

**MBA
USP
ESALQ**

UNSUPERVISED MACHINE LEARNING: Exercícios Aplicados

Prof. Dr. Wilson Tarantin Junior

*A responsabilidade pela idoneidade, originalidade e licitude dos conteúdos didáticos apresentados é do professor.

Proibida a reprodução, total ou parcial, sem autorização. Lei nº 9610/98

Contextualização

- Técnicas não supervisionadas
 - Para avaliar a relação conjunta entre as variáveis (interdependência), também são chamadas de técnicas exploratórias
 - Não há a proposição de modelos do tipo “ $y_i = x_{1i} + x_{2i} + \dots + u_i$ ”
 - Não são adequadas para fins de inferência, ou seja, não têm caráter preditivo para observações de fora da amostra
 - Se novas observações forem adicionadas ou retiradas do banco de dados, é adequado refazer a análise; o mesmo é válido se variáveis forem adicionadas ou retiradas

Análise de Cluster

- **Objetivo: agrupamento de observações**

- O objetivo é agrupar as observações em grupos homogêneos internamente e heterogêneos entre si
 - **Dentro do grupo:** observações semelhantes com base nas variáveis utilizadas na análise
 - **Entre grupos distintos:** observações diferentes com base nas variáveis utilizadas na análise

Análise Fatorial – PCA

- **Objetivo: agrupamento das variáveis métricas**

- O objetivo é o agrupamento das variáveis em fatores
- A técnica depende das correlações entre variáveis
- Algumas possíveis finalidades:
 - Obter o comportamento conjunto de variáveis, combinando-as para redução estrutural
 - Análise da validade de construtos pela identificação das variáveis alocadas aos fatores
 - Elaboração de rankings para classificação de desempenho por meio dos fatores
 - Criação de fatores ortogonais entre eles e posterior uso em modelos supervisionados

Análise de Correspondência

- **Objetivo: análise de associação entre variáveis categóricas**
 - O objetivo é verificar se existe associação estatisticamente significativa entre as variáveis qualitativas (duas ou mais de duas variáveis) e suas categorias
 - Como estão associadas as categorias das variáveis?
 - Busca-se criar o mapa perceptual para visualizar tais associações
 - Análise pela proximidade das categorias no mapa perceptual

Referência

Fávero, Luiz Paulo; Belfiore, Patrícia. (2024). Manual de análise de dados: estatística e machine learning com Excel®, SPSS®, Stata®, R® e Python®. 2 ed. Rio de Janeiro: LTC.

OBRIGADO!

[linkedin.com/in/wilson-tarantin-junior-359476190](https://www.linkedin.com/in/wilson-tarantin-junior-359476190)