Sequential Ice Hockey Events Generation using Generative Adversarial Network

Md Fahim Sikder

md.fahim.sikder@liu.se
Reasoning and Learning Group (ReaL),
Department of Computer and Information Science (IDA),
Linköping University
Overview

1. Contest Problem Statement:
   - Generate / Find pattern of events that leads to a particular outcome.

2. Swedish Hockey League data: 20 matches data provided by Sportlogiq
Motivation

1. Analytic can be helpful for the team level management for devise tactics
2. Such as: we can learn what steps to take to for a particular outcome
   - Goal, successful zone entry, analyze the whole game
   - Determine player performance
   - In our case, sequential events leads to goal and position
Solution

1. Learn a generative model (TimeGAN\textsuperscript{1}) to capture the pattern of original data
2. Generated synthetic data using the generative model
3. Showed the sequence of events that leads to a "goal"
4. Plotted the coordinates of the events in a hockey rink\textsuperscript{2}

\textsuperscript{1}Time Series Generative Adversarial Network, J Yoon, NeurIPS, 2019
\textsuperscript{2}https://github.com/the-bucketless/hockey_rink
Approach

1. Why Generative model?
   - Model learns internal pattern of the original data
   - Generated data follows the same distributions as the original data
   - Sometimes size of original data can be small and it might be difficult for other ML method to learn patterns
   - Unlimited sampling!

2. Training Approach:
TimeGAN

Losses used in TimeGAN:
1. Reconstruction Loss
2. Supervised Loss
3. Unsupervised Loss

Figure: TimeGAN Architecture
Evaluation

Beside TimeGAN, we have also implemented another two GAN architectures (LSGAN\textsuperscript{3}) and compared their results with TimeGAN.

We have evaluated the synthetic data using the following:

1. Principal Component Analysis (PCA) plot
2. Sequence Prediction tasks: Given 23 sequence predict the next one.

\textsuperscript{3}Least Squares Generative Adversarial Networks, X. Mao, ICCV, 2017
Evaluation

(a) PCA
TimeGAN
(b) PCA LSGAN
- LSTM
(c) PCA LSGAN
- GRU
Evaluation

Sequence prediction task, loss function were Mean Absolute Error (MAE) and Mean Squared Log Error (MRLE)

Table: Comparison of Three GAN models on sequence predicting task

<table>
<thead>
<tr>
<th>Models</th>
<th>MAE</th>
<th>MRLE</th>
</tr>
</thead>
<tbody>
<tr>
<td>TimeGAN</td>
<td>0.246165</td>
<td>0.053882</td>
</tr>
<tr>
<td>LSGAN-LSTM</td>
<td>0.2999977</td>
<td>0.062845</td>
</tr>
<tr>
<td>LSGAN-GRU</td>
<td>0.293644</td>
<td>0.071429</td>
</tr>
</tbody>
</table>

4https://github.com/ydataai/ydata-synthetic
Results

(a) Synthetic vs Original (bin calc method = count)

(b) Synthetic vs Original (bin calc method = mean)
Results

(a) Synthetic Goal

(b) Synthetic Goal Events

Figure: Synthetic goal plot and events
Discussions

1. Idea of using generative models and synthetic data
2. Generalized approach (can work with goal, zone entry, etc)
3. This is the first work using Synthetic data on Ice Hockey (to our best knowledge)
4. Code for this project can be found here\(^5\)

\(^5\)https://github.com/fahim-sikder/event-generation-ice-hockey
How can this solution help Hockey Analytics

1. Sample different patterns from a single model (find pattern of goal, zone entry, etc)
2. Unlimited sampling!
3. Event plotting & heatmaps!
Thank you!

www.liu.se