Types of Inference انواع استنتاج

Deductive استدلال قیاسی

این که از نتیجه های کلی به نتیجه های جزئی برسیم.

مثلا همه پسرا لاشین.

امیر پسره

پس امیر لاشیه

یا مثل قضیه هایی که توی ریاضی و هندسی اثبات میکردیم.

دراین نوع استدلال به نتیجه ی درستی میرسیم، البته به این شرط که مفروضات درست باشه (مثلا همه پسرا لاشی باشن (:)

In deductive inferences, what is inferred is necessarily true if the premises from which it is inferred are true; that is, the truth of the premises guarantees the truth of the conclusion

Inductive استدلال استقرایی

یعنی از جز به کل رسیدن. با چند مشاهده کلی یک نتیجه درباره یک چیزی بگیریم. به دلیل این که با مشاهدات محدود کار دارد خیلی به آمار مرتبط است.

مثلا مادر یک خانواده هر روز ساعت ۱۲ نهار می‌پزد. به مدت یک هفته کار او را بررسی کردیم و نتیجه گرفتیم که مادر این خانواده هر روز ساعت ۱۲ نهار می‌پزد.

Inductive inferences are based purely on statistical data, such as observed frequencies of occurrences of a particular feature in a given population

Abductive استدلال بهترین تبیین یا ربایشی

مثلا میای خونه و میبینی پنجره شکسته.

حالا واسه شکستگی پنجره چندین فرضیه وجود داره. یکی این که دزد اومده، یکی این که توفان اون رو شکسته و دیگه ای که بچه ها شکستن.

الان هرفرضیه را بررسی میکنیم. مثلا میگیم دزد بعیده اومده باشه چون که محله ی ما امنه

از طرفی توفان هم نیومده چون چند روز قبل هوا آفتابی بوده

ولی احتمالا بچه ها شکستن چون که دیروز میشنیدم توی کوچه داشتن بازی میکردن

حالا میای واسه تست کارای مختلف میکنی، مثلا دوربین مدار بسته چک میکنی یا از همسایه ها پرس و جو میکنی

استدلال آبداکتیو و ایندااکتیو از محدوده ی فرضیات خود فراتر میروند و می توانند چیزی بگویند که ممکن است درست باشد یا غلط

اما استدلال دیداکتیو به شرط مفروضات درست، حرف صد در صد درستی میزند.

مشکلی که این جا داریم این است که ماشین لرنینگ "غالبا" از استدلال ایندااکتیو استفاده میکند که لزوما درست نیست.

The problems of inductive inference:

1. What we measure is a sub-set of reality. This is called survival bias or low hanging fruit.
2. some aspects of our society and our life is only measurable indirectly
3. Societies have historical and structural inequalities
4. Spurious correlations and confounding factors

Bias

1. Historical bias: it is obvious

2. Measurement bias: when there are systematic errors in the way data is collected, measured, or recorded.

3. Representation bias (sampling bias): When the data you have fails to be a good representative of the population; so it tends to generalize. For example, in the dataset some groups are overrepresented or underrepresented. This leads to skewed representation of true distribution. This can cause that the model performs poorly on certain subgroups.

4. Omitted relevant variable's bias: When a model fails to take into account variables that are importantly related to either independent variable or independent variable. When an important variable is ignored, we will have biased estimates of the relationship between independent and dependent variable.

Suppose you want to examine the relationship between education level and income. You collect data on individuals' education levels and their corresponding incomes. Without considering any other factors, you run a simple regression analysis and find a strong positive relationship between education level and income. Based on this analysis, you conclude that higher education leads to higher income. But you missed the effect of many important factors, such as work experience.

5. Deployment bias (implementation bias): when implementation of a decision or intervention happens in a way that introduces biases.

In a study evaluating an educational program, deployment bias can occur if certain schools receive more resources or if there are variations in how the program is implemented. This can skew the results and make it difficult to assess the true effectiveness of the program. Sampling bias and contextual factors can also introduce bias. Mitigating deployment bias involves equitable resource allocation, consistent implementation, representative sampling, and accounting for contextual factors.

6. algorithmic bias: when the outcome of an algorithm exacerbates the situation for the minority.

6.1 aggregation bias: when subgroups are so different that they should be considered differently rather than one

6.2 learning bias: It arises when algorithmic design choices (e.g., the learning objective function) are not equally suited for all subgroups.

6.3 evaluation bias: it arises when evaluation data do not represent the target population, irrelevant metrics of performance are used.

- facial recognition mostly trained on white face, then used/evaluated in contexts where other skin colors exist

- a model trained with purchases record on American supermarkets then evaluated with data from African supermarkets.

- Model minimizes false negative, when false positive minimization is more important

Sources of Bias

Goodhart's law: When a measurement becomes the target, it stops being a good measurement.

Campbell's law (reflexivity problem): The more quantitative indicator is used, the more it will influence the people involved. This changes the person's behavior and corrupt the evaluation.

The Simpson's paradox: a behavior, trend etc. tends to be different in subgroups than when the subgroups are aggregated and form a supergroup.

Fairness Criteria

1. Independence: Independence refers to the notion that a sensitive attribute (such as race or gender) should be independent of the predicted outcome or decision.

Advantages:

• it can be applied at every stage of the process

•Disadvantages:

• It ignores the possible correlation between Y and A.

• It allows to have good classifications in one group, and

random classifications in another

2. Separation involves maintaining a clear separation between sensitive attributes and other variables within the model.

Advantages:

• It is compatible with R=Y

• It incentives to reduce errors uniformly in all groups

Disadvantages:

• more difficult to apply

• Does not take into account false negative rate

3. Sufficiency: Sufficiency ensures that the model has access to enough relevant and non-sensitive information to make accurate predictions or decisions.

If they say the model is calibrated, it means that sufficiency is respected.

Fairness Qualitative Assessment

In order to perform a fairness qualitative assessment, the four questions below should be answered:

1. What social conflicts arise from the design/operation of the system?

2. Which stakeholders are adversely affected?

3. What values/social interests are at stake?

4. How the conflict can be solved?

Compass

AMS

It was initially a project to identify those who need consultancy and classes to go to job market.

But it turned into a classification task by group L, H, M.

Its political goal was to spend less for training people for the job market.

There is a problem in data collection. They used data that were initially collected for other purposes, but without consent was used also for this respect.

****you can say there are three genders, binary one is discrimination (woke hard shit)

Age: building a model for all people below 30 is discriminatory.

Citizenship: very general, what does it mean non-EU?

Aggregation bias: High school very different with university

Measurement bias:

Having a care obligation only applies to women

Historical bias:

Women: when women were always marginalized, you train your data and they fall most often in the lower segments

Migration background and people above 50

Regional labor market

Obligation of care: value-laden and sexist

Occupation group: Very general

Prior occupational career

Cases with 4-year interval

Duration of cases: indicates how long did you need assistance

Measure claimed: what kind of help you sought

Duration of employment: about the last post

Regional labor market: there's issue because economy is strong in some places and weak in other places

Another minor issue is that people couldn't object to the results.

Having only TPR is not good because it only monitors TP. What about FN?

And also, TPR skewed through the minority.

When we say TPR rate is high and it's 90%, how many people it will affect?

The design of the system is to punish people rather than give them help, because assistance for job seeking was turned from universal to selective

Then explain historical bias and measurement bias

Aggregation: high school different from university, some groups have higher chance of employability

Employability changes from what you already capable of to what group characteristic you have

Objectives as stated by AMS:

1. increase the efficiency of the counselling process

    When we routinely select candidates, how the counselling process can be efficient?

2. Increase the effectiveness of the used funds

    Strict profiling of clients endangers this because there are people who need this fund and cannot have it

3. Standardize the granting of funds

    When you give a group more funds than they deserve, how it is standardizing?

Detecting social welfare in the Netherlands (SyRi)

Organization goal:

Classification (system goal):

Data was incredibly large and collected from many sources and collected many things

It was initially pseudonymized then reverted back to normal state

A complex procedure to verify that the risks involved are recognized

If you apply a system in a certain area, you would definitely find the fraud only in that area!

You have Campbell's law in the form of feedback loop

If you didn't read the national Gazette, you couldn't know you are subject to it, unless it is the end of the process; as a result, people couldn't object the results

They didn't reveal how the system works because they said people will adjust their behavior accordingly

Scientific experiment: the results weren't reproducible, they said they have identified the risks involved but the risks were random thus spurious correlations

The court said it is illegal to use this system but the damage was done

1. Privacy of people were compromised (Article 8)

2. If your privacy is intruded, they should have a good reason; which they didn't; it wasn't a fair balance because it undermines individual rights for the public good.

3. it is not only infringement of privacy, but also other rights as well

The system was a veritable disaster for example they said that if you are a single person and you have a low consumption of water, you are a welfare fraud /////: the rational was that in such case you probably live at another place and receive welfare as single. But it ignores the chance that maybe this guy travels often, or the water counter is defective, the guy is frugal etc. THIS IS FUCKING CRAZY.

In this case, they had distrust to citizens because they saw everyone of them as potential frauds.

There was a contrast between people's individual right against a public good (to understand welfare fraud)

People were angry and the prime minister resigned

Alternative possibility was to investigate people who had the right to allowance but didn't require it

The childcare-benefits affair

Tens of thousands of parents were identified fraud without the right to objection and explanation particularly in the minority

So parents who were identified frauds in this respect, were sanctioned and penalized. Their benefits suspended while they needed it. They were hostilely investigated.

The consequence was dire as families experienced debt, unemployment and forced evictions because people were no longer able to pay the mortgage. Also, there was mental health issues and personal problems and broken homes.

Coders

Qualitative Assessment

Code of Conduct

Kinds of ethics