

Simple Youtube Recommendation
engine using basic Math

6th oct 2019

AppliedAICourse.com

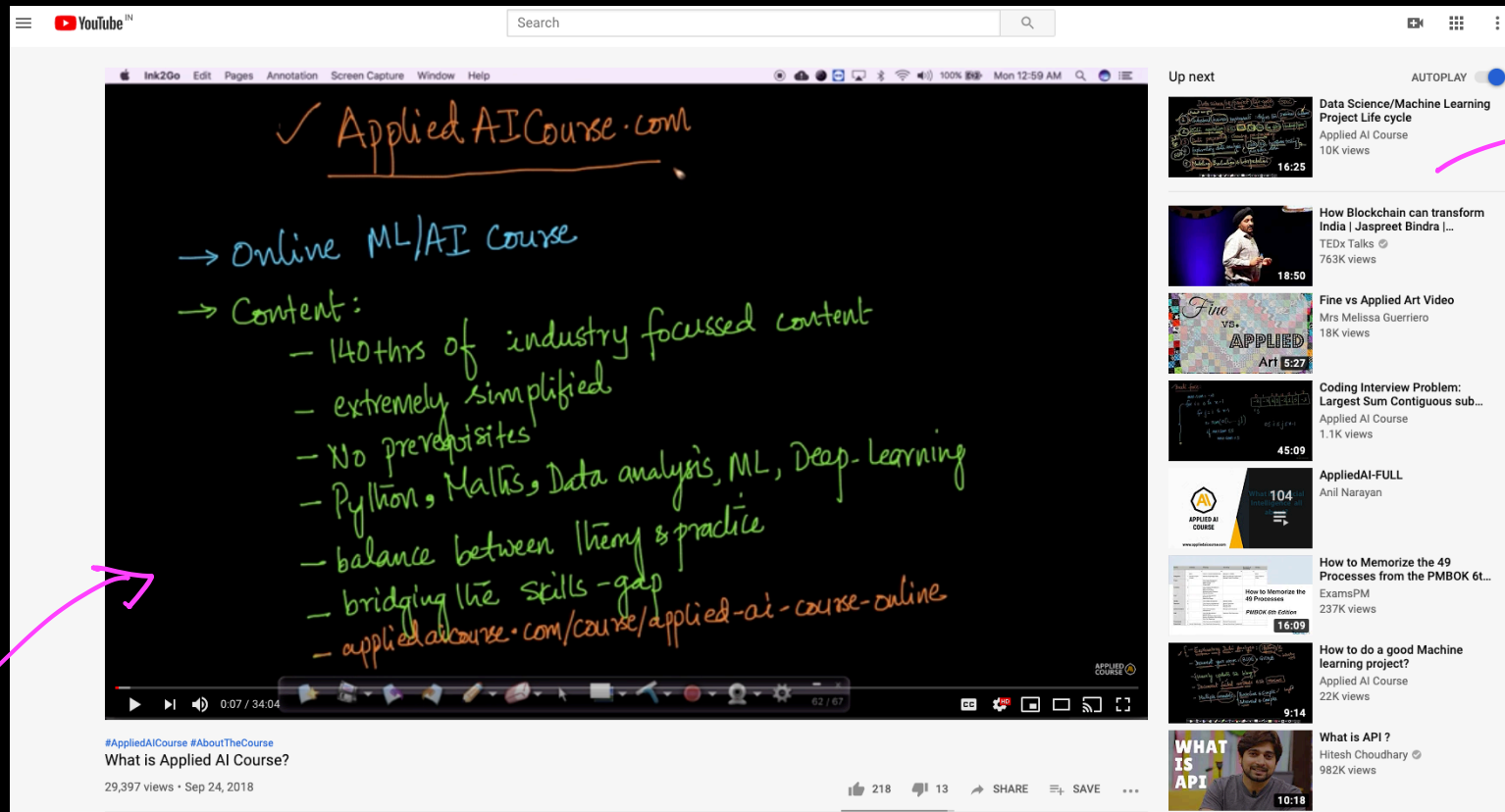
Agenda/Plan -

→ Understand the problem

→ use simple qth/10th math to build a first-cut-solution

→ Slightly advanced-math (11th & 12th)

Problem - definition.



The screenshot shows a YouTube video player with a blackboard background. The text on the blackboard is as follows:

✓ AppliedAICourse.com

→ Online ML/AI Course

→ Content:

- 140+ hrs of industry focussed content
- extremely simplified
- No prerequisites
- Python, Maths, Data analysis, ML, Deep-Learning
- balance between theory & practice
- bridging the skills-gap
- appliedai.com/course/applied-ai-course-online

Below the video player, the title "What is Applied AI Course?" is visible, along with 29,397 views and the date Sep 24, 2018. To the right of the video player, there is a list of recommended videos.

Up next

- Data Science/Machine Learning Project Life cycle Applied AI Course 10K views 16:25
- How Blockchain can transform India | Jaspreet Bindra | TEDx Talks 763K views 18:50
- Fine vs. APPLIED Art 18K views 5:27
- Coding Interview Problem: Largest Sum Contiguous sub... Applied AI Course 1.1K views 45:09
- AppliedAI-FULL Anil Narayan 104:00
- How to Memorize the 49 Processes from the PMBOK 6t... ExamsPM 237K views 16:09
- How to do a good Machine learning project? Applied AI Course 22K views 9:14
- WHAT IS API Hitesh Choudhary 982K views 10:18

→ Video-recommendations

Current video.

What data does youtube have?

Data-matrix

	v_1	v_2	v_3	v_4	v_5	\dots	v_j	\dots	v_m
u_1									
u_2									
u_3									
u_4									
\vdots									
u_i									
\vdots									
u_n									

$m = \# \text{ videos}$

$n = \# \text{ users}$

Simplest - idea: Similarity

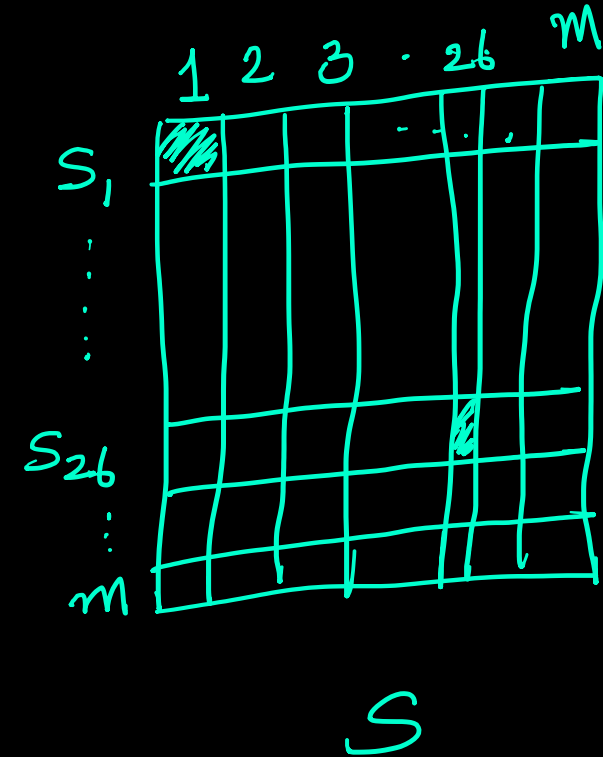
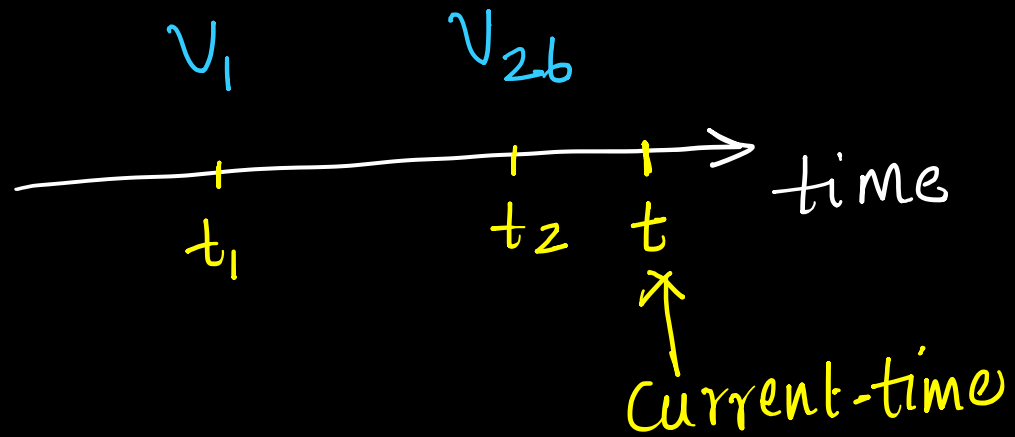
Finding similar videos using Data-matrix

$$\text{Sim}(V_i, V_j) = |V_i \cap V_j|$$

	V_i	V_j	V_k
u_1	1	1	0
u_2	1	0	0
u_3	1	1	0
u_4	0	1	1
\vdots	\vdots		
u_n	1	1	1

Recommend based on Similarity-scores.

\mathcal{U}_i



Problem/Issue: Popular-videos

$$\text{Sim}(v_i, v_j) = \text{Sim}(v_i, v_k)$$

$$\text{JS}(v_i, v_k) = \frac{|v_i \cap v_k|}{|v_i \cup v_k|}$$

$$\text{JS}(v_i, v_j) = \frac{4}{4}$$

vs

$$\text{JS}(v_i, v_k) = \frac{4}{7}$$

	v_i	v_j	v_k
	1	1	1
	1	1	1
	0	0	1
	0	0	1
	0	0	1
	0	0	0
	1	1	1
	1	1	1

popular
video

Similarity between users:

	1	2	3	4	...	m
u_i						
u_j						

Deployment & Productionization

1B users

100MM videos

New-videos-watched

⋮

Eigen-values & Eigen-vector based Approach

SVD & MF

$$A_{n \times m}$$

$$A_{m \times n}^T A_{n \times m} = S_{m \times m}$$

Diagram illustrating the SVD decomposition of matrix A into U , S , and V^T .

The matrix $A_{m \times n}^T A_{n \times m}$ is shown as a 5×5 matrix with a green row vector v_i and a green column vector v_j .

The matrix $S_{m \times m}$ is shown as a 5×5 matrix with a green row vector v_i and a green column vector v_j .

The matrix V^T is shown as a 5×5 matrix with a green row vector v_i and a green column vector v_j .