**Predicting how much attention a YOUTUBE video will receive**

*By: Amanraj Singh, Monica Sharma*

**Introduction:**

YouTube is one of the largest video-sharing websites with humongous number of videos on it. The site allows users to upload, view, rate, share and comment on videos. There is a huge possibility for analyzing data present on YouTube and getting useful insights out of it.

**PURPOSE:**

Predicting how much attention a video will receive on YouTube is of great significance to the design and service management. For example, predicting the future video popularity is useful in planning advertisements, and the earnings and the costs estimated are relevant to the views on YouTube as well. Acquiring an approximation of the popularity ahead of time enables market strategy adjustments and taking measures to change the popularity.

**OBJECTIVE:**

The objective of our project is to predict the number of likes of a video on YouTube.

**RESEARCH QUESTIONS:**

We will get better understanding of data by answering questions as below:

1. When was the video posted? What is the type of video (Music, Politics, Sports etc)?
2. What is the number of comments received so far on that video?
3. What is the channel associated with the video and how old is the channel?

**DATA SOURCE:**

We are using the YouTube API for extracting the videos for the duration 2010-2016. Currently we have downloaded approximately 22,000-24,000 videos for every category. There are 15 YouTube video categories. Total 26 attributes are present in the dataset. Some of them are as follows: like\_count, video\_length, category\_id, duration, published\_at, channel\_view\_count, country etc.

**ANALYSIS METHOD:**

We will first use Recursive Feature Elimination (RFE) technique to extract final features to train the model. We will then use Multiple Linear Regression algorithm followed by Random Forest algorithm and Support Vector Machine algorithm for predicting the like count of the videos.