Journal of Information & Knowledge Management
Vol. 22, No. 4 (2023) 2250074 (17 pages)

DOI: 10.1142/S0219649222500745

World Scientific
www.worldscientific.com

Machine Intelligence for Language Translation from Kashmiri to English

Nawaz Ali Lone*, Kaiser J. Giri† and Rumaan Bashir‡
Department of Computer Science
Islamic University of Science & Technology, Kashmir, India
**lone.nawaz@iust.ac.in*
†*kaiser.giri@islamicuniversity.edu.in*
‡*rumaan.bashir@islamicuniversity.edu.in*

Published 27 August 2022

Abstract. Machine translation (MT) is an emerging research as well as application area in the contemporary world. It is receiving significant attention from academia, industry and corporate houses. A wide range of translation techniques are being applied either individually or in combination for machine translation of different languages across the globe. However, there are still many languages that are either completely missing or poorly visible on the machine translation map. The Kashmiri language is one such language where very little or negligible work has been done related to its machine translation. This paper aims to present a Kashmiri-to-English Machine Translation System and highlight various features of the Kashmiri language. The system is based on machine intelligence having the ability to learn various translation rules from the translated set of input sentences, using Long Short-term Memory (LSTM) architecture for deep sequence learning. The paper also discusses various challenges related to machine translation from Kashmiri to English or other languages. The work presented in this paper is the first of its nature and can serve as a bedrock for research community interested to work on machine translation of Kashmiri language.

Keywords: Machine Learning; Kashmiri; Perso-Arabic; Metric; Encoder; Decoder; Machine Translation.

1. Introduction

Automatic conversion or change of a language for a certain message, while keeping the underlying content or meaning of the message intact, is defined as the process of Machine Translation (MT). It is the branch of computational linguistics which embroils many other fields such as machine intelligence, statistics, linguistics, translation theory, etc. It is one of the most efficacious application areas of Natural Language Processing (NLP) besides sentiment analysis, search engines, Natural Language Generation, voice assistants, report preparation, content understanding, etc. Machine translation tends to equip people by facilitating access to critical information and two-way information exchange in their local languages. Various approaches to machine translation range from simple "word-to-word" substitution to learning translation through a sizeable dataset (Tripathi and Sarkhel, 2010).

† Corresponding author.