World Class Machine Learning

Machine Learning Hub

<u>Menu</u>

ABOUT US

<u>Menu</u>

What is the Difference between MAR and MCAR?

Posted on October 1, 2018 by Jonathan Steele

I <u>came across a couple potentially confusing terms</u> and decided to clarify them. They are listed below:

- NMAR – Not Missing at Random
- MAR – Missing at Random
- MCAR – Missing Completely at Random

Context: Why are Missing Values Significant?

This is in the context of when you encounter a dataset that has missing data. It could be because a survey had optional fields or the survey was abandoned before completion. The problem is that, in a Machine Learning context missing data can introduce bias into the training set by omitting values. By introducing bias into the training set, you are probably going to decrease the effectiveness/accuracy of your Machine Learning Model

What are NMAR, MAR, MCAR?

This is where NMAR, MAR, MCAR come in; **It is best to try and discover any plausible explanations for why the data is missing.**

- If you find that there is a trend behind the missing data, then you would say that data is **Not Missing at Random**. If this is the case, there may be a serious constraint on the effectiveness of your study or the performance of your ML model.
- If you do not find any trend within the primary variable of interest, then you don't have any reason to believe that your project will be majorly afflicted by bias from missing data. In this case, you would say that data is **Missing at Random**.
- If you find that there is absolutely no pattern for rows with missing data, then you can confidently continue on with your study/project without bias from missing data. In this case, you would say that data is **Missing Completely at Random**.

Resources

- This Resource and my class's quiz brought on this question for me.
- Wikipedia
- How to Handle Missing Data

LANGUAGE

English ⬍

RECENT POSTS

Block Chain Killer Apps – Article Review

What is the Difference between MAR and MCAR?

Review of Andrew Ng's Machine Learning Course

F Scores

Single Numeric Evaluation Metric

CATEGORIES

Blockchain (1)

Chinese (12)

    Chinese Programming Words (12)

Machine Learning (13)

    Ideas for ML Side Projects (6)

Software Development (2)

    Intro to Python (2)

Uncategorized (2)

PAGES

About Us

🇺🇸 English