

The Rohingya refugee crisis is one of the biggest humanitarian crises of modern times with more than 700,000 Rohingyas rendered homeless according to the United Nations High Commissioner for Refugees. While it has received sustained press attention globally, no comprehensive research has been performed on social media pertaining to this large evolving crisis. In this work, we construct a substantial corpus of YouTube video comments (263,482 comments from 113,250 users in 5,153 relevant videos) with an aim to analyze the possible role of AI in helping a marginalized community. Using a novel combination of multiple Active Learning strategies and a novel active sampling strategy based on nearest-neighbors in the comment-embedding space, we construct a classifier that can detect comments defending the Rohingyas among larger numbers of disparaging and neutral ones. We advocate that beyond the burgeoning field of hate-speech detection, automatic detection of help-speech can lend voice to the voiceless people and make the internet safer for marginalized communities.

* Work done with Shriphani Palakodety and Jaime G. Carbonell,

Accepted at: AI for Social Impact (AISI) track of The Thirty-Fourth AAAI Conference on Artificial Intelligence (AAAI-20) for oral presentation.

Bio: Ashique KhudaBukhsh is currently a Project Scientist at the School of Computer Science, Carnegie Mellon University (CMU). Prior to this role, he was a postdoc mentored by Prof. Jaime Carbonell at CMU. His PhD thesis (Computer Science Department, Carnegie Mellon University, also advised by Prof. Jaime Carbonell) focused on referral networks, an emerging area at the intersection of Active Learning and Game Theory. His Master's thesis at the University of British Columbia (UBC), advised by Prof. Kevin Leyton-Brown and Prof. Holger H. Hoos, focused on automated algorithm design for combinatorial hard problems. During this period, Ashique with his collaborators developed SATenstein, a generalized, highly parameterized SAT solver which has received wide attention from both research and academia. One particular high-impact application of SATenstein was in the multi-billion dollar "incentive auction" conducted by the US Federal Communications Commission to repurpose the radio spectrum from broadcast television to wireless internet.

Ashique's current research focus is in Computational Social Science. In this field, he (with his collaborators) has published the first ML-focused large-scale social media analysis on the Rohingya refugee crisis (to appear in AAAI-2020, AI for Social Impact track).

With three published collections of poems to his credit, an experience of directing music at a New York theater play, occasional dabbling at journalism and column-writing, and a recent success at swimming 50 meters underwater (a navy seal requirement), Ashique enjoys his multiple distractions that keep him away from work.