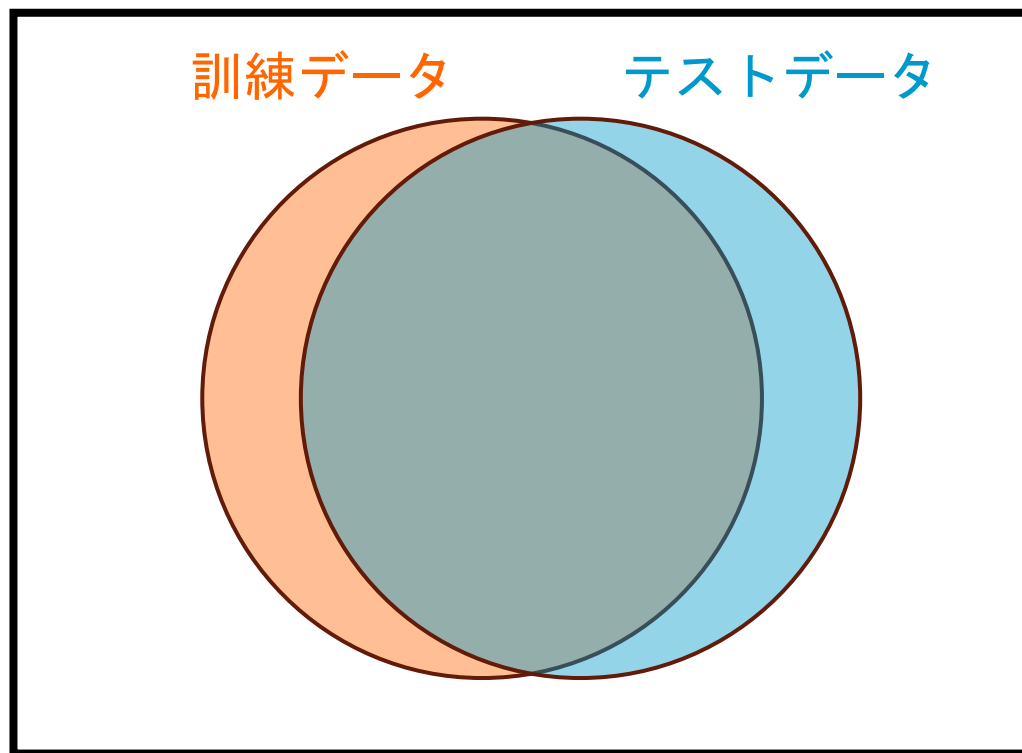


ドメインシフト

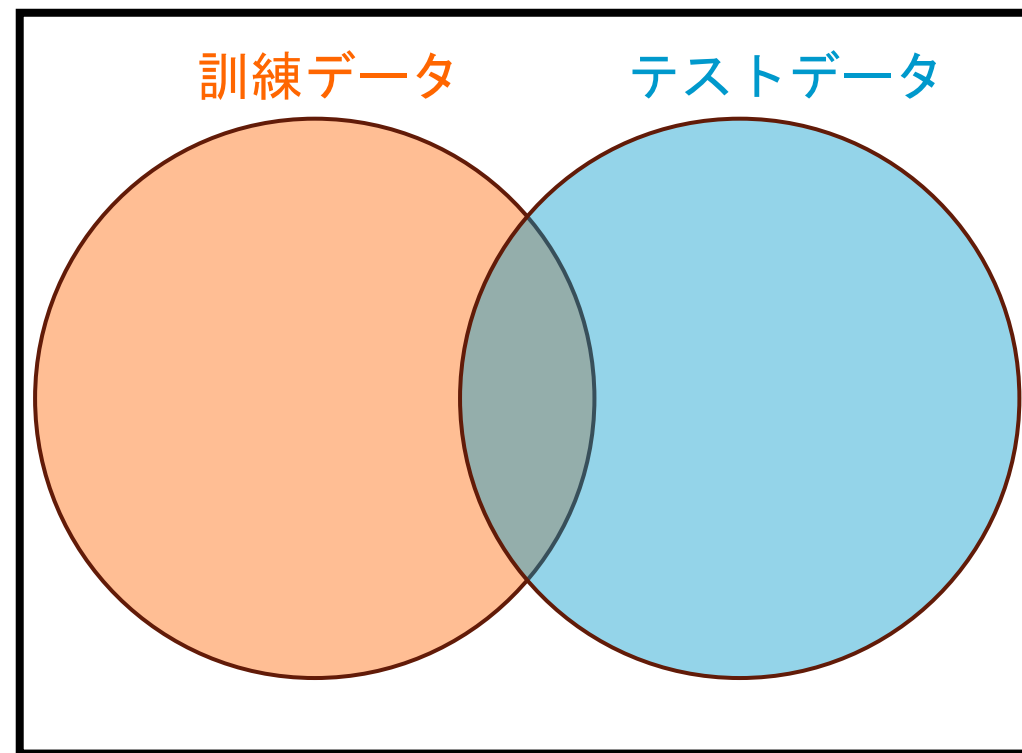
ドメインシフトは、機械学習のコンテキストにおいて、訓練データの分布とテストデータの分布が異なる現象
機械学習モデルが未見のデータに対して期待通りの性能を示さなくなる主要原因の一つ

母集団



訓練データとテストデータの分布が近い
→ テストデータの予測精度が高い

母集団

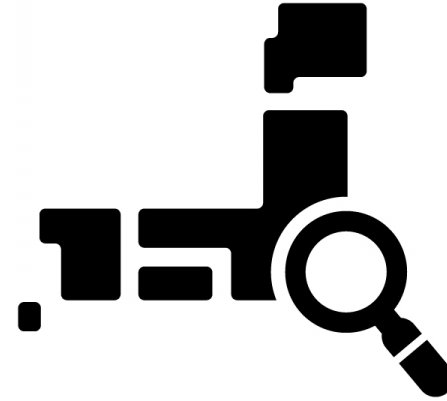


訓練データとテストデータの分布が遠い
→ テストデータの予測精度が低い

ドメインシフトの原因は、時間的変動・地域的変動・機械的変動・操作的変動



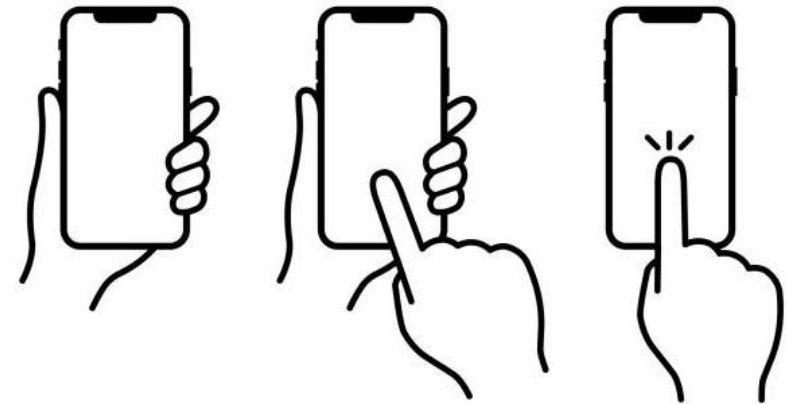
時間的変動



地域的変動



機械的変動



操作的変動

ドメインシフトの対処方法：データの再収集・データ拡張・転移学習・ドメイン適応・不変表現学習

- **データの再収集**

テストデータの分布に合うようなデータを定期的に収集し、モデルを更新

- **データ拡張**

データの多様性を増やすために、画像の回転、反転、色などを変更

- **転移学習**

大量のデータで訓練された学習したモデルを、新しいタスクに転移

- **ドメイン適応**

訓練データとテストデータのドメインの違いを考慮しながらモデルを学習

- **不変表現学習**

異なるドメインのデータでも共通の特徴表現を学習