# Data Mining and Machine Learning



# HMMs for Automatic Speech Recognition: Types of HMMs





Peter Jančovič

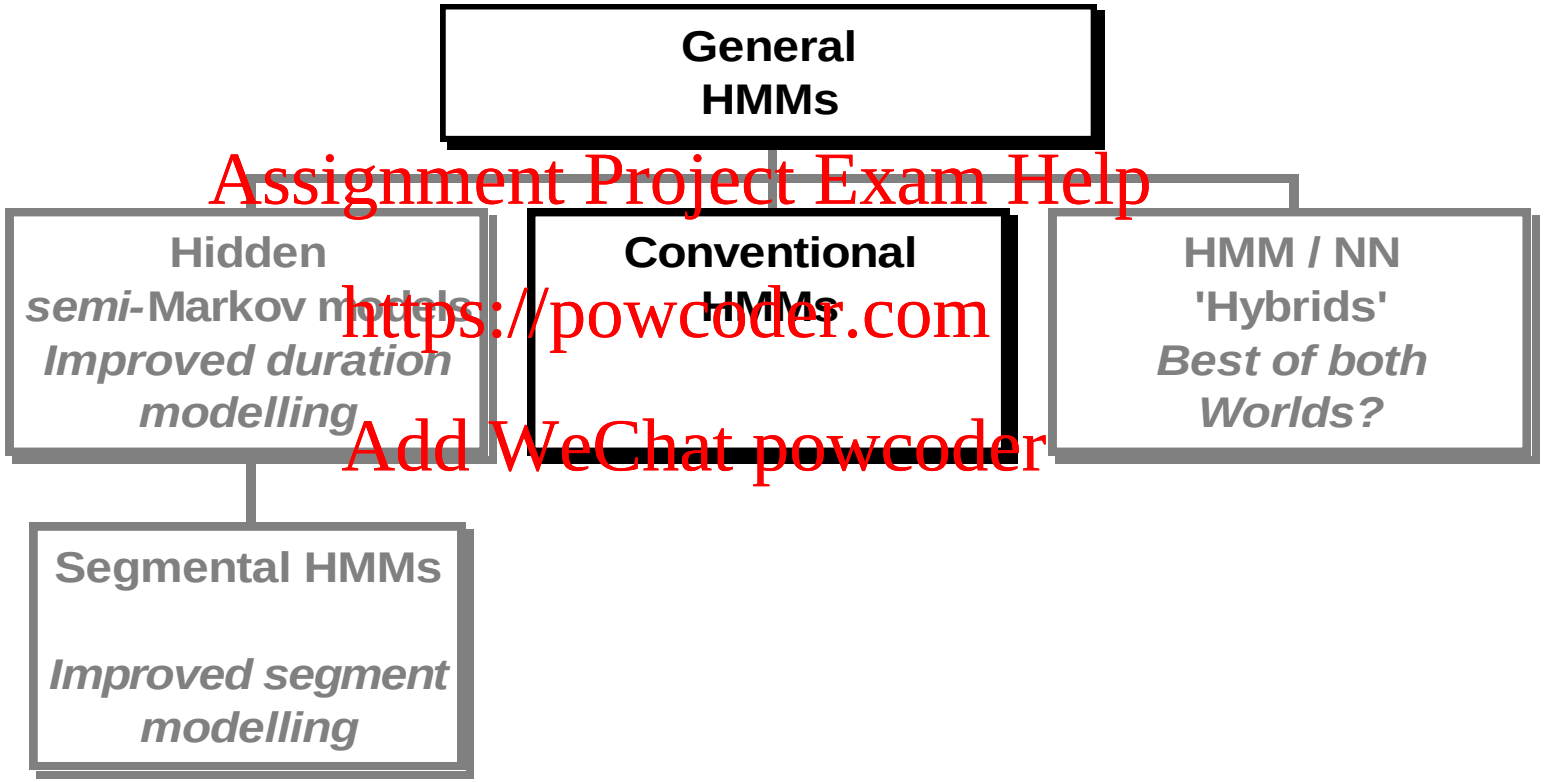
# Objectives

To understand

- Differences between types of HMMs

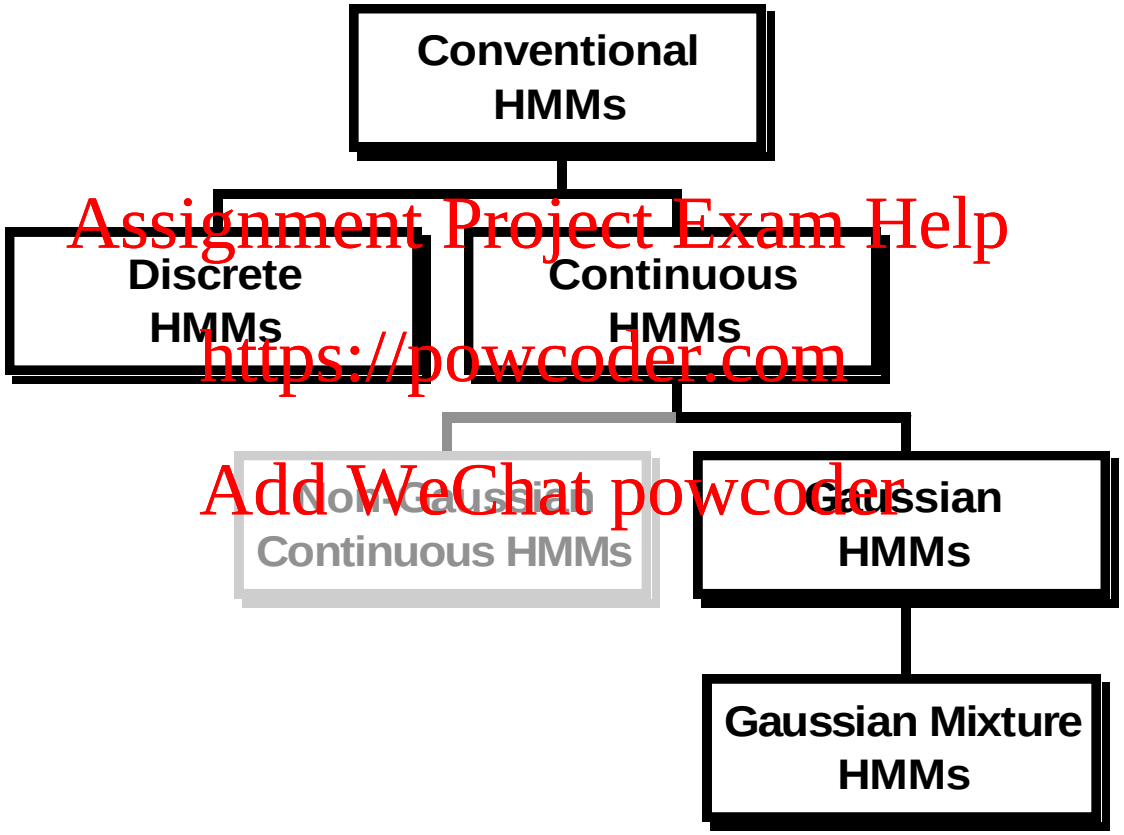


# HMM taxonomy






# Types of Conventional HMM



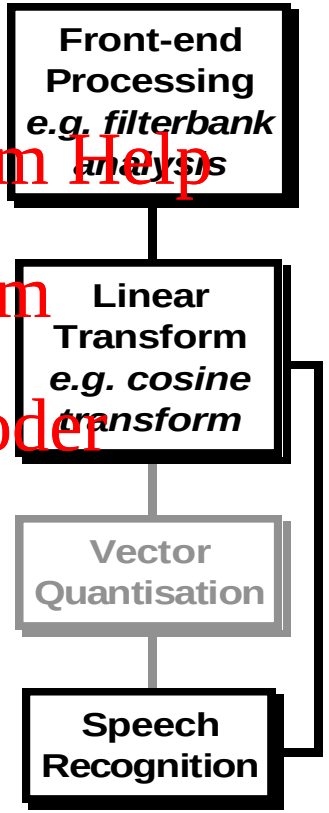



# Front-End Processing Re-Visited

Vectors in d-dimensional (continuous) space

Vectors in d-dimensional (continuous) space

Symbols from a **finite** set






# Discrete HMMs

- If VQ is used, then a state output PDF $b_i$ is defined by a **list** of probabilities

  $b_i(m) = \text{Prob}(y_t = z_m \mid x_t = s_i)$

- The resulting HMM is a **discrete HMM**
- Common in mid-1980/ early-1990s
- Computational advantages
- Disadvantages
  - VQ may introduce non-recoverable errors
  - Choice of metric $d$ for VQ?
- Outperformed by Continuous HMM



# Continuous HMMs

- Without VQ, $b_i(y)$ must be defined for any $y$ in the (continuous) observation set $S$
- Hence discrete state output PDFs no longer viable
- Use parametric continuous state output PDFs - **Continuous HMMs**
- Choice of PDF restricted by mathematical tractability and computational usefulness (see "HMM training & recognition" later)
- Most people begin with Gaussian PDFs
- Resulting HMMs called **Gaussian HMMs**



# Gaussian HMMs

- State output PDFs are multivariate Gaussian



$$b_i(y) = \frac{1}{\sqrt{(2\pi)^d |C_i|}} \exp\left\{-\frac{1}{2}(y - m_i)' C_i^{-1}(y - m_i)\right\}$$

- $m_i$ and $C_i$ are the mean vector and covariance matrix which define $b_i$

# Gaussian HMMs - Issues

- Significant computational savings if covariance matrix can be assumed to be diagonal
- In general, Gaussian PDFs are **not** flexible enough to model speech pattern variability accurately
  - In many applications (e.g. modelling speech from multiple speakers) a unimodal PDF is inadequate
  - Even if unimodal PDF is basically OK there may be more subtle inadequacies



# Gaussian Mixture HMMs

- Any PDF can be approximated arbitrarily closely by a Gaussian mixture PDF with sufficient components
- But...
  - More mixture components require more data for robust model parameter estimation
  - Parameter smoothing and sharing needed (e.g. ‘tied mixtures’, ‘grand variance’,...)
- Gaussian mixture HMMs widely used in systems in research laboratories



# Relationship with Neural Networks

- 'Classical' HMM training methods focus on fitting state output PDFs to data (modelling), rather than minimizing overlap between PDFs (discrimination)
- NNs are good at discrimination
- **But** 'classical' NNs poor at coping with time-varying data
- Research interest in 'hybrid' systems which use NNs to relate the observations to the states of the underlying Markov model
- More recently, recurrent NNs also replacing HMMs



# Summary

- Types of HMM
  - Discrete HMMs
  - Continuous HMMs – Gaussian Mixture HMMs