

プレイデータからのサッカーの速報テキスト生成

谷口 泰史[†]高村 大也^{‡§}奥村 学[‡]東京工業大学工学院[†], 東京工業大学科学技術創成研究院[‡], 産業技術総合研究所[§][†]yasufumi@lr.pi.titech.ac.jp, [‡]{takamura, oku}@pi.titech.ac.jp

1 はじめに

IoT サービスの発達により, 大規模かつ複雑なデータを処理するエッジコンピューティングへの需要が高まっている. 大規模なデータから有用な知見を得るため, 重要なデータを抽出し, 状況を説明する速報テキストを自動的に生成する必要がある. 速報テキストの自動生成では, 多種多様なデータから言及すべきデータを選択する必要があり, 状況によって着目するデータは変化する. さらに着目したデータによって表される事象を解釈し, テキスト化しなければならない. 本研究では, サッカー選手のプレイデータから速報テキストを生成する研究課題に取り組む. 現状サッカーの速報テキストはプロチームの試合などにのみ人手で付与されているが, 自動生成によりアマチュアチームの試合に速報テキストを付与することや, プロチームの試合でも選手毎や, より詳細な速報テキストを生成可能になる.

テキスト生成課題で広く用いられる encoder-decoder[6] モデルでは, 選手名やチーム名のような多様な入力データに対する低頻度な単語とプレイ内容を説明する単語を区別することなく出力している. そのためモデルは低頻度な単語に強く影響を受け, プレイデータについて正しく言及するようなテキストが出力できないという問題点があった. そこで, 提案手法では出力となる速報テキストにおける言語表現を汎化した上で学習し, それらの復元方法も同時に学習することで, プレイデータにより即した出力結果が得られた. さらに, 提案モデルでは, ルールベースでは復元が困難な表現を含む速報テキストも生成可能にした.

2 関連研究

スポーツのプレイデータから速報テキストを自動生成する研究はいくつか行われている. 例えば, Chenら [2] は, Robocup のサッカーのシミュレーションゲームによって得られるロボットのプレイデータと速報テキストデータのペアを用いて, 自動生成を行う研究に取り組んでいる. この研究で用いられているプレイデータは, 選手と, プレイの内容と発生時間という限られた入力である. 一方本研究では, 前述した情報に加えて, 選手のフィールド上での位置や, ボールのスピードなど詳細なデータが与えられる. モデルが

表 1: プレイデータが提供する情報とその具体例

種別	具体例	値
選手名	岡崎慎司	離散値
プレイ内容	パス, シュート	離散値
時間	30:20	連続値
ボールの xy 座標	10,21	連続値
プレイの成否	1 or 0	連続値
詳細情報	ボレー, ロングパス	離散値

扱う情報が増えることで, プレイ内容を詳細に言及可能になることが期待される. 今後センサの利用が普及して選手のプレイから得られる情報の種類が増加することが想定され, 多様な情報をどのように活用するかは重要な問題になってくる. さらに Chenら [2] の研究では統計的機械翻訳の手法を用いているが, 本論文では, 近年言語生成系課題で頻繁に使用されている encoder-decoder[6] を使って課題に取り組んだ.

Wisemanら [7] は様々な data-to-text 課題の難易度の調査や, 従来の研究課題よりも内容に即したテキストを生成するのが困難なデータセットを使った課題の提案を行っている. 提案されているデータセットは, バスケットボールの試合のスコアや選手のアシスト数などの情報から, その試合の要約を生成することを想定したものであり, 課題を提案し一般的な手法で比較実験を示している. この研究によれば, テキスト自動生成に最も有用であったモデルは, encoder-decoder[6] をベースにしたモデルであるが, その拡張は機械翻訳や自動要約の手法を利用したモデルであり, data-to-text のために設計されたモデルではない. それにより選手名やチーム名といった低頻度な単語の影響で, プレイデータに即した速報テキストが出力できないという課題がある. 本研究では encoder-decoder モデルに対して, 汎化した言語表現を復元する手法を取り入れたことで, data-to-text に即した速報テキスト生成手法となっている.

3 サッカーのプレイデータ概要

本研究では, Premier League¹ の 15/16 シーズンの全試合のデータを利用した. 実験で主に利用したデータセットは, 試合中にボールを持った選手のアクション (パスやシュートなど) を時系列で記述するプレイ

¹<https://www.premierleague.com/>

表 2: 汎化タグの変換例

変換前	変換後
岡崎慎司, ルーニー	player
レスター, マンチェスター・ユナイテッド	team

データと, Opta 社によって各試合に時系列に付与された速報テキストデータから成る. プレイデータから得られる情報は表 1 に示す通りであり, プレイ内容は 70 種類, 詳細情報は 298 種類の ID によって分類されている. 現状これらの情報を自動取得することは困難であるが, プレイの成否や詳細情報の一部を除けば, センサ技術の導入により自動化が期待できる. そこで本研究では, これらの情報が自動取得可能なものとし, すべてを入力として用いた. 速報テキストデータは, 速報テキストとその生成時間の 2 つの情報からなっている. 速報テキストデータにはプレイ内容に言及するテキスト以外に, 試合の開始や終了, 出場選手の説明をするテキストが含まれており, それらは取り除いた.

プレイデータと速報テキストデータの対応はデータセット中には記述されていないので, 速報テキストの生成時間の前後 5 分以内であり, 速報テキスト中に含まれる選手に関するプレイデータと速報テキストデータの対応付けを行った. 上述のルールにより, 1 つの速報テキストデータに対して, 複数のプレイデータが対応したパラレルデータとなり, 速報テキストの発話時刻, 速報テキストで言及すべき選手は所与であるという設定の課題に取り組む.

4 提案手法

本研究で用いた encoder-decoder モデルはプレイデータ列を入力系列 $X = \{x_1, x_2, \dots, x_n\}$ として encoder に入力し, その出力から速報テキスト $Y = \{y_1, y_2, \dots, y_m\}$ を decoder に予測させるモデルである. ここで, x_i はプレイデータ, y_i は速報テキスト中の単語であり, n は入力するプレイデータの総数で, m は速報テキストに含まれる単語数である. 関連研究である村上ら [9] の研究の encoder-decoder モデルに基いてモデルの構築を行う. encoder には, 村上ら [9] の研究で最も有用な結果を出している多層パーセプトロン (multilayer perceptron; MLP) を利用する. decoder には, 機械翻訳 [1] 等の decoder として用いられている再帰型ニューラル言語モデル (recurrent neural network language model; RNNLM) を利用する. さらに Luong ら [4] の研究で提案されている注意機構と input feeding を用いる.

4.1 プレイデータの encode 方法

encoder への入力列 $X = (x_1, x_2, \dots, x_n)$ を式 (1) によりベクトル表現へと変換する. x_i は表 1 に示すプレイデータを表す:

$$p_i = f(x_i). \quad (1)$$

f は, 表 1 に示した値が離散値の場合, それらに対応する埋め込み行列によって埋め込みベクトルに変換し, 連続値の場合は標準化し, それらを列方向に結合しベクトルに変換する関数である. ただし詳細情報については, Bag-of-Features で表現したベクトルを MLP に入力し, 得られたベクトルを用いる. 各 x_i を関数 f によりベクトル表現に変換したものを $P = (p_1, p_2, \dots, p_n)$ とすると, encoder の計算は以下のようになる:

$$h = MLP([p_1; p_2; \dots; p_n]). \quad (2)$$

ここで $[]$ はベクトルの列方向の結合を表す. MLP はバッチ正規化 [3] を含んだ多層パーセプトロンである. 計算された h を decoder に引き継ぎ, 速報テキストの生成を行う. Yang ら [8] が encoder に出力文の出現単語を予測させることで, 精度の向上を報告している. 本研究では, encoder の出力 h から出力文の内容語を予測させる, 式 (3) に示すモデルを構築する:

$$g_c = \sigma(U_c h + b_c). \quad (3)$$

ただし, U_c は重みパラメータ, b_c はバイアス項, $\sigma(\cdot)$ はシグモイド関数とする. 損失関数は正解ラベルを t_c とすると,

$$loss_1 = - \sum t_c \log g_c \quad (4)$$

となる交差エントロピーを計算する. $loss_1$ は全体の損失関数の一部として用いる.

4.2 注意機構と汎化タグの復元

decoder は encoder の出力 h とプレイデータのベクトル列 P から RNNLM を用いて, $Y = (y_1, y_2, \dots, y_m)$ を予測する. m は出力文の文長を表す. s_j を decoder の j 番目の隠れ層の出力, c_j を注意機構によって得られた文脈ベクトルとすると,

$$\tilde{s}_j = \tanh(W_c [s_j; c_j]) \quad (5)$$

となり, W_s を重みパラメータとして,

$$p(y_j | y_{<j}, h, P) = \text{softmax}(W_s \tilde{s}_j) \quad (6)$$

となる. 注意機構はプレイデータ x_i を f によってベクトル表現に変換した系列 $P = (p_1, p_2, \dots, p_n)$ に対して用いる. すなわち, s_j を decoder の隠れ層の出力とすると

$$a_{ij} = \frac{\exp(\text{score}(p_i, s_j))}{\sum_{k=1}^n \exp(\text{score}(p_k, s_j))} \quad (7)$$

として, プレイデータのベクトル表現と decoder の隠れ層の出力とのアライメントの確率を計算する. score 関数は Luong ら [4] の MLP² を用いる.

本研究で用いた汎化タグの復元手法の概要を図 1 に示す. 本研究での, 速報テキスト中で汎化タグに変換

²Luong らの論文では *concat* として参照されている.

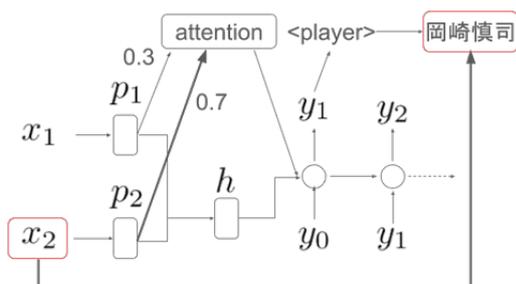


図 1: 提案手法の概要

する単語について説明する．表 2 に示す例に基づき，選手名やチーム名といった多様な入力を，速報テキスト中での出現順に番号付けし汎化タグに変換する．汎化タグを導入することで，速報テキスト中の選手名やチーム名のバリエーションが減り，学習が容易に進むことが期待される．汎化タグにより decoder はプレイ内容を説明するテキスト生成を学習し，選手名やチーム名の生成を学習する必要がなくなる．

一方で，生成文に含まれる汎化タグを選手名やチーム名に復元しなければならない．しかし，入力は複数のプレイデータから構成されており，汎化タグと置き換える候補は複数存在する．そこで本研究では，汎化タグの生成時にモデルの注意機構が着目しているプレイデータを，そのタグと最も関連の強いプレイデータとみなし，そのプレイデータが含む選手名やチーム名で汎化タグを置き換える．具体的には，decoder が汎化タグを出力した際，注意機構の a_{ij} の値が最大となるプレイデータが含む選手名やチーム名に汎化タグを置き換える．

また，汎化タグの正しい復元を学習するため，式 (8) を損失関数に加える：

$$loss_2 = \sum_{i=1}^n \sum_{k \in U} a_{ik}. \quad (8)$$

ここで， U は汎化タグに置き換えられた選手名・チーム名を含まないプレイデータに対するアライメントスコアの集合である．この損失関数の導入により，プレイデータと速報テキストのアライメントをとることなく，汎化タグの生成時に注意機構が着目すべきプレイデータを学習できる．最終的には，単語の出力の交差エントロピーに $loss_1$ と $loss_2$ を加えた損失関数を最小化する．

4.3 多様な汎化タグの導入

1 種類の汎化タグの場合，すべてのプレイ内容の速報テキストを生成する際に，モデルは同じ汎化タグを出力する必要があり，学習が困難になることが考えられる．本研究では，表 3 に示した名詞と動詞を含む速報テキスト中の選手名を，表 3 の汎化タグで置き換える．実際に使用したタグは表 3 に示す 13 種類である．

4.4 単語の結合

本研究の速報テキスト生成では，使用したプレイデータと速報テキストの総数が 17,622 件であり，学

表 3: 汎化タグの種類

汎化タグ	速報テキストに含まれる名詞/動詞
player-save-from	save from
player-fires	fires
player-cross	cross
player-free-kick	free-kick
player-forced	forced
player-caught	caught
player-release	release
player-shot	shot
player-chance	chance
player-dribbles	dribbles
player-shoots	shoots
player-goes-close	goes close
player-other	上記以外

習データセットに対して過学習することが想定される．具体的には，学習データセット中に頻出する表現のみを含む速報テキストを出力することが考えられる．実際に出力結果の中で，プレイデータの内容に関わらず，“is replaced”と “is booked”といった表現が多く見られる．“is”の頻度が高く，“replaced”と “booked”の頻度も高いので，モデルが “is” を予測すると “replaced” か “booked” が予測されている．そこでそれらをまとめることで，“is” に続く他の単語も予測されやすくなると考え，学習データセット中の頻度が上位 100 件の単語 bigram を 1 つの単語として扱う．

5 実験

5.1 実験設定

実験には，データセットの中から選手のプレイについて言及する速報テキストのみを抽出し，ルールベースによってプレイデータとアライメントをとった計 17,622 件を利用した．そのうち 14,034 件を学習データ，1,832 件をテストデータ，1,756 件を開発データとした．速報テキストの形態素解析は nltk³，モデルの実装は Chainer⁴ を使用した．実験では，速報テキストの前後 5 分以内に発生した，速報テキスト中の選手名を含むプレイデータの最大 5 件を用いた．5 件以上ある場合は，発生時間が速報テキストの生成時間に近い順に使用した．

パラメータの最適化には Adam を利用し，勾配クリッピングは 5 とした．epoch 数は，最大を 200 とし，開発データの精度が 3 回連続で下がった場合に学習を早期終了した．表 1 の離散値に対する埋め込みベクトル次元は 16，encoder の隠れ層 h の次元は 200，decoder の単語埋め込みベクトルは 128 次元とした．

評価尺度には，モデルが生成した速報テキストと正解のテキストの一致度を計る目的で BLEU[5] を用いた．自動評価の実験では，汎化タグの復元の効果を

³<http://www.nltk.org/index.html>

⁴<https://chainer.org/>

表 4: ベースモデルと拡張した汎化タグの復元モデルの BLEU スコアの比較

文長	base	tag	multi-tag	concat	both
full	0.11	0.51	0.60	0.69	0.57
20	0.23	1.83	1.84	1.90	1.93
15	0.28	3.87	3.91	4.07	4.13
10	0.45	7.46	7.23	6.79	6.76

検証するために、ベースモデル (base) と汎化タグの復元 (tag) を加えたモデルの比較を行った。またタグの多様化と単語結合の有用性を検討するため、汎化タグの復元モデルに対して、タグの多様化 (multi-tag) と単語結合 (concat) のそれぞれを加えた場合と、両方を加えた場合 (both) の実験を行った。また人手評価では、各モデルの出力の文法性と、正解文との類似度を 1-3 段階で 100 例に対して評価を行った。人手による評価には Amazon Mechanical Turk⁵を用いた。

5.2 実験結果

ベースモデルと汎化タグの復元モデルとその拡張モデルの比較結果を表 4 に示す。データセット全体での自動評価の場合精度が低いので、テストデータの文長を {10,15,20} に制限した場合の結果も示した。base モデルと tag モデルを比較すると大きな差が見られた。base モデルでは意味をなさない速報テキストが多く見られる一方で、tag モデルでは非文は少なく、速報テキストの意味を解釈できた。

表 4 に示した汎化タグの拡張モデルとの比較では、BLEU のスコアに大きな差は見られなかったが、both モデルが文長 20 と文長 15 に制限した場合に良い精度が得られた。実際の出力を比較すると速報テキストの内容は both モデルがプレイデータに即した出力となっていた。tag モデルと both モデルの速報テキストの生成例と正解文を表 5 に示す。上段の例は、Harry Kane がペナルティーエリアまでボールを運んだときの速報テキストである。tag モデルではゴールラインに対する言及は見られるが、文としての意味が理解できない文であった。一方の both モデルは、具体的な状況に対する説明はないものの、Harry Kane が得点を試みているという点に関しては正しく言及できている。表 5 の下段の例は、Odion Ighalo という選手がシュートをし、それがキーパーに止められたというプレイに対する速報テキストである。tag モデルで出力された速報テキストはプレイ内容を無視し、文としても意味をなしていない。both モデルでは、正解文と比べて表現方法が違っていたが、プレイに対しては正しく言及できていた。

人手評価の結果を表 6 に示す。tag モデルと比較した際、both モデルと multi-tag モデルについては文法性、concat モデルについては類似度に有意差があった ($p < 0.05$)。

表 5: tag モデルと both モデルの速報テキストの比較

Model	速報テキスト
正解出力	Harry Kane goes down inside the penalty area
tag	Harry Kane is the latest player to the byline
both	Harry Kane is trying to score for Tottenham Hotspur
正解出力	huge chance missed by Odion Ighalo
tag	Odion Ighalo is the ball into the box and trying
both	Odion Ighalo has a shot blocked

表 6: 汎化タグの復元モデルとその拡張の人手評価

model	文法性	類似度
tag	1.99	1.60
multi-tag	2.12	1.53
concat	2.07	1.69
both	2.13	1.67

6 結論

本研究では、スポーツの速報テキスト生成の例として、サッカーのプレイデータからの速報テキスト生成手法を提案した。スポーツの速報テキストでは、例えばゴールというプレイに対しても、選手名やチーム名の違いによって内容が異なるといった特徴がある。本研究では、プレイ内容に応じた汎化タグを導入し、その復元方法も学習することにより、多様なタグの導入を可能にした。また実験によって、復元方法も同時に学習する汎化タグ手法の有用性を示した。

謝辞

本研究は JST さきがけ JPMJPR1655 の支援を受けたものです。

参考文献

- [1] Dzmitry Bahdanau, Kyunghyun Cho, and Yoshua Bengio. Neural machine translation by jointly learning to align and translate. *arXiv:1409.0473*, 2014.
- [2] David L. Chen and Raymond J. Mooney. Learning to sportscast: a test of grounded language acquisition. In *ICML*, pp. 128–135, 2008.
- [3] Sergey Ioffe and Christian Szegedy. Batch normalization: Accelerating deep network training by reducing internal covariate shift. In *ICML*, pp. 448–456. JMLR.org, 2015.
- [4] Thang Luong, Hieu Pham, and Christopher D. Manning. Effective approaches to attention-based neural machine translation. In *EMNLP*, pp. 1412–1421, 2015.
- [5] Kishore Papineni, Salim Roukos, Todd Ward, and Wei-Jing Zhu. Bleu: a method for automatic evaluation of machine translation. In *ACL*, pp. 311–318. Association for Computational Linguistics, 2002.
- [6] Ilya Sutskever, Oriol Vinyals, and Quoc V. Le. Sequence to sequence learning with neural networks. *arXiv:1409.3215*, 2014.
- [7] Sam Wiseman, Stuart M. Shieber, and Alexander M. Rush. Challenges in data-to-document generation. In *EMNLP*, pp. 2253–2263, 2017.
- [8] Zhilin Yang, Ye Yuan, Yuexin Wu, William W. Cohen, and Ruslan R Salakhutdinov. Review networks for caption generation. In *NIPS*, pp. 2361–2369. Curran Associates, Inc., 2016.
- [9] 村上聡一朗, 渡邊亮彦, 宮澤彬, 五島圭一, 柳瀬利彦, 高村大也, 宮尾祐介. 時系列数値データからの概況テキストの自動生成. 言語処理学会 第 23 回年次大会, pp. 1117–1120, 2017.

⁵<https://www.mturk.com/>