

Analise de Rastreamento de Fitness - Brief do Projeto

Dataset

Dataset de rastreamento de progresso na academia contendo 200 dias de medicoes diarias de fitness e nutricao. O dataset inclui rastreamento de peso (kg), ingestao calorica, consumo de proteina (g), duracao de treino (minutos) e passos caminhados. Este rastreamento abrangente permite analise de relacoes entre exercicio, nutricao e mudancas na composicao corporal ao longo do tempo.

<https://www.kaggle.com/datasets/rishabhagarwal997889/gym-progress-tracking-dataset-200-days>

Objetivos do Projeto

O objetivo principal e entender os fatores que influenciam mudancas de peso atraves de analise estatistica e aprendizado de maquina. Isso inclui identificar correlacoes entre variaveis, construir modelos preditivos e descobrir padroes comportamentais que levam a resultados especificos.

Pipeline de Analise

Passo 1: Analise Exploratoria de Dados (EDA)

- Estatisticas descritivas (media, mediana, desvio padrao)
- Visualizacoes de distribuicao de cada variavel
- Identificar valores ausentes ou outliers
- Graficos de evolucao temporal (peso, calorias, proteina ao longo do tempo)
- Analise de tendencias e padroes semanais/mensais

Passo 2: Analise de Correlacoes

- Matriz de correlacao entre todas as variaveis
- Heatmap para visualizar relacoes
- Identificar quais fatores mais influenciam a perda/ganho de peso
- Analisar relacao entre consumo calorico, proteina, exercicio e peso

- Scatter plots para relacoes importantes

Passo 3: Modelo de Predicao de Peso

- Regressao Linear para prever peso futuro
- Random Forest ou XGBoost para melhor performance
- Validacao cruzada e metricas (RMSE, MAE, R²)
- Analise de importancia das features
- Testar diferentes horizontes de predicao (1 dia, 7 dias, 30 dias)

Passo 4: Clustering e Segmentacao

- K-Means para identificar padroes de comportamento
- Agrupar dias similares (ex: dias de alta atividade, dias de descanso)
- Analise de perfis de treino/nutricao
- Visualizacao com PCA ou t-SNE
- Insights sobre que combinacoes geram melhores resultados