradeoffE[y_0 - f '(x_0)]^2 = bias(f ') + var(f ') + var(epsilon)Classification ProblemModel classifier C(x) to predict class for x where class is in {1, 2, ... , L} - i. e. L classesconditional class probabilitiesp_i(x) = Pr(Y= i | X = x), i = 1, 2, ... , LBayes Optimal ClassifierC(x) = argmax_{i in 1, 2, ..., L} p_i(x)KNN (K-nearest neighbors)EquippedMisclassification errorErr_{Test} = mean_{i in Test} I[y_i neq C '(x_i)]