| | raw data | | representation | | | |
|----------------|--|-----------------|--|--|--|---|
| | in-time-based | transform-based | shape-based | model-based | feature-based | iterative-based (embedded) |
| raw data | x | X | x | X | x | X |
| transformation | | X | | | | |
| representation | | | X | X | X | |
| clustering | X | X | X | X | X | |
| example | ~ 1 | | | | | |
| | | | = + + + + + + + + + + + + + + + + + + + | \downarrow $Y_{t} = \alpha + \beta_{1} Y_{t-1} + \beta_{2} Y_{t-2} + \dots + \epsilon_{t}$ | $ \downarrow \\ \bar{x} \\ \sigma^2 \\ \hat{\rho}_k \\ \sum x_{t+1} - x_t \\ \beta_t \\ \vdots $ | Start Initial Values Expectation Maximization No Converged Yes End |
| algorithm | LKMA | DTW | DFT | VAR | tsfresh | GBTM |
| notes | low interpretability same interval (if not transformed) same length (no missing) no parameter dependence (ignore temporal order) sensitive to offset (if not transformed) sensitive to noise no shape assumed (incl. sudden changes) fast modeling readily available software cluster interpretation established in the field | | dangerous if model fit is poor assumptions violated (poor fit) too few observations (poor fit) high interpretability reduced dimensional space more acurate then raw fast modeling readily available software robust to missing data (b/c calculated on multiple observations) varying intervals varying lengths often scaleable performance (e.g., model only needs to be fitted once) relatively few observations per trajectory | | | often assume the same parametric distribution dangerous if model fit is poor (e.g., over- or under fitting) assumptions violated (fit) too few observations (fit) slower with mode complex models robust to missing data (b/c calculated on multiple observations) varying intervals varying lengths interpretatble (e.g., distinct cluster trajectories) good for prediction |