האוניברסיטה
העברית
בירושלים
THE HEBREW
UNIVERSITY
OF JERUSALEM

Artificial Intelligence in Medicine

Regression

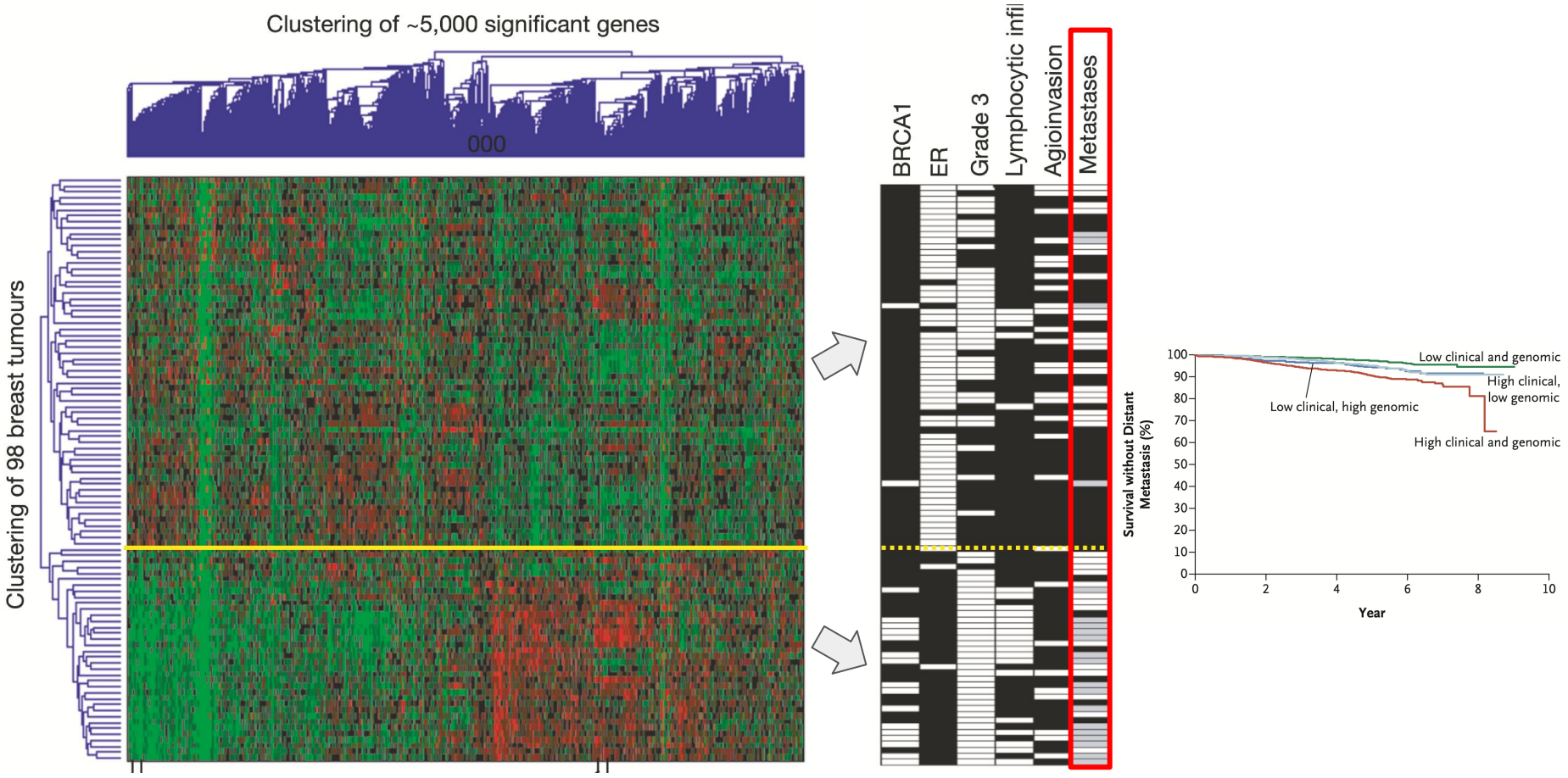
Nir Friedman and Tommy Kaplan 6/2/24

תוכן

- ניבוי רציף (כמותי) לעומת בינארי
- במקום קו, רוצים לנבא ערך, למשל תוחלת הערך הצפוי כפונקציה של פיצ'רים (ערכים אחרים)
- שערוך טעות (בינארית -> טעות ריבועית - בגלל שמניחים רעש עם התפלגות נורמלית)
- רגרסיה לינארית רגילה
- רגרסיה ריבועית, לוג, סינוס
- רגרסיה לינארית מרובה
- מורכבות ואובר-פיטינג
- רגולריזציה - רידג' L1, לאסו L2, אלסטיק-נט L1+L2
- עצי רגרסיה, יער
- מודל מבוסס קוהורט (שכנים)

Previously, we predicted metastases presence

what if we wanted to predict survival? dosage?



Regression models in medicine

- Pediatric dosage calculations (body weight)
- Chemotherapy (body surface area)

DuBois and DuBois[1] : 1.85 m^2
Equation: BSA (m^2) = 0.007184 x Height(cm)$^{0.725}$ x Weight(kg)$^{0.425}$

Gehan and George[2]: 1.85 m^2
Equation: BSA (m^2) = 0.0235 x Height(cm)$^{0.42246}$ x Weight(kg)$^{0.5145^6}$

Haycock[3]: 1.85 m^2
Equation: BSA (m^2) = 0.024265 x Height(cm)$^{0.3964}$ x Weight(kg)$^{0.5378}$

Mosteller[4]: 1.84 m^2
Equation: BSA (m^2) = SQR RT ([Height(cm) x Weight(kg)]/ 3600)

1) DuBois D, DuBois DF. A formula to estimate the approximate surface area if height and weight be known. Arch Int Med 1916;17:863-71.

2) Gehan EA, George SL. Estimation of human body surface area from height and weight. Cancer Chemother Rep 1970;54:225-35.

3) Haycock GB, Schwartz GJ, Wisotsky DH. Geometric method for measuring body surface area: A height-weight formula validated in infants, children and adults. J Pediatr 1978;93:62-6.

4) Mosteller RD. Simplified calculation of body-surface area. N Engl J Med 1987;317:1098.

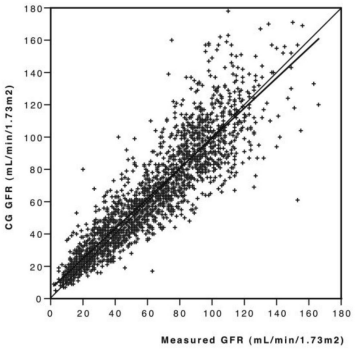
- GFR (Creatine, Creatinine, and kidney function)
  (estimated Glomerular Filtration Rate)

Prediction of Creatinine Clearance from Serum Creatinine[1]

DONALD W. COCKCROFT and M. HENRY GAULT

Departments of Medicine, Queen Mary Veterans' Hospital, Montreal, Quebec, and Memorial University, St. John's, Newfoundland

$$C_{cr} = \frac{(140 - \text{age})\,(\text{wt kg})}{72 \times S_{cr}\,(\text{mg}/100\,\text{ml})}$$



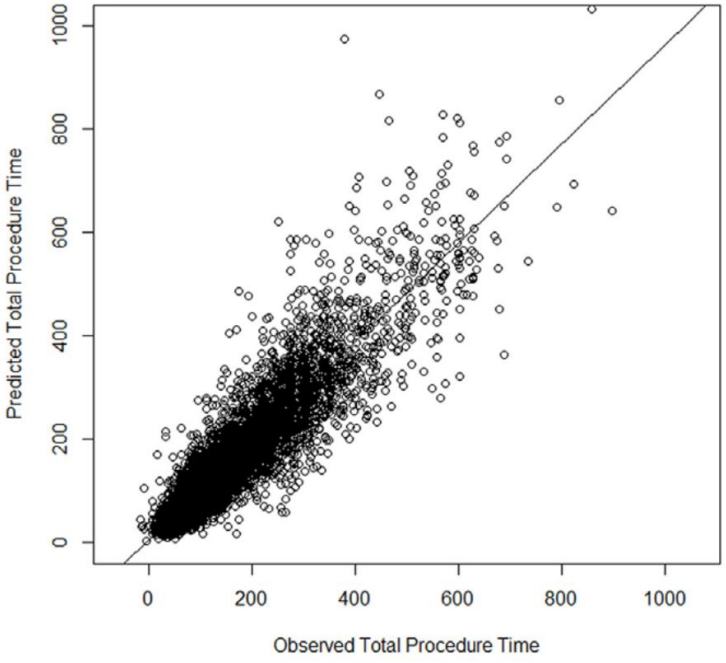
החישוב נעשה כדלהלן:
Val = [(140-**גיל**) X **משקל**) / (72 X **serum creatinine**)
אם הנבדק היא **נבדקת** יש להכפיל את הערך ב 0.85.

משמעות התוצאות

ערך	משמעות
מעל 50	תפקוד הכליות **תקין** אין צורך להתחשב בו למתן תרופות
בין 10 **ל** 50	ליקוי **בינוני** בתפקוד הכליות יש להתחשב בו במתן תרופות מסויימות
מתחת ל 10	ליקוי **חמור** בתפקוד הכליות

Improving the Prediction of Total Surgical Procedure Time Using Linear Regression Modeling

Eric R. Edelman [1], Sander M. J. van Kuijk [2], Ankie E. W. Hamaekers [3], Marcel J. M. de Korte [3], Godefridus G. van Merode [4] and Wolfgang F. F. A. Buhre [3]*



Multiple regression model to analyze the total LOS for patients undergoing laparoscopic appendectomy

Teresa Angela Trunfio[1], Arianna Scala[2*], Cristiana Giglio[3], Giovanni Rossi[4], Anna Borrelli[4], Maria Romano[5] and Giovanni Improta[2,6]

Noise is all around us

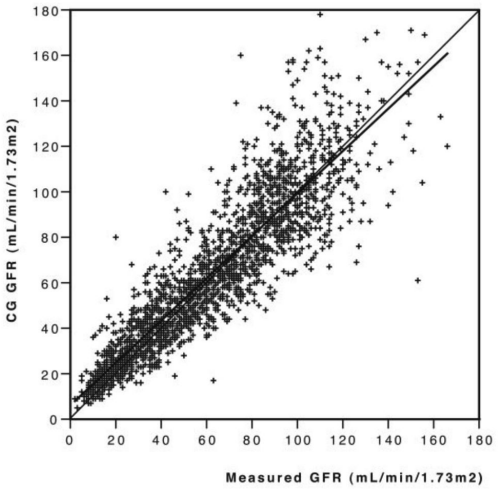
Sum of multiple effects ⇒ Central limit theorem

המשפט החזק של המספרים החלשים

In the absence of a better model or knowledge

Non-additive noise ⇒ multiplicative cases

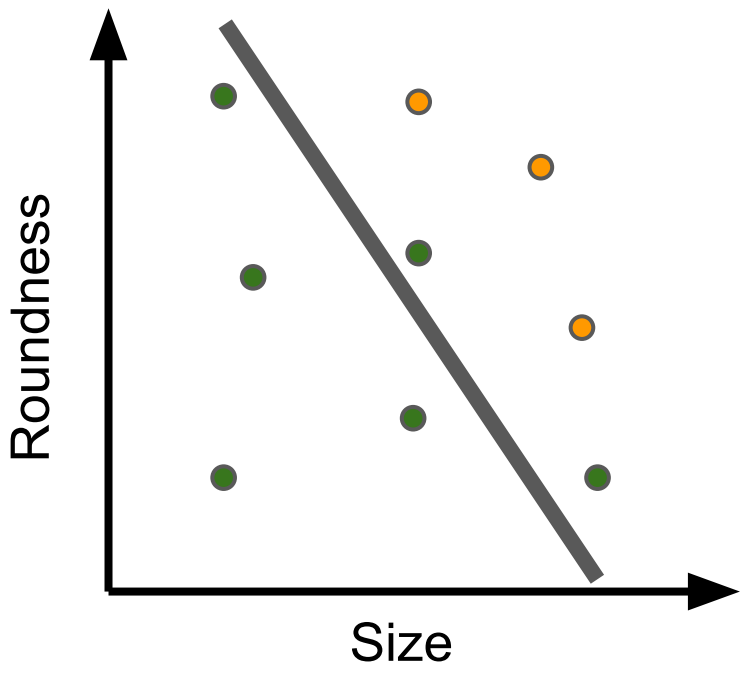
Creatinine = noise depends on body mass



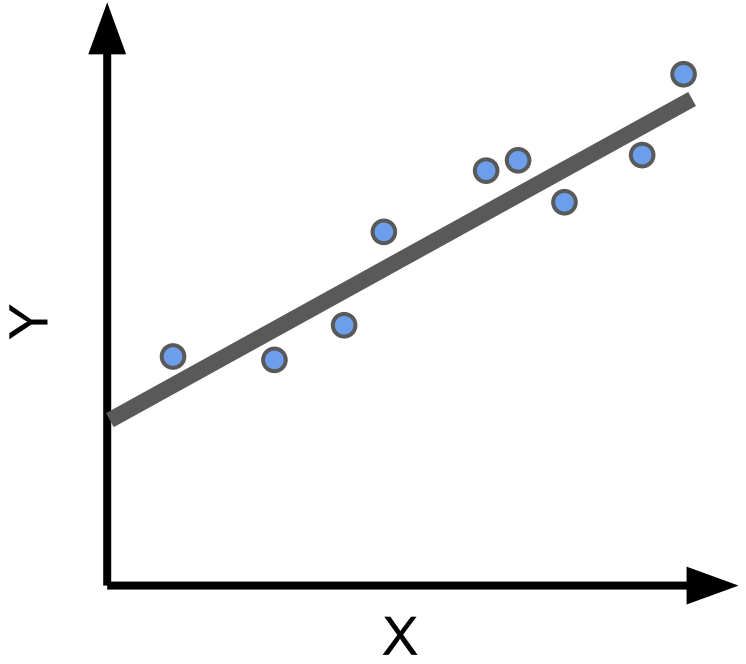
Error estimation

“No more counting dollars, we'll be counting stars”

Minimize classification errors

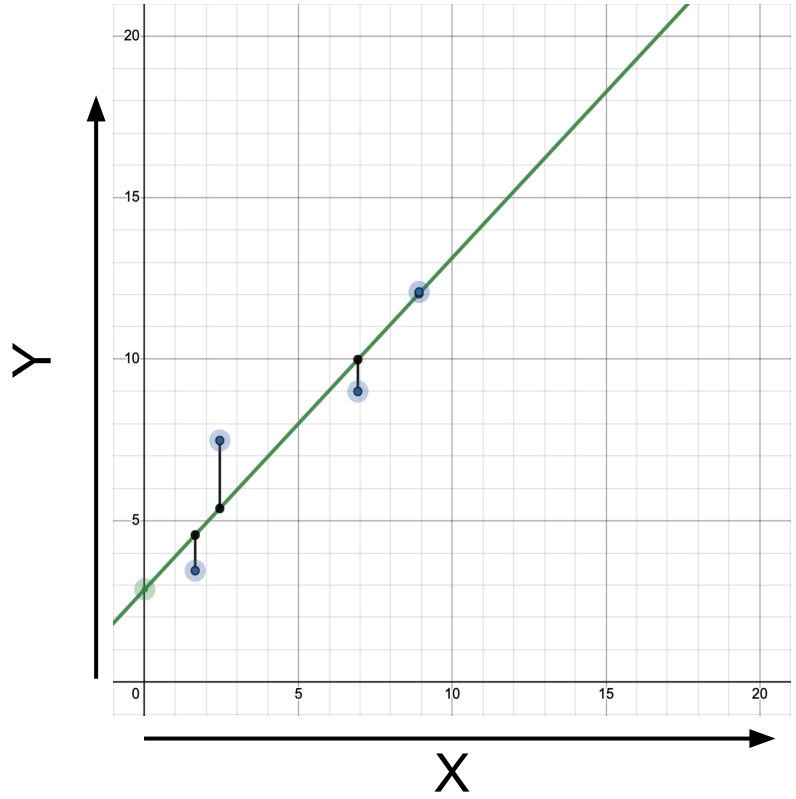


Accurately predict Y from X



Error estimation

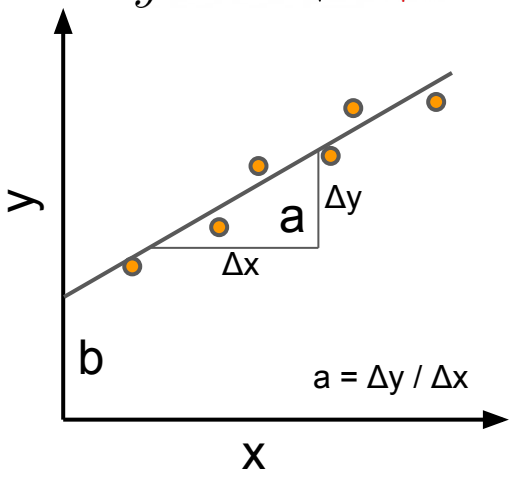
Absolute distances? Squared distances? …



Linear regression (univariate)

Least squares method

$$y = ax + b + \epsilon$$



$$\arg\min_{a,b} \sum \left(y_i - (ax_i + b)\right)^2 =$$

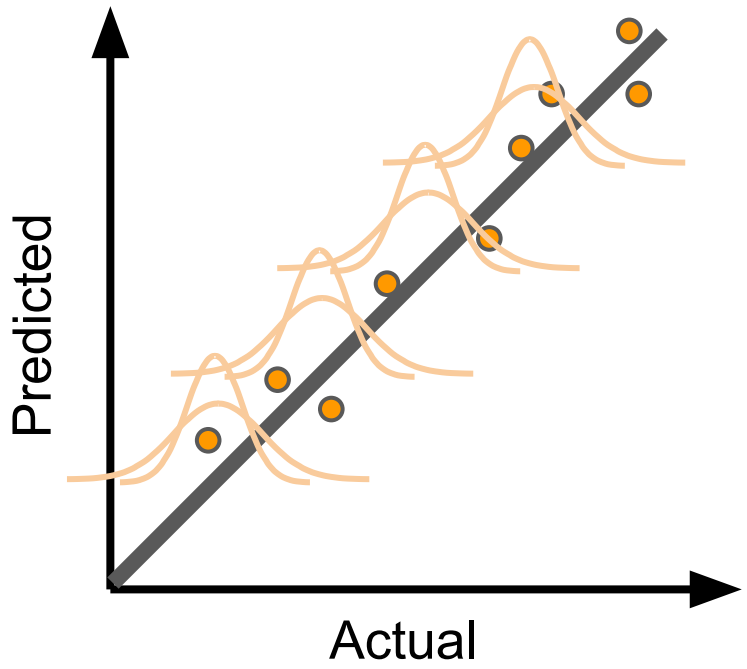
$$\arg\min_{a,b} \sum \left(y_i - \hat{y}_i\right)^2 =$$

$$\arg\min_{a,b} \sum \epsilon_i^2$$

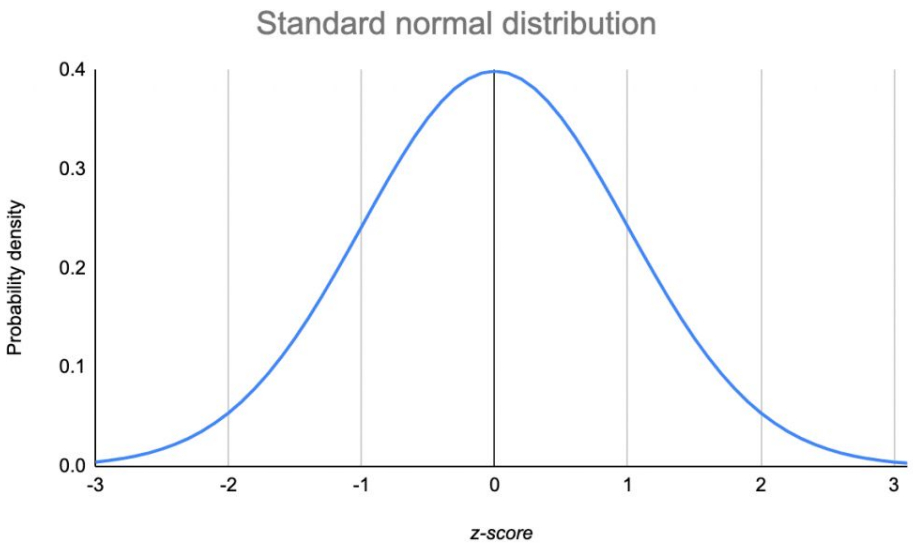
Why squared errors?

Assuming normally distributed noise

$$f(\epsilon) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}(\frac{\epsilon}{\sigma})^2}$$



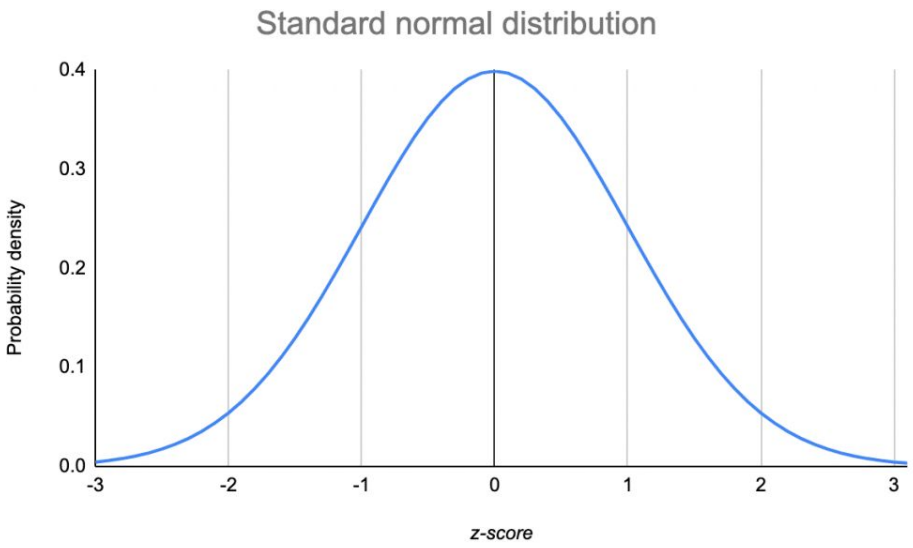
Normal Distribution



$$f(\epsilon) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}(\frac{\epsilon}{\sigma})^2}$$

Function of the square distance from the mean

Normal Distribution



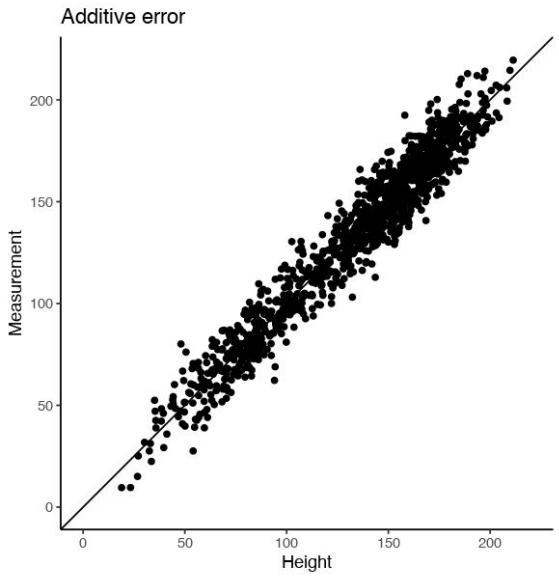
Properties:

- Symmetrical
- Central limit theorem:
  sum of small effects results in a normal distribution
- Formally, the simplest distribution with a given mean+variance
- Mathematical convenience
- Maintained under addition, shifting and rescaling

Additive vs Multiplicative Noise

Additive measurement noise:

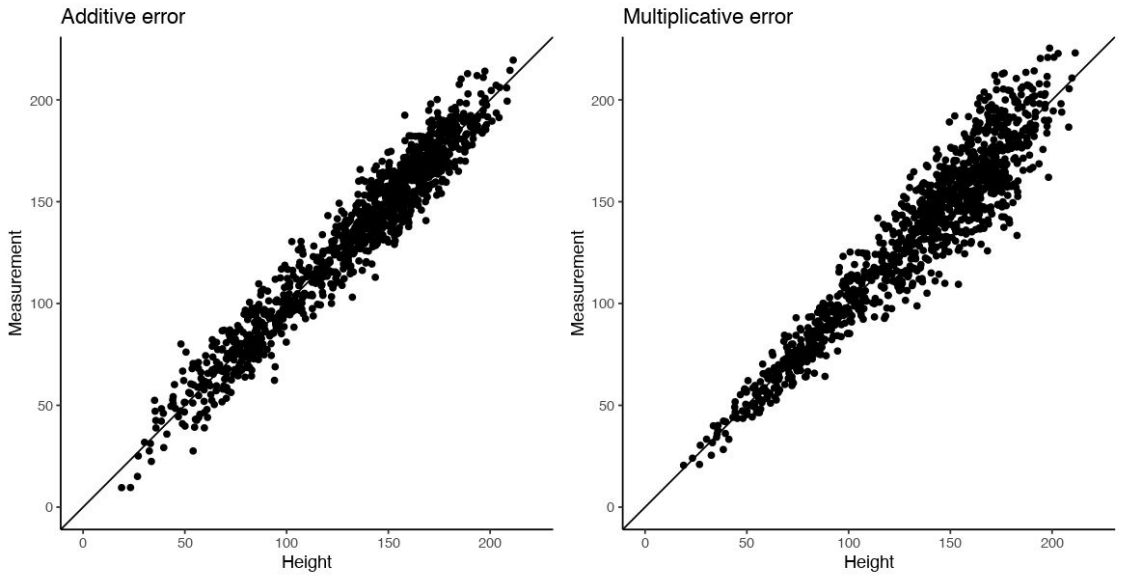
- “Height +/- 10cm”
- Error size does not depend on the actual value



Additive vs Multiplicative Noise

Multiplicative measurement noise:

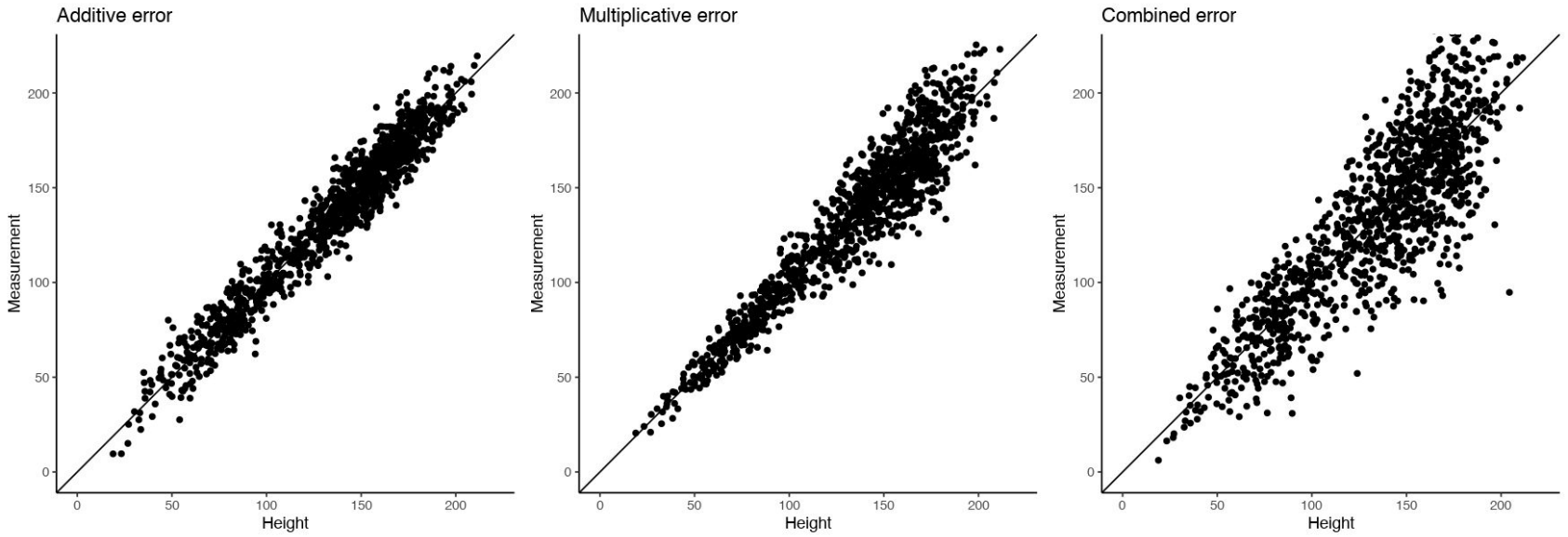
- “Height +/- 10%”
- Error size increase with actual value



Additive vs Multiplicative Noise

Combined noise

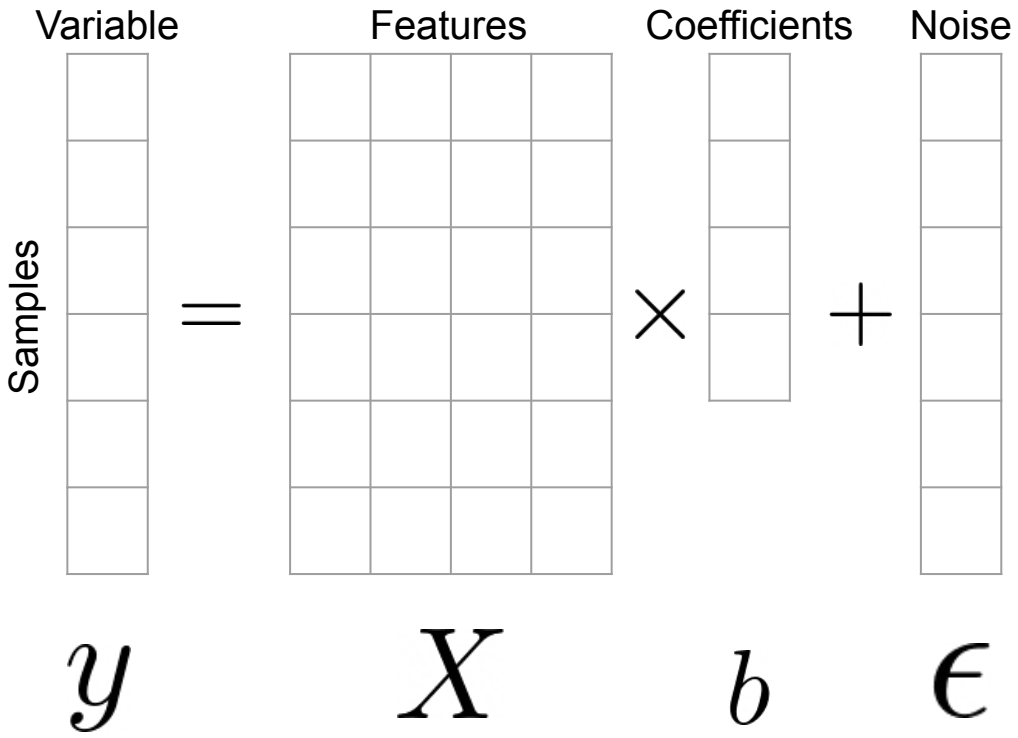
- Additive + multiplicative components



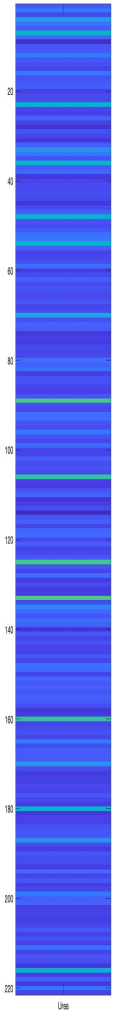
Linear regression (multivariate)

Definition, formalities, High-dimensional interpretation

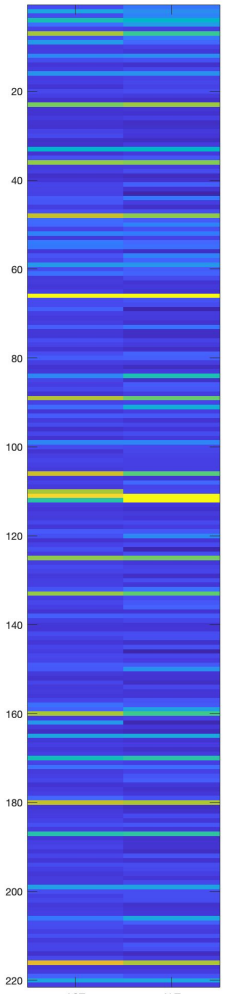
$$y = b_1 x_1 + \cdots + b_n x_n + b_0 + \epsilon$$



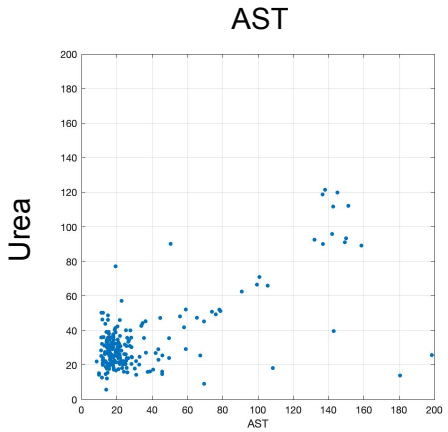
Linear regression (multivariate)

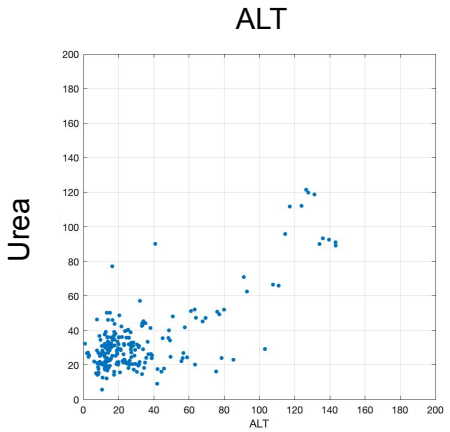


Urea



AST, ALT



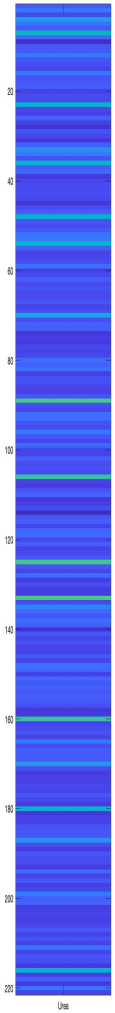


A Machine Learning Approach to Predict Creatine Kinase Test Results

Zehra Nur Canbolat [a], Gökhan Silahtaroğlu [a,b], Özge Doğuç [a*], Nevin Yılmaztürk [b]

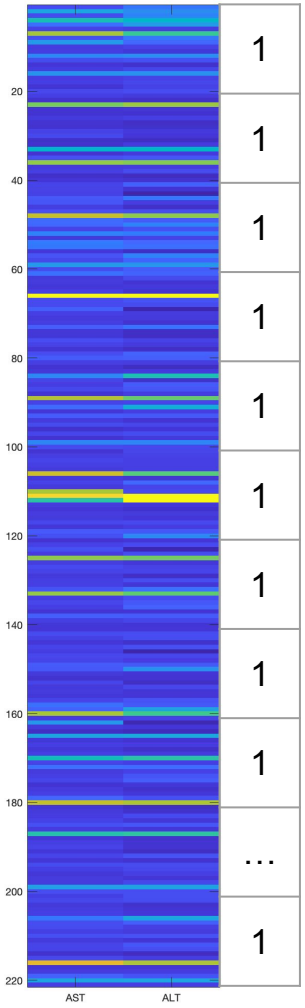
Age, Sex, AST, ALT, Urea, Glucose, etc. from 222 patients, hospitalized in Istanbul, Turkey, 2017-2019

Linear regression (multivariate)



Urea

=



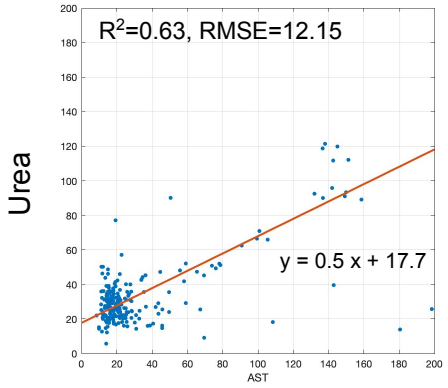
AST, ALT, 1

×

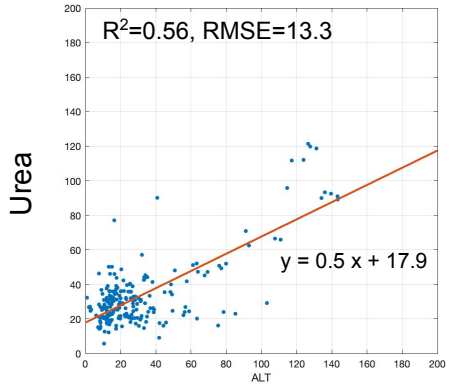
.40
.12
17

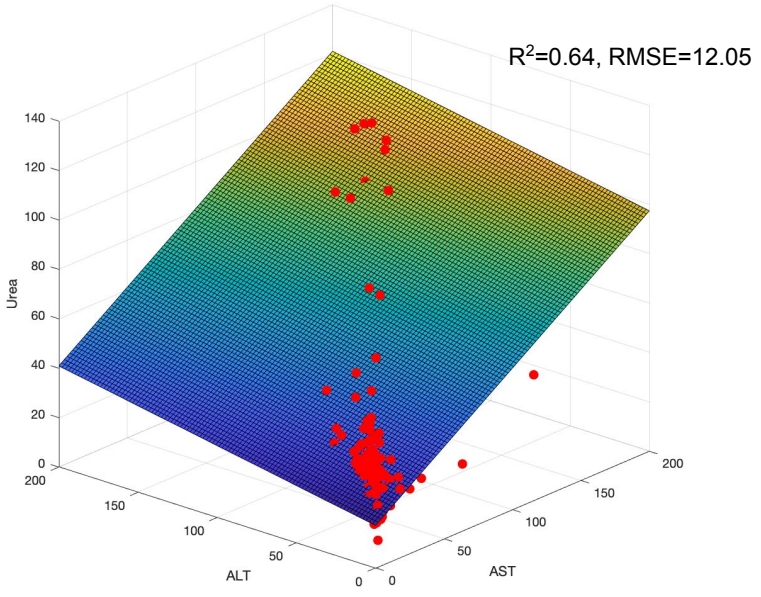
coeffs

AST



ALT





Linear regression (multivariate)

- Regularization: antagonizing “weak” features
- Feature selection: choosing the variables that matter
- Reducing complexity ⇒ more generalization

$$y = b_1 x_1 + b_2 x_2 + b_3 x_3 + \cdots + b_n x_n + b_0 + \epsilon$$

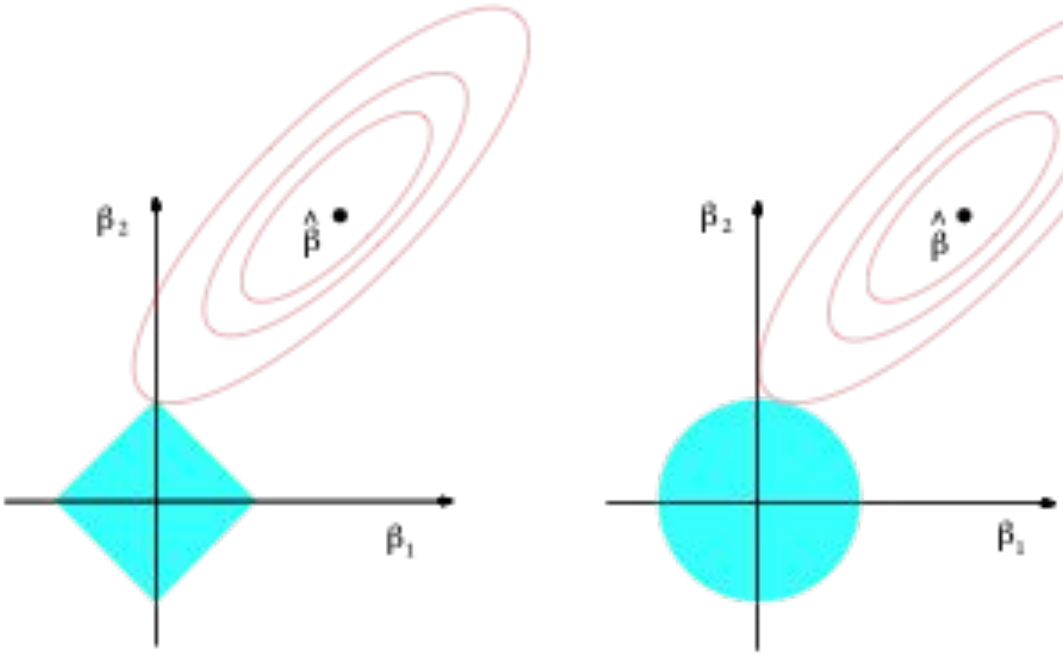
Ignoring x_2,x_3 is like setting b_2=0, b_3=0

Linear regression (multivariate)

$$\arg\min_{\beta}\left(\|y - X\beta\|^2 + \lambda_1 \sum |\beta_j| + \lambda_2 \sum |\beta_j|^2\right)$$

Error — Lasso — Ridge

Elastic net

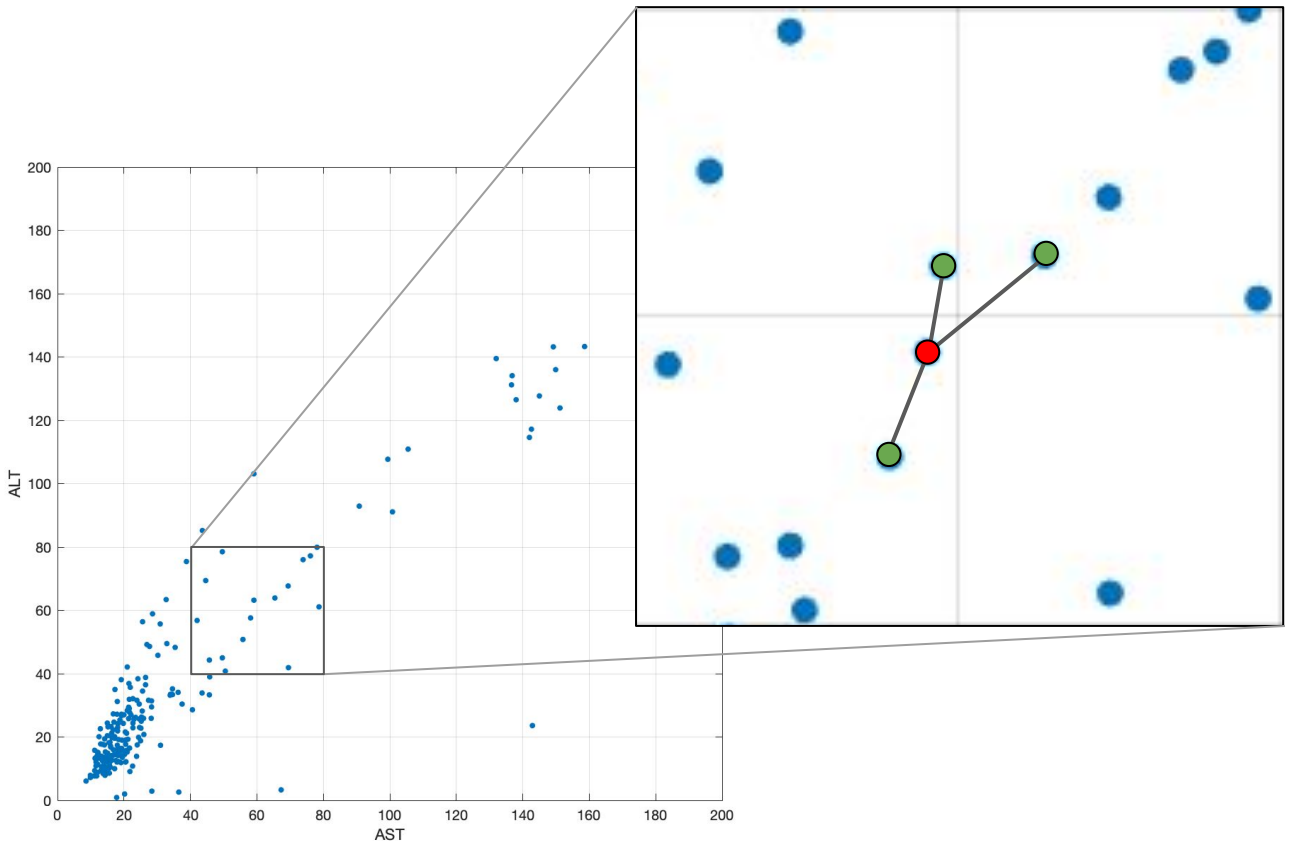


The elements of statistical learning

Cohort-based models

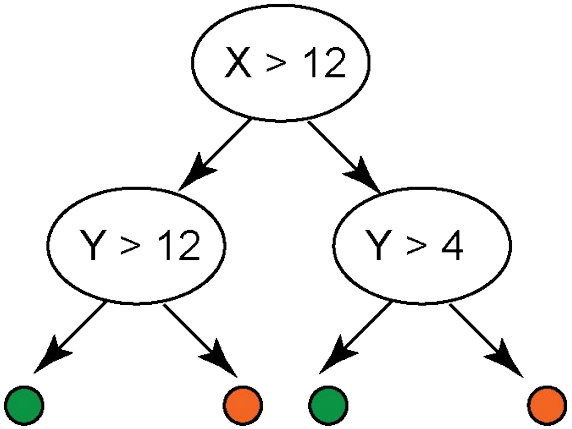
K-nearest neighbors regression models

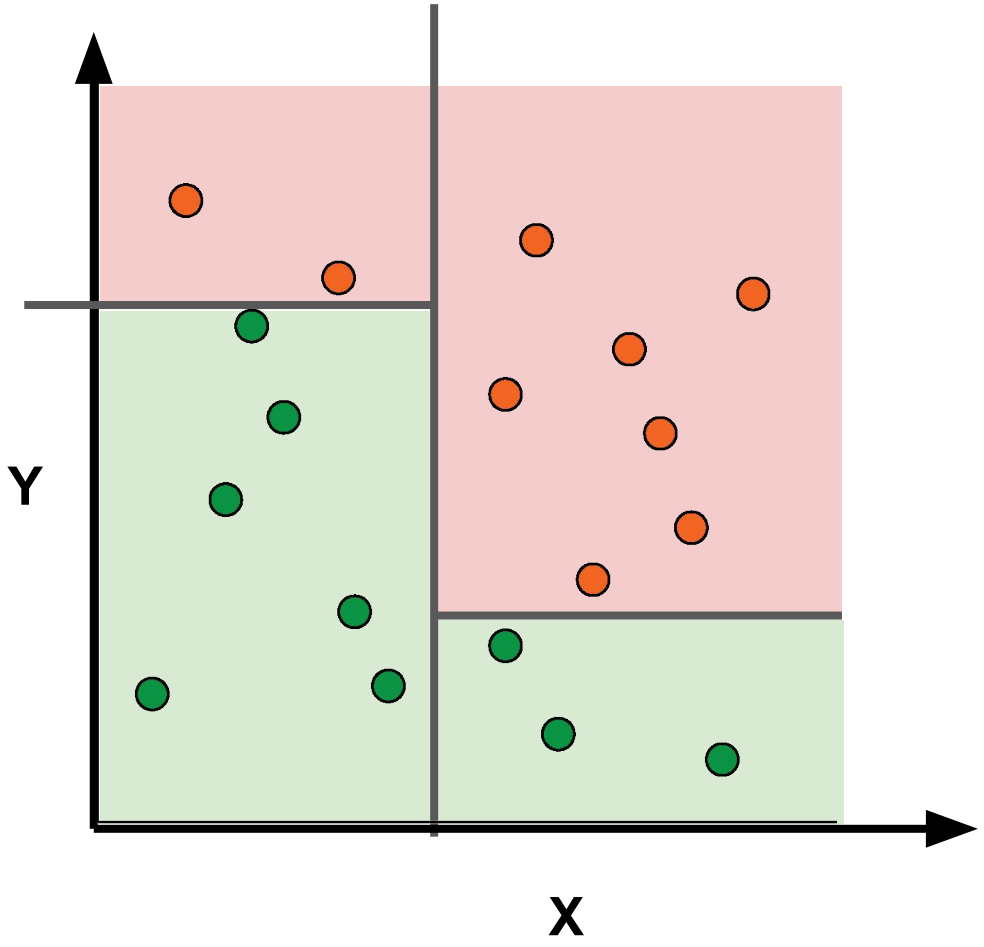
- Distances calculation
- Weight calculation



$$y_i = \frac{1}{K} \sum_{j \in N_i} y_j$$

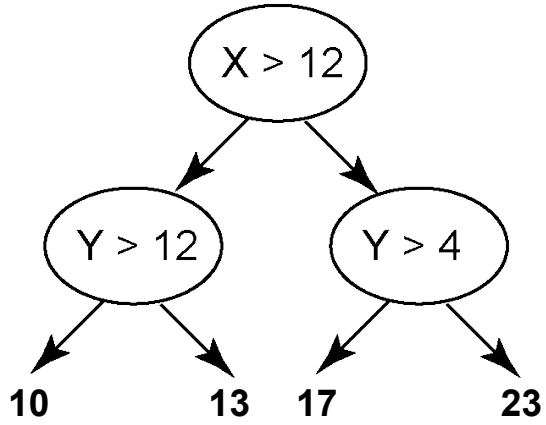
Regression trees

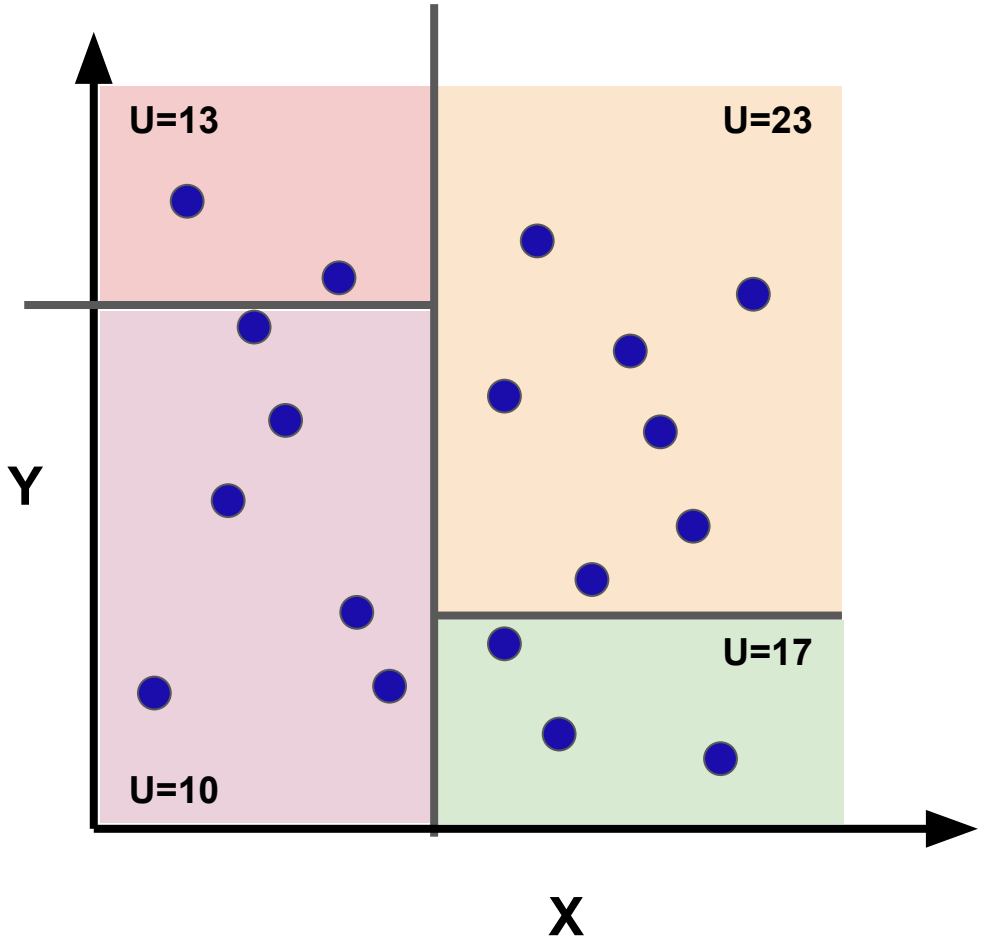




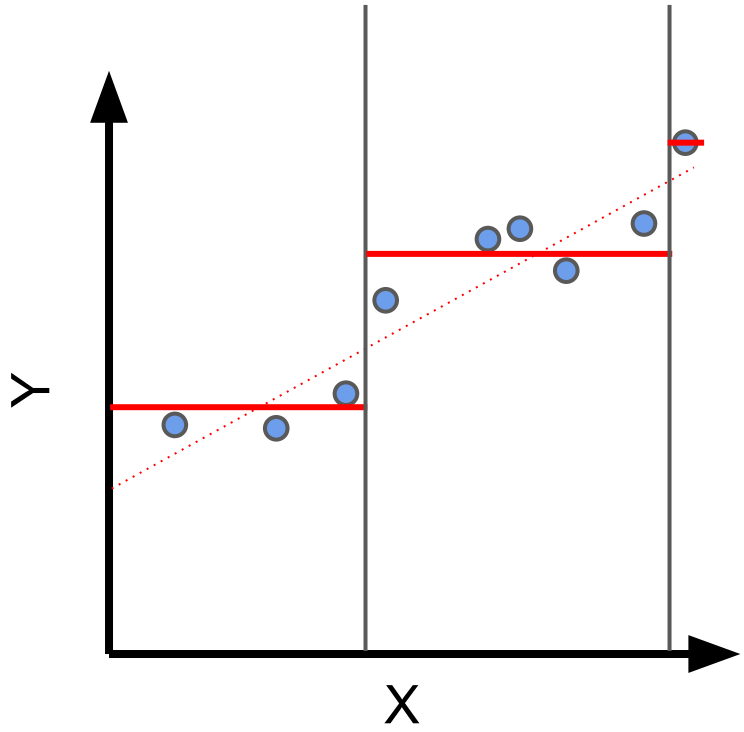
Regression trees

Predict a constant value at each leaf



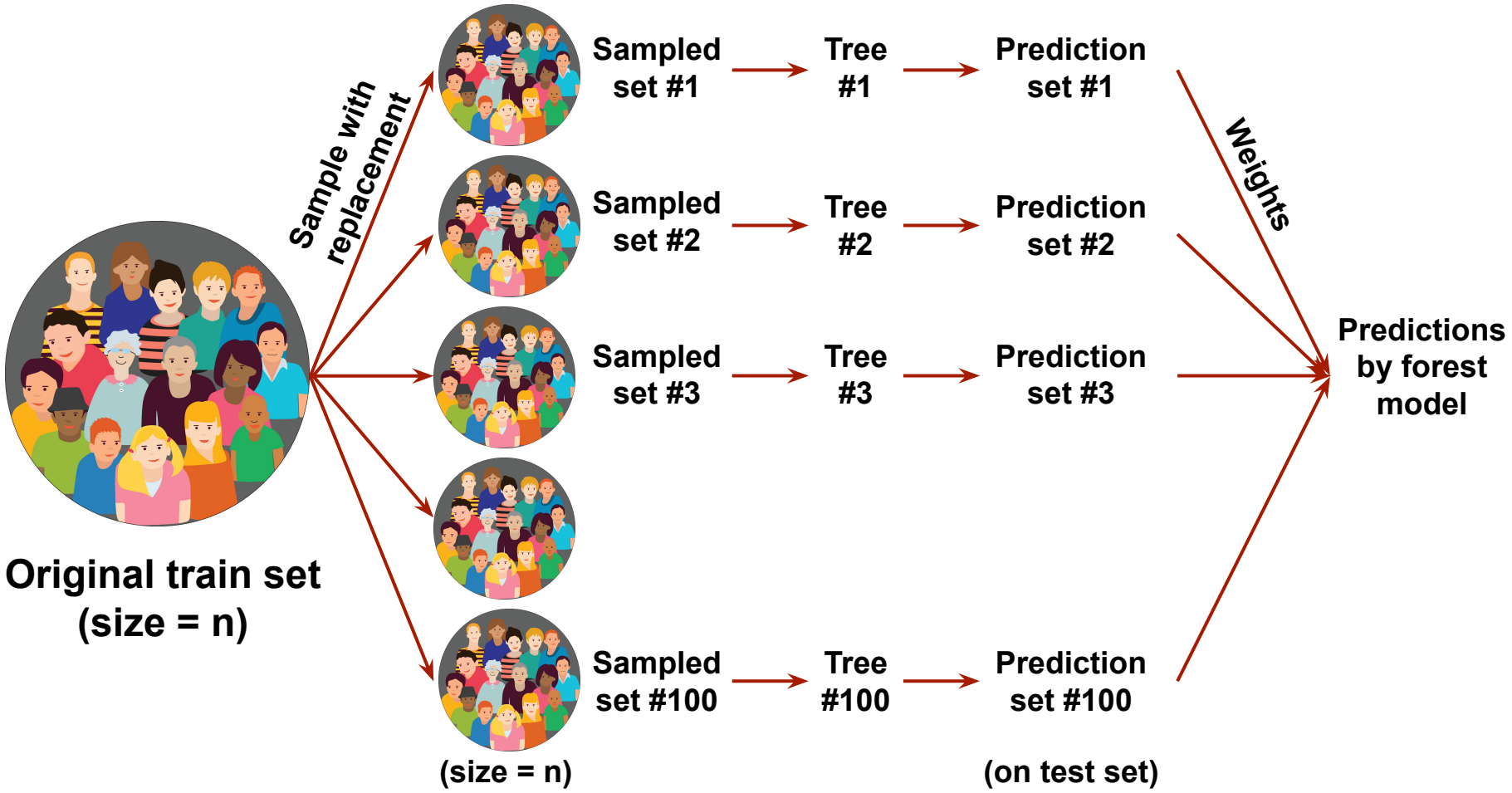


Regression tree (univariate)

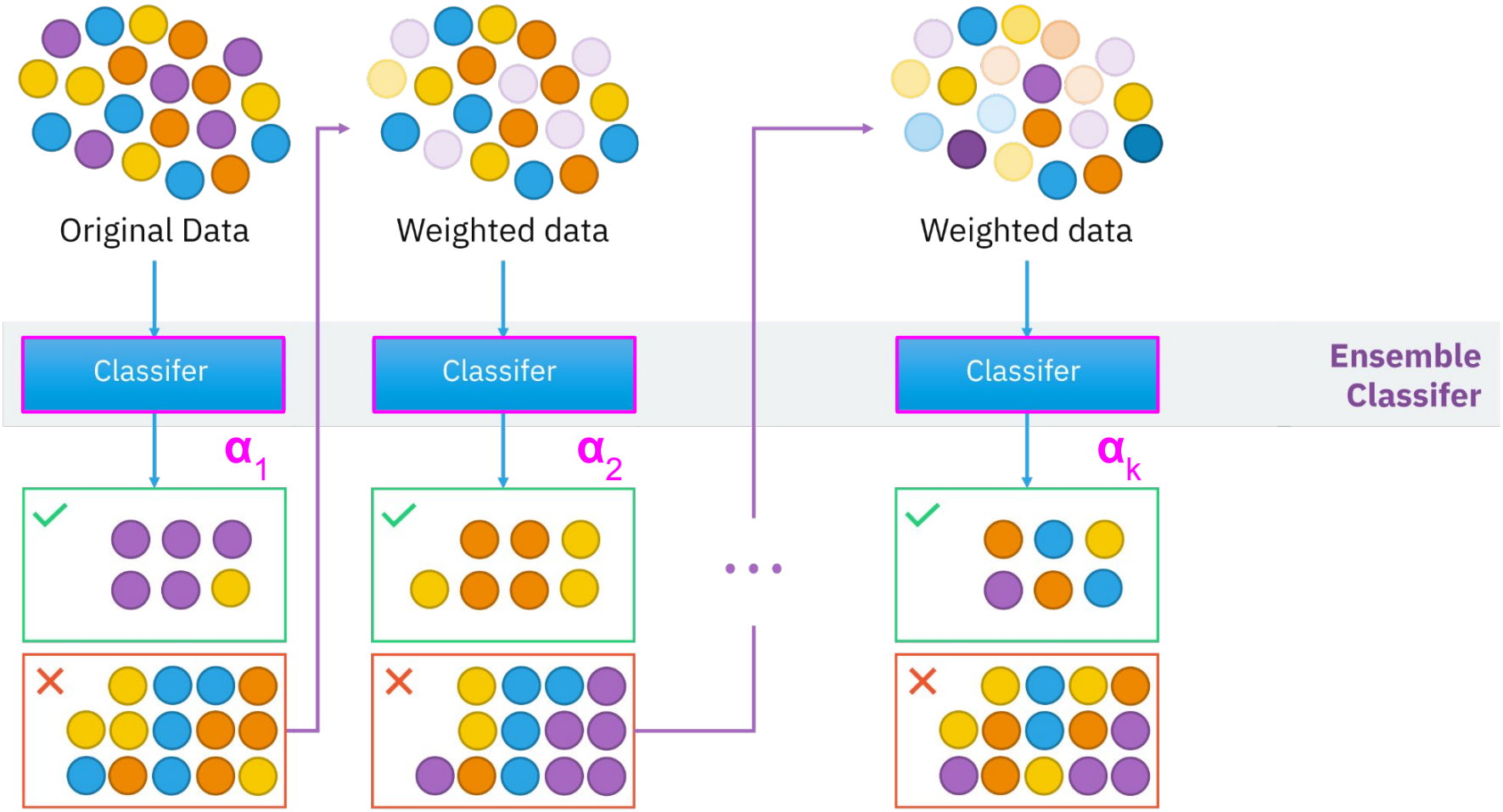


Ensemble models - forests

- Bootstrapping: sampling with replacement

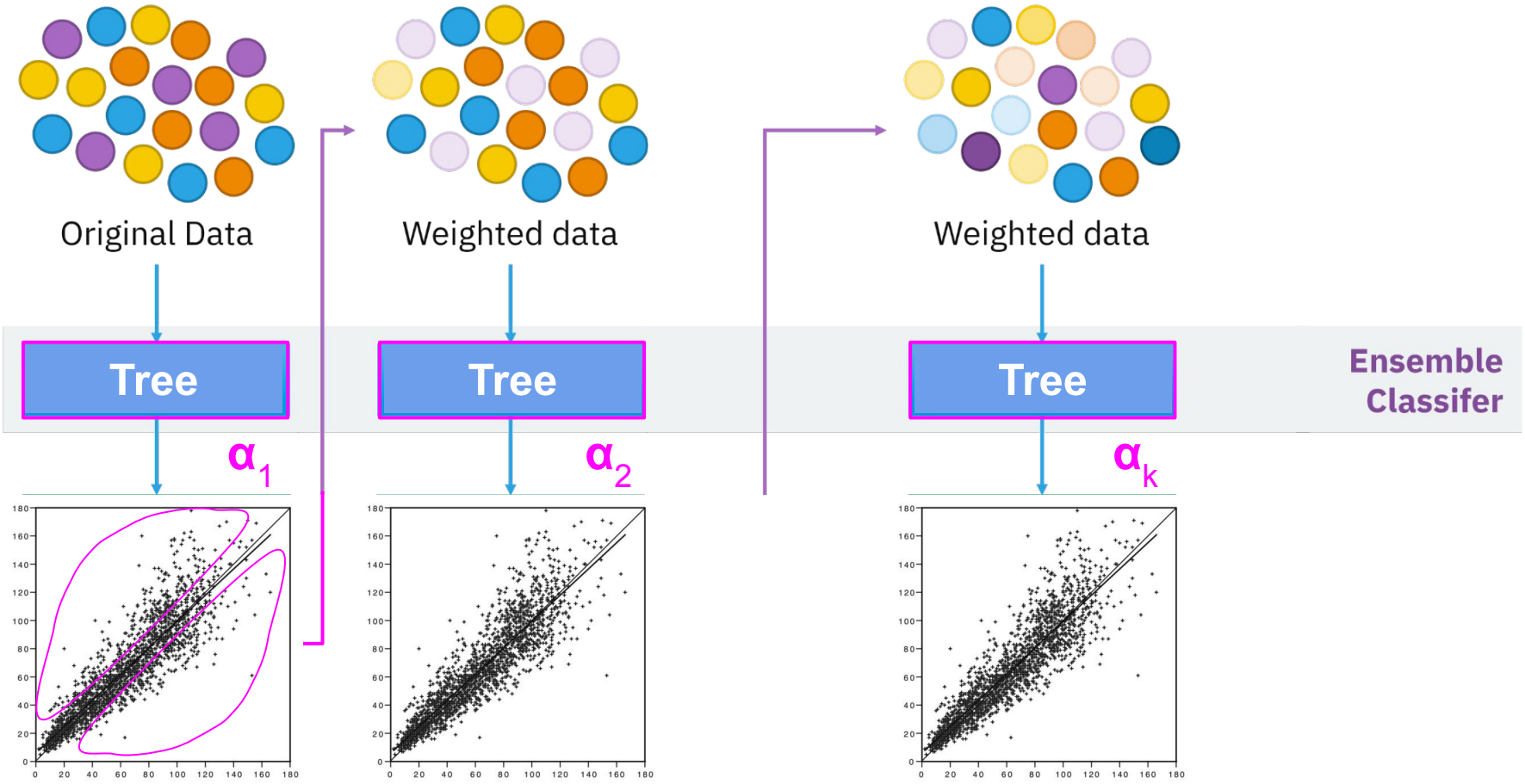


Boosting (ensemble classifier)



AdaBoost: Adaptive Boosting algorithm, Freund and Schapire, 1995

Gradient boosting trees (XGBoost)



Non-linear regression

Log-transformed?

Squared?

Sigmoid

BMI ~ Weight^2 / Height

log(BMI) ~ 2 * log(Weight) - log(Height)

Syllabus

1. Introduction
2. Classification
3. Learning 1
4. AI in ophthalmology (Prof. Itay Chowers)
5. Learning 2
6. Regression
7. Clustering
8. Visualization (and dimensionality reduction)
9. Deep learning in image analysis (Prof. Leo Joskowicz)
10. Missing data, statistical dependencies
11. Natural language in medicine (Dr. Gabi Stanovsky)
12. Decisions (utility)
13. Longitudinal Data / Project