Attention-Based Bidirectional LSTM With Differential Features For Disk RUL Prediction

Anhui Bai, Mingjin Chen, Siyuan Peng, Guojun Han, Zhijing Yang
School of Information Engineering,Guangdong University of Technology

Introduction

- The LSTM-based methods have achieved the good prediction accuracy for RUL in HDD, there are several challenges:

  1) Some of the SMART features are blank, and are nearly constant which contain minimal deterioration information that only a few properties undergoing significant change prior to failure.

  2) For the long sequence data, the LSTM-based approaches frequently lose the essential historical information,because only the output of the last time step is used when employing an LSTM network for the RUL prediction.

- The main contribution of this paper are summarized as follows:

  1) The differential features is adopted to compensate the limitation of the SMART features.

  2) A bidirectional LSTM is used to forecast the output at time step by incorporating the data change pattern before and after that time step.

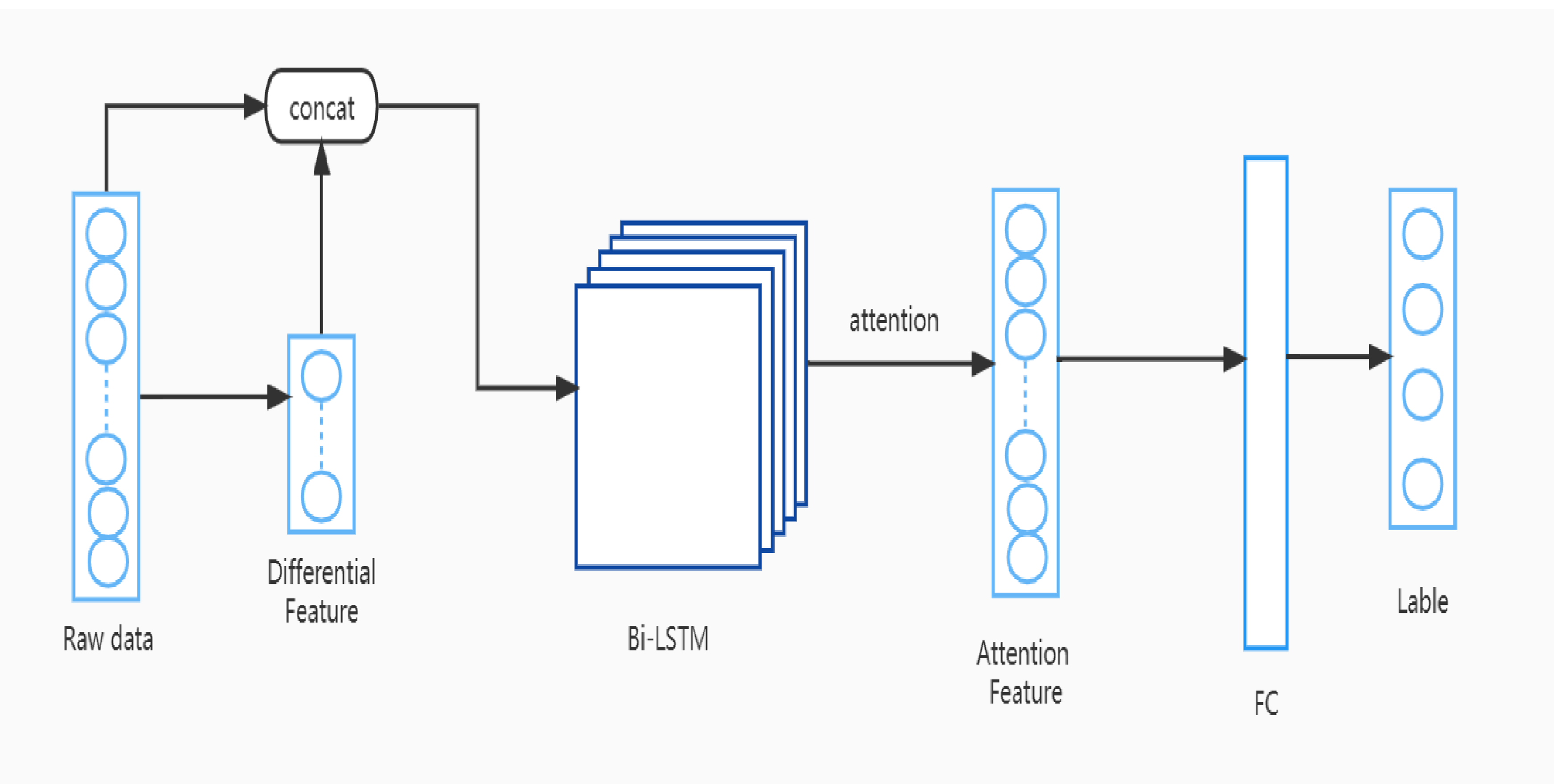
  3) Based on the attention mechanism, different weights for different features are assigned at different time step,resulting in greater attention being paid to the more significant features.

Methods

SMART ID	Feature Name
1	Raw Read Error Rate
3	Spin Up Time
5	Reallocated Sectors Count
7	Seek Error Rate
9	Power On Hours
187	Reported Uncorrectable Errors
189	High Fly Writes
194	Temperature Celsius
197	Raw Current Pending Sector Count

Feature Engineering:

- Several differential features are extracted base on the SMART features.
- The differential features can eliminate the random trends contained in time-series data to make it easier to discover the relationship between the SMART feature values and the remaining useful life.



Proposed framework:

- The raw sequence data are concatenated with the differential features as the input sequence.
- The input sequence are transferred to the bidirectional LSTM network for feature learning.
- The weighted sum of the hidden states is passed to the fully connected layers.
- The FC layers maps the feature dimension to the output size which indicates the results of the RUL prediction of HDD.

Results

- The Bi-LSTM with differential features and attention outperforms the other varients of Bi-LSTM.

Methods	ACC	ACC_G	ACC_F	ACC_G^{TOL}	ACC_F^{TOL}	FDR	FAR
Bi-LSTM	91.74%	87.51%	92.46%	96.01%	98.23%	95.22%	3.89%
Bi-LSTM+Attention	92.61%	89.10%	93.21%	96.74%	98.46%	95.68%	3.34%
Bi-LSTM+differential features	96.20%	95.79%	96.28%	98.98%	99.18%	97.64%	1.66%
Bi-LSTM+Attention+differential features	**96.32%**	**96.21%**	**96.33%**	**99.01%**	**99.20%**	**97.83%**	**1.74%**

- Our proposed method performs better than the origin LSTM under all the metrics.

Methods	ACC	ACC_G	ACC_F	ACC_G^{TOL}	ACC_F^{TOL}	FDR	FAR
LSTM	87.15%	83.37%	87.79%	93.55%	96.77%	92.46%	6.76%
Bi-LSTM+Attention+differential features	**96.32%**	**96.21%**	**96.33%**	**99.01%**	**99.20%**	**97.83%**	**1.74%**

Conclusions

- An attention-based bidirectional LSTM with differential features framework is proposed for the RUL prediction in HDD.
- The differential features and attention mechanism are useful for improving the performance of RUL prediction.
- Our method achieves a 97.83% failure detection rate (FDR) to predict RUL of HDDs up to 60 days before failure.

References

[1] G. S. Babu, P . Zhao, and X.-L. Li, "Deep convolutional neural network based regression approach for estimation of remaining useful life," inProc. Int. Conf. Database Syst. Adv. Appl., 2016, pp. 214–228.

[2] C. Xu, G. Wang, X. Liu, D. Guo and T. -Y . Liu, "Health Status Assessment and Failure Prediction for Hard Drives with Recurrent Neural Networks," in IEEE Transactions on Computers, vol. 65, no.11, pp. 3502-3508, 1 Nov. 2016.

[3] A. De Santo, A. Galli, M. Gravina, V . Moscato and G. Sperl`ı, "Deep Learning for HDD Health Assessment: An Application Based on LSTM," in IEEE Transactions on Computers, vol. 71, no. 1, pp. 69-80, 1 Jan. 2022

www.PosterPresentations.com