Trending on Youtube

BY: CODIE JAMES

The Problem

What makes a video gain viewership?

Definition of popular: 1 Million Views

Value that can be brought to a company includes:

- Maximizing Growth
- Reducing cost
- Gain Exposure

Data Collection

Found on Kaggle

- Collection time
- Preview of Data
- Structured
- Amount of data
- Categorically Seperated

Web-Scraped from YouTube

Snapshot of data

Looking at Data

Good

- Over 40,000 Rows of data
- 16 Different features
- Able to create features

Bad

- Missing Subsriber count
- Channel Views
- Number of videos

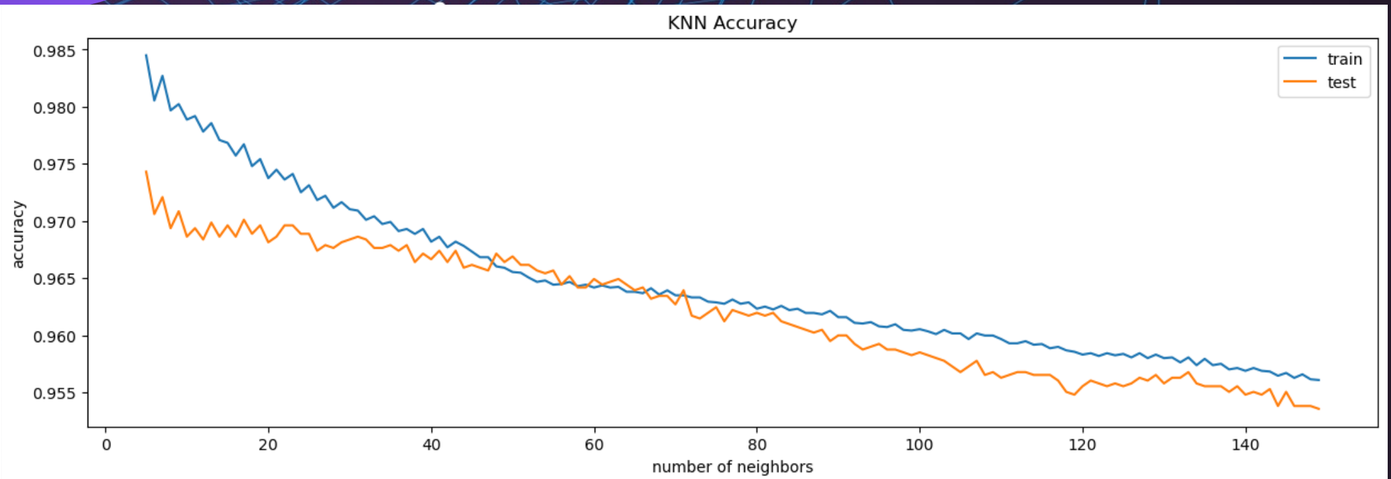
EDA + Models

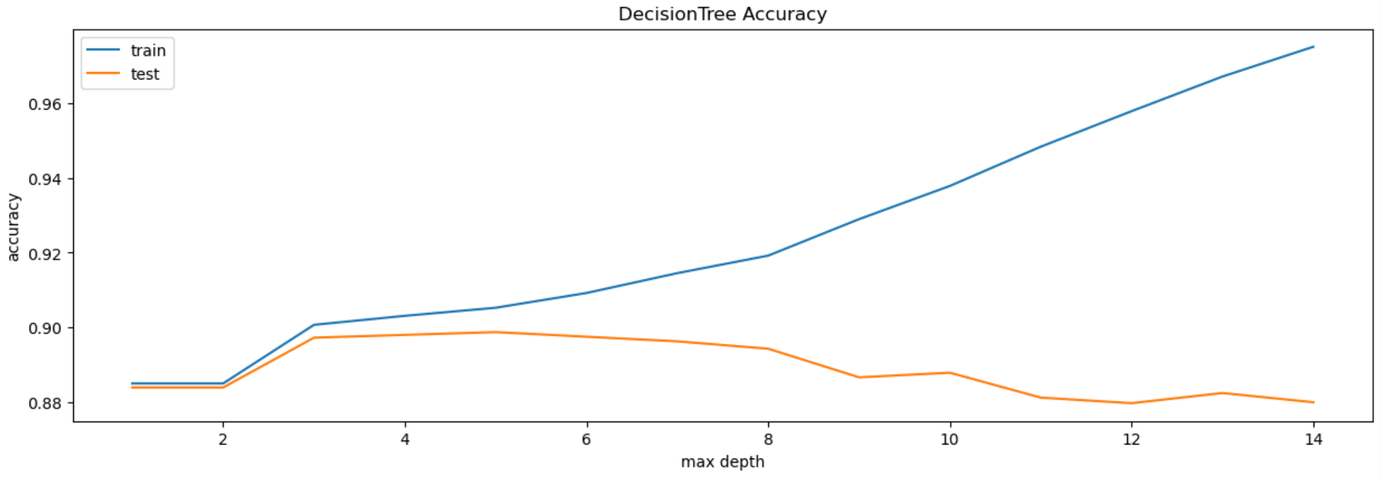
Exploratory Data Analysis

- <1300 N/A Values
- Dropping Non-Num Col.
- Converting to Num Col.

Modeling

- Hyper Parameter Optimizing
- Avoiding Data Leakage
- Logistic Reg.
- K-Nearest Neighbors
- Decision Tree Classifier

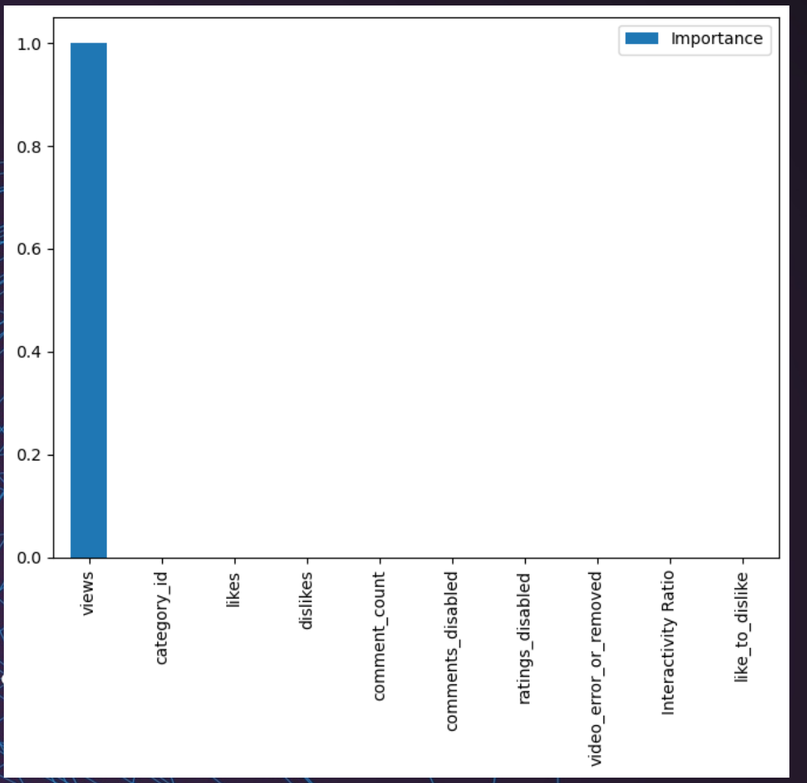


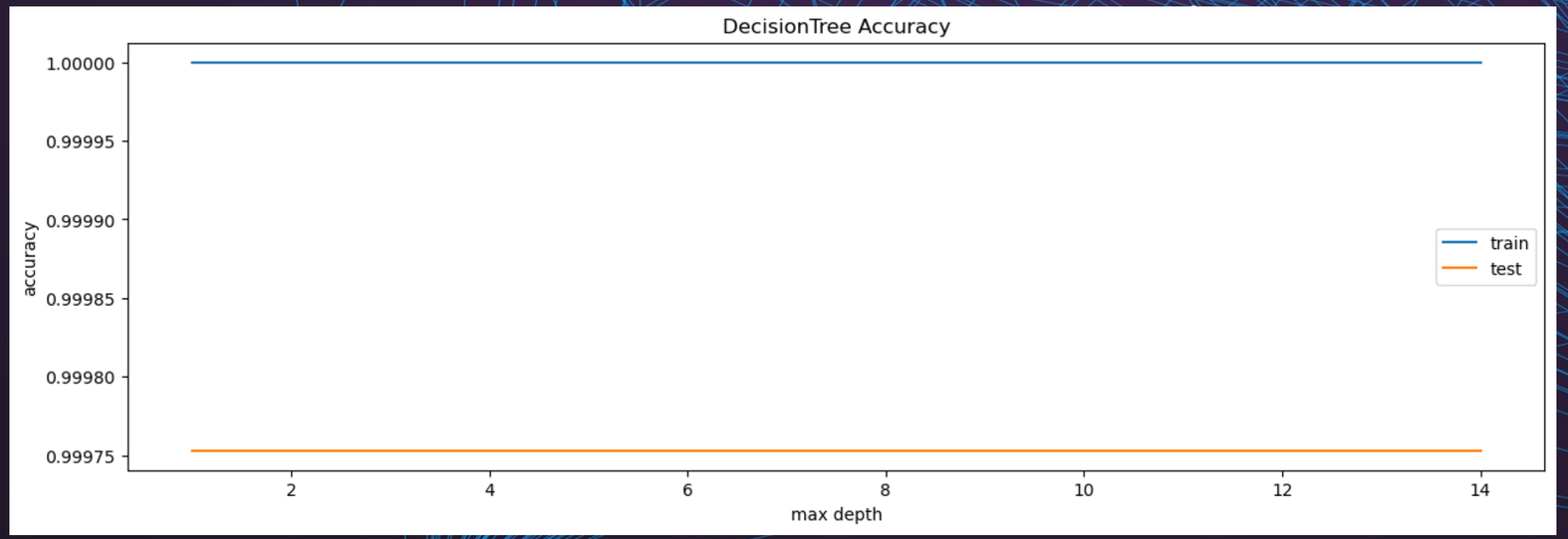


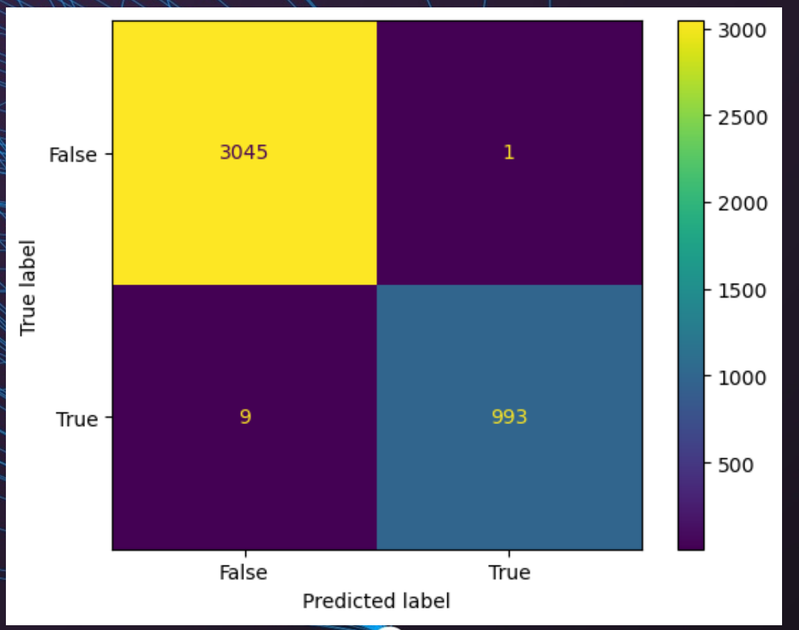
Challenges

Initial models introduced data leakage

Data imbalance of 70/30





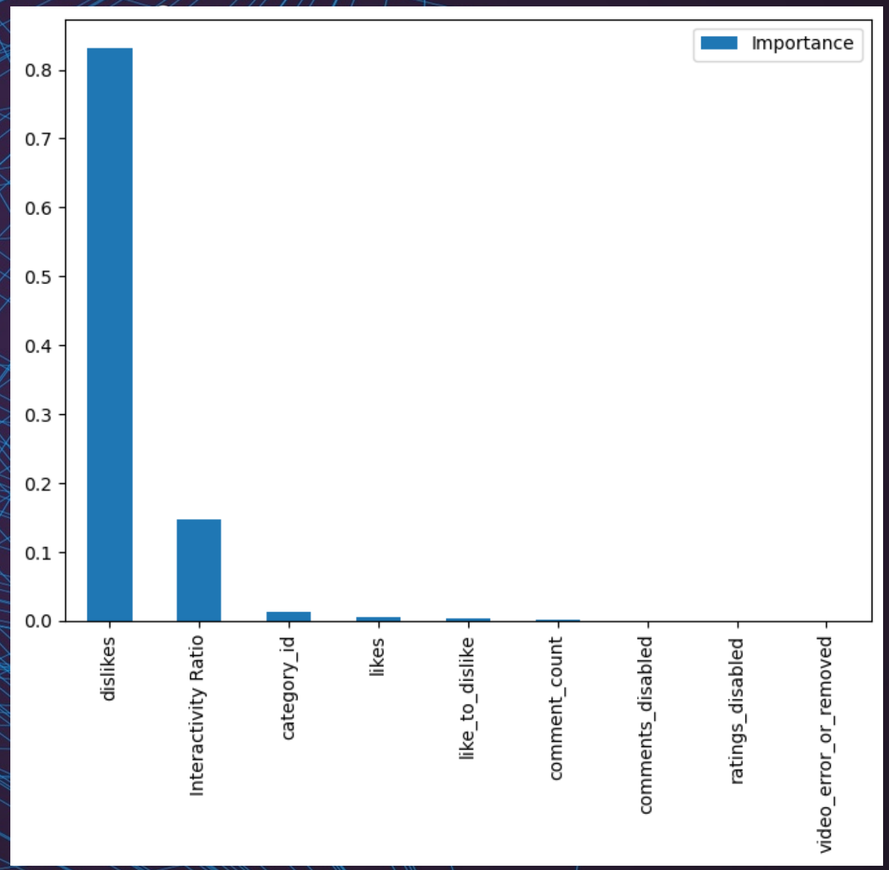


Findings

All 3 modelling methods used scored similarly

Most impactful feature was dislikes

Best Scoring model was 91% accurate



Thank you