

# ASANTE SANA

for a thousand smiles

---



**Rec Engines**  
**MBD O-1**

**Francesca**  
**Manoni**



# EXECUTIVE SUMMARY

---

Recommendation systems/engines apply **data processing techniques** and **prediction algorithms** in order to predict users' interest on a variety of items. They are nowadays utilized in several areas including e-commerce, entertainment items( movies, music, news, books) food, clothes and beauty products, people (dating apps) etc. As we can see, in all of these fields, both the "buyer" and the "seller" may take large advantage of the recommendations made by the system, that is the reason why these technologies have grown exponentially in the last years. The main challenge for recommendation systems today is that of, of course, producing correct recommendation but also that of handling huge information sets, continuously growing and changing on different levels.

The objective of the present work is to propose an innovative business idea for a recommendation system, along with creating a functional MVP prototype in R based on that idea. The use case proposed is 1) creating a recommendation system of NGO projects to potential volunteers and, on the other hand 2) a recommendation system of possible volunteers to NGOs.



# INTRODUCTION

---

The field of **recommendation system** development started with a couple of dozen researchers in the mid-1990s and as of now we count hundreds and hundreds of them. The technology behind them has also improved a lot, from being often inaccurate to being extremely sophisticated and often game-changing for companies and their business value (one need only think to the case of Netflix or Spotify). As of today, it is calculated that more than 80% of the TV shows people watch on Netflix are discovered through the platform's recommendation system. At this point we can say that recommendation engines became a big deal in today's world and probably a lot more is yet to come in this field. However, it is important to note that there are no applications of recommender systems in the field of **social responsibility**.



# INTRODUCTION

---

For this reason my proposal "ASANTE SANA" (the swahili translation for "thank you a thousand"), addresses exactly this absence, trying to provide a solution that could help the **three parties** involved in a social/volunteering activity:

- 1) the **volunteer**,
- 2) the **NGO** that is responsible for the project and, most importantly,
- 3) the **final beneficiaries** of the project (addressees of the project).

Another reason why I made this choice is that I have experience both as a volunteer and as part of the management of an NGO and I personally had to deal with the problems that might arise on both sides, therefore I strongly believe that a technology like this is needed and may be really beneficial.



1



2



3

# SOME DATA

---

An NGO can be any kind of organization provided that it is **independent** from **government influence** and is **not-for-profit**. NGOs have become essential actors in the social field, particularly in the fight against poverty and social exclusion.

To understand the important of these entities, let's look at some data:



**10 MILLION**  
**NGOS WORLDWIDE**



**1.2 BILLION DONORS IN 2011**  
**1.5 BILLION DONORS IN 2014**  
**EXPECTED 2.5 BILLION IN 2030**



**IF NGOS WERE A COUNTRY,**  
**THEY WOULD BE THE 5TH**  
**LARGEST ECONOMY**  
**IN THE WORLD**

# SOME DATA

---



WITHIN THE **EUROPEAN UNION**  
NGOS AND FOUNDATIONS  
PRODUCE AN **ANNUAL VALUE**  
OF ABOUT **€53 BILLION**

**EU**  
**€53B**



**1 OUT OF 10 PEOPLE IN THE**  
**US WORKS FOR A NGO**  
**NGO WORKFORCE IS THE**  
**THIRD LARGEST AMONG U.S.**  
**INDUSTRIES**

**US**  
**1/10**

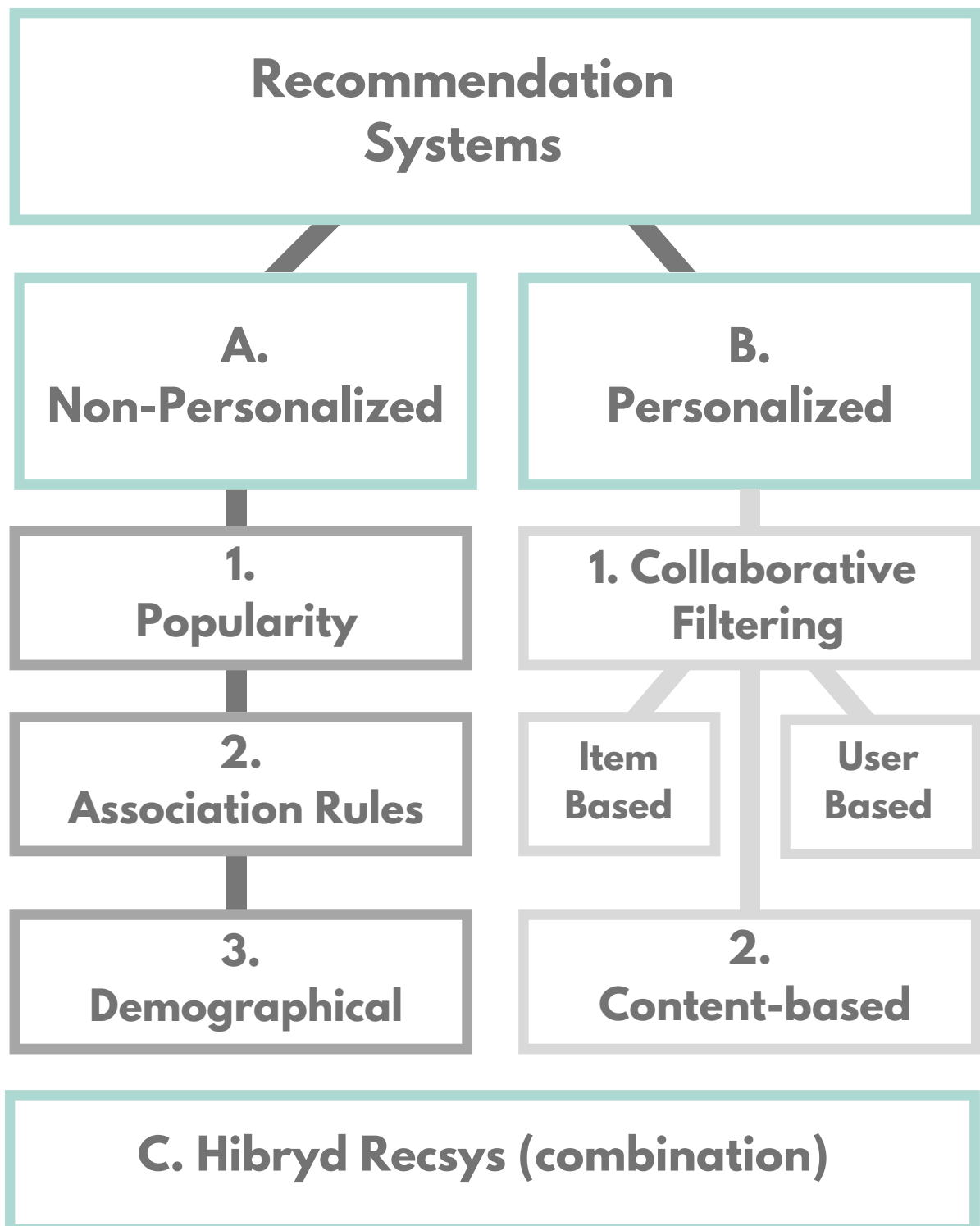


**ESTIMATED VALUE OF**  
**VOLUNTEER IS \$23.07 PER**  
**HOUR**  
**7.7 BILLION HOURS OF**  
**VOLUNTEER WORK X 62.6**  
**MILLION AMERICANS, IN 2013**  
**WAS \$173 BILLION**

**US**  
**\$173B**

# APPROACHES

---



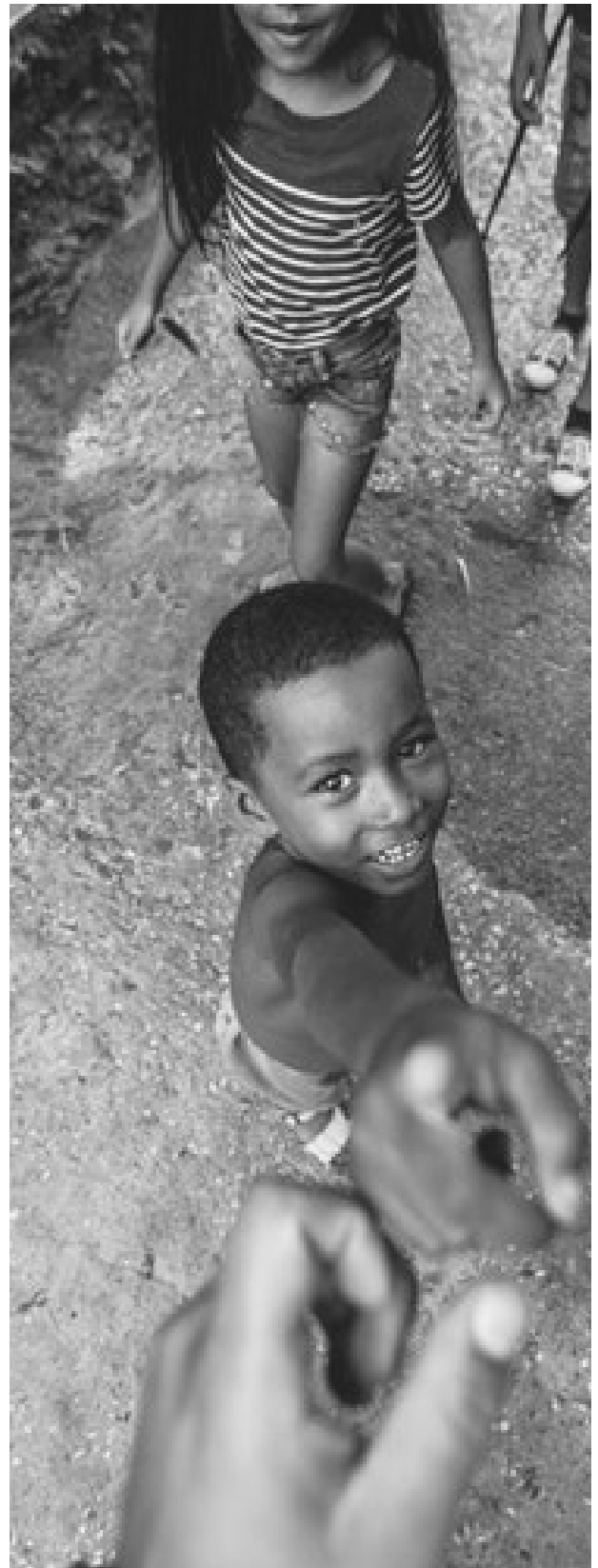
# APPROACH

---

As we can see above, we can follow different approaches while building our recommendation system. Moreover, each of the algorithms can be implemented using different hyper-parameters that will lead to different results (example: IBCF, Item-based collaborative filtering, can compute the distance among items in different ways; Association Rules can have different levels of confidence and support etc.).

In a real-life scenario, after having gathered the data that we need to work on, we would try to approach it with different algorithms and different hyper-parameters in order to see **which one performs better** and suits the most our use case. This analysis would need to be coupled with a business analysis and to reach a final decision on the algorithms to use, the two have to be taken into consideration.

Nevertheless, in the case hereby presented, the dataset has been faked for the purpose of the analysis, so the algorithm will be chosen mostly based on the business case, bearing in mind that in a real life scenario, all the data would be gathered before, and the decision on the algorithm would be a combination of the business point of view and the performance of the algorithms.





# DATASET

---

The dataset has been faked and created first as a dataframe and then converted into a Real Rating Matrix . The matrix will contain a list of names of NGOs and 1000 users.

Moreover, these are the levels:

- **LEVEL 1:** type of project: education, health, elderly care, children, homeless, prisoners, animal care, sports, people with disabilities, advocacy, micro-finance, environmental
- **LEVEL 2:** country of operation of the NGO: the focus of this project is on *Sub-Saharan African countries*



## FOCUS:

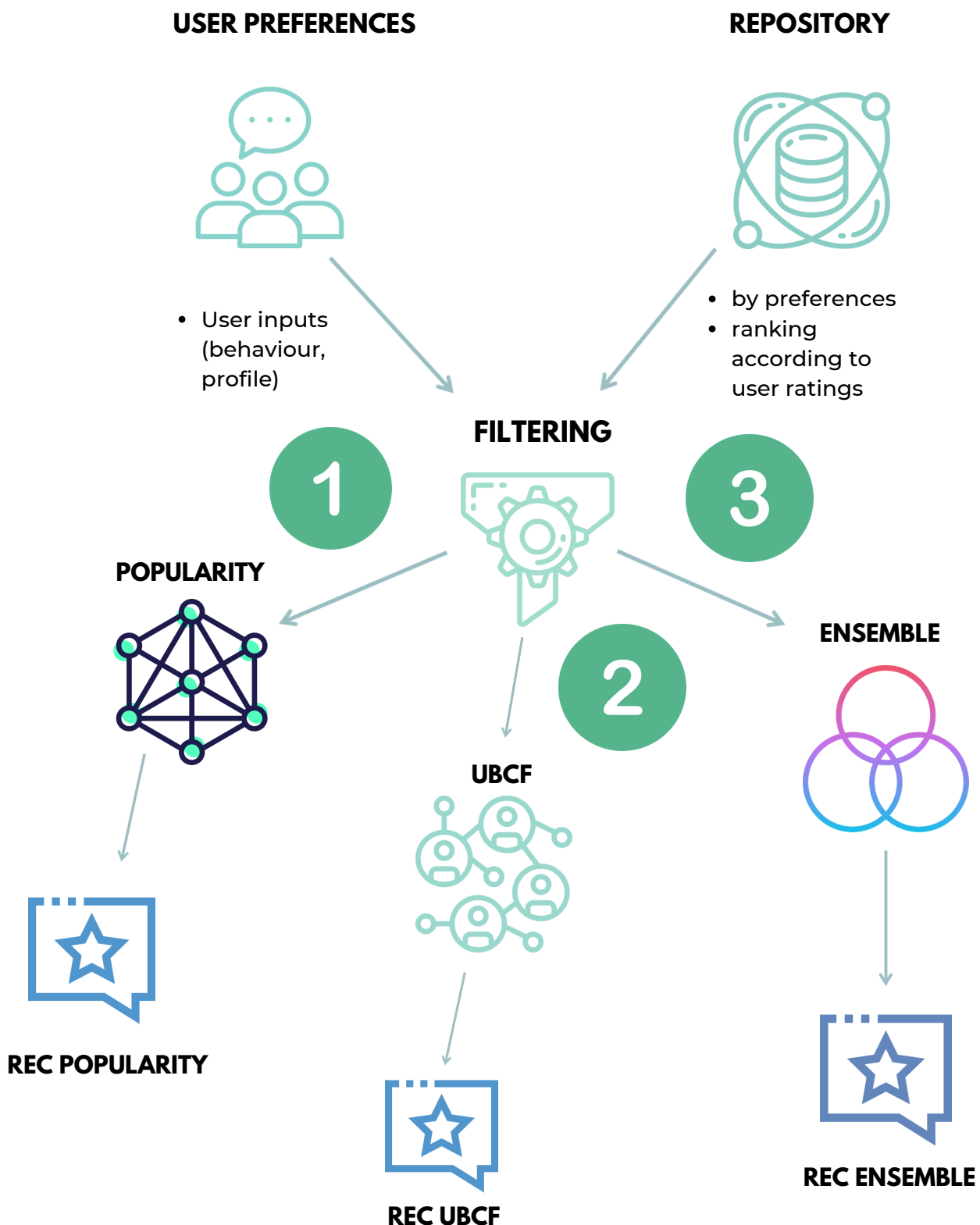
- Sub- Saharan Africa:
- all African countries excluding Maghreb (Morocco, Algeria, Tunisia, Mauritania, Libya) and Egypt and we end up with a total of **45 states**.

The prototype created would be the first version of a potentially scalable recommendation system. One of the future steps would be the inclusion of other features such as:

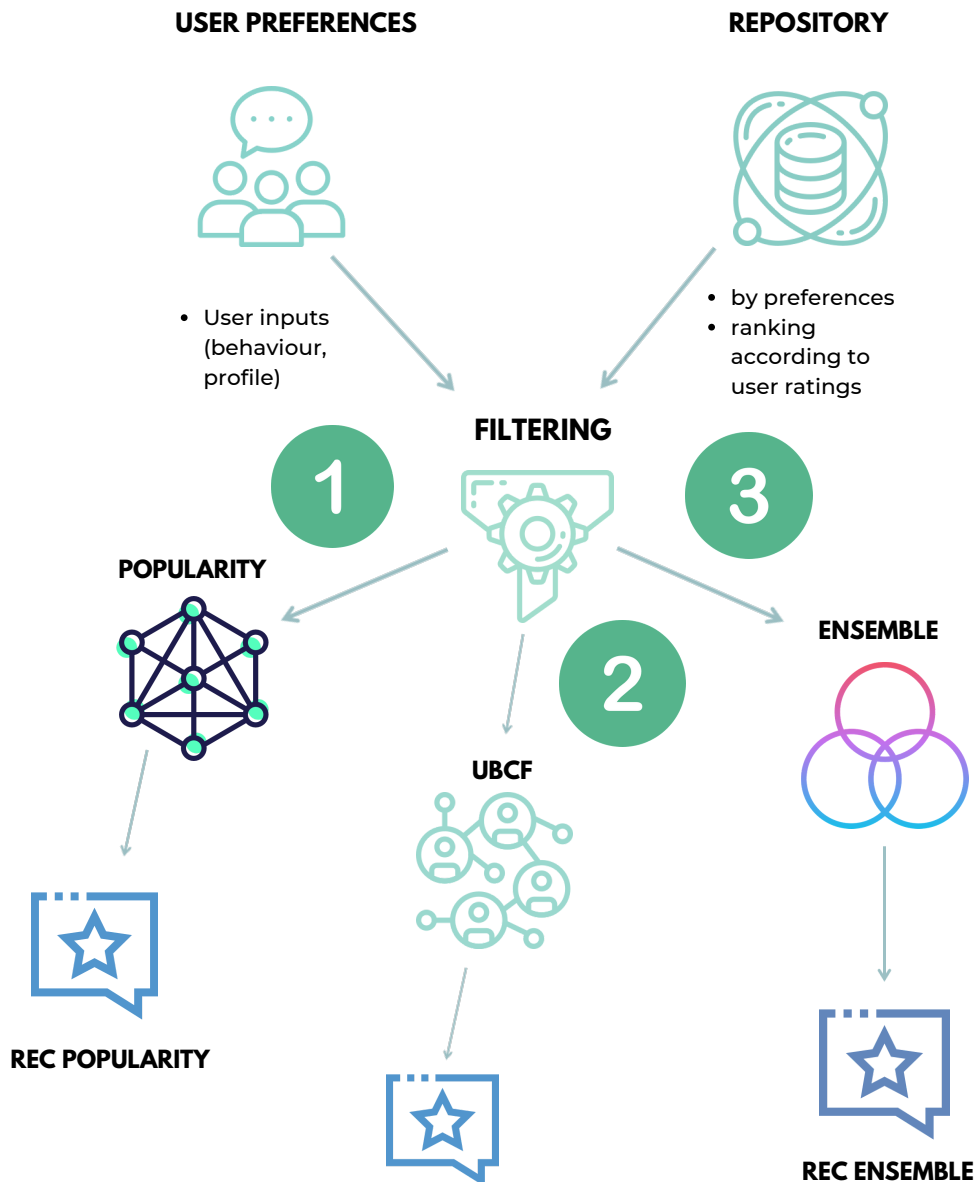
- n. of employees
- preferred languages
- length of the project: n. of weeks, months etc.
- possibility of choosing the length (free choice or fixed period?)
- n. of people participating in the projects (n. of volunteers)
- fee contribution/no fees

# ALGORITHM

After having introduced the context, let's have a look on how the algorithm would work:



# ALGORITHM



## LAST STEP IN A REAL LIFE CASE SCENARIO

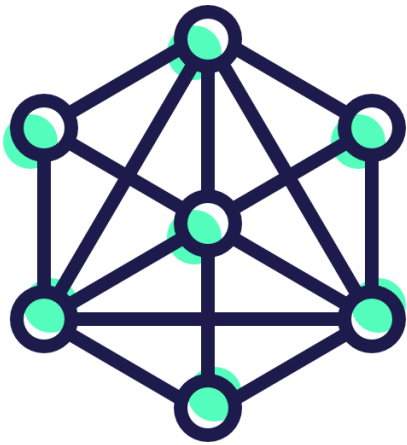
plot the ROC curves of the different algorithms in order to check which one performs the best. Employ the best one.



# ALGORITHM

---

## WHY POPULARITY?



The popularity algorithm makes recommendations according to the popularity of the items. It is a good algorithm to implement when we still don't have any information about our users.

## WHY UBCF?

UBCF makes recommendations based on the common preferences of the users, it recommends items (in our case NGOs) by finding similar users to the target user



# REVENUE MODEL

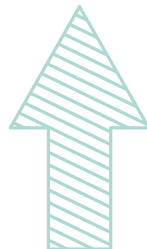
---

## VOLUNTEER

one-time fee: only in case of acceptance by the NGO



**ASANTE  
SANA**



monthly fee in order to be present on the platform

**NGO**

\_\_\_\_\_

```
[1] "Associazione Ricreativa e Culturale Italiana" "International Crisis Group"
[3] "Saferworld" "Hostwriter"
[5] "The Big Issue"
```

# NEXT STEPS

---

## SHORT TERM



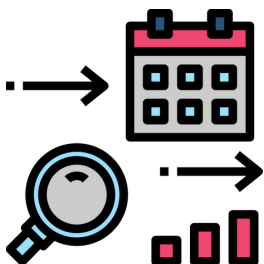
Adding **features** to the recommendation system: n.of employees, preferred language, length of the projects, n. of people participating etc.

## MEDIUM TERM



As previously mentioned, building the second recommendation system to recommend volunteers to NGOs

## LONG TERM



Expansion of the business, not only NGOs, but being able to use the recommendation system also for advising companies on CSR decisions (companies would use the recom. system in order to understand which are the **most meaningful investments**)