Multivariate Screening and Feature Engineering for Statistical Decision Making in Disease Severity Diagnostics

FEATURE ENGINEERING

Feature Space Formulation

Use statistical tests to identify sample differences

⟹

Feature Ranking

Ranking the tests according to significance level

⟹

Feature Simulation

Bootstraping attributes presenting significant statistical differences

Screened Features
Statistically significant disease indicators

Unscreened Features
Complete set of measurements

CLUSTERING

Hierarchical

Tree-based merging approach

&

K-Means

Centroid-based iterative partition

⟹

Applied to Screened and Unscreened Features

EVALUATION

Cluster Quality Evaluation

&

Screened vs. Unscreened Features