Investigating the Potential of Topic Modelling On Social Media to Support the Identification of Refugee Needs

Misbah Rizaee

Master of Science in Computer Science (Future Networked Systems)

University of Dublin, Trinity College

Supervisor: Professor Timothy Savage

It is critical that sustainable development goal 10 which ensures equal opportunity and reduces inequalities within and among countries, receives special consideration because according to newly released data from the United Nation at the end of 2021, there are approximately 89 million people who have been forcibly displaced worldwide and nearly 27 million refugees are among them. The United Nation and governments are challenged by the growing refugee population as the crisis has expanded geographically and numerically.

The reactions to refugee flow often deal with practical challenges, such as meeting the basic needs of refugees and organizing the logistics of relocation, but every refugee crisis has its own unique considerations and challenges. Identifying the refugee needs in each refugee crisis has becoming increasingly important and therefore it is a key problem to look into. Equally important is the fact that most existing solutions are post crisis analysis and evaluation which means that the majority of current analysis occurs after the event.

The research discovered that social media and topic modelling have potential to address the current issue in identifying the refugee needs in each refugee crisis as it is occurring. In the past several years, social media sites have played an important role in the area of refugees. In order to overcome the problem outlined earlier, a significant amount of Twitter data was collected through Twitter's API and pre-processed using natural language processing library. Latent Dirichlet Allocation (LDA) was used to classify and identify various topics from the tweets. It has been determined that the latent topics that have been discovered are representative of refugee needs.