

# Reverse-engineering User Selection of YouTube Video Categories

*CAPP 30255: Advanced Machine Learning for Public Policy*

Ta-Yun Yang & Patrick Lavalley Delgado\*

20 April 2020

## 1 Overview

We aim to classify YouTube videos to the categories selected by their authors using language available in the their titles, descriptions, and tags. While categories on YouTube are perhaps relics from how people used to browse videos, they do offer insight to how authors associate their content with those high-level abstractions. We hope to identify semantic associations between the distinguishing particulars of a video and the location in the corpus of user-generated content that its author intended.

---

\*Candidates, MS Computational Analysis and Public Policy, {tayuny, pld}@uchicago.edu.