Models

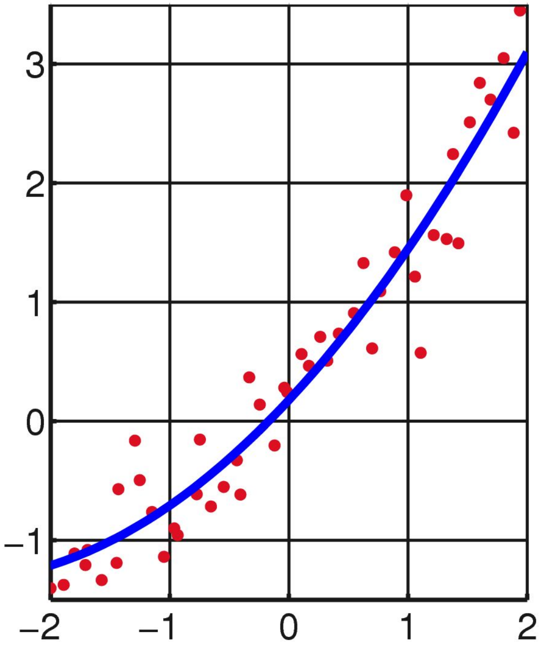
We are going to use

Means (baseline)

- Experiments with means over:
  - One area/neighbouring areas
  - Various previous time duration (last week/last month/last year)
  - Predict for various time segments (5 min/half hour/hour/3 hours)
  - Mean over demand during last X weekdays (previous Mondays/Tuesdays…)
  - Over the same time in the previous years? (January of last years/18th week of last years)
  - Anything else what comes to your mind

Linear (least-squares) regression

$$f(\vec{x};\vec{\beta}) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \ldots$$



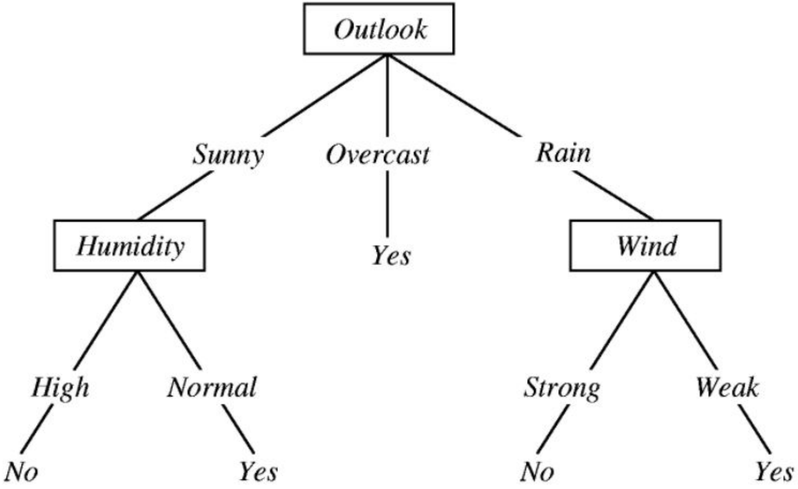
x = vector of features

β = vector of weights

- Goal is to find β
- You can apply other functions to features in order to derive ‘new features’
- i.e. x^2, log(x), x1*x2, sin(x)...
- β makes the model ‘linear’ despite using other non-linear functions

Decision trees

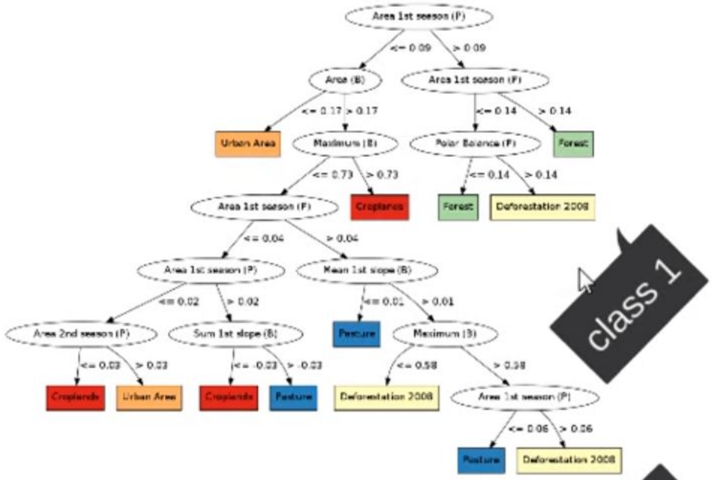
Is John going to play tennis?

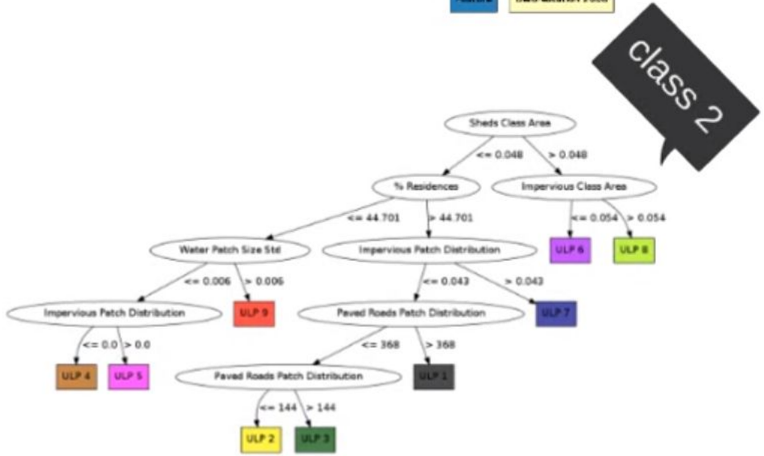


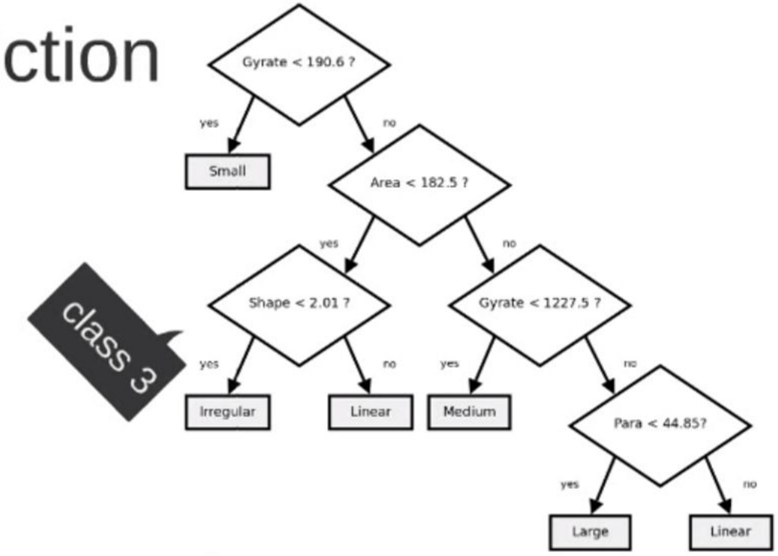
Random forest

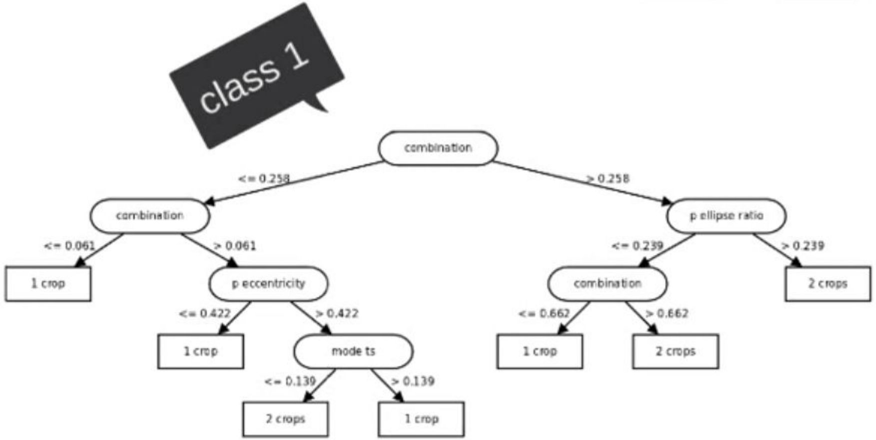
- Training
  a. Divide training data into **random** subsets
  b. Train Decision **tree** on each of the **subsets**
- Prediction
  a. Element is decided by all trees
  b. The result is what a majority of trees predicts

Class prediction





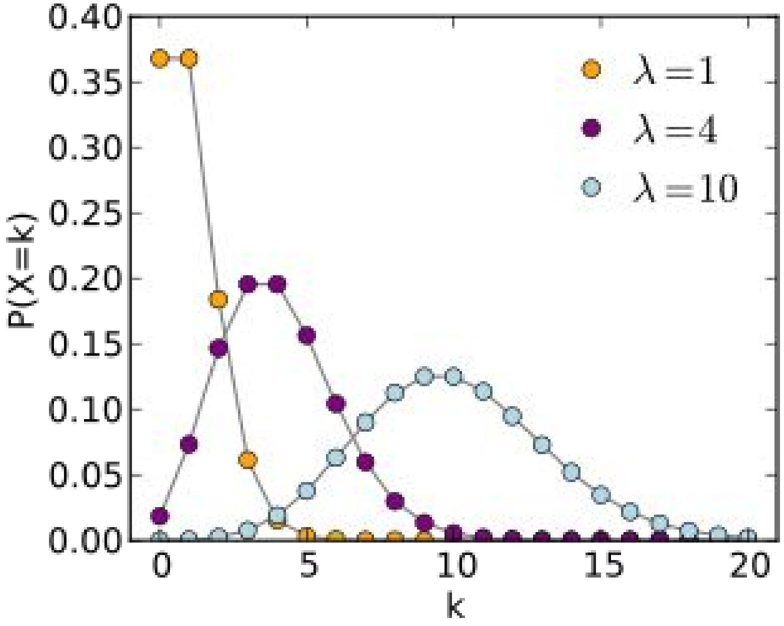




Poisson distribution

- Modelling the number of times an event occurs in a time interval
- E.g. calls in telephone center

$$P(n;\lambda) = \frac{e^{-\lambda}\lambda^n}{n!}$$



Time - Varying Poisson Model

- Different day -> different demand
- Different time of day -> -||-
- Relative change for weekday
- Relative change for time period
- Result: expected demand during a week

$$\lambda(t) = \lambda_0 \delta_{d(t)} \eta_{d(t),h(t)}$$

$$\sum_{i=1}^{7} \delta_i = 7$$

$$\sum_{i=1}^{I} \eta_{d,i} = I \quad \forall d$$

Weighted Time - Varying Poisson Model

- In previous model seasonality is not taken into consideration
- Recent data is more relevant than the older data
  - Exponential smoothing

$$\omega = \alpha * \left\{1, (1-\alpha), (1-\alpha)^2, \dots, (1-\alpha)^{\gamma-1}\right\}, \gamma \in \mathbb{N}$$

$$\mu(t)_k = \sum_{i=1}^{\gamma} \frac{X_{t-(\theta * i)} * \omega_i}{\Omega}, \Omega = \sum_{i=1}^{\gamma} \omega_i$$

ARIMA

- Widely used for modeling and forecasting univariate time-series data
- Prediction value is linear function of several past observations and random errors

$$R_{k,t} = \kappa_0 + \phi_1 X_{k,t-1} + \phi_2 X_{k,t-2} + \cdots + \phi_p X_{k,t-p}$$
$$+ \varepsilon_{k,t} - \kappa_1 \varepsilon_{k,t-1} - \kappa_2 \varepsilon_{k,t-2} - \cdots - \kappa_q \varepsilon_{k,t-q}$$

- Weights, parameters, white noise error terms
- One of the simplest models (for nonseasonal stationary time series)

Bayes networks + neural networks

- Bayes networks at our MI classes
- Neural networks at Coursera course