

Accelerating Epidemiological Investigation Analysis by Using NLP and Knowledge Reasoning: A Case Study on COVID-19

Jian Wang, PhD, Ke Wang, Jing Li, Jianmin Jiang, PhD, Yanfei Wang, PhD, Jing Mei, PhD, Shaochun, Li
IBM Research, Beijing, China

Abstract

COVID-19 is threatening the health of the entire human population. In order to control the spread of the disease, epidemiological investigations should be conducted, to trace the infection source of each confirmed patient and isolate their close contacts. However, the analysis on a mass of case reports in epidemiological investigation is extremely time-consuming and labor-intensive. This paper presents an end-to-end framework for automatic epidemiological case report analysis and inference, in which a Tuple-based Multi-Task Neural Network (TMT-NN) is designed and implemented for jointly recognizing epidemiological entities and relations from case reports, and an epidemiological knowledge graph and its corresponding inference engine are built to uncover the infection modes, sources and pathways. Preliminary experiments demonstrate the promising results, and we published a real data set of COVID-19 epidemiological investigation corpora at Github, as well as contributing our COVID-19 epidemiological knowledge graph to the open community OpenKG.cn.

Introduction

A novel coronavirus (namely “SARS-CoV-2”) was first detected by China in December 2019¹, and the disease it causes has been named “coronavirus disease 2019” (abbreviated “COVID-19”). By 14 March 2020, there were 142,539 confirmed cases in global 136 countries or territories². Besides COVID-19, SARS in 2003, H1N1 influenza in 2009 and MERS in 2012 are all epidemics with the rapid spread to a large numbers of people. These epidemic diseases are posing a serious challenge to human health in the 21st century, as well as jeopardizing social and economic activities at world wide.

Epidemiological investigation is a critical task to study the distribution, pathogenesis, transmission characteristics and