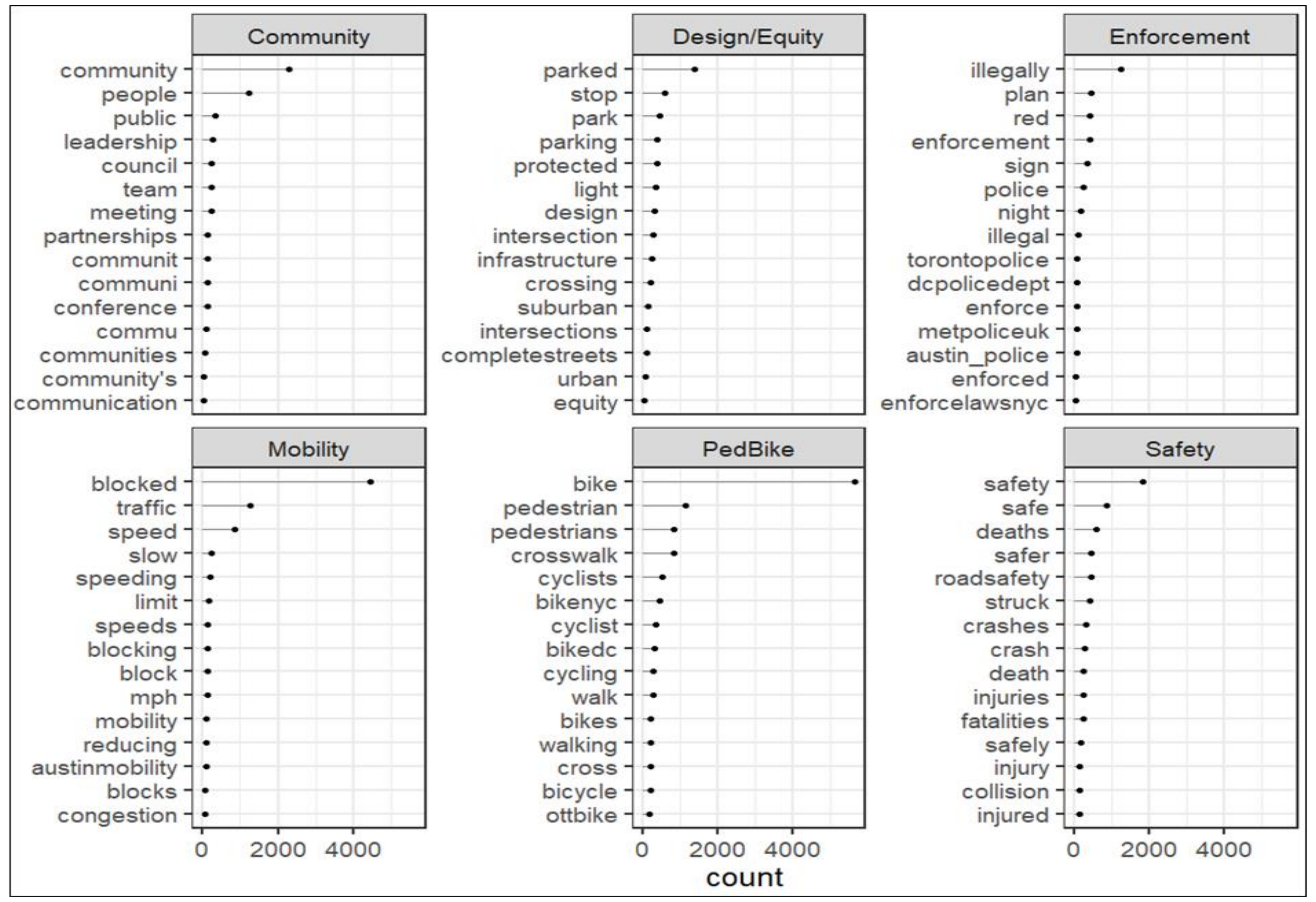
Vision Zero Hashtags in Social Media: Understanding End-User Needs from Natural Language Processing

Subasish Das, Ph.D., Associate Transportation Researcher, Texas A&M Transportation Institute (s-das@tti.tamu.edu)
Kartikeya Jha, Research Associate, Texas A&M Transportation Institute (k-jha@tti.tamu.edu)
Anandi Dutta, Senior Lecturer, Dept. of Computer Science and Engineering, Ohio State University (dutta.34@osu.edu)

Research conducted by
Texas A&M Transportation Institute
The Texas A&M University System
The Ohio State University College of Engineering

Abstract

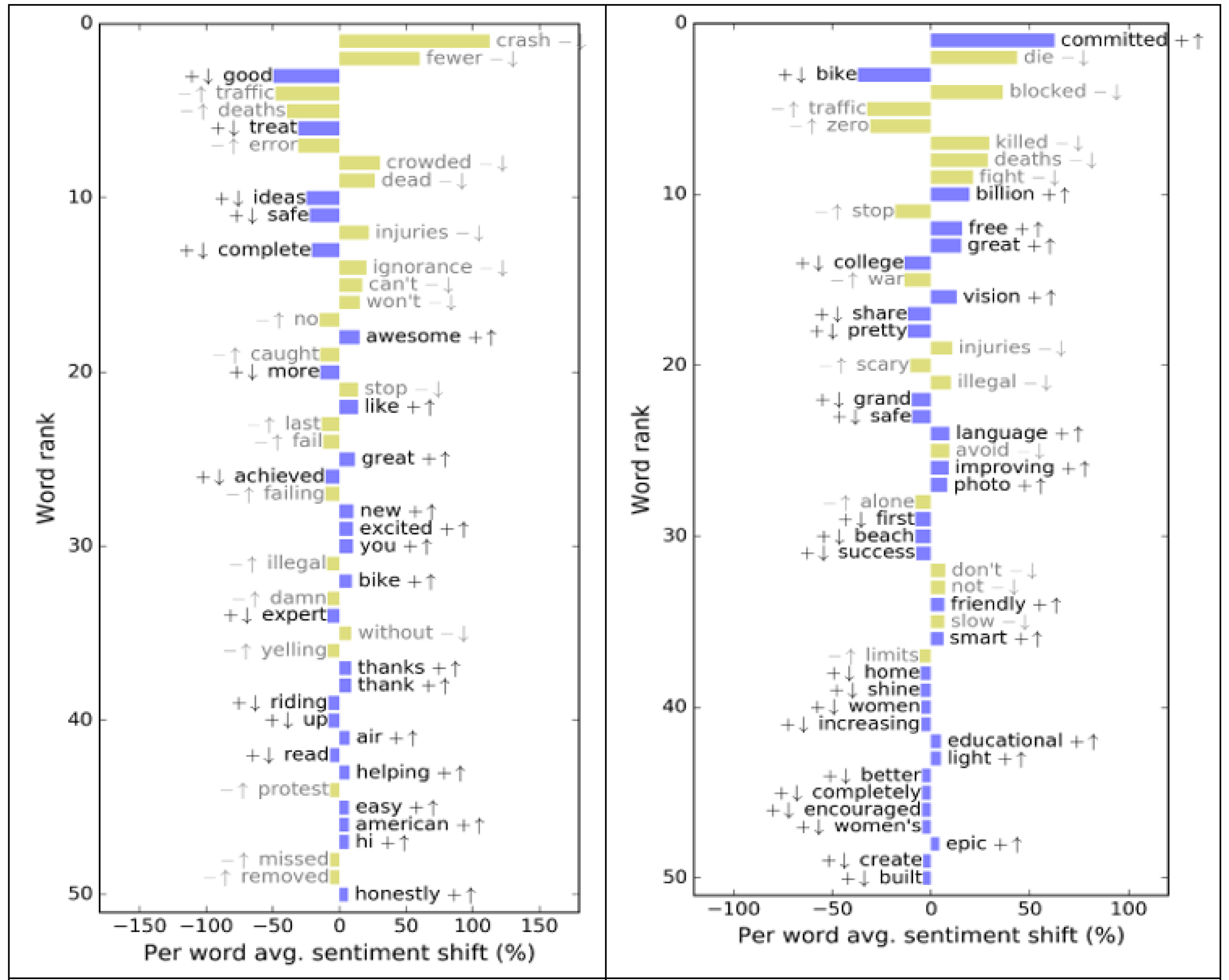
- Vision Zero (VZ) is a global traffic safety policy which promotes road safety by advocating the shared responsibility of designers, policy-makers, and road users to work toward the goal of nullifying preventable fatalities and severe injuries caused by manmade errors.
- This paper investigates the use of the hashtag #VisionZero on Twitter, a popular microblogging platform, to understand public needs and requirements, identify innovative and context-sensitive solutions, and facilitate the exchange of ideas and best practices.
- Results reveal interesting trends, contexts, and patterns with regard to user reactions to road safety needs, issues, and solutions.



Top unigrams by key groups

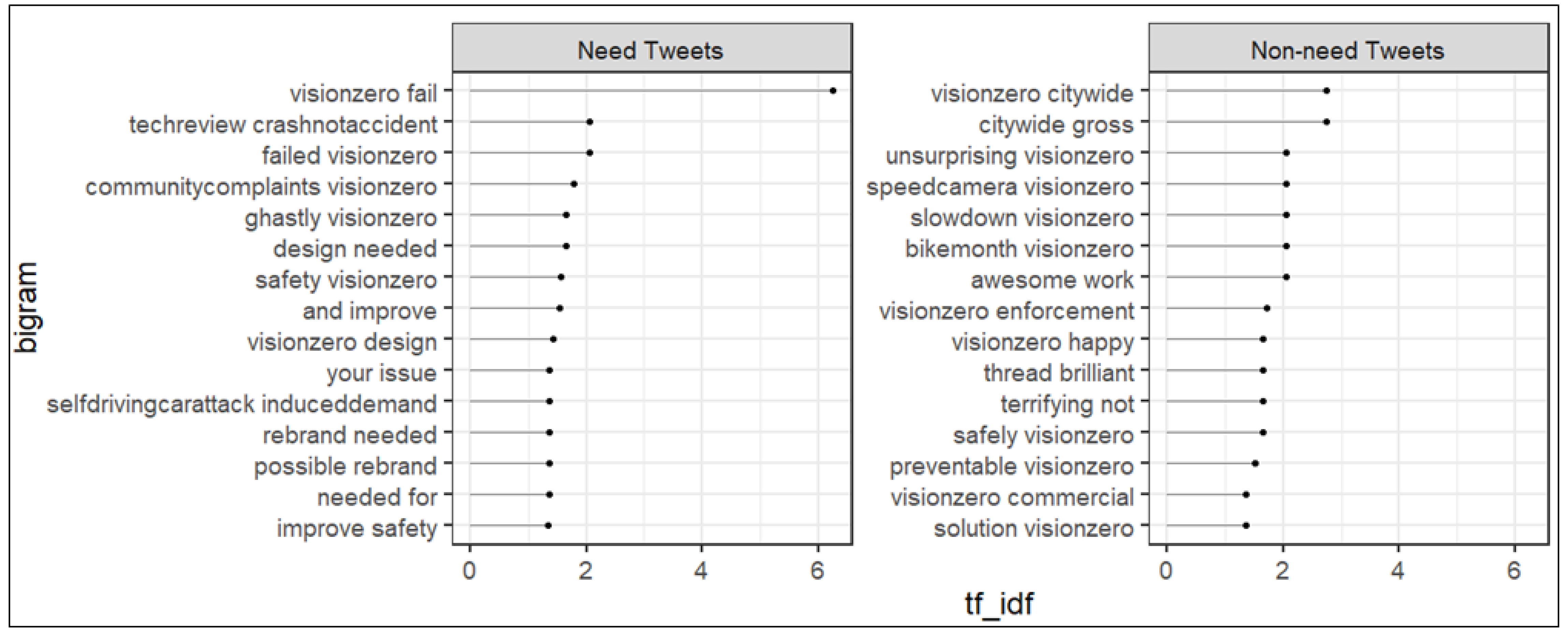
Data Collection

- To collect relevant data, several keywords have been used during the data collection process. The keywords include ‘visionzero’ and ‘zerovision.’ *Used R package: twitteR*
- The dataset was collected for four months in 2019 (March 12, 2019, to July 16, 2019).
- A total of 32,000 unique tweets were collected. These tweets are retweeted 69,450 times. The tweets are associated with approximately 18,283 Twitter handles.



(a) Need related tweets (b) Non-need related tweets

Word shift plots



Top bigrams with high tf-idf for need and non-need tweets

Methodology

- Used natural language processing (NLP) tools to do data cleaning and exploratory data analysis.
- Performed term frequency inverse document frequency (tf-idf) to distinguish between need related and non-need related tweets.
- Developed word-shift measures by using sentiment and emotion measures.

Key Findings

- Identified the significance, frequency, and patterns of word strings to examine the relevance of needs for several different focus areas within VZ.
- Results reveal interesting trends and patterns with regard to user reactions to road safety needs, issues, and solutions.
- The current findings can help policy- and decision-makers in identifying and anticipating the most relevant demands and pressing requirements for different individual areas of interest.