International Journal of Engineering Trends and Technology
ISSN: 2231–5381 / https://doi.org/10.14445/22315381/IJETT-V72I9P134

Volume 72 Issue 9, 367-383, September 2024

Original Article

Gender Biases in Professions: A Machine Learning – Powered Search Engines Analysis

Nicolás Alejandro Tirado Vilela[1], Adriana Maemi Ueunten Acevedo[1], Marcos Fernando Ruiz-Ruiz[1], Wilfredo Yushimito[2]

[1]*Faculty of Engineering, University of Lima, Lima, Perú.*
[2]*Facultty of Engineering and Sciences, Adolfo Ibáñez University, Valparaíso, Chile.*

[1]*Corresponding Author : 20181885@aloe.ulima.edu.pe*

Received: 25 August 2023 Revised: 22 November 2023 Accepted: 27 August 2024 Published: 28 September 2024

Abstract *- Machine learning is becoming increasingly important and pervasive in people's lives. Yet, when its conclusions reflect biases that support ingrained prejudices in society, many vulnerable groups' psychological wellbeing may be impacted. The study focuses on occupations to investigate if gender biases exist in image search engine algorithms that use machine learning. To do this, searches for various professions were run on Google, DuckDuckGo, and Yandex. Using web scraping techniques, a sample of images was retrieved for each selected profession and search engine. The images were then manually classified by gender, and statistical indicators and analyses were computed to detect potential biases in the representation of each gender. This analysis included a comparison between search engines, the calculation of mean, standard deviation, and coefficient of variation, a confidence interval analysis, a logistic regression analysis, and a Chi-Square test. It was discovered that there is a strong association between men and leadership positions or STEM professions, while women are predominantly portrayed in traditionally female-associated professions. For instance, it was discovered that 100% of the search results for secretaries and nurses in Yandex are female, while 94% of the search results for engineers are male. Similar statistics may be found on DuckDuckGo, where 96% of results for mathematicians were men, and on Google, where 73% of results for teachers were women. These findings illuminate novel manifestations of gender prejudices in contemporary society and their potential to affect access to particular professions.*

Keywords *- Diversity, Gender biases, Machine learning, Professions, Search engines.*

1. Introduction

The goal of Machine Learning (ML), a branch of Artificial Intelligence (AI) and computer science, is to quickly and logically develop solutions that can deliver consistent results for any given task [1], mimicking how humans learn using data and algorithms [2]. However, because people developed them, these algorithms have a propensity to embrace the same biases found in society due to the vast amounts of data needed to train the models. Harvard Business Review [3] claims that ML systems occasionally adopt prejudices depending on gender. For instance, the article examines word associations in ML algorithms and Natural Language Processing (NLP) and found that virtual assistants like Alexa, Siri, and Amazon associate words like "doctor" with "masculine" and "nurse" with "feminine." Furthermore, although the scope of this study is restricted to the male and female genders, it is crucial to understand that non-binary genders are similarly impacted by the gender biases found in ML algorithms. According to Metz [4], gender biases may result in problems with facial recognition software that uses AI, such as security applications at sporting events, concerts, and airports. Non-binary and transgender expressions would be disregarded if algorithms took "male" and "female" into account, which would impact these populations.

Many institutions currently base their decisions on ML-based AI systems. These systems help determine how much credit financial institutions provide certain customers, which job candidates organizations contact for interviews and more. Moreover, individuals are also significantly impacted by ML in their daily lives; for instance, one might see such biases in search engines when looking for jobs, doing translations, and even searching for images. The issue is that gender bias in these systems is widespread and has a negative long-term impact on women's psychological, economic, and physical well-being because it amplifies and reinforces detrimental gender stereotypes and prejudices that already exist [5]. Therefore, this research aims to explore whether it is possible to identify gender biases in ML-based search engine algorithms. Peake and Feldman [6] note that ML models have been used in a variety of societal contexts, frequently in circumstances that have an impact on social well-being. Leavy [7] agrees with this idea, stating that AI is having a growing impact on people's attitudes and actions in daily life. Although the models provide concise solutions to complex issues, Thelwall [8] contends that they may also pick up on and perpetuate societal prejudices and stereotypes.

As a result, algorithms may also be biased and treat certain groups or people differently based on factors like gender. Therefore, it is crucial to eliminate gender biases in learning models as society increasingly relies on AI to assist in decision-making. As Celis and Keswani [9] state, the problem of gender biases in search engines can be summarized as the representation or favouring of certain people in the results belonging to privileged groups regarding socially prominent attributes like gender and the overrepresentation of men in the design of these technologies could silently undo decades of progress in gender equality [7]. Search engines are widely used as a filter for the wealth of information available on the internet, especially in the image searching field [10]. Therefore, their current relevance makes them an optimal subject for the present study. Suppose any gender biases were confirmed as a visual representation of the selected occupations (including STEM field-related professions).

In that case, the large audiences that these engines cater to will likely be exposed to the reaffirmation of gender roles by associating a particular gender with the profession. Applications of AI and their advantages are becoming more and more popular in a variety of professions. With the emergence of capable models that employ AI techniques, these applications are anticipated to soon permeate all industries. Through data transformation and knowledge extraction, these algorithms can extract value from massive amounts of data for decision-making and predictive analysis [11]. In addition, Lemoine et al. [12] state that using ML systems for training enables the development of meaningful metrics for equity, demographic parity, equality of probabilities, and equality of opportunities, as well as the achievement of fair results. The main goal of this research is to analyze the responses provided by several image search engines to find any potential gender biases in these machine learning algorithms. To do this, data was collected using web scraping methods and then categorized, allowing the exploration of the extent of gender bias in search engine operation.

2. Literature Review on Machine Learning, Gender Roles, and Occupation

AI is a field of computing that focuses primarily on the transmission of anthropomorphic intelligence and thinking into machines that can assist humans in many ways, creating structures with intelligence on par with or better than that of humans [13]. A major branch of the former is ML, which focuses on leveraging data and algorithms to mimic human learning processes while gradually increasing the accuracy of the results [2]. However, considering that people are susceptible to cognitive biases, using human cognitive talents and IQ as the benchmark for ML models has several negative effects [14]. According to Howard and Borenstein [15], gender biases in AI algorithms result from pattern recognition in input data because they represent societal biases that the systems ultimately reproduce. But, according to Booth et al. [16], gender biases in machine learning algorithms are the result of decisions made by the people in charge of the model, not just the input data.

On the other hand, ML models are more likely to produce biased results based on delicate traits like gender when they are created to optimize just one parameter, such as accuracy or gains [6]. The usage of ML models in daily life must not support stereotypes or gender roles as they are employed in decision-making (either in addition to or instead of people). However, due to ongoing exposure to biased information, biases frequently appear subconsciously in decision-making models, making their mitigation difficult. The fact that biases frequently serve as a protective mechanism in high-risk situations makes them difficult to eliminate [15].

In the workplace, gender roles have a significant impact on the employment that people choose. Society frequently steers people into professions based on various traits and features associated with gender [17]. Additionally, according to Otterbacher et al. [18], receiving gender-biased results from image search engines has a big impact on how people perceive the world. This occurs as a result of users' high level of trust in search engine results, who view them as impartial and accurate. The retrievability of images that corroborate people's cognitive biases is higher than those that do not, which leads to a vicious cycle where these outcomes reinforce people's preconceptions.

Numerous research concur that the quality of input data is the primary contributor to gender biases in ML systems. In contrast, Aleyani [14] also underlines how society is accountable for propagating cognitive biases because, in the end, ML models gather information from human experience and repeat it. Similarly, image search is widely used in a variety of fields, from creating educational resources for students to communication experts creating content for both conventional and contemporary media. Considering this, acquiring and retrieving skewed results might have an impact on historically underprivileged groups and intensify discrimination [19]. Studies take into account a wide range of disciplines where these gender inequalities are replicated. For instance, data in the field of voice recognition revealed that Google Speech's algorithm was more likely to understand questions given by male voices [20]. On the other hand, Reich-Stiebert and Eyssel [21] explore the potential advantages of giving instructional robots a particular gender in order to enhance student learning. It was discovered in 2015 that the hiring process at Amazon, one of the most valuable brands in the world in terms of money, disfavoured female candidates. This was due to the fact that the resumes of past employees from 10 years were utilized to train the algorithm, and they tended to be dominated by men. The algorithm swiftly learned to penalize resumes containing the word "women's," for example, "women's chess club," and penalized women's universities as well, mirroring the male dominance in the technological industry [22].

After conducting a thorough literature analysis on gender bias in ML and AI, Shrestha and Das [23] found that 48 out of 120 studies primarily focused on exploring the existence of bias in the algorithms. Many of these works study NLP, covering dialogue generation, machine translation, text parsing, and sentiment analysis. They also found that research on gender inequalities in audiovisual media is scarce, especially when it comes to image search engines.

Among the studies found covering audiovisual media, Gutierrez [24] investigated gender bias in the field of image search by manually counting the results showing women following a search query in Google. The study discovered that only 8% of the results for the search "CEO" on the platform featured women. Although there is a notable underrepresentation of said gender in the study, it is restricted to a single search engine and does not include a more comprehensive statistical analysis.

The present study advances the previous research by using a more thorough and rigorous technique. To assess gender bias, it pulls search results from Google, DuckDuckGo, and Yandex and runs four different hypothesis tests to measure the over- or underrepresentation of the genders. This study presents strong evidence of gender bias in the image search engines analyzed by comparing multiple platforms and employing statistical analyses, offering a more thorough and methodical assessment than earlier research.

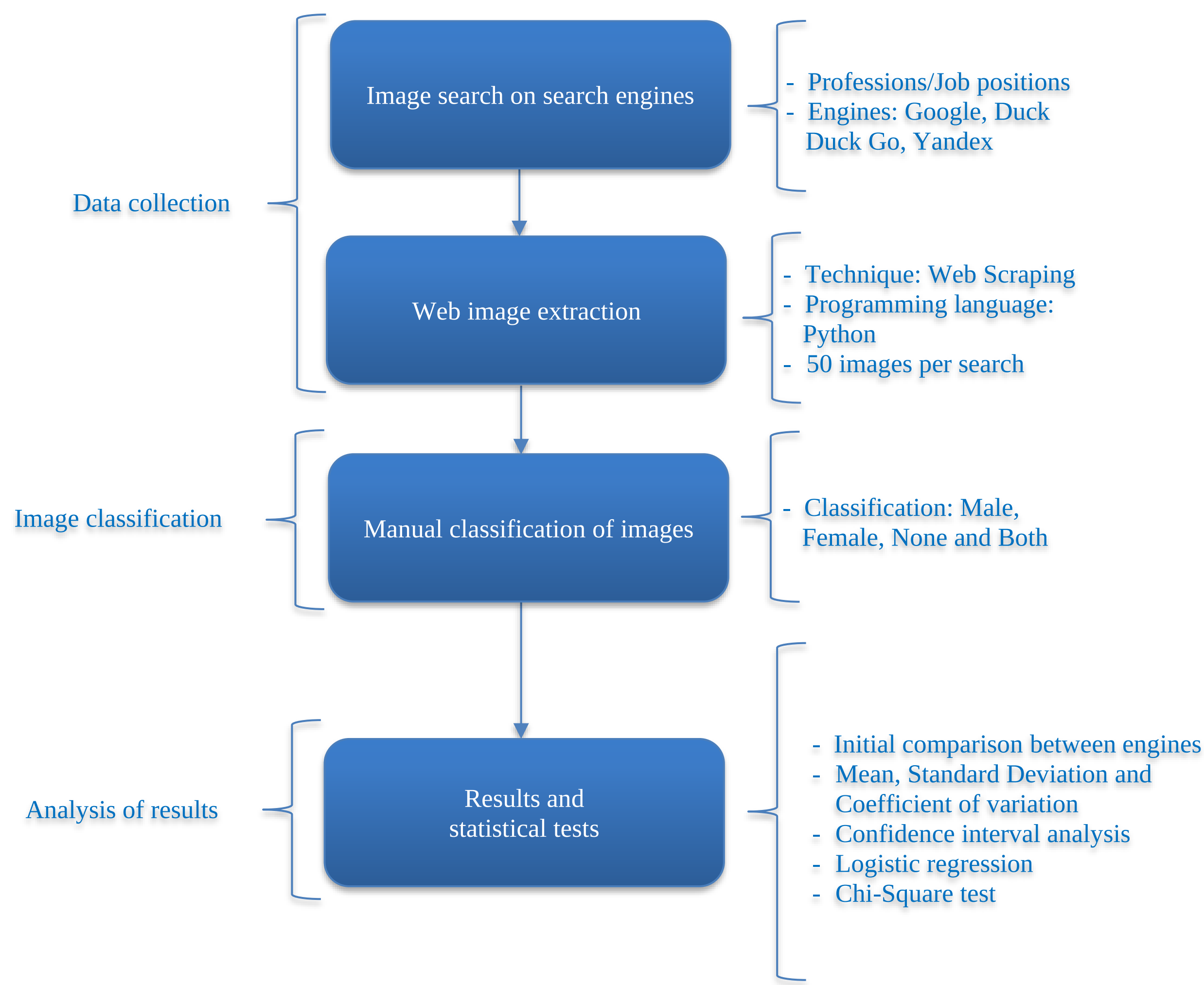

Fig. 1 Flowchart of the methodological process

3. Materials and Methods

The study was non-experimental, cross-sectional, and descriptive. It was divided into the following three stages: data collection, image classification based on the gender of the subjects, and result analysis. Google Images, DuckDuckGo, and Yandex were the search engines selected for the initial round. The flowchart in Figure 1 graphically depicts the methodological procedure.

3.1. Data Collection

This initial stage aimed to gather the sample, which consists of images from the chosen search engines. The following terms were selected for web searches: CEO, engineer, manager, lawyer, nurse, teacher, secretary, mathematician, and scientist. Engineers, mathematicians, and scientists were picked because they work in STEM subjects and are traditionally linked with the male gender. Similarly, nurses, teachers, and secretaries were chosen because they are occupations traditionally associated with the female gender. Additionally, it was thought crucial to add the CEO and manager search, as these symbolize the most significant roles in a firm, and it would be important to investigate which gender the search engines associate with this word. Finally, the lawyer was also included because, unlike the preceding situations, it does not have a strong relationship with any particular gender. It is significant to note that the terminology was developed in English due to its gender-neutrality. The sample consisted of the first 50 results shown by each search engine for each query, totalling 1050 photos to be processed. The information was extracted using a technique called web scraping, which gathers data from the internet and stores it in a file system for later processing and analysis [25]. It is crucial to note that Chrome Driver was used to prevent bias in the results from the computer's search history. This was done in the Sublime Text code editor with Python as the programming language. The subsequent step is the code utilized for web scraping of images in Google.

The code starts by importing the required modules and libraries for the script. The details of these components are the following:

- Pillow: An open-source library that enables working with picture files.
- Selenium: A testing environment that enables automated web browser download and code retrieval.
- Requests: An HTTP library that enables the downloading of information about the image.
- io: Python module for handling input-output operations.
- Time: For adding delays

```python
from selenium import webdriver
from selenium.webdriver.common.by import By
import requests
import io
from PIL import Image
import time
```

Fig. 2 Importing necessary libraries

```python
PATH = "C:\\Tesis\\Web_Scraping_Google\\chromedriver.exe"

wd = webdriver.Chrome(PATH)
```

Fig. 3 Setting up the chrome driver

```python
def get_images_from_google(wd, delay, max_images):
    def scroll_down(wd):
        wd.execute_script("window.scrollTo(0, document.body.scrollHeight);")
        time.sleep(delay)
```

Fig. 4 Defining functions

```python
    url = "https://www.google.com/search?q=ceo&rlz=1C1ONGR_enUS964US964&sxsrf=ALiCzsbB4rylciIyx_
    wd.get(url)
```

Fig. 5 Setting up Google Images

Here, the path to the Chrome WebDriver is specified, and a new instance of the driver is created.

• `get_images_from_google(wd, delay, max_images)`: This function is used to extract URLs from images. Three parameters are needed: wd (WebDriver instance), delay (the amount of time in seconds between scrolling and clicking on images), and max_images (the maximum number of images to be scraped).

• `scroll_down(wd)`: This function scrolls down the webpage to load more images. It utilizes JavaScript.

This part of the code launches the WebDriver instance at the URL that has been set up for the Google Image search with the query “CEO”. This line of the script varies depending on the search term; for instance, the Google search for “engineer” has a different URL.

```python
    image_urls = set()
    skips = 0

    while len(image_urls) + skips < max_images:
        scroll_down(wd)

        thumbnails = wd.find_elements(By.CLASS_NAME, "Q4LuWd")

        for img in thumbnails[len(image_urls) + skips:max_images]:
            try:
                img.click()
                time.sleep(delay)
            except:
                continue

            images = wd.find_elements(By.CLASS_NAME, "n3VNCb")
            for image in images:
                if image.get_attribute('src') in image_urls:
                    max_images += 1
                    skips += 1
                    break

                if image.get_attribute('src') and 'http' in image.get_attribute('src'):
                    image_urls.add(image.get_attribute('src'))
                    print(f"Found {len(image_urls)}")

    return image_urls
```

Fig. 6 Scraping images

```python
def download_image(download_path, url, file_name):
    try:
        image_content = requests.get(url).content
        image_file = io.BytesIO(image_content)
        image = Image.open(image_file)
        file_path = download_path + file_name

        with open(file_path, "wb") as f:
            image.save(f,"JPEG")

        print("Succes")
    except Exception as e:
        print('FAILED -', e)
```

Fig. 7 Defining the download function

Two variables, `image_urls` and `skips`, are initialized in this section of the code to keep track of the scraped image URLs and the number of skipped images, respectively. The code then goes into a loop, iterating through the page's thumbnails while scrolling down the webpage to load more images. It attempts to click on each thumbnail to see the larger version of the image. If successful, the image URLs are extracted and added to `image_urls`. It makes sure that no duplicate URLs are inserted, and if one is found, it increases the number of photos that are skipped.

The function called `download_image` is defined in this portion of the code and is responsible for downloading and saving images from the specified URL. `Download_path` (the directory where the images will be saved), `url` (the image URL to be downloaded), and `file_name` (the name under which the image will be saved) are the three arguments that it requires.

Inside the function:

1) It attempts to retrieve the image's content from the provided URL using the requests.get function and reads it as binary content.
2) To treat the image content as a file-like object, it generates a BytesIO object.
3) It opens the image using the Image.open function from the PIL library.
4) The download_path and file_name are concatenated to create the file path.
5) It saves the image to the specified file path using image.save, specifying the file format as JPEG.
6) If successful, it prints "Success". If an exception occurs during any step of the process, it prints "Failed" along with the error message.

The code calls the `get_images_from_google` function to retrieve a set of image URLs from Google Images through the WebDriver instance `wd`, with a delay time of 3 seconds and a maximum of 50 images. Then, it iterates over each URL using a for loop and downloads the corresponding image by calling the `download_image` function.

The file name of every image is obtained from its index in the list of URLs, and it is saved in the "CEO_images/" directory. Each search term has its own directory path; for instance, the engineer images will be downloaded in "engineer_images/".

```
urls = get_images_from_google(wd, 3, 50)

for i, url in enumerate(urls):
    download_image("CEO_images/", url, str(i) + ".jpg")
```

Fig. 8 Retrieving images

```
wd.quit()
```

Fig. 9 Quitting the WebDriver

```
PATH = "C:\\Tesis\\Web_Scraping_DuckDuckGo\\chromedriver.exe"
```

Fig. 10 DuckDuckGo WebDriver path

```
url = "https://duckduckgo.com/?q=ceo&t=h_&iar=images&iax=images&ia=images"
```

Fig. 11 DuckDuckGo web URL structure

```
while len(image_urls) + skips < max_images:
    scroll_down(wd)

    thumbnails = wd.find_elements(By.CLASS_NAME, "js-lazyload")

    for img in thumbnails[len(image_urls) + skips:max_images]:
        try:
            img.click()
            time.sleep(delay)
        except:
            continue

        images = wd.find_elements(By.CLASS_NAME, "js-detail-img-high")
        for image in images:
            if image.get_attribute('src') in image_urls:
                max_images += 1
                skips += 1
                break

            if image.get_attribute('src') and 'http' in image.get_attribute('src'):
                image_urls.add(image.get_attribute('src'))
                print(f"Found {len(image_urls)}")

return image_urls
```

Fig. 12 DuckDuckGo thumbnail and image elements

Finally, this line closes the WebDriver instance.

The code used for scraping images from DuckDuckGo is similar to the code for scraping images from Google, with a few distinctions: The primary differences with the Google code lie in the Chrome WebDriver path and the URL format for image searches. Similar to the Google code, this URL varies based on the search term used.

Within the Google code, the class names "Q4LuWd" and "n3VNCb" were used to identify the thumbnails and full-sized images, respectively. In the DuckDuckGo code, the thumbnails are identified using the class name "js-lazyload", and the full-sized images are identified using the class name "js-detail-img-high". Finally, the function for extracting the image URL is now defined as `get_images_from_DDG`, maintaining the same parameters. The final images are saved in a different directory path from the Google images.

The last browser used was Yandex. The differences in the Yandex code compared to the Google code are analogous to those highlighted in the DuckDuckGo code. These include variations in the URL structure, specific class names used for identifying thumbnails and full-sized images, the WebDriver path, the name of the function used for extracting the URLs from the images and the directory path used for saving the images. Such distinctions accommodate the unique structure and functionality of Yandex's image search page, mirroring the adjustments made for DuckDuckGo. Figures 13, 14, and 15 show the main changes in the Yandex script.

```
PATH = "C:\\Tesis\\Web_Scraping_Yandex\\chromedriver.exe"
```

Fig. 13 Yandex WebDriver path

```
url = "https://yandex.com/images/search?pos=0&from=tabbar&text=ceo&img
```

Fig. 14 Yandex web URL structure

```python
    while len(image_urls) + skips < max_images:
        scroll_down(wd)

        thumbnails = wd.find_elements(By.CLASS_NAME, "MMThumbImage-Image")

        for img in thumbnails[len(image_urls) + skips:max_images]:
            try:
                img.click()
                time.sleep(delay)
            except:
                continue

            images = wd.find_elements(By.CLASS_NAME, "MMImage-Origin")
            for image in images:
                if image.get_attribute('src') in image_urls:
                    max_images += 1
                    skips += 1
                    break

                if image.get_attribute('src') and 'http' in image.get_attribute('src'):
                    image_urls.add(image.get_attribute('src'))
                    print(f"Found {len(image_urls)}")

    return image_urls
```

Fig. 15 Yandex thumbnail and image elements

While the core functionality of scraping images remains the same across all three scripts, the main differences arise in the specifics of each browser's webpage structure. The URLs used for initiating image searches are unique to each browser, reflecting their respective search engines' query formats and parameters. Also, the class names used for identifying thumbnails and full-sized images vary across browsers. Furthermore, differences in WebDriver paths illustrate the need for tailored configurations to interact with each browser. Overall, these differences show the adaptability of web scraping techniques to accommodate the various interfaces and functionalities provided by different search engines.

3.2. Gender Classification in Images

Following the acquisition of the images, a manual classification was done to determine the gender portrayed in each image. The leading factor in the categorization criteria was the gender of the individual depicting the occupation. However, it was discovered that there were photographs in the entire sample where it was impossible to determine the gender of the subject with certainty; as a result, these elements were labelled as 'None'. Images of objects, cartoons, and sketches that did not depict actual people were also included in this category. The category 'Both' was added because of the collection of numerous images that portrayed both genders. In summary, the scraped images were categorized as either Male, Female, None, or Both.

3.3. Analysis of the Results

The information was combined in the Results section after the images were divided into four categories, enabling a deeper examination of gender representation in the executed searches. The discussion that followed focused on the use of the statistical tools described in Figure 1 in order to determine gender biases in the results.

Table 1. Gender-based image classification results by profession and search engine

	Search Engine	Male	Female	None	Both	Total
CEO	Google	31 (50%)	18 (29%)	8 (13%)	5 (8%)	62 (100%)
	Duck Duck Go	35 (73%)	12 (25%)	1 (2%)	0 (0%)	48 (100%)
	Yandex	44 (94%)	2 (4%)	1 (2%)	0 (0%)	47 (100%)
Engineer	Google	23 (40%)	12 (20%)	14 (24%)	10 (17%)	59 (100%)
	Duck Duck Go	29 (60%)	15 (31%)	0 (0%)	4 (8%)	48 (100%)
	Yandex	32 (76%)	2 (5%)	3 (7%)	5 (12%)	42 (100%)
Lawyer	Google	19 (35%)	21 (38%)	11 (20%)	4 (7%)	55 (100%)
	Duck Duck Go	29 (58%)	6 (12%)	14 (28%)	1 (2%)	50 (100%)
	Yandex	33 (77%)	6 (14%)	1 (2%)	3 (7%)	43 (100%)
Manager	Google	22 (42%)	16 (30%)	12 (23%)	3 (6%)	53 (100%)
	Duck Duck Go	28 (61%)	5 (11%)	1 (2%)	12 (26%)	46 (100%)
	Yandex	27 (66%)	10 (24%)	0 (0%)	4 (10%)	41 (100%)
Nurse	Google	1 (2%)	39 (74%)	5 (9%)	8 (15%)	53 (100%)
	Duck Duck Go	7 (14%)	38 (78%)	2 (4%)	2 (4%)	49 (100%)
	Yandex	0 (0%)	45 (90%)	3 (6%)	2 (4%)	50 (100%)
Secretary	Google	1 (2%)	50 (82%)	8 (13%)	2 (3%)	61 (100%)
	Duck Duck Go	0 (0%)	47 (96%)	2 (4%)	0 (0%)	49 (100%)
	Yandex	0 (0%)	49 (100%)	0 (0%)	0 (0%)	49 (100%)
Teacher	Google	10 (18%)	27 (48%)	17 (30%)	2 (4%)	56 (100%)
	Duck Duck Go	7 (14%)	39 (80%)	3 (6%)	0 (0%)	49 (100%)
	Yandex	7 (16%)	34 (77%)	1 (2%)	2 (5%)	44 (100%)
Mathematician	Google	42 (72%)	7 (12%)	4 (7%)	5 (8%)	58 (100%)
	Duck Duck Go	46 (96%)	2 (4%)	0 (0%)	0 (0%)	48 (100%)
	Yandex	41 (84%)	6 (12%)	2 (4%)	0 (0%)	49 (100%)
Scientist	Google	17 (50%)	10 (29%)	2 (6%)	5 (15%)	34 (100%)
	Duck Duck Go	22 (55%)	15 (38%)	1 (3%)	2 (5%)	40 (100%)
	Yandex	24 (52%)	8 (17%)	0 (0%)	14 (30%)	46 (100%)

4. Results

The results were acquired after the manual classification. The final sample size for each search varies from 34 to 62 photographs since certain images with display faults had to be eliminated, or when the algorithm ran across repeated numbers, the sample size was extended. Table 1 displays the image classification for each profession.

Men clearly outnumber women in the occupations of mathematician, scientist, engineer, lawyer, manager, and CEO, according to results from the three search engines used. In contrast, women outnumber men in the professions of nurse, secretary, and teacher. The low percentage of men who are portrayed as secretaries and nurses stands out. The proportional difference between male and female representation was graphically represented when search engines were compared (Figures 16, 17, and 18). It is vital to note that Google is the only platform that predicts gender-balanced results for the management and lawyer professions. One gender clearly dominates each occupation on Yandex and DuckDuckGo, whereas the nurse and secretary professions are significantly unbalanced on all three platforms. The best gender equity in representation among the three engines is seen in the scientific field, while there is still a definite imbalance that favours men. Figure 16, Figure 17, and Figure 18 show the proportional disparities in situations like nurse and secretary, where there were almost no male gender findings at all, and are far bigger than the differences in CEO, engineer, manager, mathematician, or scientist.

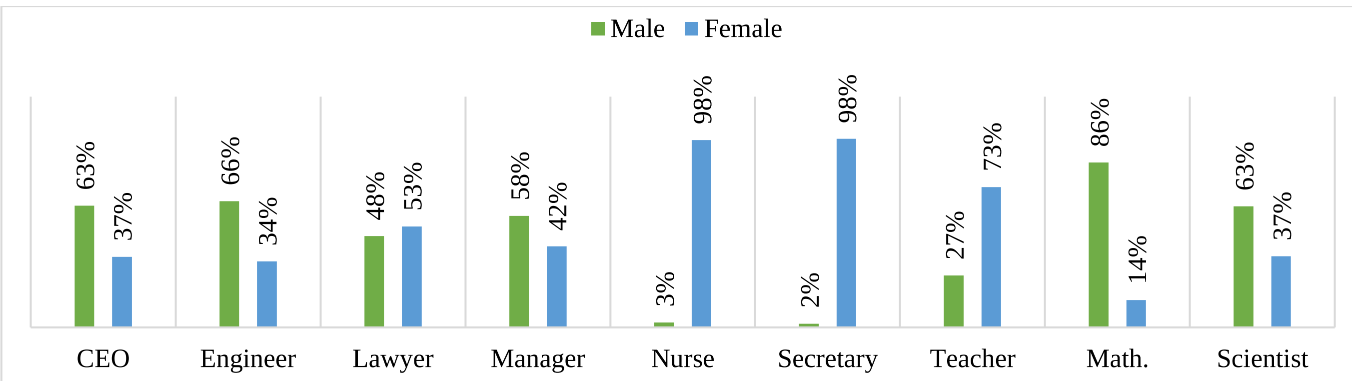

Fig. 16 Gender proportions by profession in Google

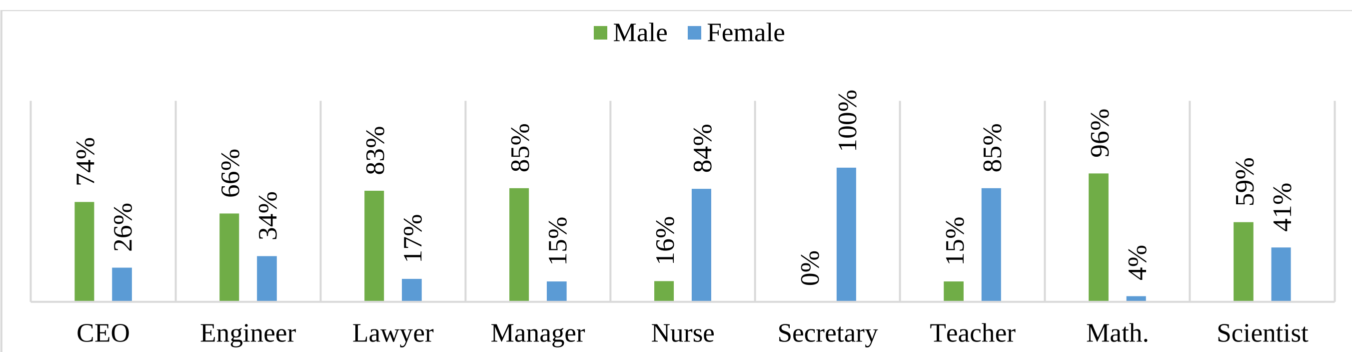

Fig. 17 Gender proportions by profession in DuckDuckGo

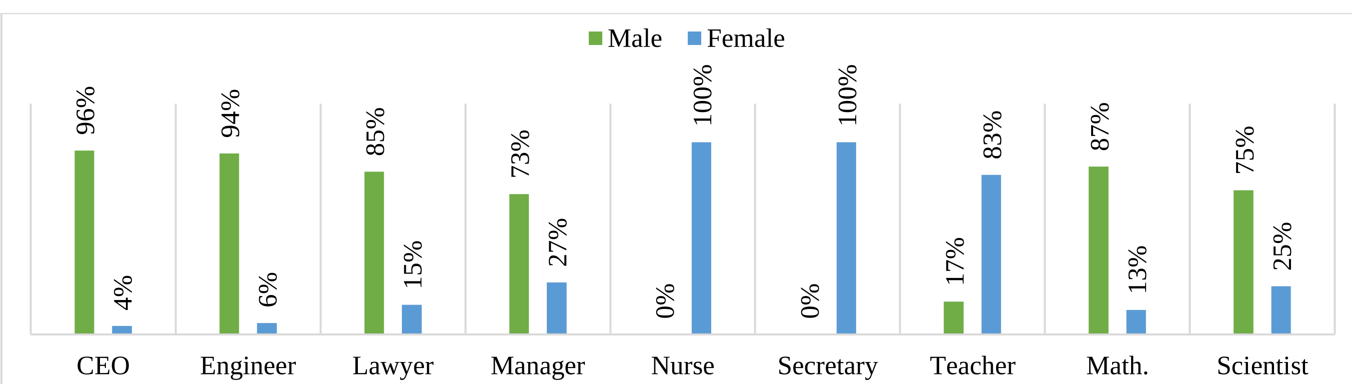

Fig. 18 Gender proportions by profession in Yandex

Table 2. Gender ratio of results: Male vs Female predominance

Profession	Google	DDG	Yandex
CEO	0.72	1.92	21
Engineer	0.92	0.93	15
Lawyer	-0.1	3.83	4.5
Manager	0.38	4.6	1.7
Nurse	-0.97	-0.82	-1
Secretary	-0.98	-1	-1
Teacher	-0.63	-0.82	-0.79
Mathematician	5	22	5.83
Scientist	0.7	0.47	2
Mean Bias	**1.15**	**4.04**	**5.87**

[a)]((For instance, the sample search results for CEOs on Google showed 18 images associated with women and 31 images related to men. In this situation, the ratio is $\frac{31}{18}-1=0.72$. *This indicates that 72% more male than female images were scraped in this search));*

Table 2 was constructed to check for bias in the results. Positive numbers indicate a stronger male representation, while negative values indicate a greater female representation, demonstrating the extent to which male outcomes outweigh female results. In the context of this research, the presence of bias will be proven if the value is more than or equal to 0.6 when taking the results' absolute value into account. Thus, the results show that all searches for all professions and search engines are biased, except for the results for managers and lawyers in the Google search engine and scientists in DuckDuckGo. Additionally, it was determined which search engine has the most bias by calculating the mean value of the occupations' absolute value ratio for each engine. Yandex presents the highest gender bias among the search engines, with a ratio of 5.87. As seen in Table 1, there was only female representation in two of the nine searches made using this engine (nurse and secretary); there were no male images. Yandex also shows a substantial bias towards the male gender in the CEO and engineering professions, showing in proportion 21 (2100%) and 15 (1500%) more people of the male gender than the female gender, respectively.

On the other hand, Google, with a 1.15 larger average representation of one gender than the other, suggests it is the search engine with the least bias. This is the only engine that offers at least one result for each gender in all queries and has the least bias in jobs that are traditionally held by men, such as CEO, engineer, manager, and lawyer (in the last two instances, no bias is apparent).

While it is true that some search engines have more bias than others, it is clear that bias is higher in professions linked with the feminine gender than those associated with the male gender in all circumstances. Transitioning from the analysis focused on comparing the search engines, the subsequent analyses will shift their attention to a comprehensive examination and comparison of the gender results regardless of the search engine from which the data came, starting with an in-depth evaluation of the mean, standard deviation, and coefficient of variation to provide further insights into the data's central tendency, dispersion, and relative variability.

With an average of 36.7 compared to 10.7, the representation of men in the CEO images is noticeably higher than that of women. Furthermore, the coefficient of variation (CV) for women is significantly higher (75.8%) than for men (18.2%), suggesting that there is more variation and inconsistency in how women are represented in this field. A similar pattern is observed in the manager profession, where there is lower and more inconsistent female representation.

Table 3. Mean, standard deviation and coefficient of variation results

Profession	Gender	Google	DuckDuckGo	Yandex	Mean	Standard Deviation	Coefficient of Variation
CEO	Male	31	35	44	36.7	6.7	18.20%
	Female	18	12	2	10.7	8.1	75.80%
Engineer	Male	23	29	32	28	4.6	16.40%
	Female	12	15	2	9.7	6.8	70.40%
Lawyer	Male	19	29	33	27	7.2	26.70%
	Female	21	6	6	11	8.7	78.70%
Manager	Male	22	28	27	25.7	3.2	12.50%
	Female	16	5	10	10.3	5.5	53.30%
Nurse	Male	1	7	0	2.7	3.8	142.00%
	Female	39	38	45	40.7	3.8	9.30%
Secretary	Male	1	0	0	0.3	0.6	173.20%
	Female	50	47	49	48.7	1.5	3.10%
Teacher	Male	10	7	7	8	1.7	21.70%
	Female	27	39	34	33.3	6	18.10%
Mathematician	Male	42	46	41	43	2.6	6.20%
	Female	7	2	6	5	2.6	52.90%
Scientist	Male	17	22	24	21	3.6	17.20%
	Female	10	15	8	11	3.6	32.80%

This variability may reflect lower recognition and visibility of women in business leadership roles, potentially impacting the perception of women's ability to occupy these positions. Similarly, in the engineering profession, the average representation for men is 28.0, while for women, it is 9.7; and the CV for women (70.4%) is more than quadruple than that for men (16.4%), suggesting greater dispersion in search results for female engineers.

The inconsistency in female representation may reflect persistent stereotypes and barriers in the engineering field that hinder the perception of women as professionals in this area. On the other hand, occupations that have historically been held by women, like nursing and secretary, exhibit a high and consistent percentage of female representation. The average representation of female secretaries is 48.7, with a relatively low variance of 3.1%, and the average representation of female nurses is 40.7, with a variation of 9.3%.

These statistics show a strong correlation in the perception between these occupations and women, reinforcing gender norms that restrict diversity in these fields. Even though the teaching profession presents the closest CV for males and females among the other professions (21.7% and 18.1%, respectively), it still shows a much greater mean for female images (33.3) than for male images (8).

Finally, the scientist profession is considered the most balanced among all the searched professions because, although the average representation for men (21.0) is higher than for women (11.0), it is the smallest difference among all the means for both genders. Additionally, this profession has the second smallest difference in CVs after the teaching profession. This indicates similar dispersion in the representation of both genders, although the lower average for women reflects a small bias towards men in this profession. In summary, the results indicate a significant gender bias in search results, with notable inconsistency in the representation of women in traditionally male-dominated professions. The large inconsistency in the representation of women suggests that their visibility in these roles is highly variable, which could be influenced by gender stereotypes and structural barriers that limit their recognition and representation in these areas.

Following the prior analysis, a confidence interval analysis was conducted using a 95% confidence level (Z = 1.96). The primary objective was to assess whether the observed gender proportions significantly differ from 50%. For this purpose, the following hypotheses were established:

- Null Hypothesis (H_0): The observed proportion $\hat{p}$ is not significantly different from 50% (p_0).
- Alternative Hypothesis (H_a): The observed proportion $\hat{p}$ is significantly different from 50% (p_0).

It is important to note that to calculate the observed proportion, the results categorized as "both" and "none" were excluded from the analysis. The focus was placed exclusively on gender results, allowing for a more precise and relevant evaluation of the stated hypotheses. The formulas used to calculate the standard error and confidence intervals, respectively, are as follows:

$$SE_p = \sqrt{\frac{p * (1 - p)}{n}}$$

Where:

- p is the observed ($\hat{p}$) or the expected (p_0) proportion
- n is the sample size

$$CI = p \pm Z * SE_p$$

Where:

- p is the observed ($\hat{p}$) or the expected (p_0) proportion
- Z is the critical value of the standard normal distribution (1.96 in this case for a 95% confidence level)
- SE_p is the standard error for the observed ($\hat{p}$) or the expected (p_0) proportion

In 8 of the 9 professions analyzed, the Confidence Intervals (CI) did not overlap, which implies the rejection of the null hypothesis. This provides sufficient evidence to conclude that the observed proportions differ considerably from 50%. In other words, a confidence interval-based analysis greatly suggests the existence of notable gender biases in most of the professions under investigation. Suppose the CI of a gender's observed proportion is greater than the CI of the expected proportion in situations where the CIs do not overlap. In that case, the gender is considered to be overrepresented. Conversely, gender is considered to be underrepresented if the Confidence Interval (CI) of its observed proportion is lower than the CI of the expected proportion. This type of analysis does not only reveal the presence of gender biases in various professions but also highlights how these biases vary between the professions. By comparing technical professions with those related to caregiving and teaching, a consistent pattern is observed: women tend to be underrepresented in technical and leadership roles, while they are overrepresented in caregiving and teaching professions.

For example, women are underrepresented in the professions of CEO, Engineer, Lawyer, and Manager, with observed proportions of 22.5% (CI: [15.663%, 29.407%]), 25.7% (CI: [17.610%, 33.717%]), 28.9% (CI: [20.622%, 37.273%]), and 28.7% (CI: [20.172%, 37.236%]) respectively; that is, with observed proportions significantly lower than 50%. This underrepresentation in leadership and technical roles can perpetuate gender inequality, limiting professional development opportunities for women and negatively affecting their influence in business and technological decision-making. In contrast, women are overrepresented in the professions of Nurse, Secretary, and

Teacher, with proportions of 93.8% (CI: [89.715%, 97.977%]), 99.3% (CI: [97.991%, 100%]), and 80.6% (CI: [73.691%, 87.599%]) respectively; that is, with observed proportions significantly higher than 50%. This overrepresentation in caregiving and teaching roles reinforces traditional gender stereotypes, where women are thought to be more suitable for caregiving and support jobs. This pattern not only limits professional opportunities for men in these areas but can also influence social perceptions of the capabilities and appropriate roles for each gender.

The only profession in this test that did not exhibit a significant bias was Scientist. Men's confidence intervals (CI: [56.124%, 75.126%]) and women's confidence intervals (CI: [24.874%, 43.876%]) overlapped. This points to a more equitable representation in this field and may represent a step toward gender parity in some scientific domains. Nonetheless, it is important to investigate further to understand the factors that contribute to this balanced representation and how they can be replicated in other professions. To further explore the relation between gender representation and different professions, a logistic regression analysis was performed.

For this analysis, it is essential to use a reference category to avoid multicollinearity issues, ensuring the stability and robustness of the model. In the context of this article, "Scientist" will be used as the reference profession, as previous analyses have shown that it has a more balanced observed gender proportion. Additionally, according to the confidence interval analysis, there is insufficient evidence to conclude that its observed proportion is significantly different from 50%. By using "Scientist" as the reference category in the logistic regression analysis, a clear view of how genders are represented across various professions, compared to a relatively balanced profession, is obtained.

Table 4. Calculation of the confidence interval for the observed proportion

Profession	Gender	Total	Observed proportion ($\hat{p}$)	Standard Error for $\hat{p}$	Confidence Interval for $\hat{p}$	
					Lower Limit (α)	Upper Limit (β)
CEO	Male	110	77.5%	3.506%	70.593%	84.337%
	Female	32	22.5%		15.663%	29.407%
Engineer	Male	84	74.3%	4.109%	66.283%	82.390%
	Female	29	25.7%		17.610%	33.717%
Lawyer	Male	81	71.1%	4.248%	62.727%	79.378%
	Female	33	28.9%		20.622%	37.273%
Manager	Male	77	71.3%	4.353%	62.764%	79.828%
	Female	31	28.7%		20.172%	37.236%
Nurse	Male	8	6.2%	2.108%	2.023%	10.285%
	Female	122	93.8%		89.715%	97.977%
Secretary	Male	1	0.7%	0.678%	0%	2.009%
	Female	146	99.3%		97.991%	100.000%
Teacher	Male	24	19.4%	3.548%	12.401%	26.309%
	Female	100	80.6%		73.691%	87.599%
Mathematician	Male	129	89.6%	2.546%	84.594%	94.573%
	Female	15	10.4%		5.427%	15.406%
Scientist	Male	63	65.6%	4.848%	56.124%	75.126%
	Female	33	34.4%		24.874%	43.876%

[a]*((The lower limit of the Male Secretary CI was originally -0.649%, but since the analysis is based on proportions, it will be considered 0%. Similarly, the upper limit of the Female Secretary CI was originally 100.649%, but since the analysis is based on proportions, it will be considered as 100%));*

Table 5. Calculation of the confidence interval for the expected proportion

Profession	Standard Error for expected proportion (p_0= 50%)	Confidence Interval for p_0=50%	
		Lower Limit (θ)	Upper Limit (λ)
CEO	4.196%	41.776%	58.224%
Engineer	4.704%	40.781%	59.219%
Lawyer	4.683%	40.821%	59.179%
Manager	4.811%	40.570%	59.430%
Nurse	4.385%	41.405%	58.595%
Secretary	4.124%	41.917%	58.083%
Teacher	4.490%	41.199%	58.801%
Mathematician	4.167%	41.833%	58.167%
Scientist	5.103%	39.998%	60.002%

Table 6. Determining overlap of confidence intervals

Profession	Gender	Confidence Overlap ($\alpha \leq \lambda$ and $\theta \leq \beta$?)	Conclusion	Representation
CEO	Male	No overlap	H_0 rejected: $\hat{p}$ is significantly different from p_0	Overrepresentation
	Female	No overlap	H_0 rejected: $\hat{p}$ is significantly different from p_0	Underrepresentation
Engineer	Male	No overlap	H_0 rejected: $\hat{p}$ is significantly different from p_0	Overrepresentation
	Female	No overlap	H_0 rejected: $\hat{p}$ is significantly different from p_0	Underrepresentation
Lawyer	Male	No overlap	H_0 rejected: $\hat{p}$ is significantly different from p_0	Overrepresentation
	Female	No overlap	H_0 rejected: $\hat{p}$ is significantly different from p_0	Underrepresentation
Manager	Male	No overlap	H_0 rejected: $\hat{p}$ is significantly different from p_0	Overrepresentation
	Female	No overlap	H_0 rejected: $\hat{p}$ is significantly different from p_0	Underrepresentation
Nurse	Male	No overlap	H_0 rejected: $\hat{p}$ is significantly different from p_0	Underrepresentation
	Female	No overlap	H_0 rejected: $\hat{p}$ is significantly different from p_0	Overrepresentation
Secretary	Male	No overlap	H_0 rejected: $\hat{p}$ is significantly different from p_0	Underrepresentation
	Female	No overlap	H_0 rejected: $\hat{p}$ is significantly different from p_0	Overrepresentation
Teacher	Male	No overlap	H_0 rejected: $\hat{p}$ is significantly different from p_0	Underrepresentation
	Female	No overlap	H_0 rejected: $\hat{p}$ is significantly different from p_0	Overrepresentation
Mathematician	Male	No overlap	H_0 rejected: $\hat{p}$ is significantly different from p_0	Overrepresentation
	Female	No overlap	H_0 rejected: $\hat{p}$ is significantly different from p_0	Underrepresentation
Scientist	Male	Overlap	H_0 is not rejected: There is not enough evidence to conclude that $\hat{p}$ is significantly different from p_0	-
	Female	Overlap	H_0 is not rejected: There is not enough evidence to conclude that $\hat{p}$ is significantly different from p_0	-

The variables used in the logistic regression model are defined as follows:

- Dependent Variable: Gender
- Independent Variable: Profession

And the following hypotheses are tested:

- Null Hypothesis (H_0): The probability that an image corresponds to a female is not significantly different from the probability for a Scientist.
- Alternative Hypothesis (H_a): The probability that an image corresponds to a woman is significantly different from the probability for a Scientist.

As seen in Table 7, it is necessary to convert the dependent variable into binary form. In this case, "male" is represented by 0 and "female" by 1. This binary coding allows for the logistic regression model to predict the probability of the outcome being 1 (female) for each observation based on the independent variables.

The logistic regression model used in this analysis follows the structure shown in this formula:

$$\log\left(\frac{p}{1-p}\right) = \beta_0 + \beta_1 * X_1 + \beta_2 * X_2 + \cdots + \beta_n * X_n$$

Where:

- p is the probability that the dependent variables equals 1 (female)
- β_0 is the model intercept
- $\beta_1, \beta_2, \dots, \beta_n$ are the coefficients of the independent variables $X_1, X_2, \dots, X_n$

The goal is to estimate the independent variables coefficients, which represent the strength and direction of the relationship between the independent variables (professions) and the likelihood of the dependent variable (genders) being 1 (female).

Python was employed to determine the coefficients, standard errors, p-values, and confidence intervals for the model. The script was the following:

The code begins by importing the necessary libraries for the analysis:

- Pandas: A Python library used for data manipulation and analysis.
- Statsmodel: A Python library used for estimating statistical models.

```
import pandas as pd
import statsmodels.api as sm
```

Fig. 19 Libraries import

```
data = {'Profession': ['CEO', 'CEO', 'Engineer', 'Engineer', 'Lawyer',
        'Lawyer', 'Manager', 'Manager', 'Nurse', 'Nurse', 'Secretary',
        'Secretary', 'Teacher', 'Teacher', 'Mathematician', 'Mathematician',
        'Scientist', 'Scientist'], 'Gender': [0, 1, 0, 1, 0, 1, 0, 1, 0, 1,
        0, 1, 0, 1, 0, 1, 0, 1], 'Total': [110, 32, 84, 29, 81, 33, 77, 31,
        8, 122, 1, 146, 24, 100, 129, 15, 63, 33]}

df = pd.DataFrame(data)
```

Fig. 20 Creatin of dataframe

```
df['Profession'] = df['Profession'].astype('category')
df['Profession'] = df['Profession'].cat.reorder_categories(
    ['Scientist', 'CEO', 'Engineer', 'Lawyer', 'Manager', 'Nurse',
     'Secretary', 'Teacher', 'Mathematician'], ordered=True)
```

Fig. 21 Converting string variables to categorical and reordering

Next, a dataframe is created, which is a data structure in pandas that stores data in a table format. In this case, the dataframe consists of the obtained results, with 0 being male and 1 female.

The professions are then converted from a string (text) variable to categorical variables (which represent categories or groups). Subsequently, the categories are reordered to place "Scientist" first, as it will be used as the reference category.

Afterwards, the first category ("Scientist") is removed because it is the reference profession, and the remaining categories are converted into dummy columns (0 and 1), forming a new dataframe. Then the new dataframe is expanded so that each row represents an image. Thus, the "Total" row is removed as it is no longer relevant.

```
df = pd.get_dummies(df, columns=['Profession'], drop_first=True)
df = df.loc[df.index.repeat(df['Total'])].reset_index(drop=True)
df = df.drop(columns=['Total'])
```

Fig. 22 Removing the reference category and expanding the dataframe

```
for col in df.select_dtypes(['bool']).columns:
    df[col] = df[col].astype(int)

print(df.dtypes)
print(df.isnull().sum())

df = df.dropna()
```

Fig. 23 Converting boolean to integers and eliminating null values

The boolean columns are converted into integers to ensure compatibility with Python functions and models (such as statsmodel). Also, the null values of each column are identified and removed because regression models cannot handle missing data properly. Although the dataset should not contain null values, it is good practice to ensure data integrity before any statistical analysis.

The dependent variable (gender) and the independent variable (professions) are defined, and the constant for logistic regression is added. In this case, the independent variable is defined by simply removing the gender column, leaving the professions. Finally, the logistic regression model is specified and fitted to the data. After running the script, the following results were obtained.

```
y = df['Gender']
X = df.drop(columns=['Gender'])
X = sm.add_constant(X)

model = sm.Logit(y, X)
result = model.fit()

print(result.summary())
```

Fig. 24 Defining independent and dependent variables and the model

```
                           Logit Regression Results
==============================================================================
Dep. Variable:                 Gender   No. Observations:                 1118
Model:                          Logit   Df Residuals:                     1109
Method:                           MLE   Df Model:                            8
Date:                Tue, 23 Jul 2024   Pseudo R-squ.:                  0.3797
Time:                        20:30:08   Log-Likelihood:                -480.32
converged:                       True   LL-Null:                       -774.36
Covariance Type:            nonrobust   LLR p-value:                 8.531e-122
```

Fig. 25 Model summary

Table 7. Logistic regression analysis results

Profession	Gender	Gender (Coded)	Total	Coefficient	Standard Error	Z	P-value	CI (95%)	Conclusion
CEO	Male	0	110	-0.5881	0.294	-1.999	0.046	[-1.165, -0.012]	The null hypothesis is rejected
	Female	1	32						
Engineer	Male	0	84	-0.4169	0.304	-1.37	0.171	[-1.013, 0.179]	The null hypothesis is not rejected
	Female	1	29						
Lawyer	Male	0	81	-0.2513	0.298	-0.843	0.399	[-0.835, 0.333]	The null hypothesis is not rejected
	Female	1	33						
Manager	Male	0	77	-0.2632	0.302	-0.87	0.384	[-0.856, 0.329]	The null hypothesis is not rejected
	Female	1	31						
Nurse	Male	0	8	3.3712	0.424	7.96	0.000	[2.541, 4.201]	The null hypothesis is rejected
	Female	1	122						
Secretary	Male	0	1	5.6302	1.026	5.487	0.000	[3.619, 7.641]	The null hypothesis is rejected
	Female	1	146						
Teacher	Male	0	24	2.0737	0.313	6.63	0.000	[1.461, 2.687]	The null hypothesis is rejected
	Female	1	100						
Mathematician	Male	0	129	-1.5051	0.347	-4.334	0.000	[-2.186, -0.825]	The null hypothesis is rejected
	Female	1	15						
Scientist	Male	0	63	Profession of reference					
	Female	1	33						

The findings of the logistic regression analysis show that the professions of CEOs and Mathematicians have negative coefficients, correspondingly negative confidence intervals, and p-values less than 0.05. This suggests that, in comparison to Scientist, the images linked to these occupations are much less likely to feature a woman.

Thus, this implies an underrepresentation of women in technical fields and high-level management roles, which may indicate gender barriers in leadership positions and technical areas. The Mathematician case should be highlighted because it showed a p-value that was extremely close to 0, which indicates very strong evidence against the null hypothesis and highlights the significant underrepresentation of women in this field.

On the other hand, the professions of Nurse, Secretary, and Teacher obtained a positive coefficient, and their respective confidence intervals were also positive, with a p-value of less than 0.05. This suggests that, in comparison to Scientists, the images linked to these occupations are much more likely to portray a woman. Hence, this implies that women are overrepresented in these fields, which is consistent with gender stereotypes relating women to support, caregiving, and educational roles.

Furthermore, it is observed that the absolute coefficients in these typically female-dominated professions are of greater magnitude compared to other professions. It is also important to note that these professions had a p-value extremely close to 0, which presents very strong evidence against the null hypothesis. This indicates that the observed differences in gender representation are highly significant and that the images associated with these professions have an extremely high probability of corresponding to a woman compared to a Scientist.

Some professions did not show significant differences in gender representation compared to Scientists, as was the case with Engineers, Lawyers, and Managers. P-values for these professions were higher than 0.05, suggesting insufficient evidence to claim that the probability of an image representing a woman differs from that of a scientist. Therefore, based on this logistic regression analysis, these professions are the least biased against women, suggesting either a more balanced representation of genders or a less pronounced gender bias than in other professions.

The final statistical analysis performed in this study is a chi-squared test, with the aim of evaluating the association between professions and the gender represented in the images. This analysis helps determine whether there is a statistically significant relationship between these two categorical variables. The hypotheses tested are as follows:

- Null Hypothesis (H_0): There is no significant association between profession and gender.
- Alternative Hypothesis (H_a): There is a significant association between profession and gender.

The analysis begins by constructing the contingency table with the observed frequencies:

Then, the expected frequencies are calculated for each case using the following formula:

$$Expected\ frequency = \frac{Row\ Total * Column\ Total}{Grand\ Total}$$

Finally, the Chi-Square statistic is calculated according to the following formula:

$$\sum \frac{(Observed\ frequency - Expected\ frequency)^2}{Expected\ frequency}$$

Table 8. Contingency table with observed frequencies

Professions	Male	Female	Total
CEO	110	32	142
Engineer	84	29	113
Lawyer	81	33	114
Manager	77	31	108
Nurse	8	122	130
Secretary	1	146	147
Teacher	24	100	124
Mathematician	129	15	144
Scientist	63	33	96
Total	577	541	1118

Table 9. Expected frequencies

Professions	Male	Female	Total
CEO	73.29	68.71	142
Engineer	58.32	54.68	113
Lawyer	58.84	55.16	114
Manager	55.74	52.26	108
Nurse	67.09	62.91	130
Secretary	75.87	71.13	147
Teacher	64.00	60.00	124
Mathematician	74.32	69.68	144
Scientist	49.55	46.45	96
Total	577	541	1118

Table 10. Chi-Square statistic calculation

Professions	Male	Female	Total
CEO	18.39	19.62	38.01
Engineer	11.31	12.06	23.37
Lawyer	8.35	8.91	17.26
Manager	8.11	8.65	16.76
Nurse	52.05	55.51	107.56
Secretary	73.88	78.80	152.68
Teacher	25.00	26.66	51.66
Mathematician	40.23	42.91	83.14
Scientist	3.65	3.90	7.55
Total	240.97	257.01	**497.98**

The p-value for a Chi-Square statistic of 497.98 with 8 degrees of freedom is $1.911 * 10^{-102}$, which is an extremely low value. Since the p-value is significantly lower than the considered significance level of 0.05, the null hypothesis is rejected, suggesting that there is a significant association between professions and gender. Upon examining the observed frequencies table and based on the previous analysis, it can be concluded that the reason for this extremely low p-value is the strong association between search engine images and traditionally female professions, such as Nurse, Secretary, and Teacher, while men are more prominently associated with leadership positions and the majority of STEM professions. This finding reinforces the presence of gender biases in the representation of the analyzed professions.

It is important to note that this Chi-square test examined all professions globally, as opposed to examining each one separately. This approach was chosen because previous statistical analyses had already evaluated the professions and the search engines separately. By consolidating the data, the Chi-square test provides a general and holistic perspective on the association between gender and profession, highlighting the overall pattern of gender representation across all professions and underscoring the nature of gender biases in professional representations.

5. Discussion

The majority of the searches that were conducted showed gender bias, supporting the points made in the literature. As noted in the introduction chapter, it might be harmful to women if search engine results have a significant impact on their career decisions. This is especially concerning for occupations with a predominately male workforce because they tend to offer higher remuneration in Latin America. For instance, civil engineers in the region earn an average of 1490 USD monthly [26], lawyers earn an estimated 2500 USD [27], and Managers, particularly those in finance and businesses with revenues over 100M USD, earn between $4400 and 5891 USD per month [28]. Coincidentally, all these professions have a higher male representation in the studied search engines. Conversely, professions with predominantly female representation tend to receive lower salaries in the region. For example, nurses earn an average salary of 745 USD [29], and university teachers earn an average of 850 USD [30]. These biases in search engines have ethical implications if they show a greater representation of men in professions with greater attractiveness or social value and, therefore, an undervaluation of predominantly female jobs. This is especially problematic because developing a critical view of search engine results is often challenging [31]. There is a chance that results from algorithms like those frequently used in search engines may reflect pre-existing social biases, promoting and maintaining gender-based social and economic inequalities. According to this viewpoint, ML decision tools may mirror preexisting social biases in algorithmic results, which in turn influence human decisions that perpetuate inequality, creating an inequality loop.

The information above is consistent with that of Vlascenau and Amodio [32], who experimentally showed that people who are exposed to biased results may have cognitive distortions, which can affect important decisions like hiring. The authors also claim that these cognitive changes can lead to judgments that are prejudiced against of those who do not fit the mold of typical members of a social group. Similar to this, Wijnhoven and Van Haren's research from 2021 [33] supports the results mentioned earlier. They found that job-related search engine inquiries for women and men, respectively, exhibited biases toward stereotypically feminine and male jobs. The existence of algorithmic biases is undisputed regardless of variations in how algorithms are implemented and how much they mirror societal biases and have an impact on users. Additionally, their influence on people, even in subtle ways, can have a real effect. The gender norms that still exist in today's professions and occupations are perpetuated in search engine results, which might affect how women view themselves in the workplace and make them doubt their skills [34].

In addition, how women are treated in fields with a predominance of men in the workforce and academia can be influenced by gender norms. For instance, according to previous study the female participation in fields related to science and engineering is still low, and the few women who choose these careers display different behaviors from men, such as higher expectations for their academic performance and unfair treatment from professors who often underestimate them. Seneviratne [35] agrees with Ruiz Ruiz et al.'s findings and claims that gender roles related to professions are mostly noticeable in STEM careers, particularly in spatially and mathematically intensive fields, which still exhibit significantly higher male representation than female representation. Furthermore, compared to professions with a stronger male representation, STEM fields with underrepresented women tend to pay higher earnings. As these biases may come from a variety of reasons, it is crucial to point out that this study does not suggest that search engines are solely accountable for the results provided. However, individuals should be conscious of these biases and any potential hazards they may entail [36]. When it comes to the function of science in social contexts, Walls [37] emphasizes that engineers have a problem with neutrality, which involves attaching moral weight to the uses of tools created by engineers rather than to their work. Professionals disassociate themselves in this way from the societal ramifications of the technologies they produce. This was evident with search engines; ideally, engineers would be aware of how they are now contributing to negative preconceptions and prioritize fostering a just society. Moreover, this study also shows how engineering methods can be used to expose social inequities and help create a more conscious and just society. Furthermore, the sexual objectification of female imagery in the nursing and secretary professions is an additional noteworthy discovery. A significant number of images for these two categories during the manual classification of the data were sexist in nature, with women being reduced to their bodies or individual body parts in these representations, which is not the case for occupations that are associated with men. This portrayal amply illustrates societal divides about the value placed on occupations connected with women, as well as how this is represented in frequently used search engines.

6. Conclusion

The primary goal of the research, which was to examine gender biases in machine learning algorithms, notably in search engines, was successfully accomplished. The findings demonstrated the existence of biases that uphold gender norms in numerous professions, suggesting that in contemporary culture, some occupations or career paths are still strongly linked with a particular gender. The sexualization of women in the secretary and nursing professions, which showed that gender roles and prejudices present in society might be recreated in ML algorithms, was another major discovery.

Furthermore, the existence of a real bias that views jobs like secretary or nurse as representations of a sexualized female nature severely undermines the desire of men to pursue these careers constricts their options, and reinforces gender roles that categorize various professions to a single gender. This study emphasizes that, as ML algorithms are increasingly incorporated into our daily lives, tackling pervasive problems like occupational gender segregation requires addressing identified biases and delivering more equitable results that do not negatively impact historically marginalized groups. In conclusion, eliminating the objectification of women in some professions and aiming for a future of greater accessibility for people of the female gender in leadership roles, STEM careers or job positions currently dominated by people of the male gender require a comprehensive change in the environment to which marginalized groups are exposed to.

To provide an equitable and just environment for future generations, it is crucial to raise awareness of the ethical and social implications that must be considered when creating new programming algorithms. These findings should be taken into account alongside their main limitations. The analysis was based on a set of images representing 9 professions, and expanding the dataset to cover a broader range of professions and incorporating images from different search engines would enhance the understanding of gender bias in machine learning algorithms. Additionally, while this study classified images manually, future research could explore automating this process to improve efficiency and consistency, enabling the analysis of larger datasets while minimizing potential human bias.

Funding Statement

The research was conducted without any external funding or financial support.

Acknowledgements

The Industrial Engineering Faculty at the University of Lima has been extremely helpful throughout the project, and the authors would like to express their profound gratitude for that.

References

[1] CSU Global, Colorado State University, 2023. [Online]. Available: https://bit.ly/3P4yRvy

[2] What is machine learning (ML)?, IBM, 2022. [Online]. Available: https://bit.ly/3M4EVCY

[3] Katherine J. Igoe, Algorithmic Bias in Health Care Exacerbates Social Inequities — How to Prevent It, Harvard T.H. Chan. [Online]. Available: https://bit.ly/3ph4tng

[4] Rachel Metz, CNN Business, 2019. [Online]. Available: https://bit.ly/42tRNHC

[5] Genevieve Smith, and Ishita Rustagi, Standford Social Innovation Review, 2021. [Online]. Available: https://bit.ly/41gUknX

[6] Tal Feldman, and Ashley Peake, "End-To-End Bias Mitigation: Removing Gender Bias in Deep Learning," *arXiv*, pp. 1-9, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[7] Susan Leavy, "Gender Bias in Artificial Intelligence: The Need for Diversity and Gender Theory in Machine Learning," *Proceedings of the 1st International Workshop on Gender Equality in Software Engineering*, Gothenburg Sweden, pp. 14-16, 2018. [CrossRef] [Google Scholar] [Publisher Link]

[8] Mike Thelwall, "Gender Bias in Sentiment Analysis," *Online Information Review*, vol. 42, no. 1, pp. 45-57, 2018. [CrossRef] [Google Scholar] [Publisher Link]

[9] L. Elisa Celis, and Vijay Keswani, "Implicit Diversity in Image Summarization," *Proceedings of the ACM on Human-Computer Interaction*, vol. 4, pp. 1-28, 2020. [CrossRef] [Google Scholar] [Publisher Link]

[10] Gregory Smyth, The Importance of Search Engines, InetAsia. [Online]. Available: https://www.inetasia.com/resources/articles-the-importance-of-search-engines.html

[11] Vishnu Vandana Kolisetty, and Dharmendra Singh Rajput, "A Review on the Significance of Machine Learning for Data Analysis in Big Data," *Jordanian Journal of Computers and Information Technology*, vol. 6, no. 6, pp. 41-57, 2020. [CrossRef] [Google Scholar] [Publisher Link]

[12] Brianhu Zhang, Blake Lemoine, and Margaret Mitchell, "Mitigating Unwanted Biases with Adversarial Learning," *Proceedings of the 2018 AAAI/ACM Conference on AI, Ethics, and Society*, New Orleans LA USA, pp. 335-340, 2018. [CrossRef] [Google Scholar] [Publisher Link]

[13] Ashlyn S. Pothen, *Artificial Intelligence and its Increasing Importance*, Learning Outcomes of Classroom Research, L' Ordine Nuovo Publication, pp. 74-81, 2021. [Google Scholar]

[14] Salem Alelyani, "Detection and Evaluation of Machine Learning Bias," *Applied Sciences*, vol. 11, no. 14, pp. 1-17, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[15] Ayanna Howard, and Jason Borenstein, "The Ugly Truth about Ourselves and Our Robot Creations: The Problem of Bias and Social Inequity," *Science and Engineering Ethics*, vol. 24, pp. 1521-1536, 2018. [CrossRef] [Google Scholar] [Publisher Link]

[16] Brandon M. Booth et al., "Bias and Fairness in Multimodal Machine Learning: A Case Study of Automated Video Interviews," *Proceedings of the 2021 International Conference on Multimodal Interaction*, Montréal QC Canada, pp. 268-277, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[17] Vivek K. Singh et al., "Female Librarians and Male Computer Programmers? Gender Bias in Occupational Images on Digital Media Platforms," *Journal of the Association for Information Science and Technology*, vol. 71, no. 11, pp. 1281-1294, 2020. [CrossRef] [Google Scholar] [Publisher Link]

[18] Jahna C. Otterbacher et al., "Investigating User Perception of Gender Bias in Image Search: The Role of Sexism," *The 41st International ACM SIGIR Conference on Research & Development in Information Retrieval*, Ann Arbor MI USA, pp. 933-936, 2018. [CrossRef] [Google Scholar] [Publisher Link]

[19] Mykola Makhortykh, Aleksandra Urman, and Roberto Ulloa, "Detecting Race and Gender Bias in Visual Representation of AI on Web Search Engines," *Advances in Bias and Fairness in Information Retrieval: Second International Workshop on Algorithmic Bias in Search and Recommendation*, Lucca, Italy, pp. 36-50, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[20] Rachael Tatman, Making Noise and Hearing Things. [Online]. Available: https://bit.ly/3nHjD4P

[21] Natalia Reich-Stiebert, and Friederike Eyssel, "(Ir)Relevance of Gender? on the Influence of Gender Stereotypes on Learning with a Robot," *2017 12th ACM/IEEE International Conference on Human-Robot Interaction*, Vienna, Austria, pp. 166-176, 2017. [CrossRef] [Google Scholar] [Publisher Link]

[22] Jeffrey Dastin, Insight - Amazon Scraps Secret AI Recruiting Tool that Showed Bias against Women, Reuters, 2018. [Online]. Available: https://bit.ly/3VElXWW

[23] Sunny Shrestha, and Sanchari Das, "Exploring Gender Biases in ML and AI Academic Research through Systematic Literature Review," *Frontiers in Artificial Intelligence*, vol. 5, pp. 1-17, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[24] Miren Gutierrez, "New Feminist Studies in Audiovisual Industries| Algorithmic Gender Bias and Audiovisual Data: A Research Agenda," *International Journal of Communication*, vol. 15, pp. 439-461, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[25] Bo Zhao, "Web Scraping," *Encyclopedia of Big Data*, pp. 1-3, 2017. [CrossRef] [Google Scholar] [Publisher Link]

[26] AAU, How Much Does a Civil Engineer Earn in Latin America 2024?, American Andragogy University, 2024. [Online]. Available: https://bit.ly/3Mt2jJn

[27] La Republica, 2023. [Online]. Available: https://bit.ly/3BNXIfE

[28] Sebastian Osorio Idárraga, How Much Will Managers of Large Companies in LatAm Earn in 2022?, Bloomberg Linea, 2022. [Online]. Available: https://bit.ly/3WsYpVs

[29] Fede Sarmiento, 2000 Carreras, 2021. [Online]. Available: https://bit.ly/3BL6w6g

[30] Find Out the Minimum Salary for Teachers in Latin America, Instituto de Altos Estudios de Derecho, 2021. [Online]. Available: https://bit.ly/3MKqMuO

[31] Andrew J. Flanagin et al., "Mitigating Risk in Ecommerce Transactions: Perceptions of Information Credibility and the Role of User-Generated Ratings in Product Quality and Purchase Intention," *Electronic Commerce Research*, vol. 14, pp. 1-23, 2014. [CrossRef] [Google Scholar] [Publisher Link]

[32] Madalina Vlasceanu, and David M. Amodio, "Propagation of Societal Gender Inequality by Internet Search Algorithms," *Psychological and Cognitive Sciences*, vol. 119, no. 29, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[33] Fons Wijnhoven, and Jeanna van Haren, "Search Engine Gender Bias," *Frontiers in Big Data*, vol. 4, pp. 1-12, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[34] Dina Gerdeman, Working Knowledge, Harvard Business School, 2019. [Online]. Available: https://bit.ly/3B57xWa

[35] Prathi Seneviratne, Are Women Reaching Parity with Men in STEM?, EconoFact, 2022. [Online]. Available: https://bit.ly/3B00SwC

[36] Robert Epstein et al., “Suppressing the Search Engine Manipulation Effect (SEME),” *Proceedings of the ACM on Human-Computer Interaction*, vol. 1, pp. 1-22, 2017. [CrossRef] [Google Scholar] [Publisher Link]

[37] Dan Walls, “The Responsibility of Engineers: Decouple Engineering and Oppression,” *International Journal of Engineering, Social Justice and Peace*, vol. 8, no. 2, pp. 86-116, 2021. [CrossRef] [Google Scholar] [Publisher Link]