Machine Learning

Dr Changjiang He, Dr Kuo-Ming Chao
Computer Science| School of Art
University of Roehampton

Lesson 11.1

Bias in Dataset

Bias in ML

- Data bias in machine learning is a type of error in which certain elements of a dataset are more heavily weighted and/or represented than others.

- A biased dataset does not accurately represent a model's use case, resulting in skewed outcomes, low accuracy levels, and analytical errors.

Bias in ML

University of Roehampton London

- In general, training data for machine learning projects has to be representative of the real world.
- This is important because this data is how the machine learns to do its job.
- Data bias can occur in a range of areas, from human reporting and selection bias to algorithmic and interpretation bias.

Sample Bias

- Sample bias occurs when a dataset does not reflect the realities of the environment in which a model will run.

- An example of this is certain facial recognition systems trained primarily on images of white men.

- These models have considerably lower levels of accuracy with women and people of different ethnicities.

- Another name for this bias is selection bias.

Exclusion Bias

- Exclusion bias is most common at the data pre-processing stage. Most often it's a case of deleting valuable data thought to be unimportant.

- However, it can also occur due to the systematic exclusion of certain information. For example, imagine you have a dataset of customer sales in America and Canada. 98% of the customers are from America, so you choose to delete the location data thinking it is irrelevant.

- However, this means you model will not pick up on the fact that your Canadian customers spend two times more.

Measurement Bias

- This type of bias occurs when the data collected for training differs from that collected in the real world, or when faulty measurements result in data distortion.

- A good example of this bias occurs in image recognition datasets, where the training data is collected with one type of camera, but the production data is collected with a different camera.

- Measurement bias can also occur due to inconsistent annotation during the data labelling stage of a project.

Recall Bias

- This is a kind of measurement bias and is common at the data labelling stage of a project.
- Recall bias arises when you label similar types of data inconsistently. This results in lower accuracy.
- For example, let’s say you have a team labelling images of phones as damaged, partially-damaged, or undamaged. If someone labels one image as damaged, but a similar image as partially damaged, your data will be inconsistent.

Observer Bias

- Also known as confirmation bias, observer bias is the effect of seeing what you expect to see or want to see in data.
- This can happen when researchers go into a project with subjective thoughts about their study, either conscious or unconscious.
- We can also see this when labellers let their subjective thoughts control their labelling habits, resulting in inaccurate data.

Racial Bias

University of Roehampton London

- Though not data bias in the traditional sense, this still warrants mentioning due to its prevalence in AI technology of late.
- Racial bias occurs when data skews in favour of demographics.
- This can be seen in facial recognition and automatic speech recognition technology which fails to recognize people of colour as accurately as it does Caucasians.

Association Bias

University of Roehampton London

- This bias occurs when the data for a machine learning model reinforces and/or multiplies a cultural bias.
- Your dataset may have a collection of jobs in which all men are doctors, and all women are nurses. This does not mean that women cannot be doctors, and men cannot be nurses.
- However, as far as your machine learning model is concerned, female doctors and male nurses do not exist. Association bias is best known for creating gender bias.