

Representation Bias

Bias that occurs when training data fails to properly reflect actual population distribution



Data Imbalance Issues

- Underrepresentation of specific races/ethnicities
- Gender imbalance
- Age distribution distortion
- Lack of rare disease data
- Geographic concentration



Resulting Impact

- Lower accuracy in minority groups
- Increased misdiagnosis rates
- Treatment recommendation bias
- Widening health disparities
- Decreased reliability



Correction Strategies

Resampling

Over/Under sampling, SMOTE

Weight Adjustment

Class weights, Sample weights

Data Augmentation

Augmentation, Synthetic data