A Gamified Approach to Naïve Bayes Classification: A Case Study for Newswires and Systematic Medical Reviews

Giorgio Di Nunzio, Maria Maistro, Federica Vezzani
University of Padua, Padua, Italy

Augmenting Intelligence with **Hum**ans-in-the-**L**oop
April 24th 2018

Summary

- Creation of a ground-truth, or golden standard, is usually very expensive as it requires a manual labelling of the objects by experts in the field.

- Crowd-sourcing?

- How do we motivate people?

QuickDraw

Quick, Draw! The Data

Get the data Play the game

Now visualizing: book

Randomize

You are looking at 111,205 book drawings made by real people... on the internet.

If you see something that shouldn't be here, simply select the drawing and click the flag icon.

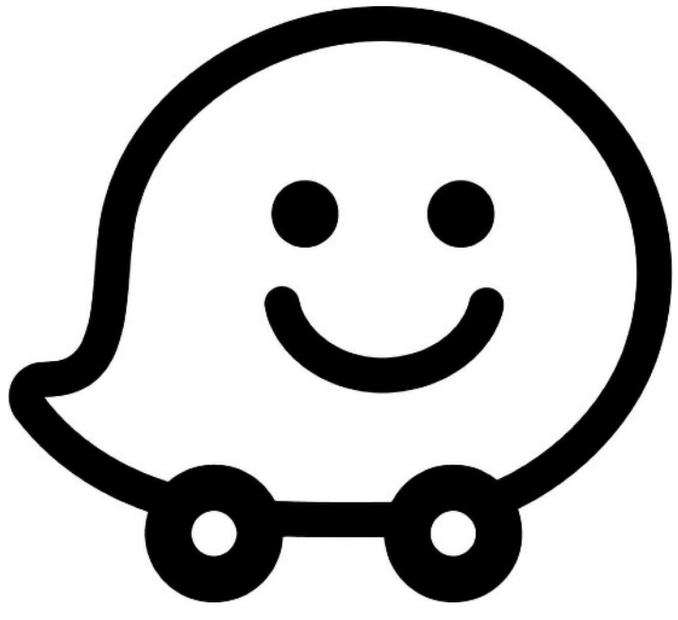
It will help us make the collection better for everyone.

Outline

- Gamification vs Interpretability
- The Classification Game vs the RPG
- Experimental Results
- Ongoing Work

Gamification

"The use of game design elements
in non-game context"



Benoit_Dup0nt

52 contributions
43 avis 1 note 8 photos

Continuez à explorer »

Style de voyage

- Urbain
- Gourmet
- Comme les habitants
- Amateur d'art et d'architecture

Envoyer un message

46 Villes **15** Pays

Mon Profil

Depuis juin 2014

Habite à : Bruxelles, Belgique

Homme de 35-49 ans

Ingénieur dans l'électronique, journaliste aussi et accessoirement enseignant à l'université, je voyage essentiellement, mais beaucoup, pour le travail.

Badges

- 43 avis (Auteur confirmé)
- 13 avis sur les hôtels
- 30 avis sur les restaurants
- Avis dans 21 villes
- 8 votes utiles
- 20 avis d'explorateur

45 591 km parcourus au total

9% du monde parcouru

Day | Week | **Month** | Year

Mar 04 - Apr 03, 2013

Clear Measurable Performance Data

Activity

Steps | Distance | Active Score | Floors | Time Active

Month graph

Friend's Leaderboard

Cals In vs Out

Total Burned
74447 cals

Total Eaten
0 cals ↓ 100%

Top Badges | My Achievements

See all badges

Motivational Badges

Friends

Rankings based on 7 day step total

Vishy
72,360
1

You
52,830
2

See All (1) | Find Friends

12:03

Back

Update price

Close

GAS STATION INC.

2709 TEAGARDEN ST

$4.30	$4.40	$4.70	$4.35
Regular	Midgrade	Premium	Diesel

Last update 1 day ago

Are these prices correct?

Yes

No

12:03

Back

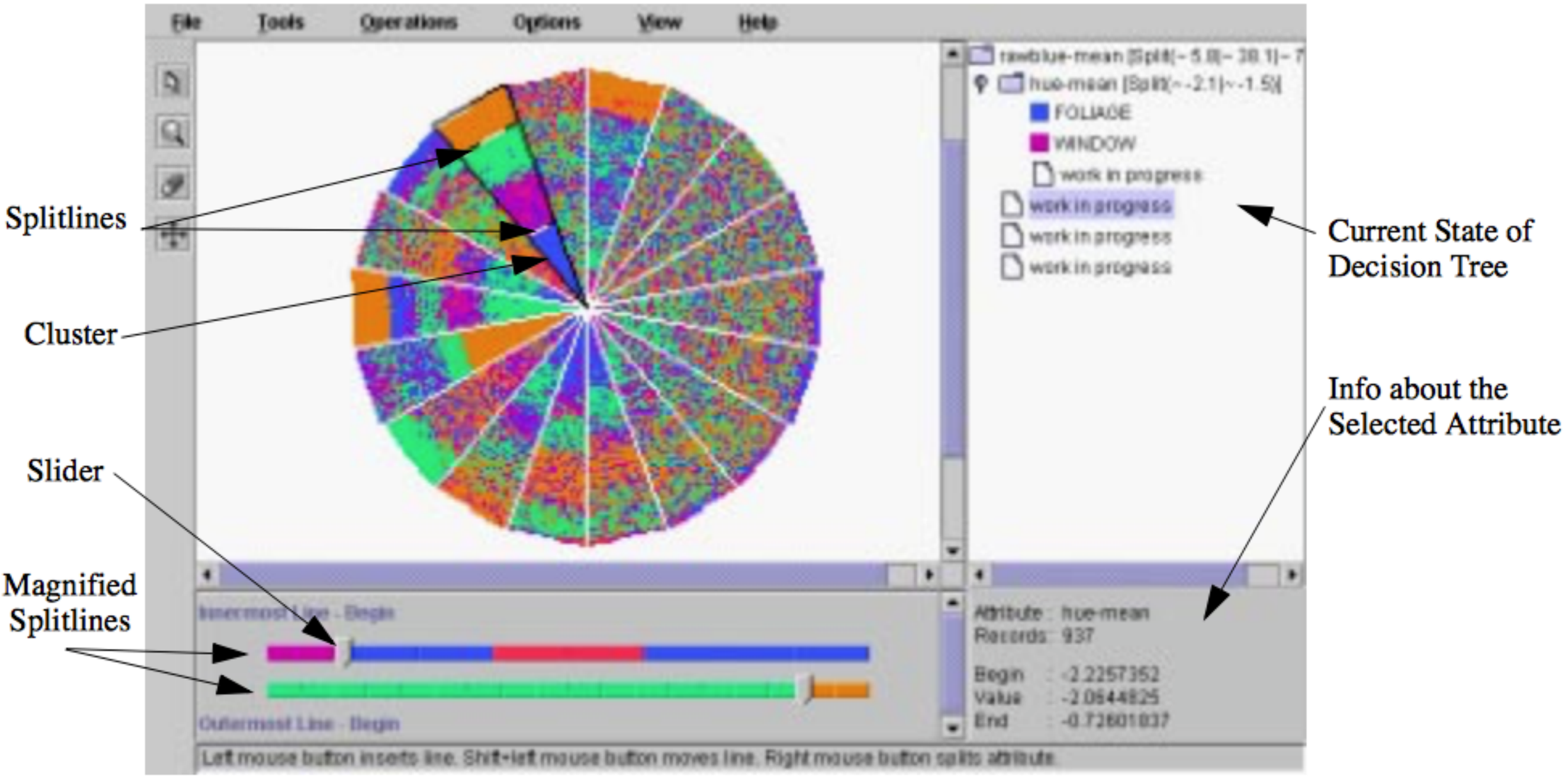
Scoreboard

My | Friends | Everyone

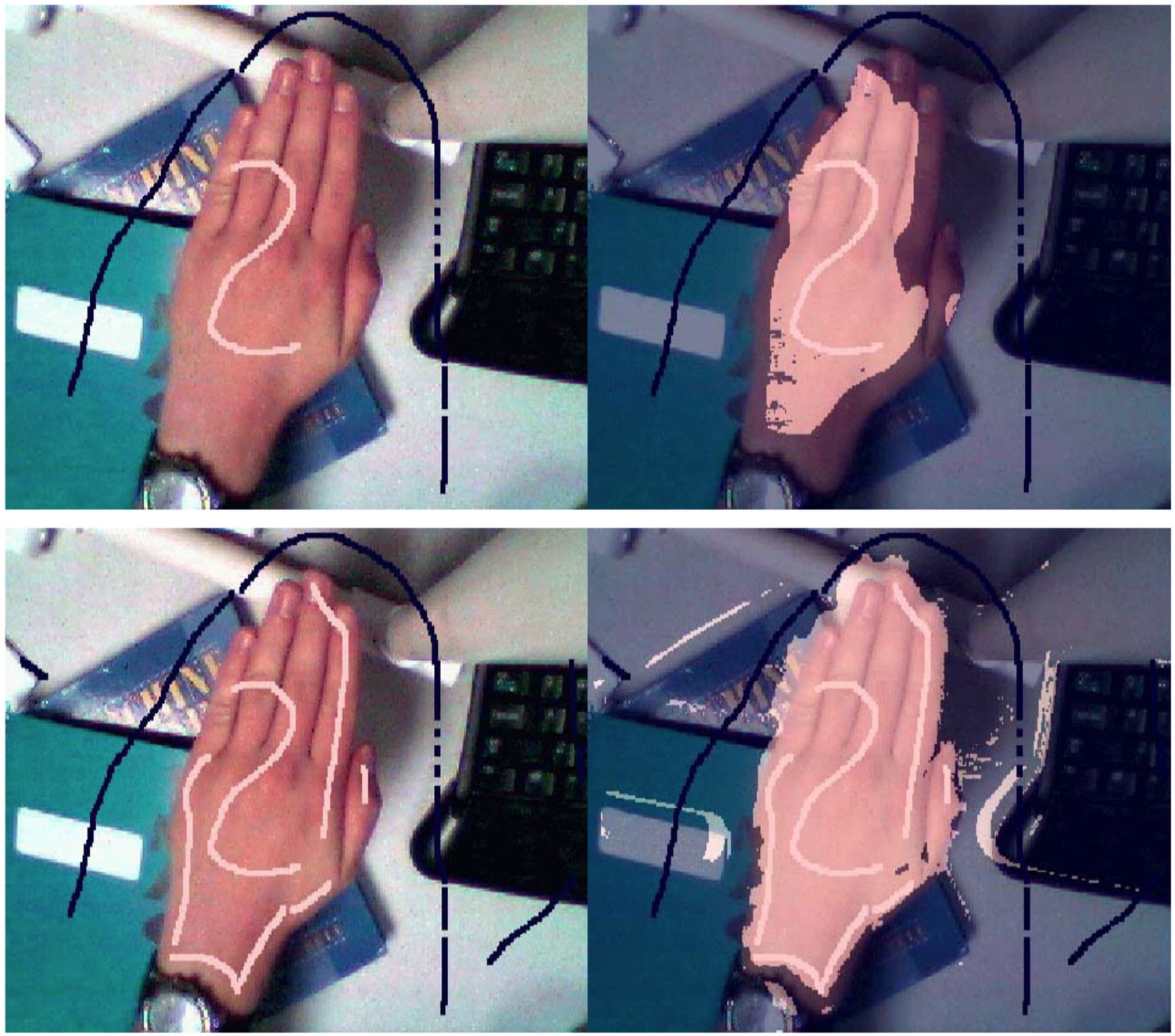
My level:
Waze Knight, 2345 pts.
360 pts. to unlock next level

My rank: #456
Here's how you can move up the ranks...

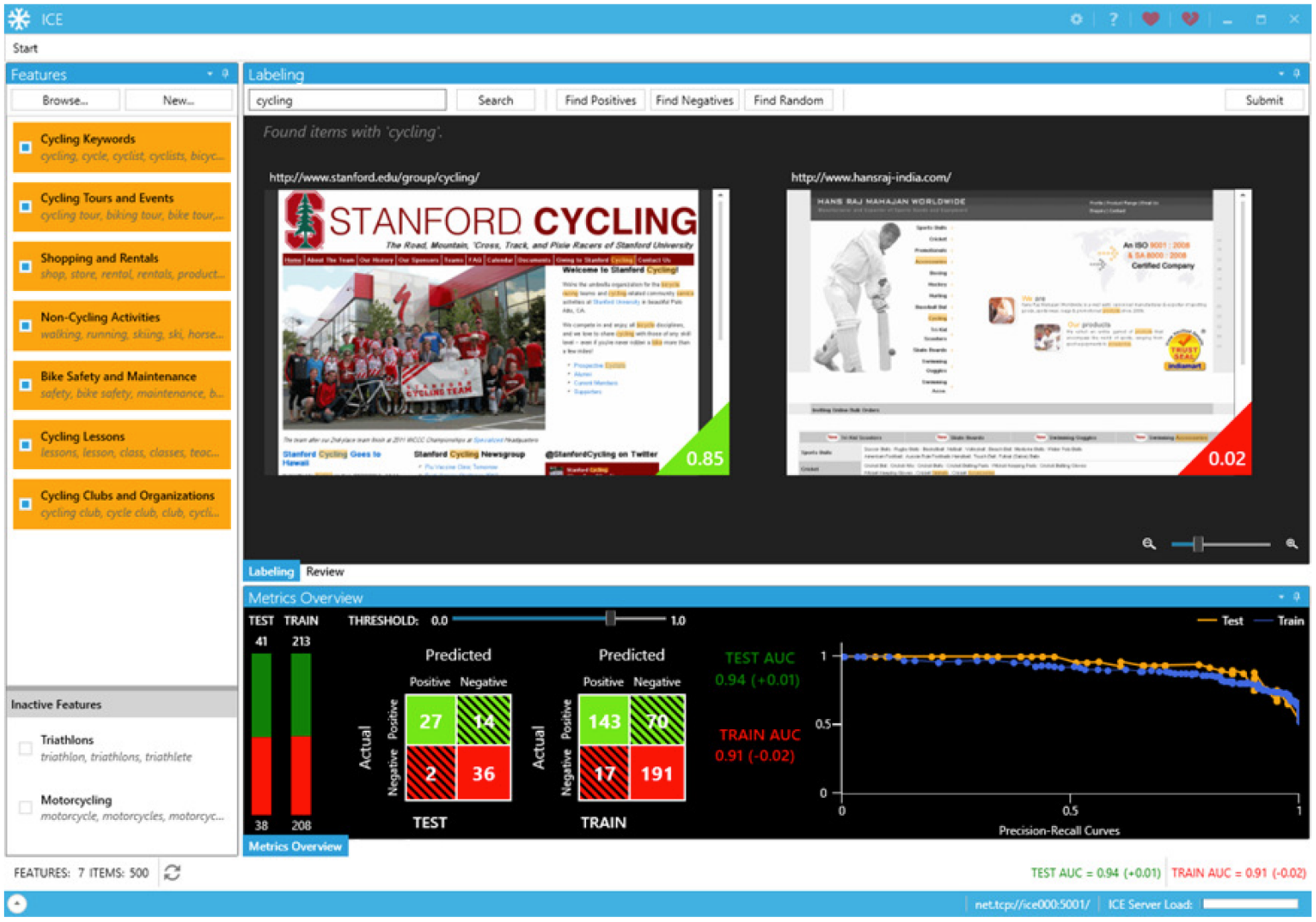
Interpretability
(in Interactive Machine Learning)



Mihael Ankerst, Christian Elsen, Martin Ester, and Hans-Peter Kriegel.
Visual classification: an interactive approach to decision tree construction.
In *Proceedings of the ACM SIGKDD 1999*



Jerry Alan Fails and Dan R. Olsen, Jr..
Interactive machine learning.
In Proceedings of the ACM IUI 2003.



Saleema Amershi, Max Chickering, Steven M. Drucker, Bongshin Lee, Patrice Simard, and Jina Suh.
ModelTracker: Redesigning Performance Analysis Tools for Machine Learning.
In *Proceedings of the ACM* CHI 2015

KEEP

CALM

AND

LET'S PLAY

A GAME

Math Background

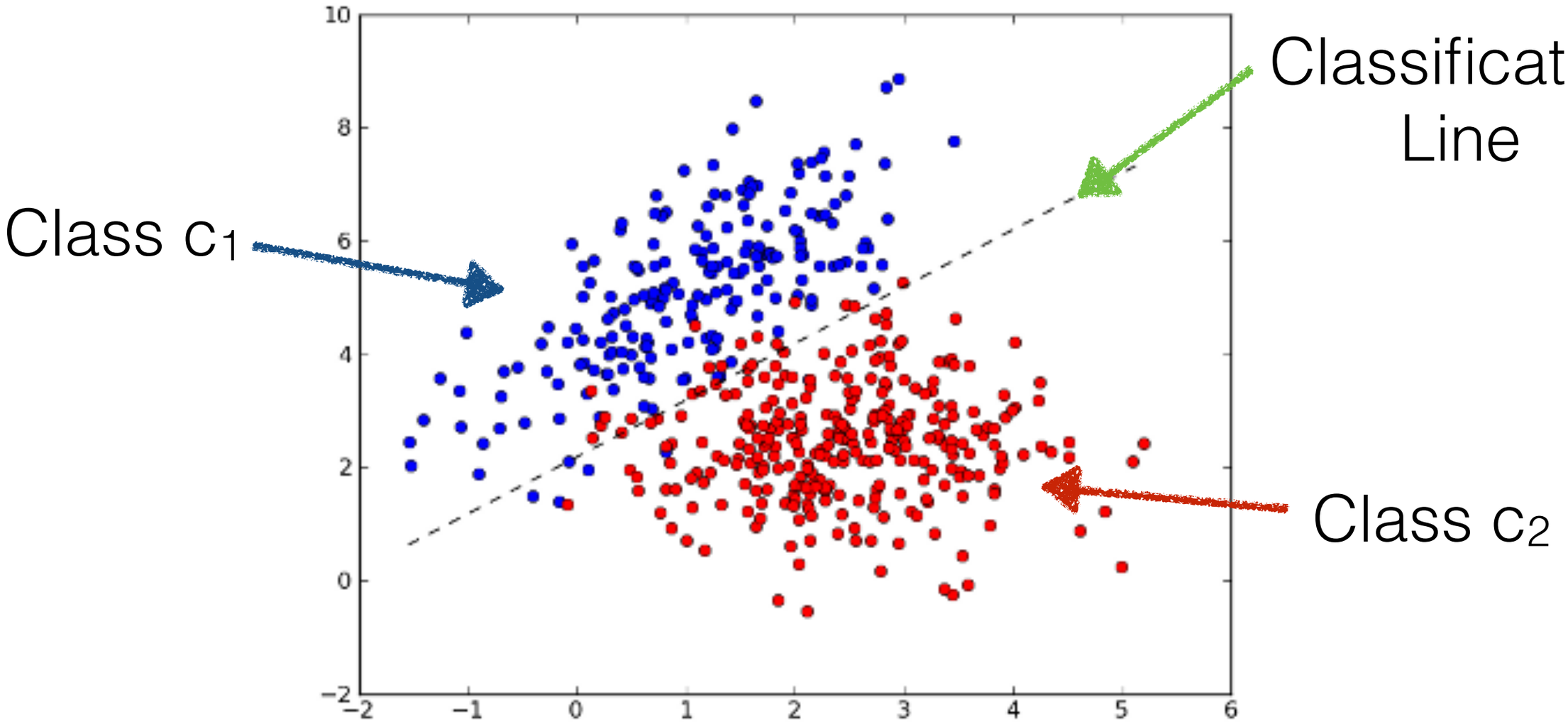
Given two classes c_1 and c_2, an object o is assigned to category c_1 if the following inequality holds

$$\underbrace{P(o|c_2)}_{y} < m\,\underbrace{P(o|c_1)}_{x} + q$$

Parameters that the user can modify through the game

y and x axis represented in the interface

Classification on a two Dimensional Plot



Experiments

1. Pilot experiment: PhD and PostDoc students

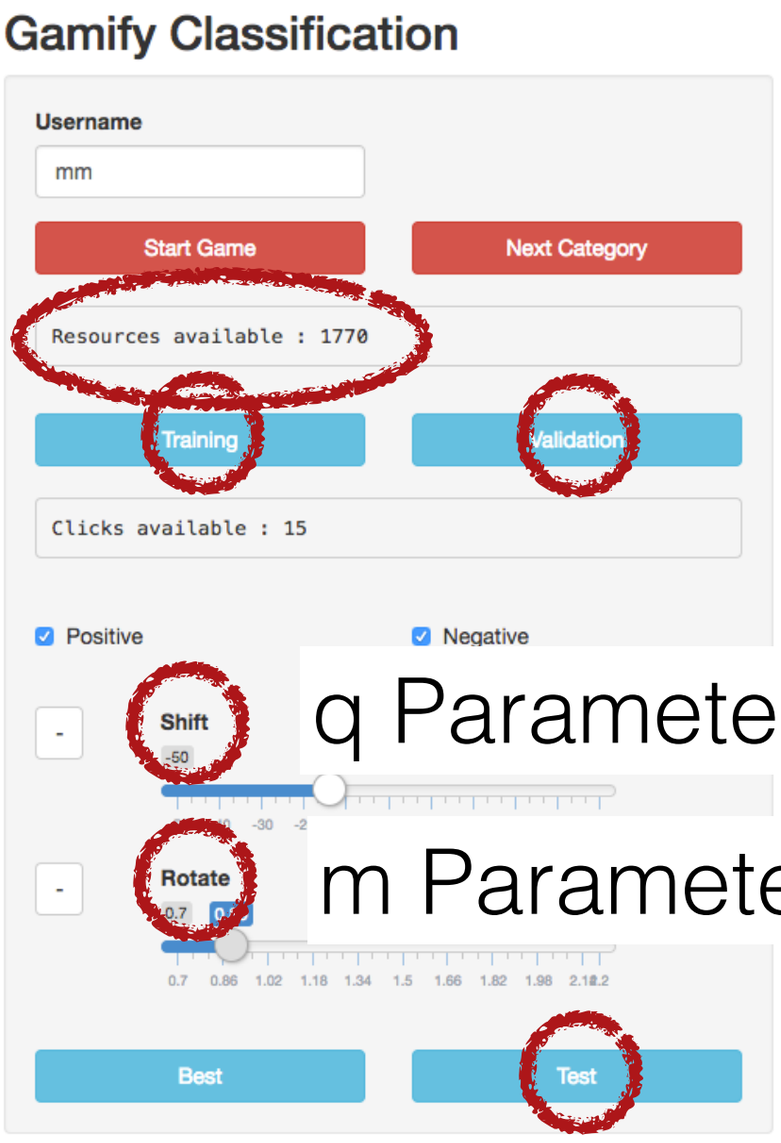
2. European Researcher's Night 2016: kids of primary and secondary schools

3. Banca d'Italia exhibition for the brand new 50 euro note: people from different ages and background

4. European Researcher's Night 2017

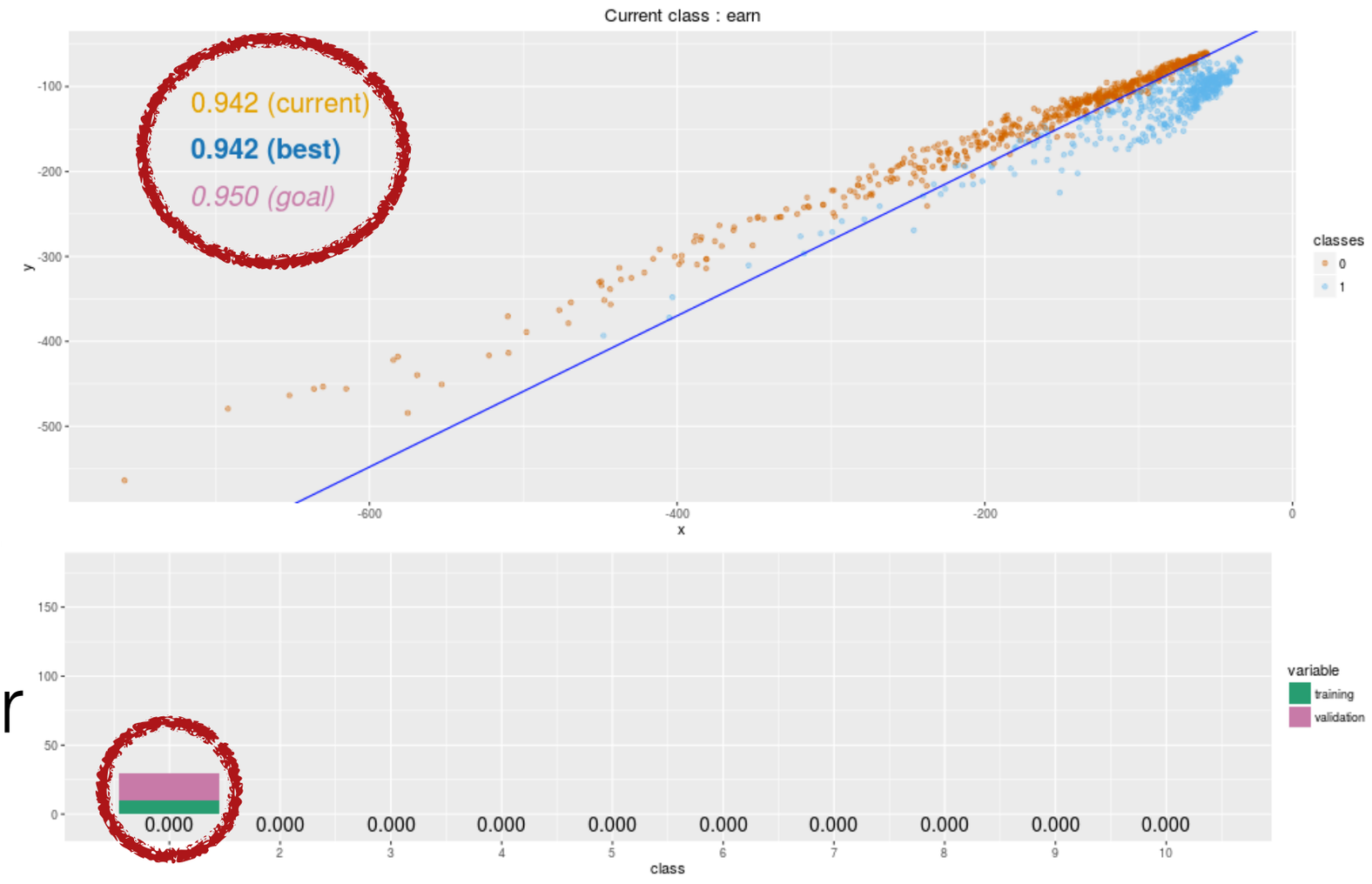
5. Role Playing Game with the Students of a Master Degree in Languages

First Interface

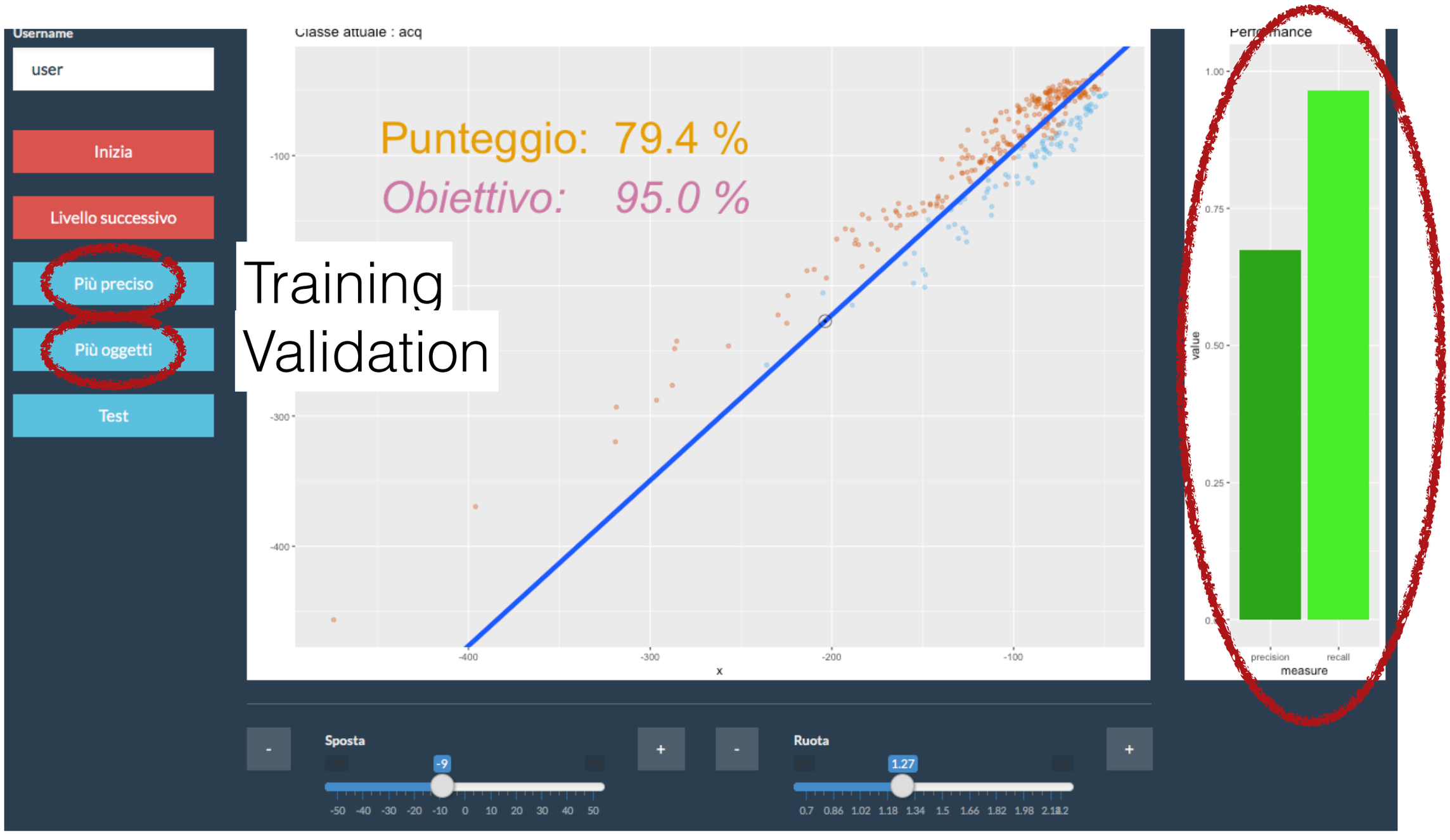


q Parameter

m Parameter

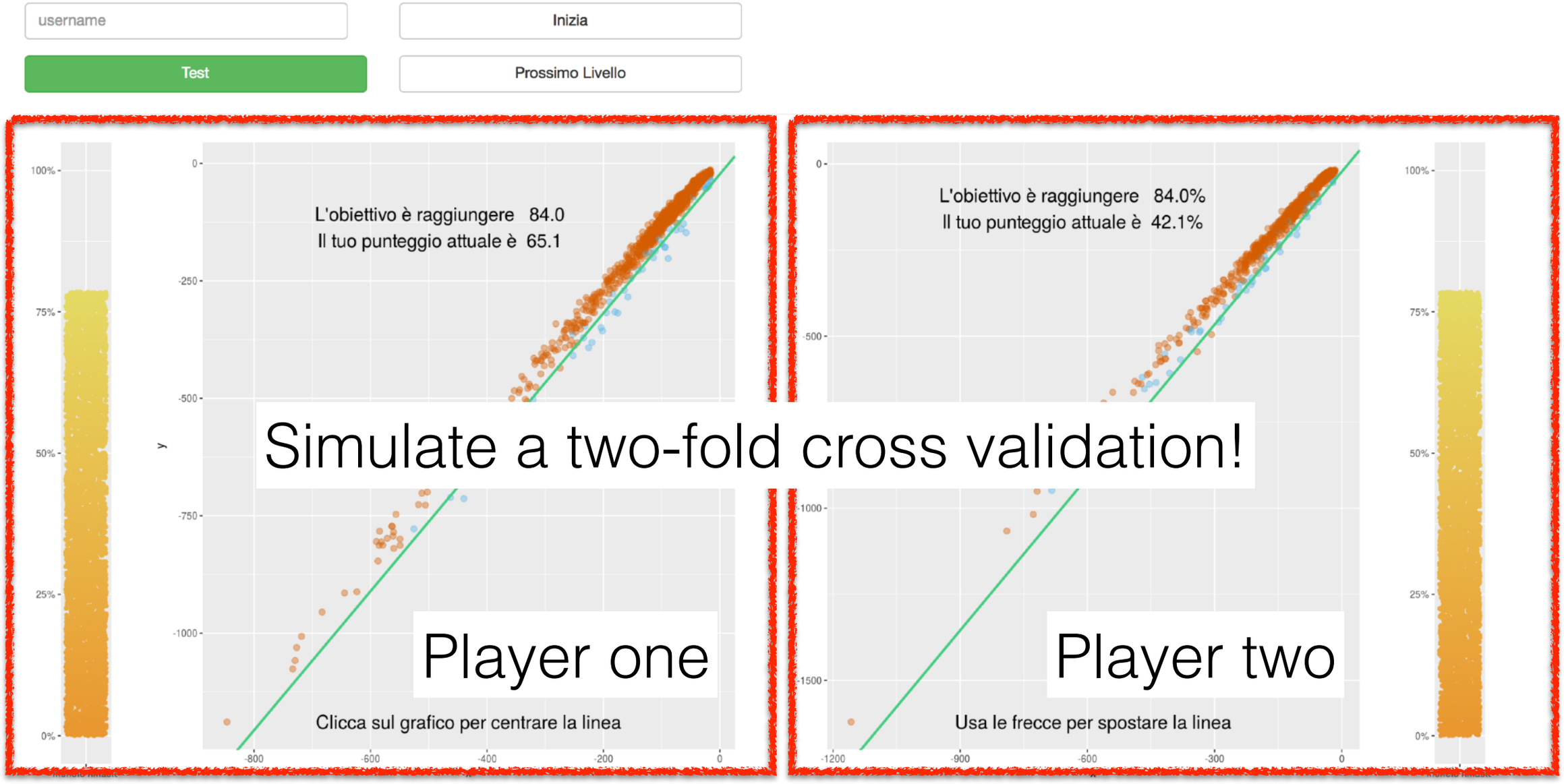


Second Interface



Third Interface

Fourth Interface



Objective of the Study

- Compare
  - How many resources the player "thinks" s/he needs to win the game
  - Player vs Algorithm
    (trained on the same data)
  - Algorithm vs Algorithm
    (trained on the subset of data and on the whole dataset)

Experimental Results

- Players used 20-30% of the training data.
- Their performance was almost identical to a Naïve Bayes classifier trained on the same dataset.
- A NB classifier trained on the subset of data selected by the player is almost identical to a NB classifier trained on the whole dataset.
    - The same for a Support Vector Machine

Role Playing Game (for Systematic Reviews)

Manual Query Rewriting Task

- Objective: Systematic Medical Reviews (100% Recall)

- Make available to non-expert users an interactive system allowing to enter the query reformulations.

- Reformulate iteratively the query in order to obtain more relevant documents for a specific topic.

Role Playing Game
Doctor "Who"?

Professional Translator

Expert in Medicine

Role Playing Game

- The physician (professor) asks to
- The project manager of a translation agency (PhD student, a professional translator) the translation of some medical abstracts.

- The project manager gives indications to the in-house translators (students).

Outcome

- 90 students, divided into 30 groups.

- 28 groups completed the task for a total of
  i) 28 list of keywords
  ii) 28 human-readable reformulation
  iii) 66 individual reformulations

- This experiment has also produced a set of terminological records following the model implemented in an eHealth linguistic resource.

- The system trained on the query reformulated by the players is as good as the system trained on the original queries formulated by the physician.

Final Remarks & Future Works

- Ongoing work on gamification for text classification of newswires and medical documents.

- Get feedback and collect enough data to study how to design the game.

- Extend the game to other tasks (clustering), investigate different game mode and design a new interface for mobile devices.

Thank you for your attention