Bias Detection & Mitigation

편향 유형

- **인구통계 편향**: 특정 인종/성별 과소표집
- **선택 편향**: 비무작위 샘플링
- **측정 편향**: 측정 도구 차이
- **라벨 편향**: 주석자 편견

공정성 메트릭

- **Demographic Parity**
- **Equalized Odds**
- **Disparate Impact**
- **Individual Fairness**

완화 기법

- 재샘플링 (over/under sampling)
- 가중치 조정
- 공정성 제약 조건 추가
- 편향 완화 후처리