Predicting Soccer Highlights From Spatio-temporal Match Event Streams





Tom Decroos

Vladimir Dzyuba

Jan Van Haaren Jesse Davis {firstname.lastname}@cs.kuleuven.be http://dtai.cs.kuleuven.be/sports

Motivation: Highlight channels



Q: How do broadcasters choose which soccer match to display?

A: Broadcasters switch to a match **AFTER** an interesting event has happened in that match.

Research question: Can we switch **BEFORE** an interesting event happens?

Problem statement

Given: A soccer match at time T.

Predict: The probability of a goal being scored in the next δ seconds.

Challenges

- Goals are rare. Not enough data to learn a conditional distribution.
- How to handle the spatio-temporal aspect of soccer gameplay?

Spatio-temporal match event streams



Our approach: POGBA (Prediction Of Goals By Assessing phases)

Key ideas

- Goals are almost always preceded by shots.
- Use a generative model to estimate goal probability.

P(*goal*, *shot*, *phase*)

mostly on shot quality.

Predicting highlights using POGBA

Interesting events: high probability that a goal will be scored in the near future.



Experiments

- Event streams of 69 matches provided by professional soccer club. \bullet
- Highlights of 25 matches from sports website used for evaluation. \bullet

9	HOME	MATCH CENTER	VIDEOZONE	EINC
	LIVE VOE	TBAL LIVE WIE		Eincle De eerste tooper van deze zondag in Ben efficient Club Brugge veegoe oa een vroege actinestand weg, maar in tenen voor na rood voor Duarte. Ge overwicht - ondanks een paar leuke is schot van Raftma liet gilboer. Beloe slot nog de zege, maar het bleeftig z
	EINDE	li "į	KAAGENT	Pollet net niet Ook Gent maakt nog aanspraak op o naam al op het scorebord, maar hij m
JPL - SPEI	ELDAG 12 26	10/2014 14:30	••	Seis moet zich in een deverend sict n month een polition van Meneraatto m
De Sutte	0-1	Milicevic	13	
De Sutte	x 2-1			
	2-2	Refinbe	64	Het blijft een dubbettje op zijn kant. O eentije open te peuken, maar potst op

	Temporal aspect	Indirect estimation	Precision	Recall	F_1
POGBA	\checkmark	\checkmark	0.44	0.61	0.51
SpatTempDir	\checkmark	_	0.43	0.59	0.50
SpatIndir	—	\checkmark	0.36	0.46	0.40
SpatDir	_	_	0.34	0.39	0.36
Final 4th	_	_	0.22	0.72	0.33
Random	—	—	0.05	0.05	0.05

Conclusions

- Handling the spatio-temporal aspect of the data is essential.
- POGBA can also be applied to other domains with important events (industrial machines, hospitals, ...).
- POGBA can generate advanced stats for soccer analytics.