**Sid 5 A surey on fairness**

**Measurement Bias.**

Mätning, eller rapportering, bias uppstår från hur vi väljer, använder och

mäta särskilda egenskaper [140]. Ett exempel på denna typ av bias observerades i re-

verktyget COMPAS för att förutsäga cidivismrisk, där tidigare arresteringar och vänner/familjer var arresteringar

används som proxyvariabler för att mäta nivån av "risk" eller "brottslighet" - som i sig kan vara

ses som felmätta ombud. Detta beror delvis på att minoritetssamhällen

kontrolleras och poliseras oftare, så de har högre arresteringsfrekvens. Dock en

bör inte dra slutsatsen att eftersom personer som kommer från minoritetsgrupper har högre arresteringar

priser, därför är de farligare, eftersom det är skillnad på hur dessa grupper är

bedöms och kontrolleras [140].

**(2) Omitted Variable Bias.**

Utelämnad variabel bias4 uppstår när en eller flera viktiga variabler

lämnas utanför modellen [38, 110, 127]. Ett exempel för detta fall skulle vara när någon

utformar en modell för att med relativt hög noggrannhet förutsäga den årliga procentsatsen vid vilken

kunder kommer att sluta prenumerera på en tjänst, men märker snart att majoriteten av användarna

säger upp sitt abonnemang utan att få någon varning från den designade modellen.

Föreställ dig nu att anledningen till att avbryta prenumerationerna är utseendet på en ny stark

konkurrent på marknaden som erbjuder samma lösning, men för halva priset. Utseendet

av konkurrenten var något som modellen inte var redo för; därför anses det

vara en utelämnad variabel.

**(3) Representation Bias.**

Representationsbias uppstår från hur vi tar prov från en population under

process för insamling av data [140]. Icke-representativa urval saknar mångfalden av befolkningen.

tion, med saknade undergrupper och andra anomalier. Brist på geografisk mångfald i dataset

som ImageNet (som visas i figurerna 3 och 4) resulterar i påvisbar partiskhet mot västerländska kulturer.

**(4) Aggregation Bias.**

Aggregationsbias (eller ekologisk felslutning) uppstår när falska slutsatser är det

dras om individer från att observera hela befolkningen. Ett exempel på denna typ av partiskhet kan ses i kliniska hjälpmedel. Överväg diabetespatienter som har skenbar sjuklighet skillnader mellan etniciteter och kön. Specifikt HbA1c-nivåer, som används i stor utsträckning att diagnostisera och övervaka diabetes, skiljer sig på komplexa sätt över kön och etnicitet.

Därför kommer en modell som ignorerar individuella skillnader sannolikt inte att vara väl lämpad för alla etniska grupper och könsgrupper i befolkningen [140]. Detta gäller även när de är representerade skickas lika i träningsdata. Alla allmänna antaganden om undergrupper inom befolkning kan resultera i aggregeringsbias.

1. **Simpsons paradox.**

Simpsons paradox är en typ av aggregeringsbias som uppstår i

analys av heterogena data [18]. Paradoxen uppstår när en association observeras

i aggregerad data försvinner eller reverseras när samma data delas upp i sin

underliggande undergrupper (Figur 2(a)). Ett av de mer kända exemplen på typen av

en paradox uppstod under rättegången mot UC Berke-

ley [16]. Efter att ha analyserat antagningsdata för forskarskolor verkade det som om det fanns en partiskhet

mot kvinnor, av vilka en mindre del antogs till forskarutbildningen

jämfört med sina manliga motsvarigheter. Men när antagningsdata separerades och

analyserat över institutionerna hade kvinnliga sökande jämställdhet och i vissa fall till och med

en liten fördel gentemot män. Paradoxen inträffade när kvinnor tenderade att ansöka om

avdelningar med lägre antagningsgrad för båda könen. Simpsons paradox har varit ob-

tjänstgjorde inom en mängd olika domäner, inklusive biologi [37], psykologi [81], astronomi [105],

och beräkningssamhällsvetenskap [91].

**(b) Modifiable Areal Unit Problem**

är en statistisk bias i geospatial analys, som uppstår vid modellering av data på olika nivåer av rumslig aggregering [56]. Denna partiskhet resulterar i olika trender man lär sig när data aggregeras på olika rumsliga skalor. ACM Computing Surveys, vol. 54, nr 6, artikel 115. Publiceringsdatum: juli 2021.