APPLIED DATA SCIENCE PROJECT

31.01.2022

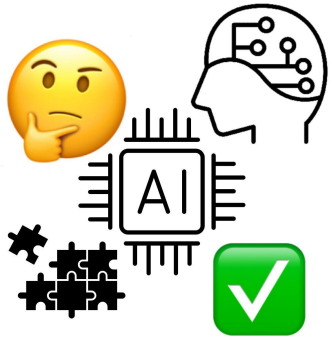
Politecnico di Torino

inda

Countless Applications

Online & Paper format

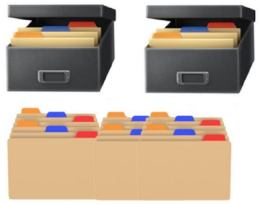
Spending a lot of time on repetitive, routine tasks



INDA provides AI solutions for recruitment and talent acquisition

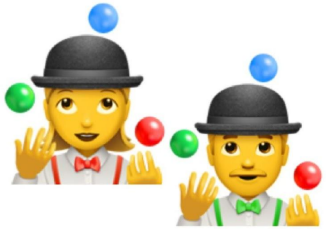
PROJECT AND GOALS

Job Advertisements



Enhanced Job Advertisements

&



Messy and inaccurate skill sets

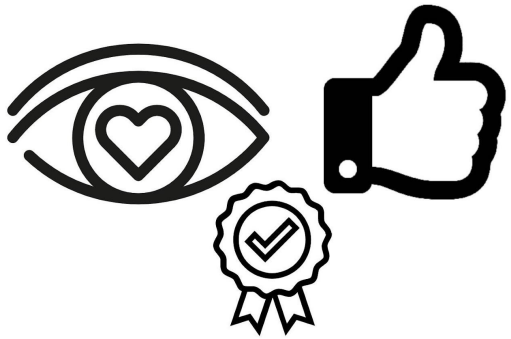
Correct skill sets

miro

PROJECT AND GOALS

Our project is in line with the 8th United Sustainable Development Goal which is Decent Work and Economic Growth.

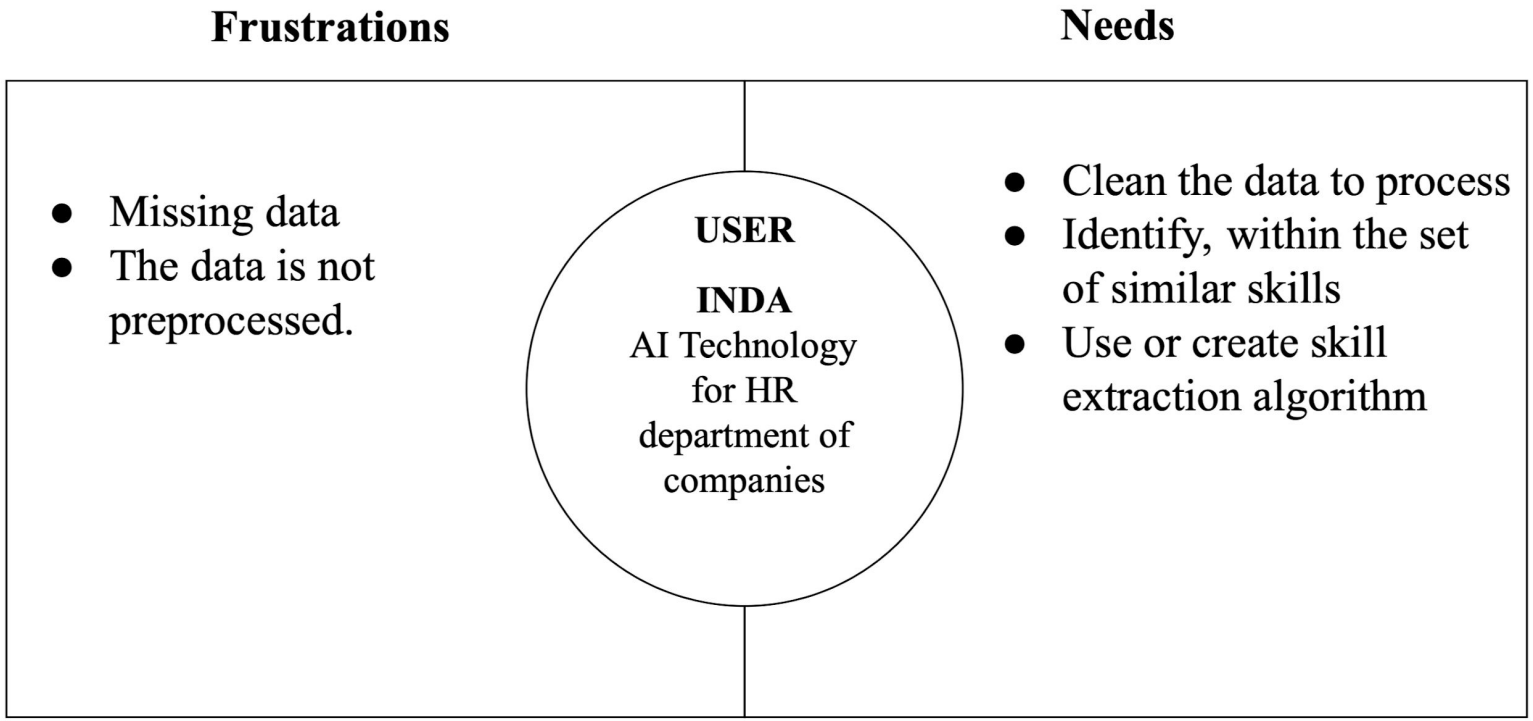
What is the attractive skill?



miro

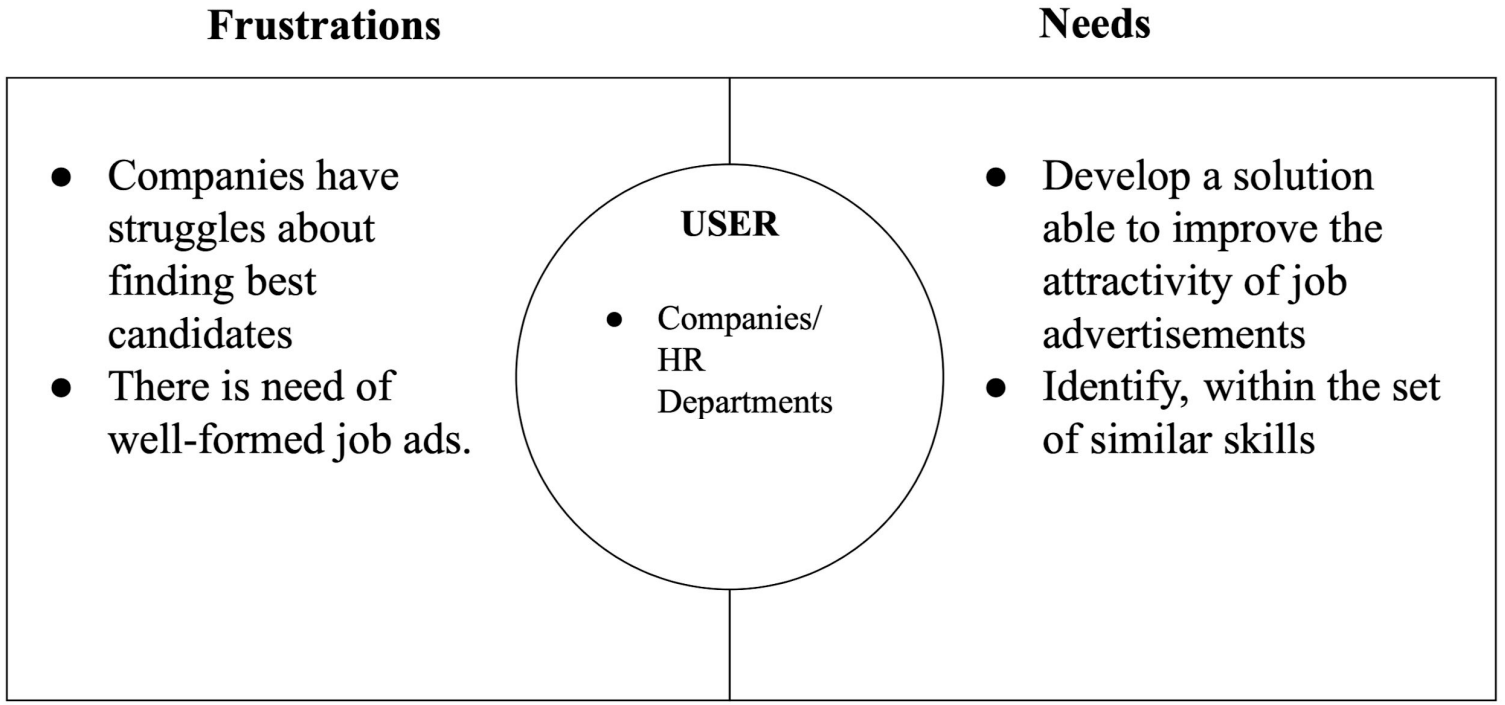
Persona Canvas

Our first target group is Inda Company. We aim to complete our task to develop a solution and improve our results.



Persona Canvas

Our project owner's target group is Human Resources departments of companies.



TOOLS

- ✔ Asana creates organization for team projects
- ✔ Google Collaboratory is a platform to develop algorithms
- ✔ Miro is a website to create flowcharts and diagrams

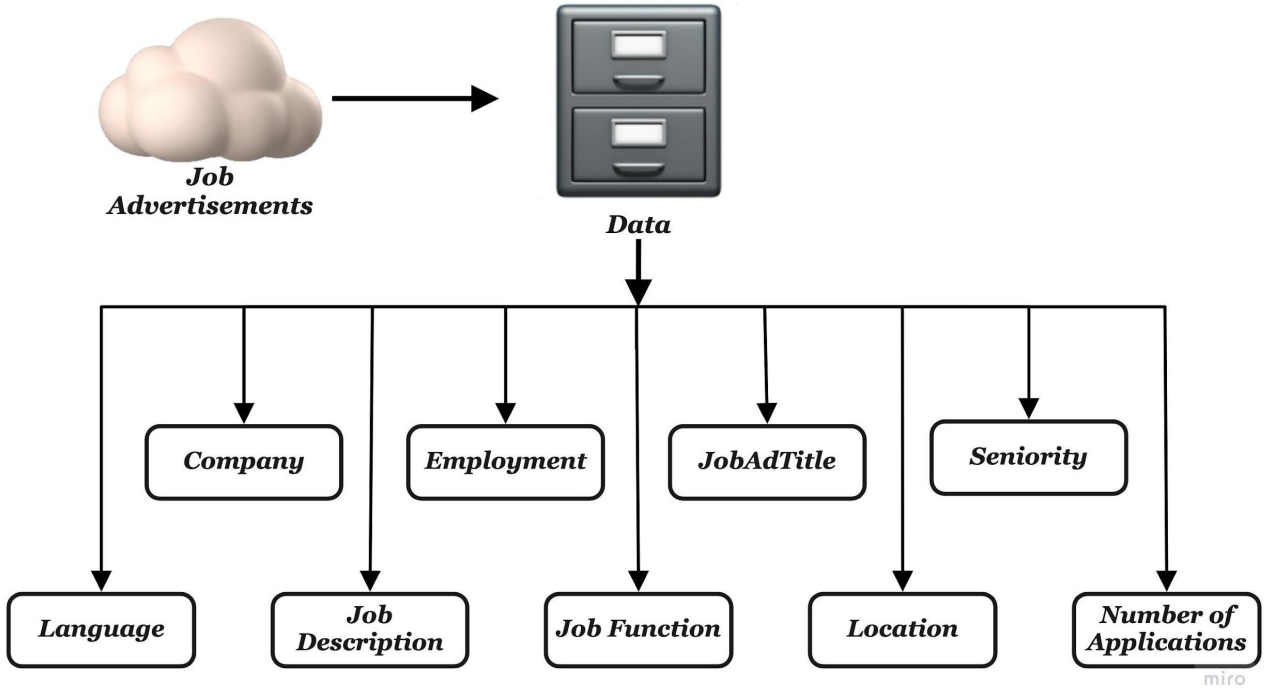
GANTT CHART

Date	25.10.21	01.11.21	08.11.21	17.11.21	24.11.21	01.12.21	08.12.21	15.12.21	22.12.21	29.12.21	05.01.22	12.01.22
Weeks	W1	W2	W3	W4	W5	W6	W7	W8	W9	W10	W11	W12
Part 1. Preprocessing												
Project Management												
Data Collection												
Data Preprocessing												
Skill Extractor												
First Checkpoint												
Part 2. Statistical Overview												
Create New Features												
Attractivty Metric (0.1)												
Statistical Overview												
Second Checkpoint												
Part 3. Complete the Algorithm												
Improve Attractivity Metric												
Calculate Similarity Metric												
Algorithm Test												
Third Check Point												

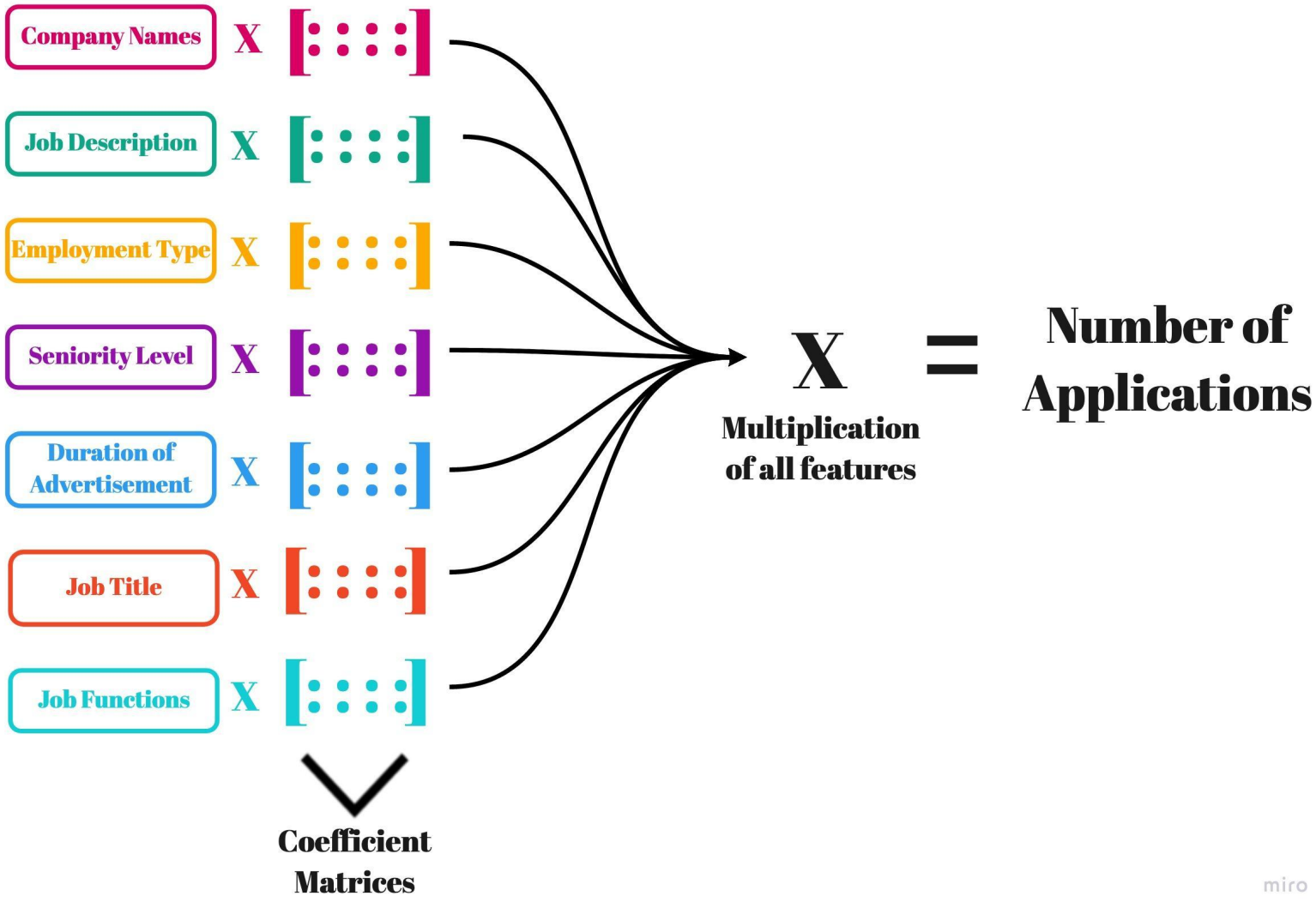
DATA STRUCTURE

- ✔ *Recent Job Advertisement Data* published on online platforms.
- ✔ *30 000 Job Advertisements*

Data Analysis Diagram

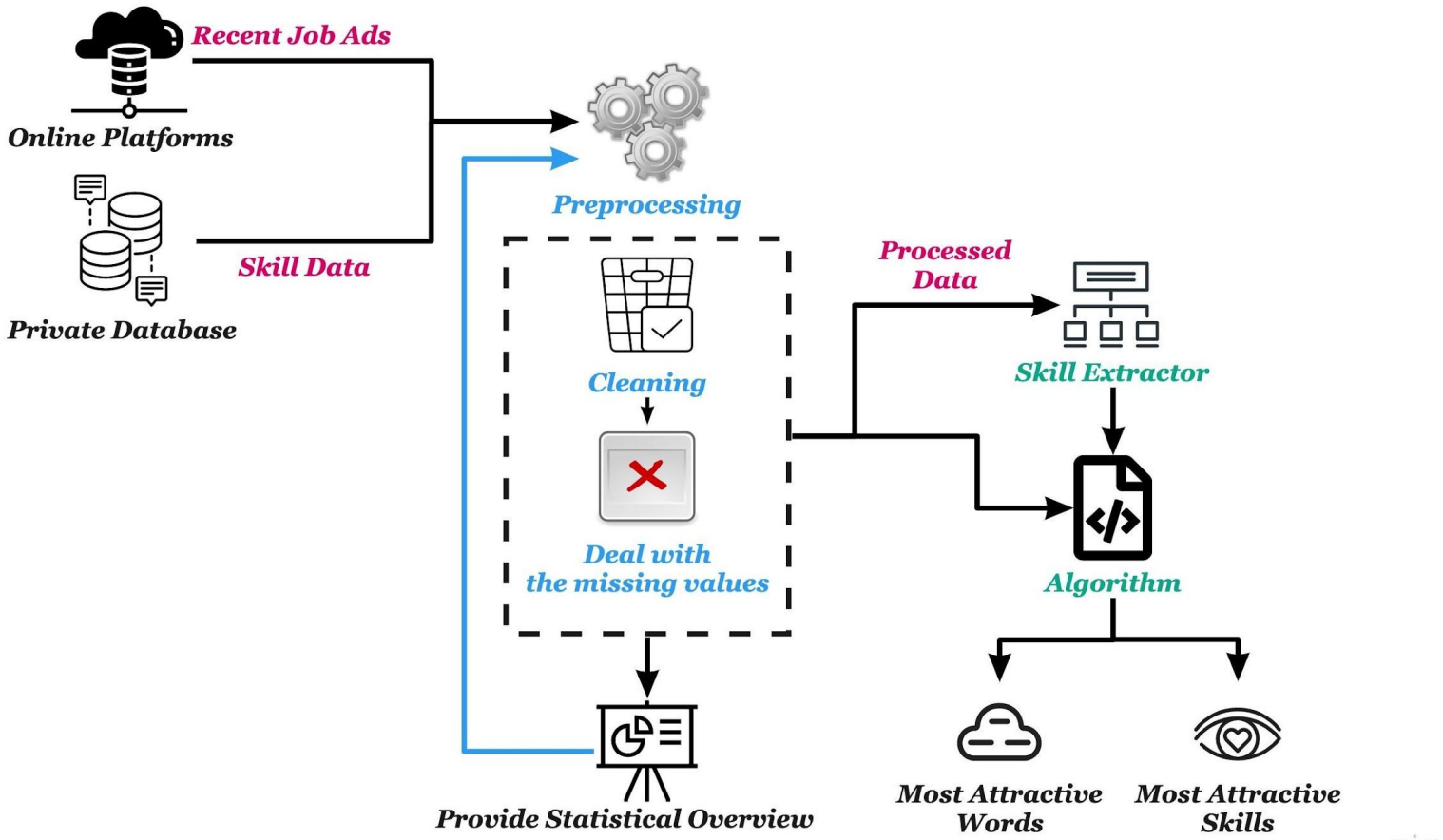


Prediction of Attractiveness Calculation Before Pre-processing



Attractiveness Calculation After Pre-processing

Functional Diagram



Attractivity Metric

$$\text{Attractivity} = \text{Number of applications} * \frac{1}{\text{Duration}} * \left\{ \frac{\text{Frequency of Employment Type}}{\text{Number of application per Employment Type}} \right\}_{\text{Normalized}} * \dots$$

MULTIPLICATION GOES FOR EVERY FEATURE

miro

Attractivity Metric

$$\text{Attractivity} = \text{Number of applications} * \frac{1}{\text{Duration}} * \left\{ \frac{\text{Frequency of Employment Type}}{\text{Number of application per Employment Type}} \right\} \text{Normalized} * \dots$$

NUMBER OF DAYS THE JOB AD STAYS OPEN ON A PLATFORM

MULTIPLICATION GOES FOR EVERY FEATURE

miro

Attractivity Metric

$$\text{Attractivity} = \text{Number of applications} * \frac{1}{\text{Duration}} * \left\{ \frac{\text{Frequency of Employment Type}}{\text{Number of application per Employment Type}} \right\}_{\text{Normalized}} * \dots$$

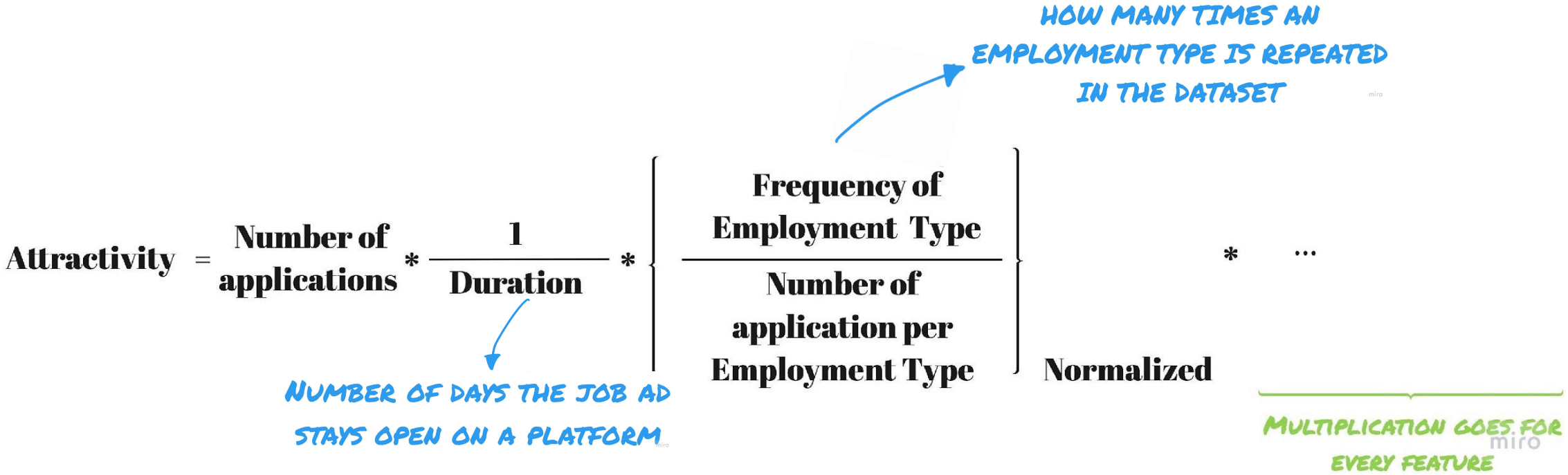
Number of days the job ad stays open on a platform

For example, 14 days

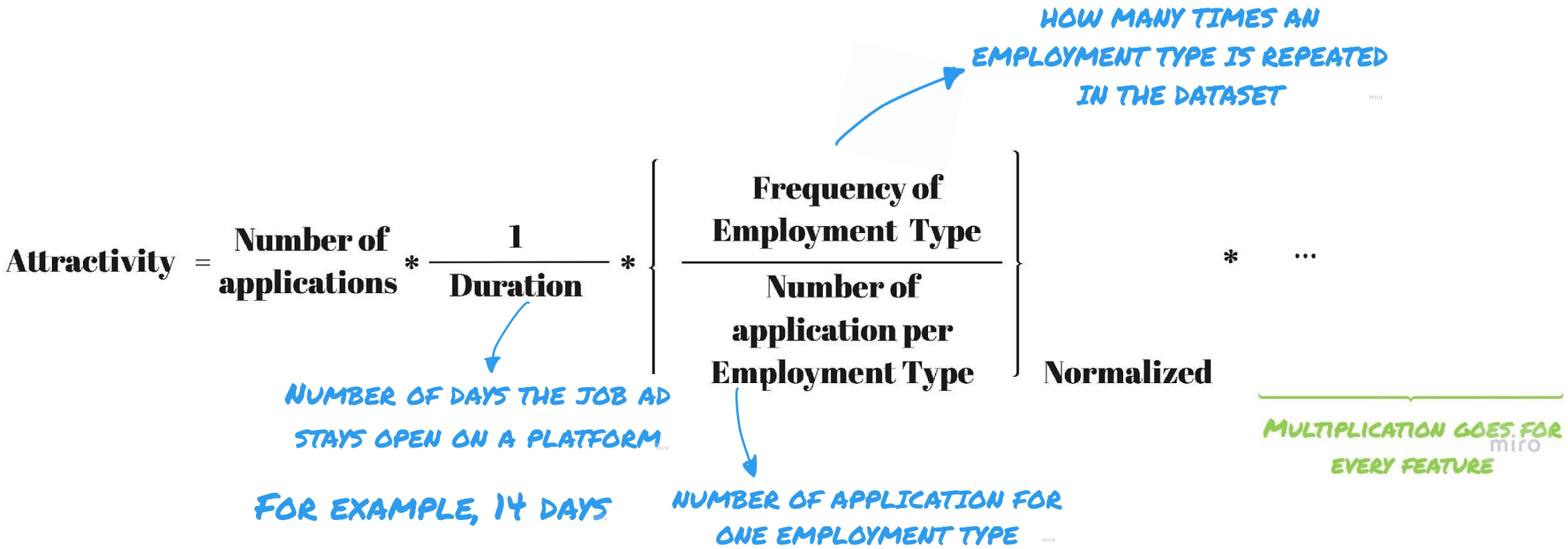
Multiplication goes for every feature

miro

Attractivity Metric



Attractivity Metric



Attractivity Metric

$$\text{Attractivity} = \text{Number of applications} * \frac{1}{\text{Duration}} * \left\{ \frac{\text{Frequency of Employment Type}}{\text{Number of application per Employment Type}} \right\} \text{Normalized} * \cdots$$

HOW MANY TIMES AN EMPLOYMENT TYPE IS REPEATED IN THE DATASET

NUMBER OF DAYS THE JOB AD STAYS OPEN ON A PLATFORM

FOR EXAMPLE, 14 DAYS

NUMBER OF APPLICATION FOR ONE EMPLOYMENT TYPE

THE DIVISION OPERATION ALLOWS THE VALUES TO BE NORMALIZED

MULTIPLICATION GOES FOR EVERY FEATURE

Results

According to the our algorithm the most attractive job advertisement includes;

- Possibility to work remotely
- Work 2 months as part-timer with a paid holiday in a European country
- 1 month vacation in a year
- Regular cultural and social activities organised by a party committee

Results

```
word = "leadership"
print('Initial skill: ',word)
```

```
Here is a more attractive skill:  organization similarity:  1.0
```

```
word = "management"
print('Initial skill: ',word)
```

```
Here is a more attractive skill:  strategies similarity:  1.0
```

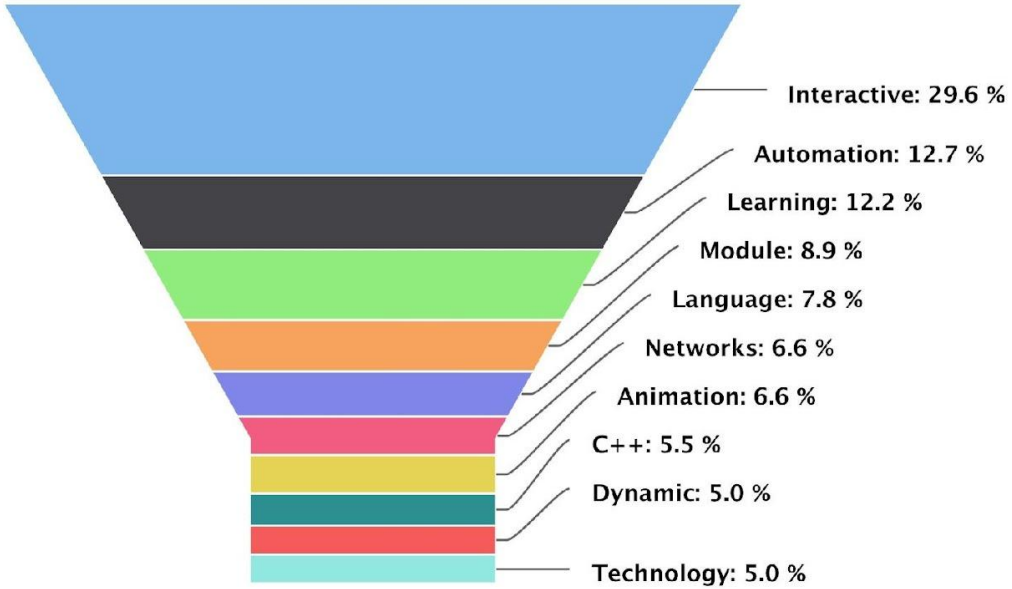
```
word = "experience"
print('Initial skill: ',word)
```

```
Here is a more attractive skill:  ['world', 'learning'] similarity:  1.0
```

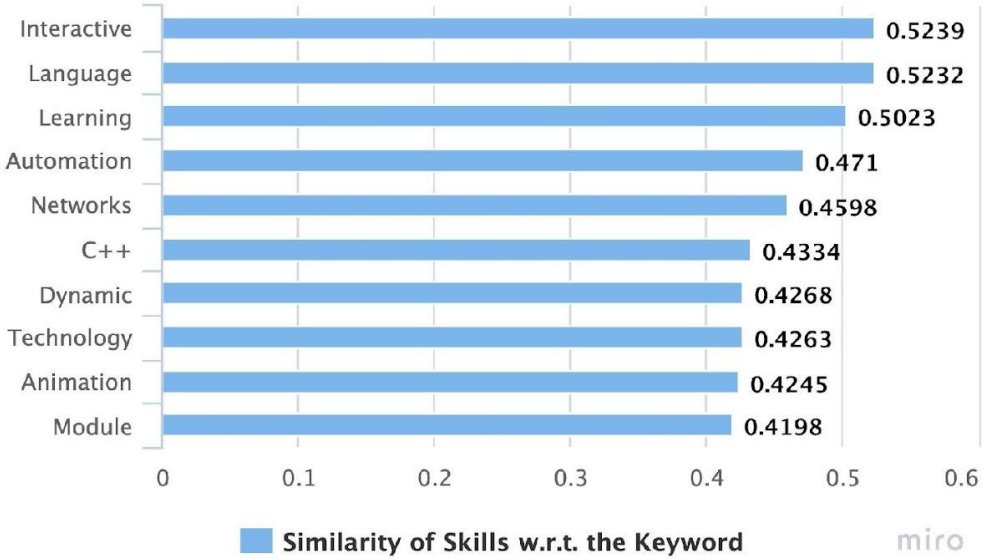
Similarity

Results

Keyword : Programming



Attractiveness of the Skills



Conclusion

- Data Pre-processing
- Data Analysis
- Attractiveness Metric

Thank you!

Questions?

Kemal Alkin Gunbay
Ecem Ture
Olga Bekchanova