Information a(symmetry)

Data and the professional sphere

Data is everywhere. However, it is quite challenging to **measure the impact it has in our life**. Ben Jones proposes to focus on three areas to understand its impact: the professional area, the personal and the public one.

This paper will focus on the usage of data in the **professional sphere**; however, it will also cover its impact on the other two. Moreover, the analysis will be based on **HireVue's algorithm**, the video interview platform used to select candidates contracted by more than one third of Fortune 100 companies, such as JP Morgan, Goldman Sachs, Unilever, Vodafone…

Many papers have published research and advice on succeeding in video interviews. Nonetheless, **few truly understand how candidates are being evaluated** and, while ML models are the source of many biases on their own, we find that information asymmetry is an even greater source of bias. The main goal of the paper is to demonstrate the relative strength of bias in the interview process. I will consolidate my findings on AI algorithms where I will conclude that, while there is no immediate solution, information symmetry is the best way to achieve a fairness.

Information:

The **information flow** regarding video interview process is the following: first, we have the information that the platform (HireVue in our case) has, second, the information they offer to companies (their main revenue source) and third, we have the information that reaches candidate. Our focus here will be on the latter, given that, it is where the greatest asymmetry is created.

I will begin by explaining the information an average candidate can normally accesses, then I will compare it to what the algorithm considers, and I will evaluate the differences in bias coming from the ML model vs those coming from not shared information.

Candidate information:

As a candidate, one wants to prepare at best for their interview, and the normal process would start by googling: ***'How to prepare for a video interview**

HireVue'*.

Here, the most common and cited papers one could find are the following: HireVue Interview Guide How to prepare for a HireVue interview (CFI), How to succeed in a HireVue Video Interview posted by the Duke Career Hub *(Asmita, 2022)*, A face-scanning algorithm increasingly decides whether you deserve the job *(Harwell, 2019)*.

Moreover, a candidate can find numerous videos on youtube that summarise how to 'hack' the video interview.

In all these videos and articles we find the following **patterns of repeated information**: they all recommend to prepare answers to common interview questions ***(verbal behaviour)*** and to make a good usage of body language ***(NVB: non-verbal behaviour).***

- **Examples include:**
  - From the **Duke Career Hub** [Published on January 11, **2022**]
    - No rapid movement: Don't move your hands/ shoulders too much. Constant body movement indicates nervousness and unpreparedness. Keep your body in a calm and composed state.
    - Voice modulation: When answering, be sure to emphasize important words and take pauses when needed.

  - From the **corporate finance institute** *(Vipond, 2023)*[Published December 1, **2018** and Updated October 5, **2023**]
    - Movement: HireVue interviews are video-based and allow a company's recruiter to see non-verbal cues – such as facial expressions, eye-movements, body movements, details of clothes, and nuances of voice.

  - From the video **HireVue Video Interview: 5 MISTAKES You Need to AVOID**: https://youtu.be/J2VnJOw5Cd0?si=awJX4PpVG4rsTV8y [(1.2M views]
    - This video suggests among other things to ensure the light is the best and that there is nobody in the surroundings – to understand the size of the audience)

Therefore, as a candidate, **the path to success seems very straightforward**: all one has to do is must develop a good response, saying many times the words from the job description and smiling a lot but not too much as it might be a sign of nervousness.

It is quite bizarre; however, that while we have all this information from blogs and secondary sources, ***how difficult*** it is to find any information **regarding technical information from the *platform itself***.

A possible reason could be that given that the company has conducted more than 33M interviews and it has clients such as JP Morgan, Goldman Sachs etc. they have a **different privacy agreement** with each one of the clients and thus they cannot disclose too much information.

How the algorithm works (*hidden* information):

Nonetheless, after extensive research I was able to find the source that explains how the algorithm that they have is used: `HV_AI_Short-Form_Explainability_3152022.pdf` *(Zuloaga, 2023)* and the last update I found was from 2022.

Here one can read about ***everything needed to **understand*** why the information available to the candidate is distorting the information and it is making applicants to focus on the wrong issues.

The bias that is normally discussed:

There are many possible **sources of bias** one could find in the paper, for instance:

- ***Who is responsible?*** - **shifting accountability to companies**:
  - *we are acting as a 'data processor' and are collecting and processing candidates' personal information on their behalf and in accordance with their instructions. That is an important distinction because it means that the majority of the obligations** under the EU and UK's GDPR **are required to be fulfilled by our customers**, and not HireVue.*

- **The test sample "inclusivity" or record date**:
  - We see 78% of the training data was from Northern America, one can make its own conclusions for non-US applicants.

```
* *Only shown in the last page of the appendix*
```

And I am confident that if one reads the full statement, there are many more issues one could call into question.

The real source of bias (*never discussed*):

However, here I want to focus on the issue of NVB (nonverbal behaviour) analysis as it is emphasised by most papers we have available to date.

HireVue clearly states: *"Importantly, our AI relies only on what is said by the candidate and* **does not use any video analysis or other audio characteristics*** (meaning that we do not assess a candidate's facial expressions, body language, their background and surroundings, or tone-of-voice)".*

Excluding the 'greatest indicator' for success:

The decision to not include NVB and intonation analysis in the interviews was taken by **HireVue CEO in March 2020**. This means that all the papers we see above (with a date after 2022) must be based on papers before 2020 which display distorted information.

There are **two main** sources of bias in NVB that might have led HireVue to decide that verbal analysis would have been better than a combination of the both verbal and non-verbal analysis. *(Renier et al., 2021)*

1. Biased ground truth
2. Unsupervised learning bias

The first type of bias is generated because the quality of the output (ie good candidate vs not) using nonverbal social sensing depends on the extent to which the data coded by the algorithm **resemble** the data on which the algorithm had been trained (aka ground truth).

Here is a very clear example to understand this type of bias: *As with many other things in this new industry, that sounds good until you think about it; then it becomes replete with problems. Given the best performers of the past, the algorithm will almost certainly include white and male as key variables. If it's restricted from using that category, it will come up with attributes associated with being a white male, such as playing rugby.** *(Cappelli, 2019)*.

The second type is explained by the fact that algorithms learn to detect patterns by themselves (i.e., unsupervised learning). For instance, algorithms might learn by themselves to discriminate women during the recruitment process (e.g.,

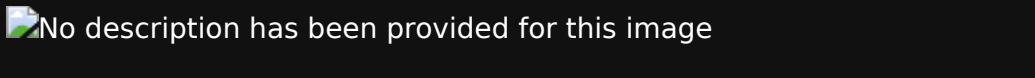
Dastin, 2018; Lambrecht and Tucker, 2019) without the developers or users being aware of this bias. *(Renier et al., 2021)*

Why is this relevant?

The **real problem** is that this issue not only overlaps with the public sphere, but it also aggravates the **personal sphere** by creating a sense of **loss of locus of control and unfairness**. We can observe is that those candidates who prepare more and search for advice on google will be penalized in comparison to candidates that do the interview without making any research. This is because they will be so centered on the non-verbal language and other things the model is not even considering that during the interview, they will lose focus.

As explained in the article by the Washington Post: *The inscrutable algorithms have forced job seekers to confront a new kind of interview anxiety. Nicolette Vartuli, a University of Connecticut senior studying math and economics with a 3.5 GPA, said she researched HireVue and did her best to dazzle the job-interview machine. She answered confidently and in the time allotted. She used positive keywords. She smiled, often and wide. But when she didn't get the investment banking job, she couldn't see how the computer had rated her or ask how she could improve, and she agonized over what she had missed.** *(Harwell, 2019)*.

Moreover, all the hours they spent in reading articles and videos about NVB and intonation in interviews has been towards nothing, because even though companies argue that they can watch the video of the candidates selected, there is always that pool of 33% of candidates that will be passed directly into `low tier` and under very few circumstances they will ever by viewed by a human.



In fact, if we go back to those 33M interviews they have recorded up to date, we can see that **the effort of 10,890,000 students** that not only submitted a normal application as a candidate (with CV and Cover Letter...), but were also 'selected' for continuing to the first step, and took the time to prepare and record the video in the best environment with the best light **has been, in the best scenario, used to improve the algorithm of the company (*for free*).**

Who should we blame?

Is it HireVue's responsibility for not making the information more accessible or is it the company's responsibility that should give this information to the

candidates? The reality is that, if we were to start looking on who to blame, we would end up in a cycle with no end and most importantly, with no real solution.

Video interviews are certainly **adding value** for companies and candidates. Candidates can do applications from wherever they are and at any time available and companies do not have to spend all the time in the process of finding new hires.

Moreover, **bias will always be present** in any process we have. Nobody can prove that a human interviewer is going to be less biased than the algorithm during the process as the algorithm in the end mimics what the human does. So if the algorithm's output is biased is because, by nature, we are biased as well.

Thus eliminating the HireVue process would not only be a sub-optimal solution, but it would also hinder innovation.

What can be done?

However, **there is something that we can do**.

While the bias arising from the model is as much controllable as the bias coming from humans, we find that the real concern we should have is the additional bias created from not shared information.

Here is how I would represent it:


No description has been provided for this image

Therefore, the one true and possible solution we can find is to **de-centralize information**, and not only make it *"accessible"*, but really add the effort so that it is easy for the user to actually **gain knowledge** from that information.

Companies like Amazon are already doing this. When you receive an application from them, they will explain to you exactly what they expect and as a candidate there is not this feeling of insecurity that is created in most of the other processes.

In conclusion, this paper suggests a **different view on bias**, one that is normally avoided because of all the noise that algorithms and data create. The reality is that the model is not generating new sources of bias, we are. Thus, **the only solution we should aim for is to recover symmetry**.