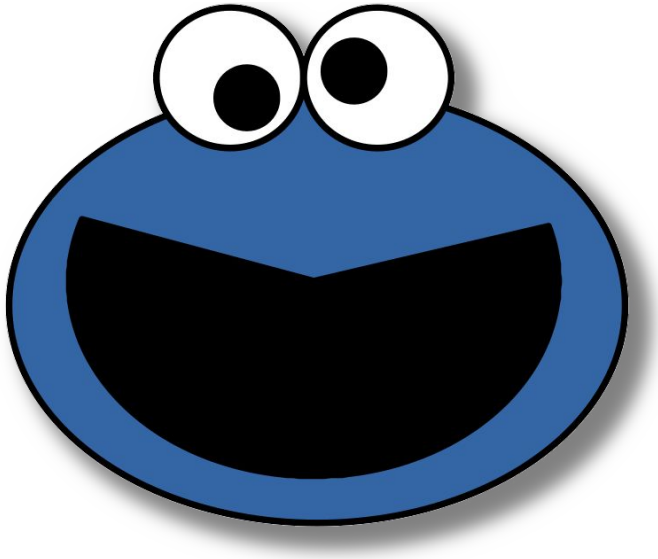
TrikAI



Learning AI with TrikAI

the CookAI Monster

Democratizando la formación

José Francisco Caro Martínez @jfrcaro , Enero 2.022

Saturdays AI

Intro

… Imaginemos que quieres aprender sobre algo nuevo

¿Por **dónde** empezarías?

Intro

… Nunca había sido tan **fácil y accesible** aprender de cualquier cosa a través de internet !!!

Hay millones de contenidos de excelente calidad…

¡¡¡ y muchos de ellos gratuitos !!!

Intro

Youtube se está convirtiendo en una de las mejores "**Universidades**" gratuitas

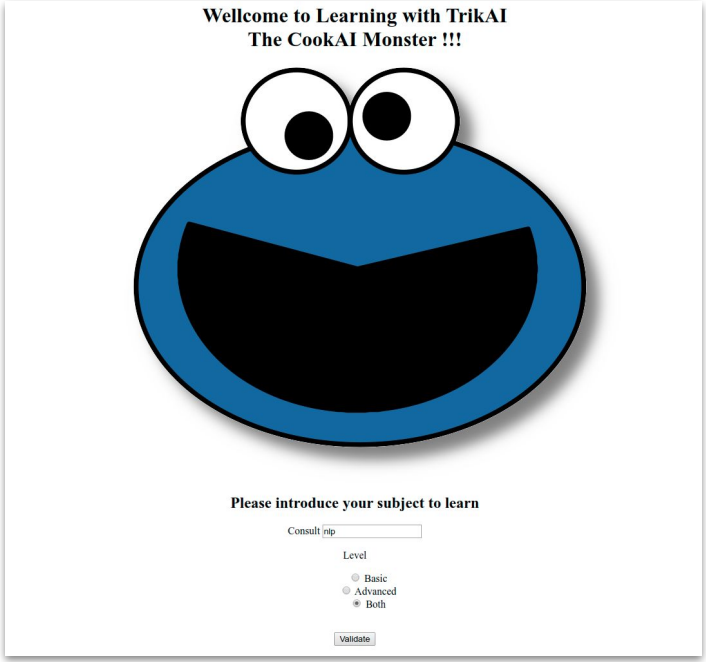
+

… pero el contenido se encuentra **disperso** y **lleva tiempo encontrar** material de calidad

→ **Recomendador**

Descripción

TrikAI - “Sistema que recolecte (1), sintetice (2), ordene (3) y recomiende (4) contenido de formación en Inteligencia Artificial (eng y YouTube) en función de la consulta realizada”



Bloques TrikAI

DAIogenes

1) Recolectador web para obtención de contenido de aprendizaje (Videos - YouTube)

Texto de transcripciones
+
metadatos

CookAI Monster

2)Sintentizador de documentos que obtiene su esencia (reducción de dimensionalidad)

Agrupación - Cluster
basada en palabras frecuentes
TF-IDF

Agrupación - Tópicos
basada en detección de temas
LDA

Similitud
basada en vectorización
Doc_To_vect + cosine similarity

Dificultad
basada en entrenamiento
modelo Spacy + textcategorizer

Nube de palabras
de la transcripción
WordCloud

NLP

Biblioteca

3)Almacén ordenado del contenido y su esencia

Ficheros con información obtenida
info(DAIogenes + CookAI Monster)

Visualización
Muestra de volumetría biblioteca, Clusters y Topics

Recomendador

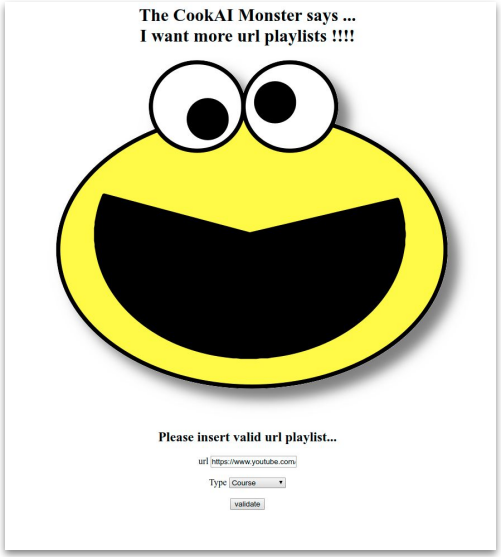
4)Recomendador de contenidos en base a consulta realizada

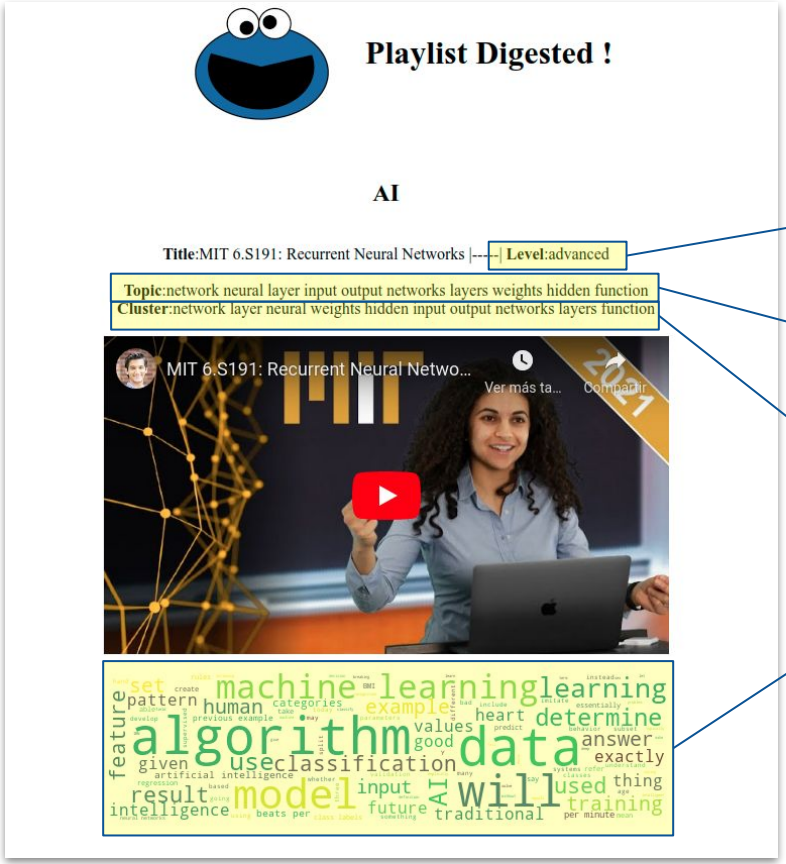
Búsqueda + Visualización
basada en palabras del título+Cluster+Topic+Similitud

Saturdays AI

CookAI Monster

2)Triturador de documentos que obtiene su esencia





Dificultad
basada en entrenamiento
modelo Spacy + textcategorizer

Agrupación - Tópicos
basada en detección de temas
LDA

Agrupación - Cluster
basada en palabras frecuentes
TF-IDF

Nube de palabras
de la transcripción
WordCloud

NLP

Saturdays AI

Biblioteca

3)Almacén ordenado del contenido y su esencia

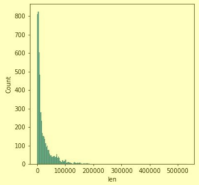
Global
Número total de videos, longitud e histogramas

TrikAI library

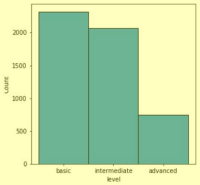
Total Videos: 5139

Total Length: 121758692

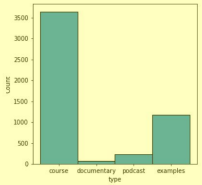
Length



Level



Type



Clusters y Tópicos
Número de videos en cada categoría y palabras clave

Topics and number of videos

counts	topic links
362	network neural layer input output networks layers weights hidden function
324	user results project users technique papers support thanks research videos
250	word words text language sentence sequence attention embedding english token
244	sort basically maybe yeah pretty paper models got doesn idea
237	instance simulation robot solution paper possible yes quite papers real
197	training train accuracy validation class create test random classifier models
196	dot column columns frame ahead values value equal rows index
168	state policy action reward game agent value actions states reinforcement
166	value values function zero equals error equal algorithm line minus
160	regression tree linear decision logistic predict classification trees line fit
151	intelligence science artificial human world technology today course systems talk
151	matrix vector vectors linear transpose times product matrices dimensional row
150	maybe yeah question trying got idea doesn probably said world
149	task neural tasks models deep network networks paper human language
145	image images pixel pixels style resolution color face map picture
145	test plot points line testing looking color percent looks tests
142	python code library jupiter notebook programming libraries type write numpy
142	gonna got stuff probably pretty cool looking maybe talk goes
133	file dot create page open save post user server directory
130	training output input examples train loss points high supervised classification
115	gradient theta minus plus descent equal step direction compute gradients
107	probability distribution given log likelihood gaussian probabilities random variance normal
102	computer program deep idea computers years human world game research
95	music company person business companies team understand questions experience level
86	function list functions array kernel write element loop memory return
85	yeah human world sort sense life interesting humans ideas conversation
84	basically particular guys suppose understand getting respect value write consider
76	life said world years power love money believe got war
70	encoder dot labels generator self basically label samples auto pass
65	features feature transform fourier movie house frequency maybe apply values
65	graph node variables noise nodes models variable latent sample knowledge
63	cluster distance clustering clusters pipeline points group component groups analysis
52	tensor flow convolution input tensorflow information size memory operation output
51	yeah yes got great talk stuff didn cool said thank
47	energy positive negative score false recall true correct threshold classifier
47	brain cells cell neurons systems human neuron body intelligent processing
37	app product database build team software order application create api
30	object objects box grid detection star question week grab doesn
27	space theory sort physics quantum universe dimensional dimensions idea field
25	loss derivative respect hat chain sub rule gradient minus cross
24	car cars driving speed light red control vision blue real
21	cloud google gpu notebook tensorflow api running job training tool
15	form query block email information pattern content simply table match
3	lambda squared norm sum smaller noise plot frequency equals regularization
3	series date price day year days future wanted format order
2	represents lambda words squared equals sum physics api word quantum

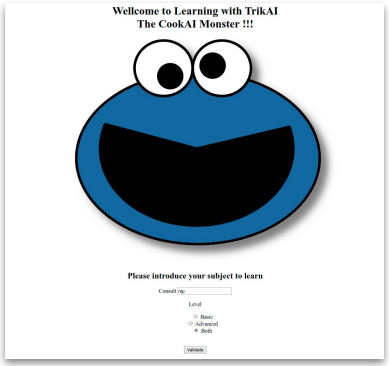
Clusters and number of videos

counts	Cluster links
721	data regression going line just let like answer probability number
440	just like going okay data gonna right know actually want
302	data model set training just going use let test like
298	like data just know people really going actually time think
289	model just file going data let python tensorflow ll like
241	network layer neural weights hidden input output networks layers function
239	papers scholars fellow technique paper video minute time károly zsolnai
188	image images layer right network filter just convolutional convolution going
183	right like uh okay know model um just going attention
174	uh um like know just okay right basically yeah going
155	know like think people just yeah kind really right things
144	gradient ok right derivative function just loss descent going respect
141	matrix pca svd principal data vector singular going transpose vectors
138	word words sentence text vectors right sequence vector sentences model
135	simulation simulations fluid papers light technique simulate paper time materials
132	data okay pandas frame column columns just dot going want
127	image images neural papers network technique style paper scholars fellow
124	decision tree trees data split entropy node leaf boundary let
110	uh like know um yeah think people just right don
101	list let print function dot say equals want value element
101	going spacey text gonna right data alright model dot let
98	learning machine data supervised algorithms unsupervised model learn like regression
95	um uh like know sort just okay right model kind
87	policy reward state action agent value just learning actions reinforcement
84	ai intelligence artificial human like humans world intelligent think people
81	game games ai reinforcement learning agent human play player deepmind
80	theta function right uh know cost going gradient regression vector
59	fourier transform okay gonna going frequency fft coefficients kind hat
56	cluster clusters clustering centroids centroid points data centers going means
16	cameron sarah height weight 115 chris plus 140 175 pounds

Saturdays AI

Recomendador

4)Recomendador de contenidos en base a consulta realizada



TrikAI recommends you this special content...

Title: Artificial Intelligence - Rise Of The Machines (It Will Kill Us)

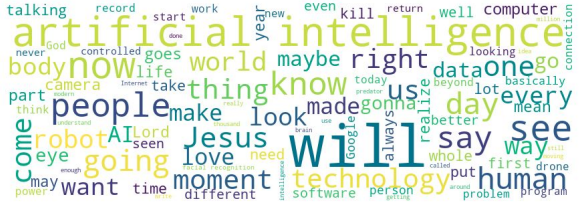
Playlist: Artificial Intelligence

Author: Time To Repent **Views:** 2427.0 **Rating:** 4.968504

Level: basic **Length:** 31913.0 **Type:** documentary

Topics: intelligence science artificial human world technology today course systems talk

Cluster: ai intelligence artificial human like humans world intelligent think people



Metadados
Título, playlist, author, vistas, valoración, longitud y tipo

Relacionados
Cluster, Topic y Doc_To_vect + cosine similarity

Selection in the same TOPIC

Author	Link	Rating	Level
Enthought	Day 1	Welcome	SciPy 2019	Serge Rey	5	intermediate
PyData	A Data Science Approach to Systemic Risk - Nikolai Nowaczyk	5	intermediate
PyData	Aileen Nielsen- Open Source Fairness	PyData Global 2020	5	intermediate
CodeEmporium	5 Data Science Myths you shouldn't believe	5	basic
Simplilearn	FREE Data Analytics Course	Data Analytics Course	Data Analytics For Beginners	Simplilearn	5	basic
Udacity	PCA for Facial Recognition - Intro to Machine Learning	5	basic

Selection in the same CLUSTER

Author	Link	Rating	Level
Yannic Kilcher	On the Measure of Intelligence by François Chollet - Part 1: Foundations (Paper Explained)	4.9927139	advanced
Two Minute Papers	OpenAI Outperforms Some Humans In Article Summarization! 📜	4.9739647	intermediate
CodeEmporium	How to keep up with AI research?	4.9716816	advanced
Time To Repent	Artificial Intelligence - Rise Of The Machines (It Will Kill Us)	4.968504	basic
IBM Technology	Can computers program computers?	4.9658117	basic
edureka!	How to become an AI Engineer	Artificial Intelligence Roadmap	AI Career Path	Edureka	4.9615383	basic

Selection similar CONTENT

Author	Link	Rating	Level
World Science Festival	Making Room for Machines: Getting Ready for AGI	4.7404523	basic
Lex Fridman	Avi Loeb: Aliens, Black Holes, and the Mystery of the Oumuamua	Lex Fridman Podcast #154	4.7976055	basic
Lex Fridman	Sebastian Thrun: Flying Cars, Autonomous Vehicles, and Education	Lex Fridman Podcast #59	4.9164019	basic
Spark	The Rising World Of Building AI Human Clones	Artificial Intelligence	Spark	4.838665	basic
Lex Fridman	Max Tegmark: AI and Physics	Lex Fridman Podcast #155	4.844152	basic
Lex Fridman	Rob Reid: The Existential Threat of Engineered Viruses and Lab Leaks	Lex Fridman Podcast #193	4.9038501	basic

Saturdays AI

Posible evolución

General

Mejoras generales de TrikAI

- **Otros idiomas** (p.e. ES)
- Gestión de usuarios

DAIogenes

1)Recolectador web para obtención de contenido de aprendizaje

- Automatizar la recolección de videos
- Recolectar **otro tipo de contenido** (papers, blogs, …)
- Ampliar a otro tipo de materias además de AI

CookAI Monster

2)Triturador de documentos que obtiene su esencia

- Más categorías de nivel (p.e. intermedio)
- Optimización de la clusterización, obtención de tópicos
- Reentrenamiento de niveles con opinión de usuarios

Biblioteca

3)Almacén ordenado del contenido y su esencia

- Uso de base de datos y no fichero pandas
- Mejora en cálculo del rating (learning rate)

Recomendador

4)Recomendador de contenidos en base a consulta realizada

- **Generador de cursos** en base a índice de contenidos
- Evolución del aprendizaje (por usuario)
- Propuesta de **plan de formación** por usuario

Conclusiones

1. Nunca había sido tan **fácil y accesible aprender** gracias al contenido disponible en internet.

2. Utilizando técnicas de **NLP** sobre cualquier tipo de contenido textual, se pueden realizar **potentes búsquedas semánticas** (Agrupación-Clusters, Tópicos, similitud entre documentos, predicción de dificultad y otras)

3. El proyecto realizado (**Trikai**) puede ser **evolucionado** en múltiples funcionalidades (Nuevos idiomas, nuevas materias, gestión de usuarios, generador de planes de formación, …)