Table based KNN for Categorizing Words

Taeho Jo

Department of Computer and Information Engineering, Inha University, 100 Inharo Namgu, Incheon, South Korea

tjo018@inha.ac.kr

***Abstract*—In this research, we propose the table based KNN as the approach to the text categorization. In previous works, we discovered that encoding texts into tables improved the performance in the text categorization, so in this research, become to consider the possibility of encoding words into tables as well as texts. In this research, we encode words into tables where entries are texts and their weights, and apply the table based version of the KNN to the task of word categorization. As the benefits from this research, we expect the better performance and more stability than the traditional version of the KNN, by doing so. Therefore, the goal of this research is to provide the improved approach to the word categorization task.**

***Keyword*—Word Categorization, Table based KNN, Table Similarity**

Taeho Jo (M'97–AM'12) This author became a Member (M) of IEEE in 1997, and an Associate Member (AM) in 2012. He was born in 1970, South Korea. He received his Bachelor degree from Korea University in 1994, his Master degree from Pohang University of Science and Technology in 1997, and his PhD degree from University of Ottawa in 2006. His research area spans mainly over text mining, neural networks, machine learning, and information retrieval. He has the four year experience of working for industrial organizations and ten year experience of working for academic ones. So his research is characterized as the connection from fundamental researches for creating theories and to applied ones for developing products, by his experience of working for both sides.