

Ciência de Dados (Python e Estatística) - Modelagem Estatística

2. Modelagem Estatística

2.1 Aplicação de Técnicas Estatísticas Avançadas Utilização de técnicas estatísticas avançadas para modelagem dos dados.

Exemplo:

```
from sklearn.linear_model import LinearRegression

# Separar variáveis independentes e dependentes
X = data[['idade', 'IMC']]
y = data['glicemia']

# Treinar modelo de regressão linear
model = LinearRegression()
model.fit(X, y)

# Coeficientes do modelo
print(f"Coeficientes: {model.coef_}")
print(f"Intercepto: {model.intercept_}")
```

2.2 Avaliação da Adequação dos Modelos Estatísticos Avaliação da adequação dos modelos estatísticos aos dados.

Exemplo:

```
from sklearn.metrics import mean_squared_error

# Fazer previsões
y_pred = model.predict(X)

# Avaliar a performance do modelo
mse = mean_squared_error(y, y_pred)
print(f"Erro Quadrático Médio (MSE): {mse}")
```

2.3 Implementação de Modelos Preditivos Implementação de modelos preditivos utilizando Python.

Exemplo:

```
from sklearn.tree import DecisionTreeClassifier
from sklearn.model_selection import train_test_split
from sklearn.metrics import accuracy_score

# Separar variáveis independentes e dependentes
X = data.drop('complicacao_diabetica', axis=1)
y = data['complicacao_diabetica']

# Dividir os dados em conjunto de treinamento e teste
X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.2,
random_state=42)
```

```
# Treinar modelo de árvore de decisão
model = DecisionTreeClassifier()
model.fit(X_train, y_train)

# Fazer previsões
y_pred = model.predict(X_test)

# Avaliar a acurácia do modelo
accuracy = accuracy_score(y_test, y_pred)
print(f"Acurácia: {accuracy}")
```

2.4 Avaliação da Performance dos Modelos Preditivos

Avaliação da performance dos modelos preditivos utilizando métricas apropriadas.

Exemplo:

```
from sklearn.metrics import classification_report

# Imprimir relatório de classificação
print(classification_report(y_test, y_pred))
```

2.5 Comparação entre Diferentes Abordagens de Análise Preditiva

Comparar diferentes abordagens de análise preditiva para determinar a mais adequada ao problema.

Exemplo:

```
# Comparar a acurácia de diferentes modelos
accuracy_decision_tree = accuracy_score(y_test, y_pred_decision_tree)
accuracy_random_forest = accuracy_score(y_test, y_pred_random_forest)

print(f"Acurácia Árvore de Decisão: {accuracy_decision_tree}")
print(f"Acurácia Random Forest: {accuracy_random_forest}")
```