

## پاسخ به سوال نیمکت

### مجید فرزانه

با توجه به زمان باقیمانده، از توانایی خارج است که بتوانم کد اصلی را بنویسم. اما روند اجرا به صورت زیر خواهد بود:

ابتدا از مجموعه داده‌ی `RawEventData` موقعیت‌های شوت را جدا میکنیم. در این مجموعه داده 24 موقعیت شوت ثبت شده است. اطلاعات مهمی در این بخش وجود دارد، مختصات شوت ( $X$  و  $Y$ )، و زمان شوت (دقیقه و ثانیه) است. از این رو ۴ ورودی از مدل را می‌توان از همین جدول استخراج نمود. لازم به ذکر است که مقادیر  $X$  و  $Y$  در دیتاست جدید نرمال‌سازی شده‌اند که باید به ترتیب در مقادیر ۱۰۵ و ۶۸ ضرب شوند.

برای به دست آوردن دیگر ورودی‌ها لازم است شماره فریم‌های آغاز شوت زدن یادداشت شده و در مجموعه داده‌های بعدی یعنی `TrackingData` بررسی شود که مقادیر ویژگی‌های دخالت هم‌تیمی، دخالت تیم مقابل، و تداخل با شوت‌زن محاسبه شود. برای این کار می‌توان محاسبه کرد که در فاصله‌های نزدیک به شوت‌زن چند بازیکن هم تیمی و چند بازیکن حریف وجود دارد.

نکته مهم اینجاست که در دیتاست مشخص نیست که شوت زن با کدام قسمت بدن شوت زده است. از این رو یکی از ویژگی‌های مدل قابل محاسبه نمی‌باشد. مگر اینکه از ابتدا مدل را بدون این ویژگی آموزش بدهیم. همچنین تشخیص `PlayType` مقداری دشوار است و شاید بتوان کل مقادیر را «در جریان بازی» قرار داد.

با مشخص شدن ورودی‌های مدل برای هر موقعیت شوت، می‌توان امتیاز هر شوت را بر اساس احتمال گل شدنش محاسبه نمود (همانند رابطه ۱ در گزارش قبلی). اما به جای اینکه امتیازها به ازای هر بازیکن محاسبه شود، این بار به ازای هر تیم محاسبه می‌شود. درنهایت هر تیمی که امتیاز بالاتری بگیرد (از نظر شوت‌زنی) برتر بوده است.