



مسابقه تحلیل داده فوتبالیست‌ها	ارائه جهت
بهار ۱۴۰۰	کوئرا کانکت

فرآیند طراحی سوالات

- شروع کار از اواسط اسفند ۱۳۹۹
- آزمون استخدامی City Football Group
- موقعیت شغلی Junior Data Scientist

راه حل سوالات

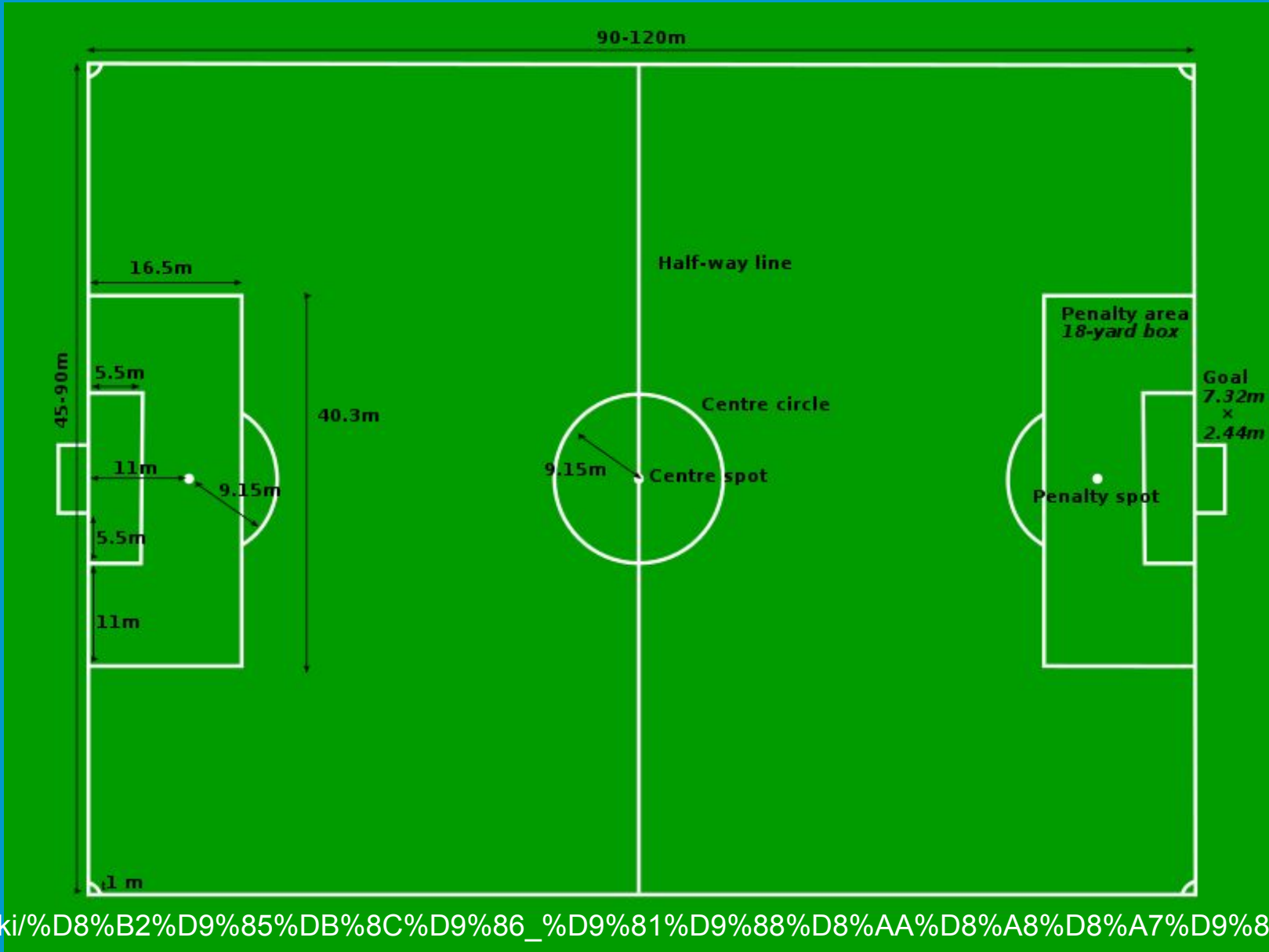


<https://colab.research.google.com>

سوال ۱

سوال ۲

مهندسی ویژگی



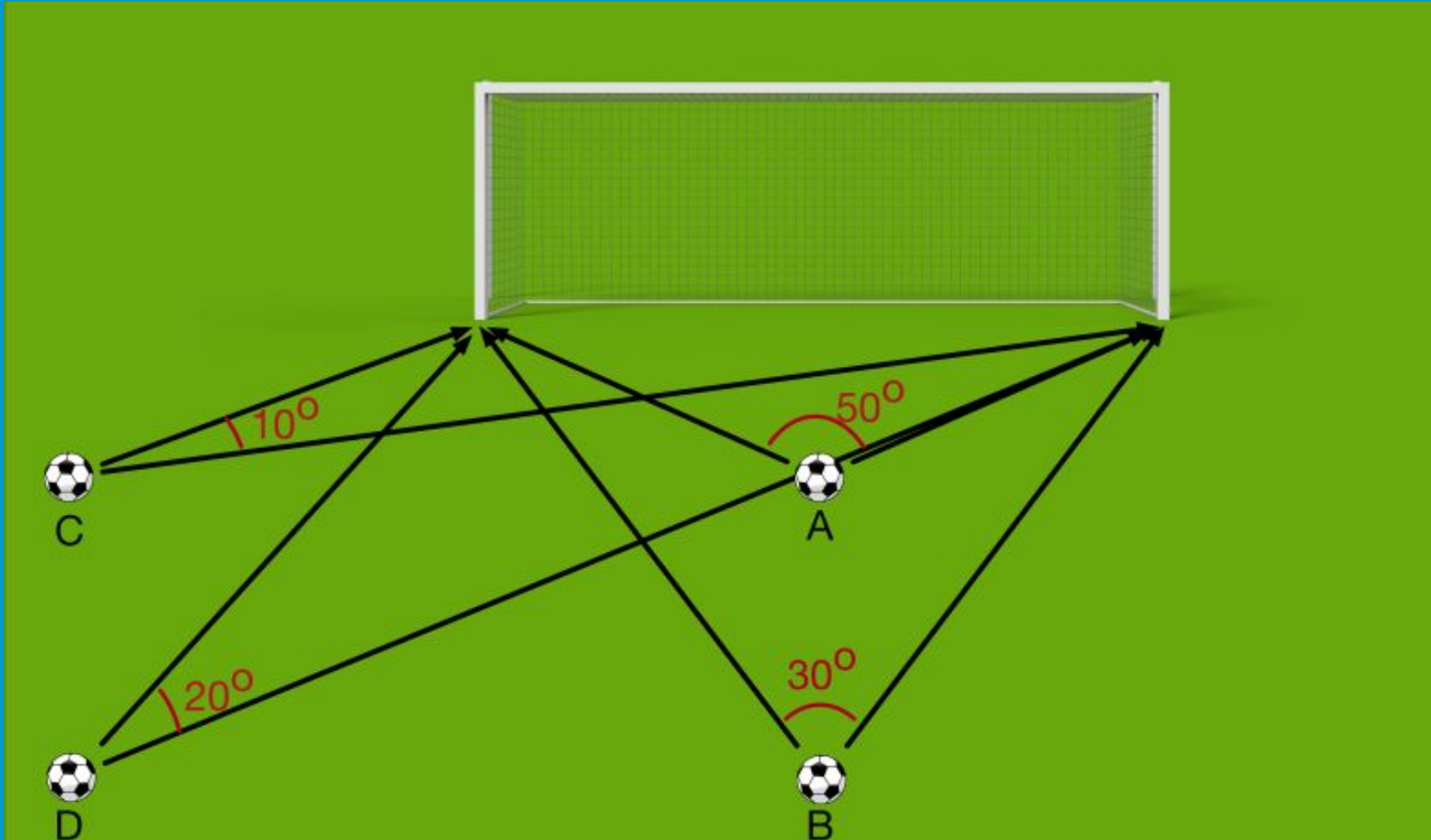
مهندسی ویژگی

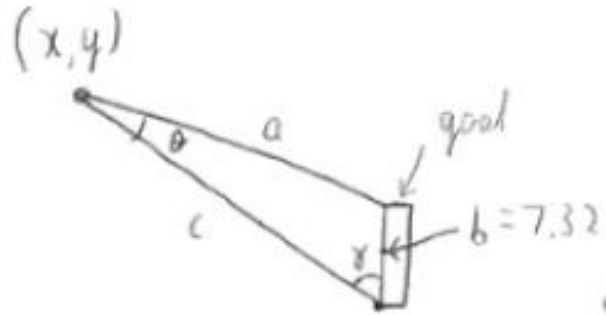
● فاصله از مرکز دروازه

$$\sqrt{x^2 + y^2}$$

مهندسی ویژگی

- زاویه توپ با تیرهای دروازه:





$$a^2 = x^2 + (y - \frac{b}{2})^2$$

$$c^2 = x^2 + (y + \frac{b}{2})^2$$

$$\sin(\gamma) = \frac{x}{c}$$

Sine law: $\frac{a}{\sin \gamma} = \frac{b}{\sin \theta}$

$$\Rightarrow \frac{ac}{x} = \frac{b}{\cos \theta \tan \theta}$$

$$\cos \theta = \frac{bx}{ac \tan \theta}$$

(1)

Angles of triangle

$$b^2 = a^2 + c^2 - 2ac \cos(\theta)$$

Note that $a^2 + c^2 = 2x^2 + 2y^2 + \frac{b^2}{2}$

and using (1),

$$b^2 = 2x^2 + 2y^2 + \frac{b^2}{2} - 2 \cancel{ac} \frac{bx}{\cancel{ac} \tan \theta}$$

$$\Rightarrow 2x^2 + 2y^2 - \frac{b^2}{2} = \frac{2bx}{\tan \theta}$$

$$\Rightarrow \tan \theta = \frac{bx}{x^2 + y^2 - (\frac{b}{2})^2} = \frac{7.32x}{x^2 + y^2 - (\frac{7.32}{2})^2}$$

$$\tan \theta = \frac{7.32x}{x^2 + y^2 - (\frac{7.32}{2})^2}$$

مهندسی ویژگی

- زاویه تا دروازه
- تعداد بازیکنان مانع
- زمان از ابتدای بازی
- ...

ویژگی‌های طبقه‌بندی شده

- موقعیت بازی (playType)
- بخش بدن (bodyPart):
 - گروه‌بندی
- فشار روی شوت‌زننده (interferenceOnShooter)
- One-hot encoded

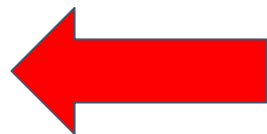
یک مثال



$$x = 6.125m, y = 5.95m, angle = 36.96^\circ$$

یک مثال

```
"timestamp" : "00:48:53.993",
"minute" : 93,
"second" : 53,
"type" : {
  "id" : 16,
  "name" : "Shot"
},
"possession" : 138,
"possession_team" : {
  "id" : 797,
  "name" : "Iran"
},
"play_pattern" : {
  "id" : 3,
  "name" : "From Free Kick"
},
"team" : {
  "id" : 797,
  "name" : "Iran"
},
"player" : {
  "id" : 5226,
  "name" : "Mehdi Taremi"
},
"position" : {
  "id" : 16,
  "name" : "Left Midfield"
},
"location" : [ 113.0, 33.0 ],
"duration" : 0.44,
"related_events" : [ "2076dca5-0f21-479d-ad99-94c3e9625d37" ],
"shot" : {
  "one_on_one" : true,
  "statsbomb_xg" : 0.35391924,
```

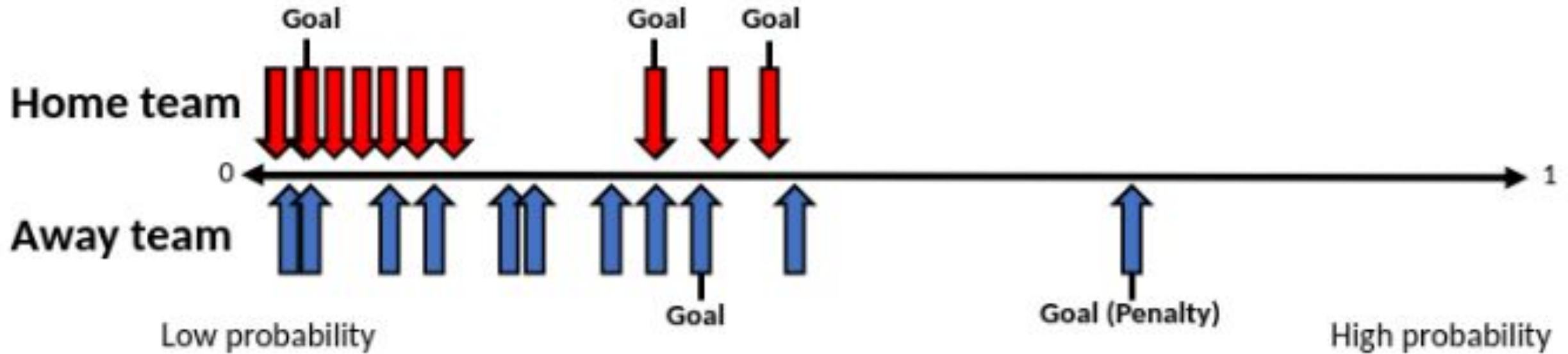


سوال ۴

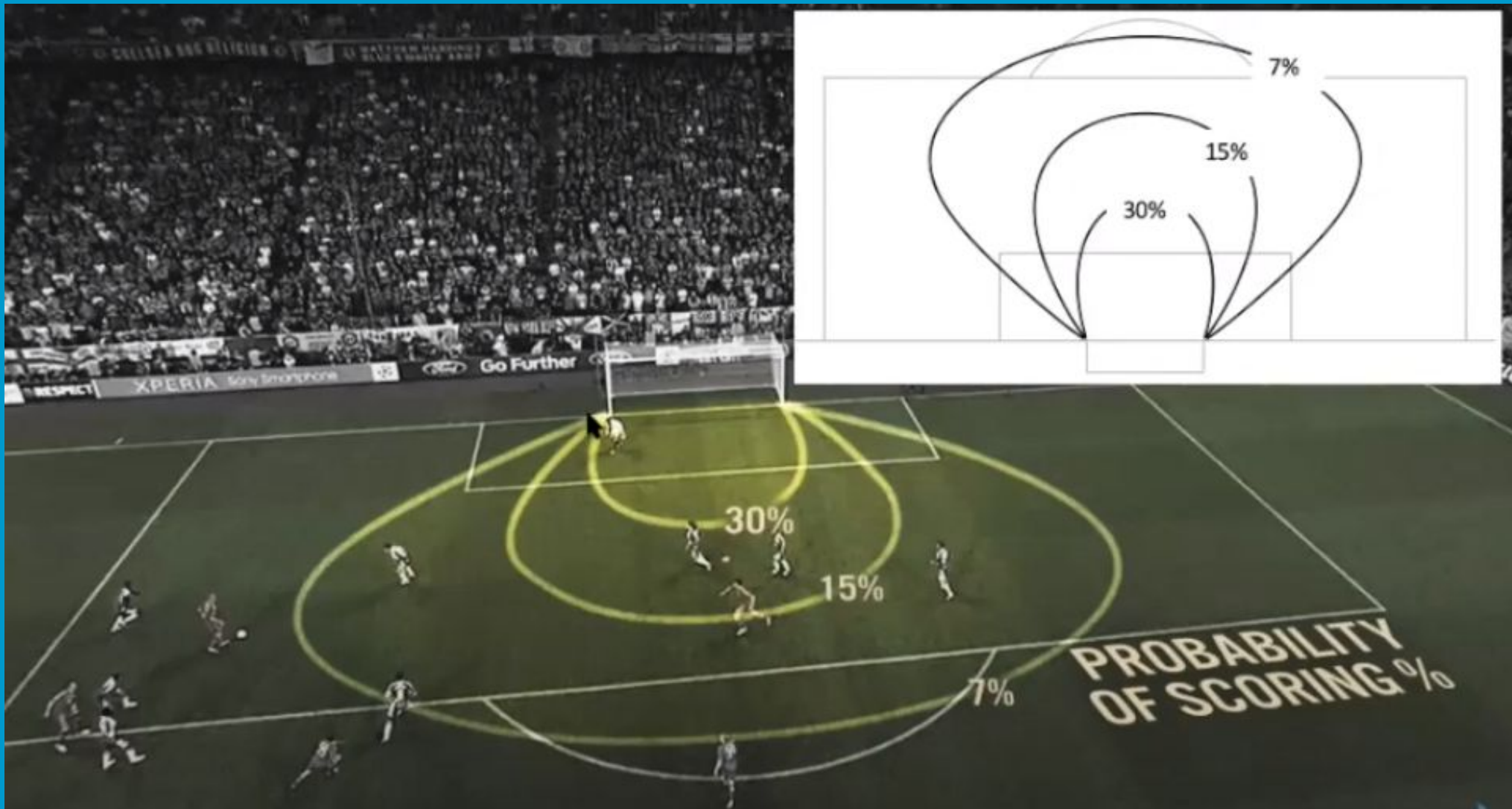
مقایسه کیفیت موقعیت‌های گل و تعداد گل

	isGoal	xG
team		
Away	2	2.81
Home	3	2.36

توزيع احتمال گل



کیف‌سازی موقعیت‌های گل



کیفیت‌سازی موقعیت‌های گل

Table 4: Chance Quality

Quality	xG per Shot
Poor	less than 0.07
Fair	between 0.07 and 0.15
Good	between 0.15 and 0.30
Very good	above 0.30

مثال کیفی سازی موقعیت گل



$$xG \approx 0.35$$

یک موقعیت عالی

کیفی سازی موقعیت های گل

Table 5: Chance qualities of each team

Team	Poor	Fair	Good	Very good	Total
Away	2	2	3	4	11
Home	3	5	1	4	13

با تشکر از توجه شما



جامعه برنامه‌نویسان ایران



quera.ir



@quera.ir



@quera_ir



quera