Application of NLP Technology in the Information Extraction of Bridge Management and Maintenance Documents

Lu-yang Zhang, Ye Xia

Department of Bridge Engineering, Tongji University, Shanghai 200092, China

Limin Sun

State Key Laboratory for Disaster Reduction in Civil Engineering, Tongji University, Shanghai, 200092, China

Contact: lmsun@tongji.edu.cn

Abstract

Bridge management and maintenance document is the accumulation of authoritative technical data to record the history of bridge operation, current technical status and management and maintenance process, which contains substantial information to support bridge maintenance decision. With more and more extensive application of bridge health monitoring in civil engineering industry, bridge management and maintenance documents become more common and quantitative. Therefore, a large number of these reports need to be manually read and analyzed to obtain effective information, which will waste a lot of effort. In order to improve this situation, this paper develops a frequency state analyzer using LSTM neural network to classify these documents automatically, and improve the efficiency of bridge management and maintenance work.

Keywords: natural language processing; bridge management and maintenance documents; text categorization; LSTM

1 Introduction

Bridge structure management and maintenance is a necessary link to ensure the safe and normal operation of existing bridges. Its management and maintenance information (such as bridge overview, operating environment characteristics, structural safety assessment, special load action assessment, conclusions, suggestions, etc.) is mostly recorded in the document in the form of natural language. The content of the document has the characteristics of comprehensiveness, specialization, and standardization: the content framework is similar, a large number of professional words are used, and the whole content as well as the potential safety hazards are summarized and pointed out in the final chapter. However, with the substantial increase in the number of bridges and management and maintenance documents, the traditional manual reading and analysis method is difficult to meet the engineering needs. Therefore, it is proposed to use natural language processing technology, an emerging technology in the field of computer science, to analyze a large number of management and maintenance documents and

https://doi.org/10.2749/nanjing.2022.1224
Distributed by structurae