

Seithiyalan – An iPhone Application for Cricket Summary Generation in Tamil

S. Thendral, R. Subhashni, V. Madhan Karky

Abstract--- Cricket plays vital role in Asian sports. Thus, there is a need for an application which summarize cricket in Natural Language. The objective of Seithiyalan application is to auto analyze the score card of the cricket match and generate summary in Tamil. The data analyzed based on the parameter Interestingness to select relevant pattern. Based on that pattern only the summary will generate.

Keywords: Cricket Summary, NLG, Machine Learning, Tamil Computing.

I. INTRODUCTION

Nowadays people interested in read the news or summary through their mobile apps only instead of view in papers or websites. Few applications are available with cricket summary, but they are providing only in English. No application available now in Appstore with Tamil summary. Thus, this Seithiyalan application provide solution by generating summary in Tamil for cricket match in mobile application.

II. SUMMARY GENERATING COMPONENTS

The major components of the Seithiyalan application are as follows:

Data mining and modelling

Particular selected cricket matchdata in internet can be gathered and modeled as feature vectors which are predefined. Then the association rules identified by using the Apriori algorithm. The consistency of player is calculated by obtain correlation of variance against average of that player.

Calculating Interestingness

The interestingness calculated depend on the weightage given to the factors like Team chasing score, fall of wickets, Team previous history, Individual players records, Team records, Team run rate, Individual players high run rate, Player's centuries, remaining overs and wickets of winning team, quick fifties or centuries or wickets.

Generating sentence

The morphological generator generates the relevant Tamil words based on nannool rules. It takes root word from input as pakhuthi and relevant idainilai, santhi, saariyai and vikhuthi areadded. Then the sandhi checker used to correct

the sandhi errors based on sandhi rules. The spell checker used to check spellings based on Tholkapiyam and nannool Tamil grammar rules and existing words in Tamil dictionary. Then based on calculated interestingness factor, the events are considered to add in the summary. Based on that events the relevant sentence generated. That sentences formed the summary.

III. EVALUATION

The summary generated by Seithiyalan is evaluated by calculate the similarity score with the summaries written by human beings. Both the summaries are compared with the mentioned event and noun parameters in each sentence. The noun, numeral and action in both human and automated summaries are compared and the score given based on that. The noun may be person or team. The numeral may be the run or wicket or over, that is any numeric value. The action is the event conducted by the noun to connect the numeral.

IV. IMPLEMENTATION

The home screen of Seithiyalan application consists of three options. The user can choose recent games, match between two countries or all matches. If the user selects recent games, then the duration list will appear. It contains up to last seven days, last month, last year options. If the user chooses the option, then the relevant list of matches will appear. After selecting the particular match, the cricket summary generates and appear in Tamil in output window. If the user selects match between two countries, then all countries list for team A will appears. After one country selects by user, the remaining countries list again appears for team B. After selects both the countries, the available matches between that countries list will appear. After selects the particular match, the summary will generate in the output window. If the user selects all matches option in home screen, the next window with two options will come. One is duration and the second is particular countries. The user can select any one or both or none of the options. If the user select duration or between two countries means the relevant windows as same as home screen will display. If the user chooses both means the next screen contains both duration list, team A country list and team B country list. After select all and submit, the matches list between selected countries for selected period will appear. After select the match, particular URL can access *espnricinfo* website for obtain score card as input. The input handled in the XCode to design feature from the database. The realm used to handle the data in iPhone.

Revised Version Manuscript Received on 22 February, 2019

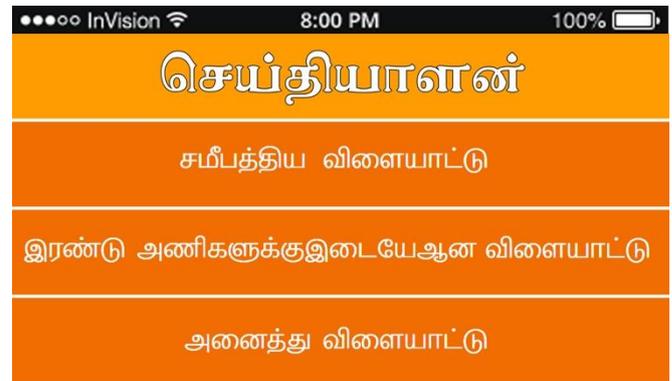
S. Thendral, Research Scholar, School of Computing, Sathyabama Institute of Science and Technology, Chennai, Tamil Nadu, India
(e-mail: thendraldots@gmail.com)

R. Subhashni, Professor, School of Computing, Sathyabama Institute of Science and Technology, Chennai, Tamil Nadu, India (e-mail: subhaagopi@gmail.com)

V. Madhan Karky, Karky Research Foundation, Chennai, Tamil Nadu, India
(e-mail: inbox@karky.in)

The needed sentence pattern chooses from database for relevant noun and event and interestingness score. The interesting score for particular noun and event is save in the array which is sorted based on score. Thus, the high score statements displayed as the output. Then the user can share the summary in social media like Facebook, Twitter or google plus.

V. SCREENSHOTS OF APP



இங்கிலாந்தும் வெஸ்ட் இண்டீஸும் மோதிய ஒருநாள் போட்டி வெஸ்ட் இண்டீஸில் 20/02/2019 அன்று நடந்தது . டாஸ் வென்ற வெஸ்ட் இண்டீஸ் அணி முதலில் பேட் செய்து 360 ரன்களை 8 விக்கெட்கள் இழந்து எடுத்தது . கெயில் புதிய சாதனையுடன் சதம் அடித்தார். இங்கிலாந்து அணி 364 ரன்களை எடுத்து 6 விக்கெட்கள் வித்தியாசத்தில் வென்றது. ஜேசன் ராய் மற்றும் ஜோ ரூட் சதம் அடித்தனர்.



Figure 2: Home screen

VI. CONCLUSION

We proposed an application which generate cricket summary in Tamil automatically. It will enhance by adding more Interestingness factors or updating more sentence patterns. It will further be extended to implement in multiple other languages. The application's next level can be generating summary to multiple sports.

REFERENCES

1. Tholgappiar, "Tholgappiam," KazhagaVeliyeedu, Madras, 1973.
2. Alice Oh and Howard Shrobe, "Generating baseball summaries from multiple perspectives by reordering content," in Proc. 5th International Natural Language Generation Conference, 2008, pp. 173-176.
3. Bavananthymunivar, "Nannool," Pary Nilayam, Madras.



Figure 1: Sample summary text

