Presentation of project 3

Antonio Sutera
a.sutera@uliege.be

Institut Montefiore, University of Liège, Belgium

ELEN062-1
Introduction to Machine Learning
November 2020

Goal: Pass prediction during football matches

Predict **the next player** who will receive the ball via a pass **based on the position of all players and the ball at a given time**.

Illustration (1)

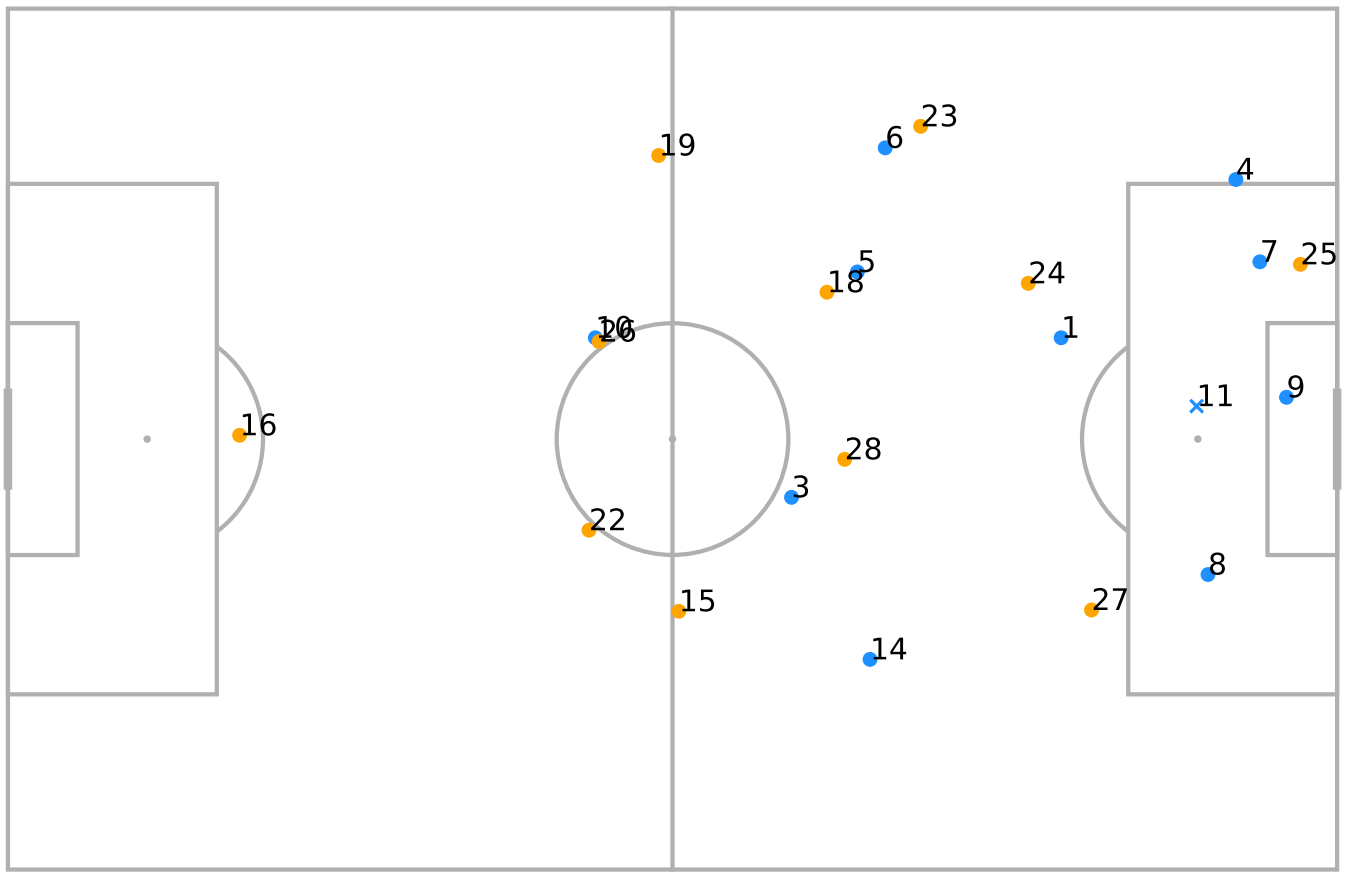
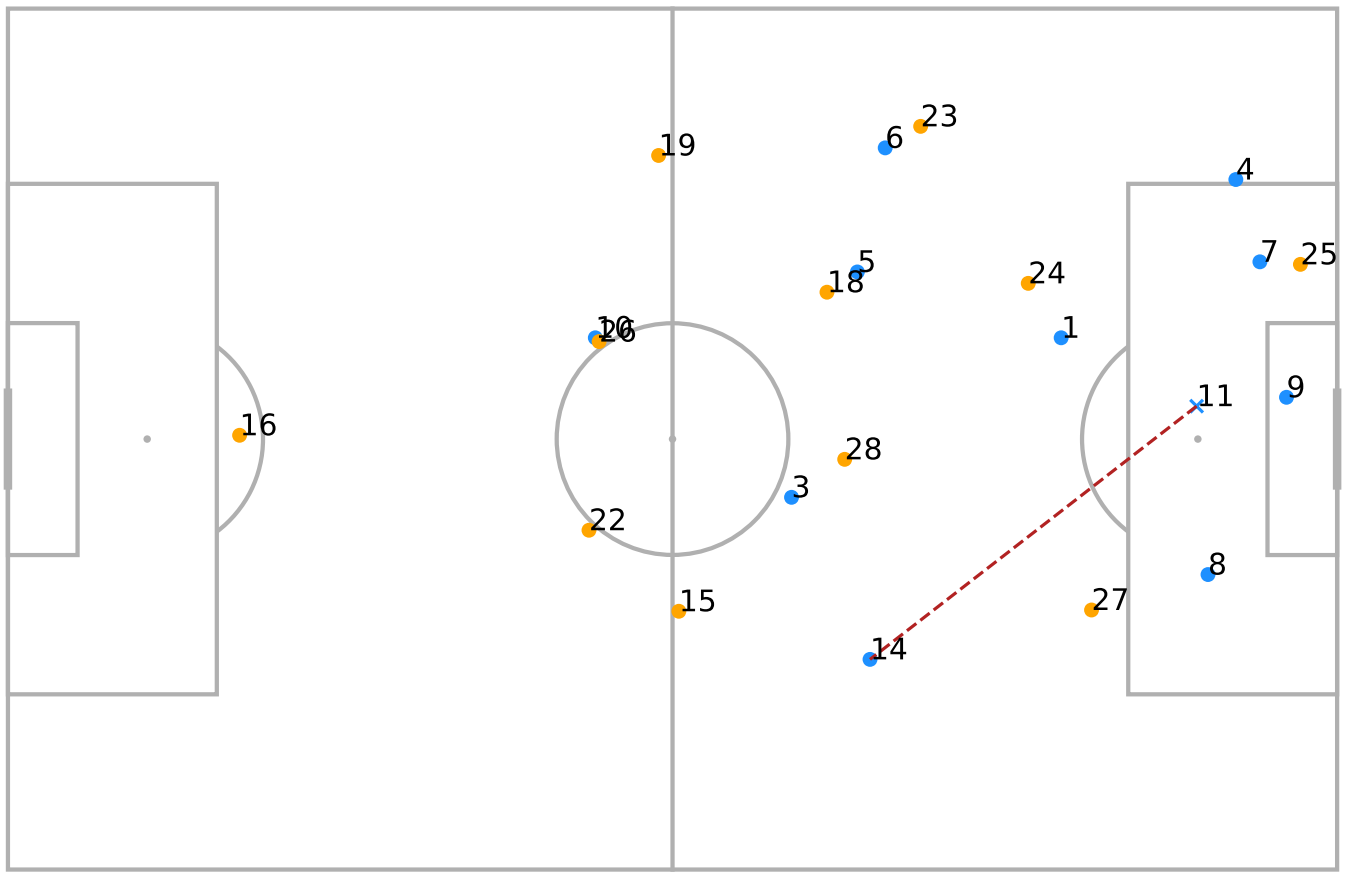


Illustration (2)



Dataset

To do that, you have a training set of 8682 snapshots (samples) of the situation when the ball is passed.

The input variable are

- the **sender**: the player who has the ball $\{1, \ldots, 22\}$;
- the coordinates (x, y) in $[cm]$ of all players;
- the time in $[ms]$ since the beginning of the concerned half-time period.

The output variable *(to predict)* is

- the **receiver**: the player who receives the ball $\{1, \ldots, 22\}$.

Organization

- Same guidelines as the previous projects:
  - A written report!
  - Your (reproducible) codes.
- Organised as a competition (see the rules),
- Three deadlines:
  0- By 27/12: Join a team (on both the Kaggle platform and Submission plateform!) and submit the toy submission.
  1- **On 12/12**: End of the competition.
  2- **By 13/12**: Submission of the project on Montefiore Submission Plateform.
- If you have questions...
  - Via **Slack**: Please ask **general** questions (i.e., related to available files, given data, seen theoretical techniques or algorithms) in public channels (some dedicated channels will be created) and ask **specific** questions in private channels (i.e., related to your ideas / solutions).
  - Some Q&A sessions can be also organized (e.g., after theoretical lectures).

Instructions of the challenge

1. Create a personal Kaggle account **(with your real name!)**
2. Create or join a team (on Kaggle, the Submission platform, and Slack)
3. Download the data and toy python script.
4. Submit a "toy submission"
   `python toy_example.py` and then submit one of the generated files.
5. **Coding, debugging, thinking, . . .**
6. Submit your submission
7. Repeat 5 and 6 until the end (of time).

AND, in parallel, write your report!

8 simple rules

Rule 1 Privately sharing code or data with other teams is not permitted.
It is a competition after all...

Rule 2 Team size: two or three.
Ask us permission. We do not want to let someone behind.

Rule 3 Submission per day: 5.
What does it implies?

Rule 4 Final submission: 2.
Why?

Rule 5 Respect the deadlines.
Elementary, my dear Watson

Rule 6 You are not allowed to use any external data.
Pole position $\neq$ Best mark.

Rule 7 No plagia.
Give your (legal) sources.

Rule 8 Have fun.
It may also helps you to better understand many notions and techniques from the theoretical courses.

The Kaggle platform

- You will receive a link and detailed instructions will be on the webpage of the practicals.
- Public and private leaderboards (more details on Kaggle):
  - **Public** leaderboard is computed on 1000 samples. Note that you should not overfit this ranking. Available immediately for every submissions in order to give you an idea.
  - **Private** leaderboard is computed on 2000 samples on your **two** selected submissions. This is the final ranking and the one that matters.
- More rankings will be computed after the end of the competition.
  - **Task 2** You also have to provide an estimate of your (private) score.
  - **Task 3** You have to provide the probability that each player receives the ball.
- 16/12/20 (TBC): presentations.