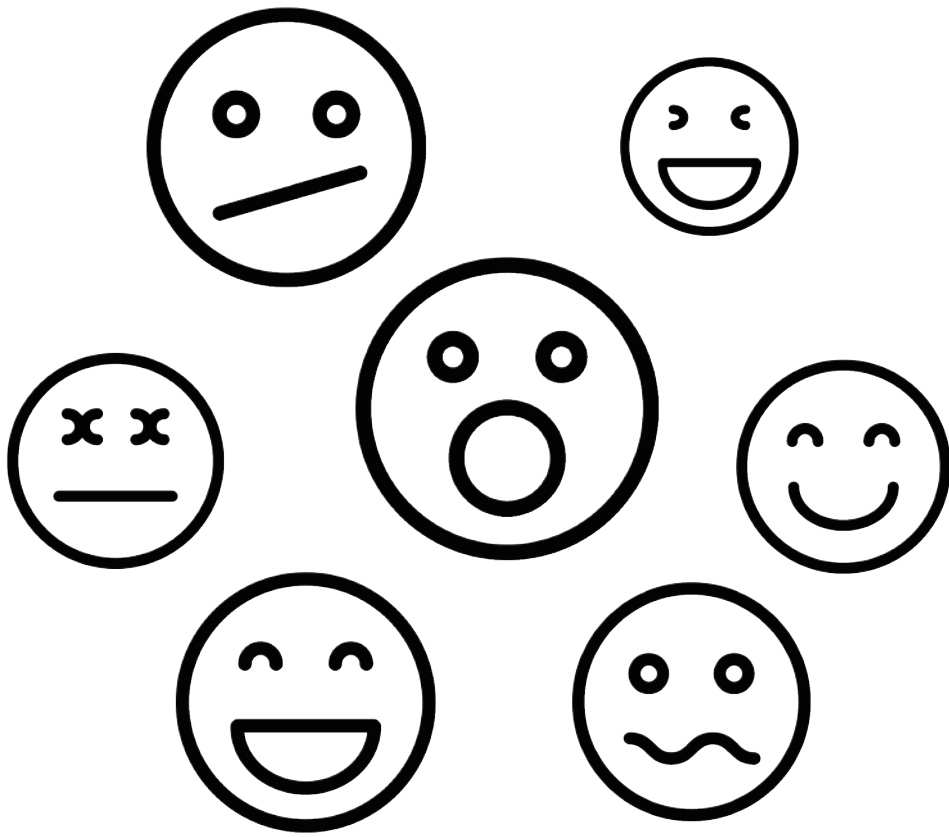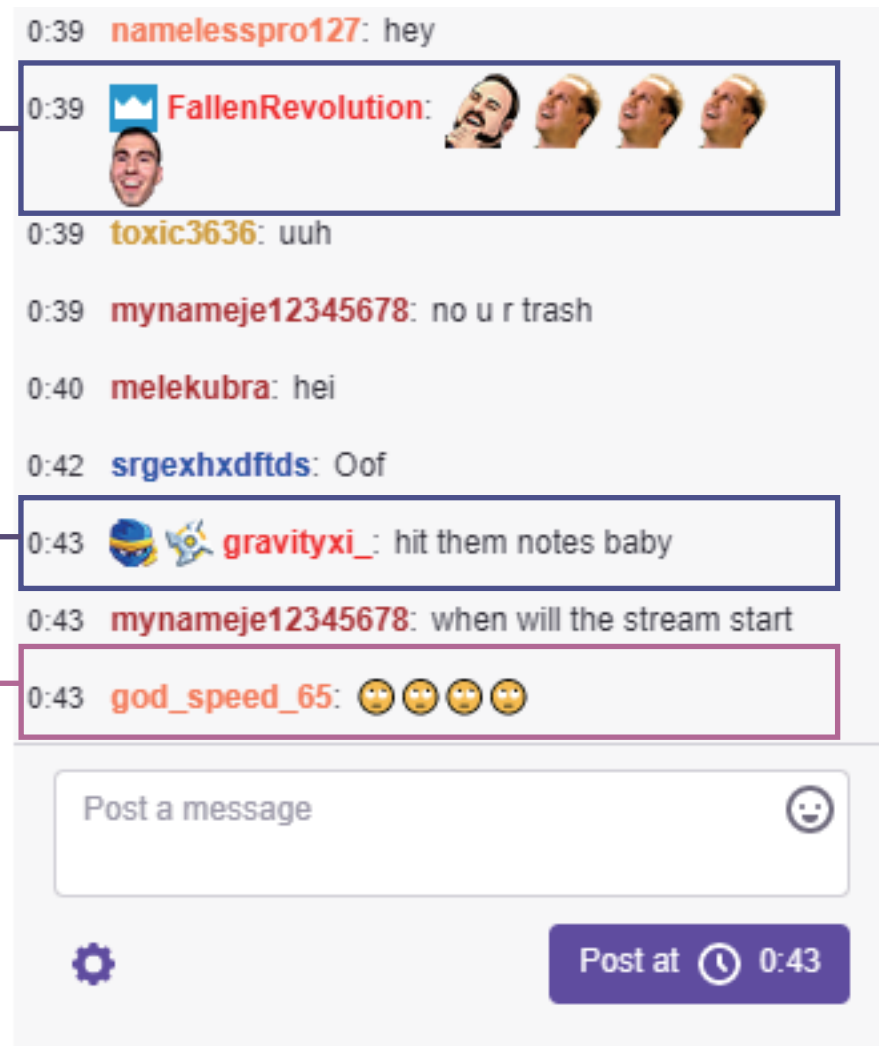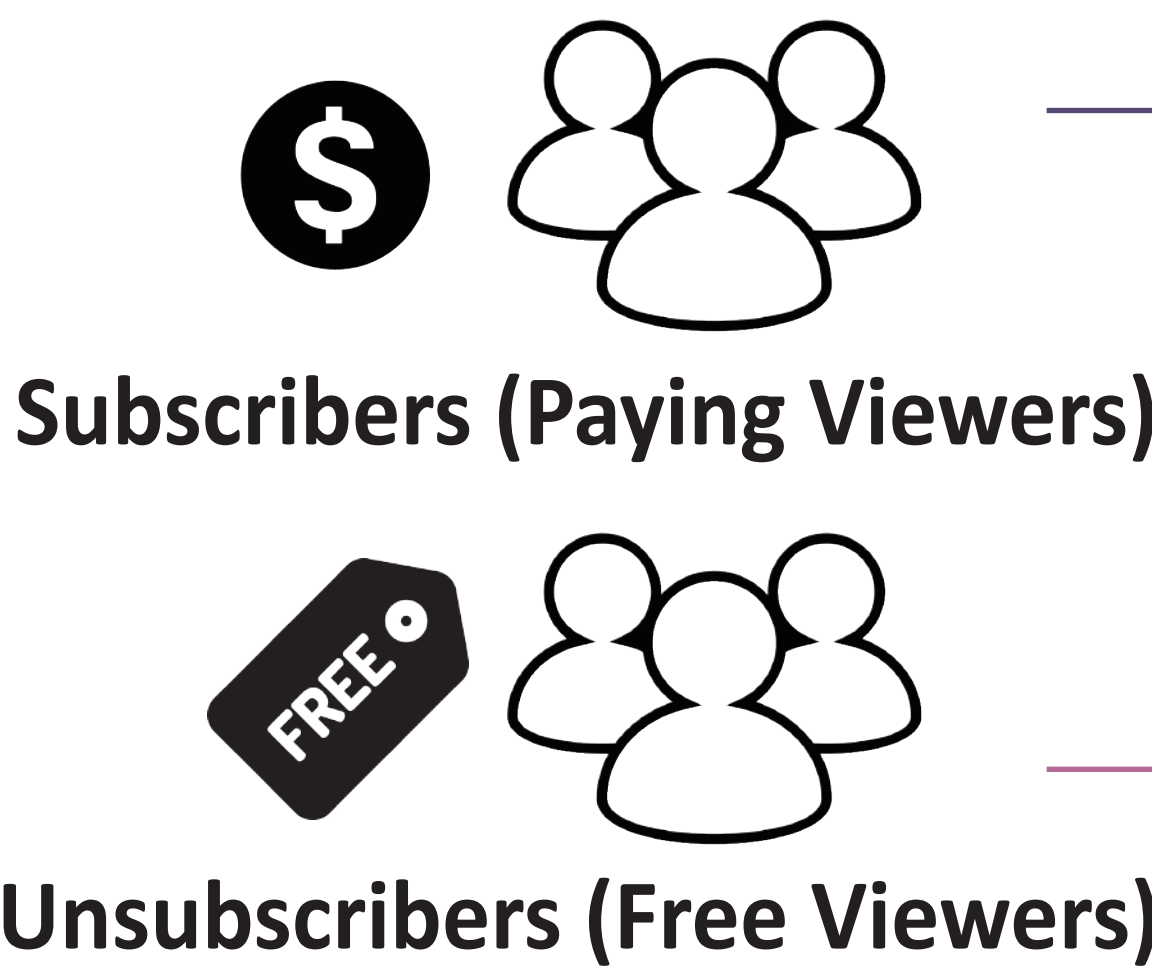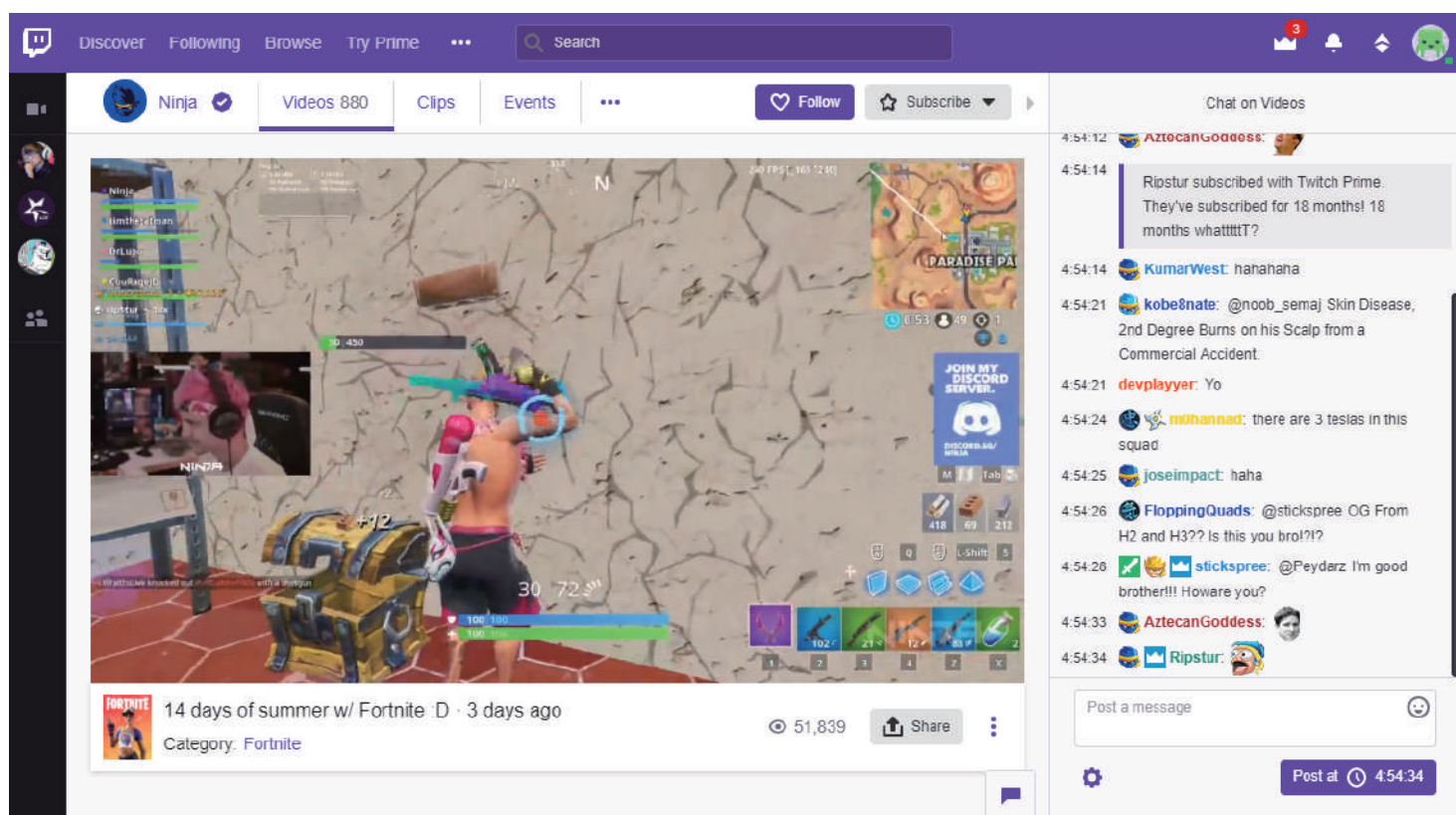# Who will subscribe to my streaming channel?: The case of *Twitch*

Jina Kim[1], Kunwoo Bae[1], Eunil Park[1], Angel P. del Pobil[1, 2]

[1] Department of Interaction Science, Sungkyunkwan University, Seoul, Korea
[2] Computer Science and Engineering Department, Jaume I University, Castellon, Spain

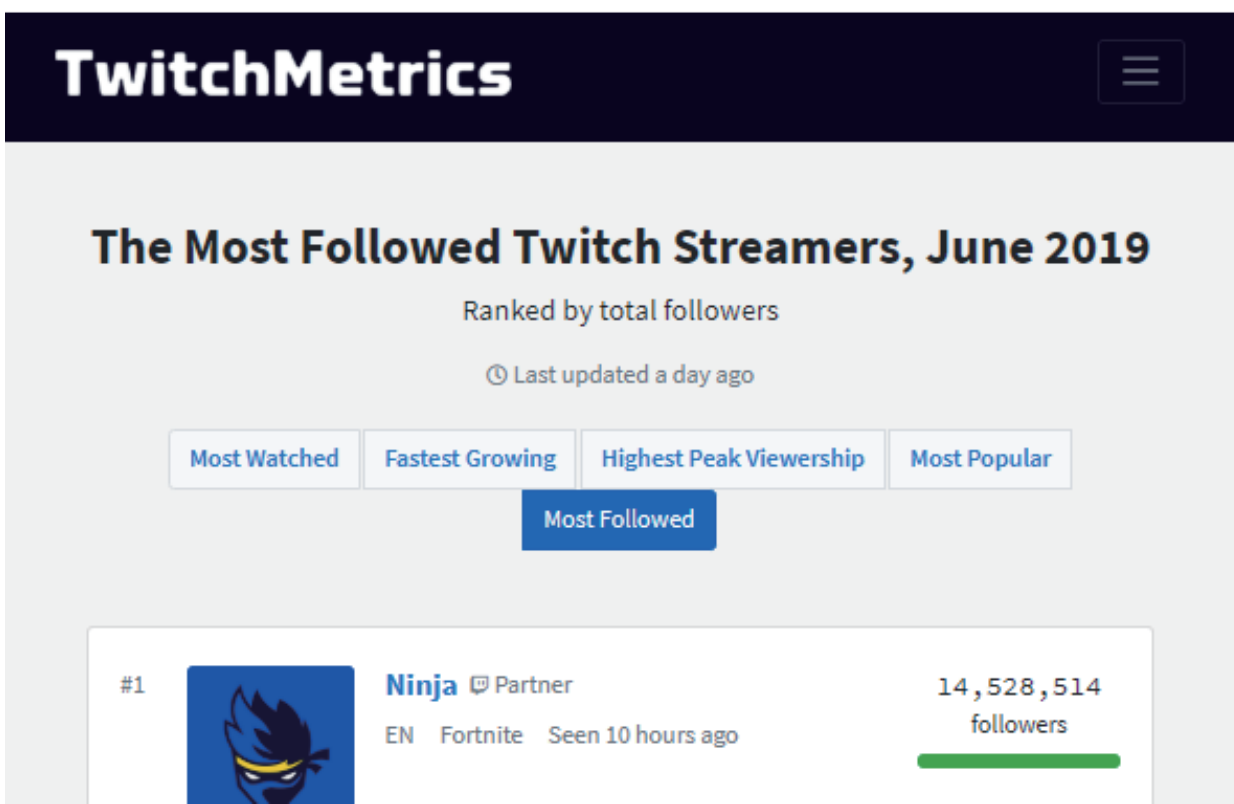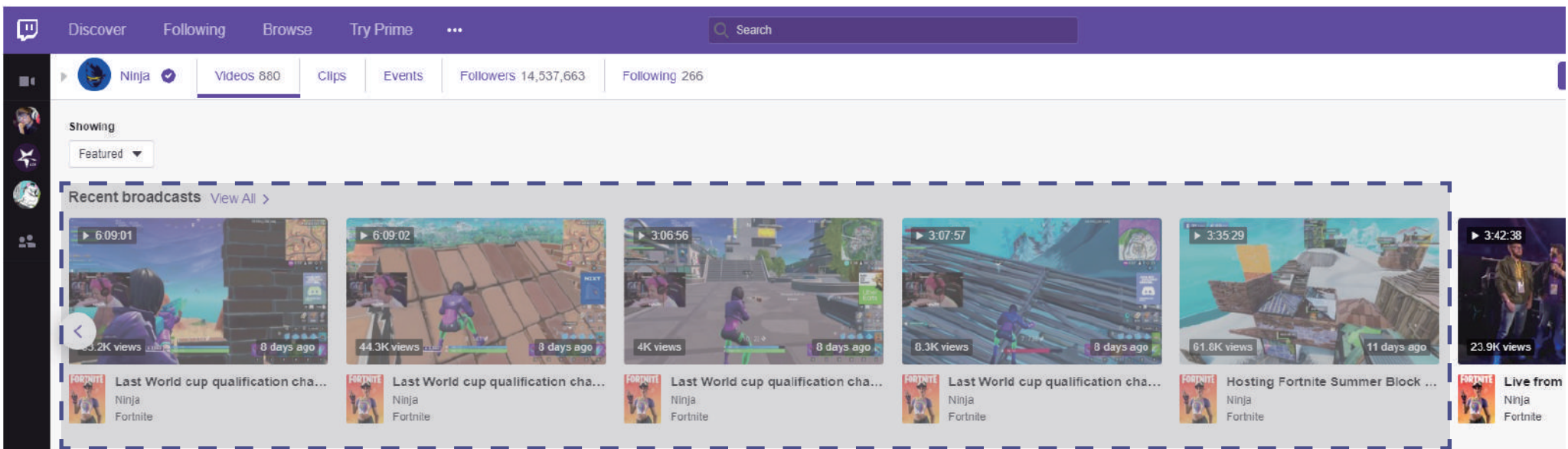✉ jinakim@g.skku.edu, eunilpark@skku.edu 🏠 http://dsl.skku.edu

Introduction

Subscribers (Paying Viewers)

Unsubscribers (Free Viewers)

RQ1. Are there **sentimental differences** of chat messages between **subscribers** and **unsubscribers**?
RQ2. Can **machine learning models** identify whether a viewer subscribes to a specific channel?

Data Collection

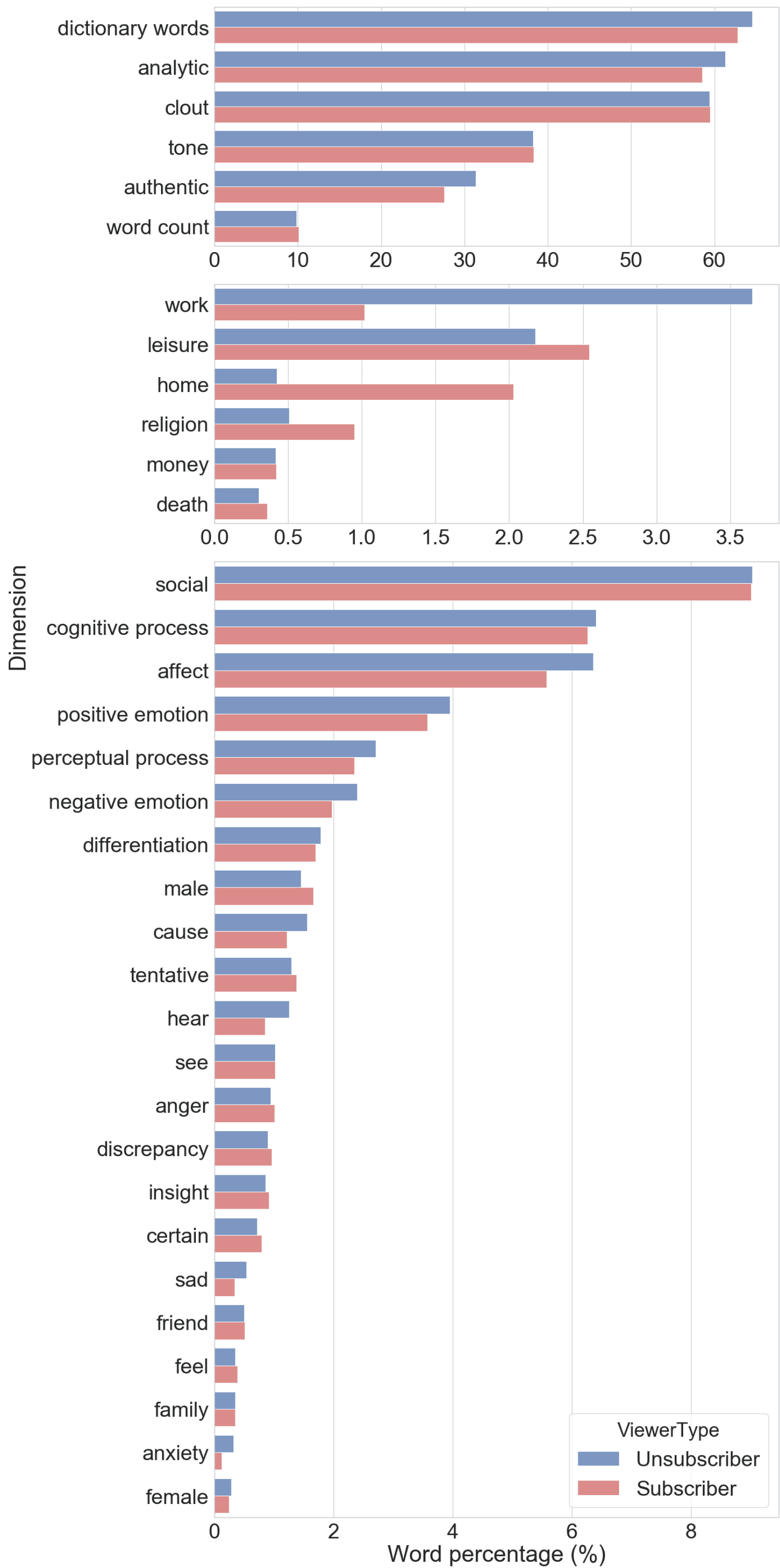The most followed *Twitch* streamers list
from *TwitchMetrics* (http://www.twitchmetrics.net)

Five recent streaming videos

Dataset

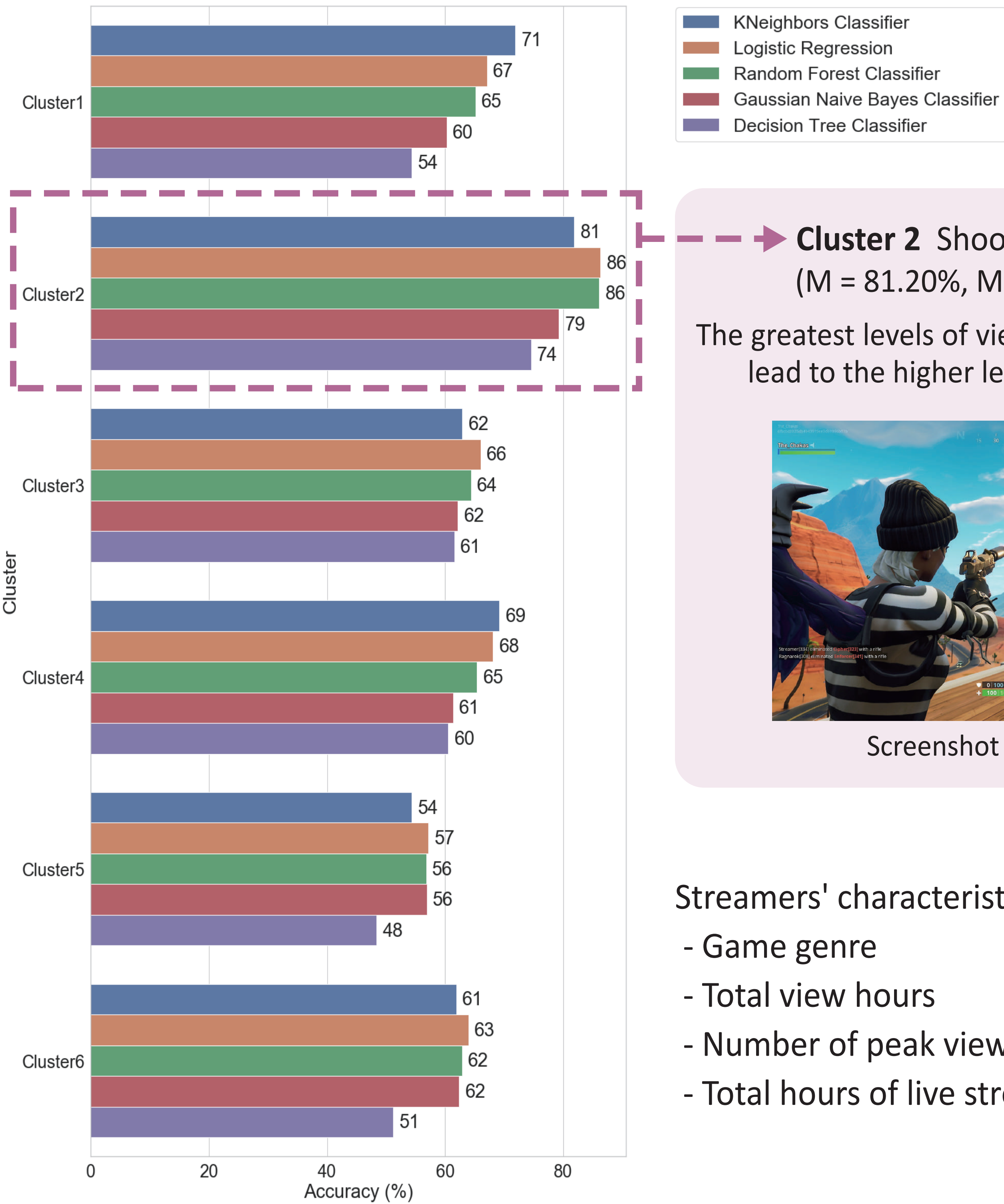|	
Stream Videos	Five recent streaming videos of top 20 streamers on *Twitch*
Chat Messages	2,934,857
	⬇ *Preprocessing*
Employed Messages	781,798
Data Categorization	457,474 *Subscriber* ⬇ 324,324 *Unsubscriber*

Study 1

Differences of Sentimental Dimensions between Unsubscribers and Subscribers

Study 2

Identification of Subscribers by Machine Learning Models

Legend:
- KNeighbors Classifier
- Logistic Regression
- Random Forest Classifier
- Gaussian Naive Bayes Classifier
- Decision Tree Classifier

Cluster 2 Shooting game (*Fortnite*)
(M = 81.20%, Max = 86%, Min = 74%)

The greatest levels of viewing hours and peak viewers lead to the higher level of livestreaming hours

Screenshot of *Fortnite* game

Streamers' characteristics as features of clustering:
- Game genre
- Total view hours
- Number of peak viewers
- Total hours of live streaming

Discussion

▪ **Significant differences** of **sentimental dimensions** between subscribers and unsubscribers

- Subscriber < Unsubscriber (perceptual, emotional)
- Unsubscriber < Subscriber (analytic, authentic)

▪ **86% accuracy** on identifying subscribers using machine learning models

Key Contributions

▪ Identification of potential subscribers by using machine learning models
▪ Recommendations of customized contents to online streamers

Future Work

▪ Considerations of socio-demographic information of both streamers and viewers