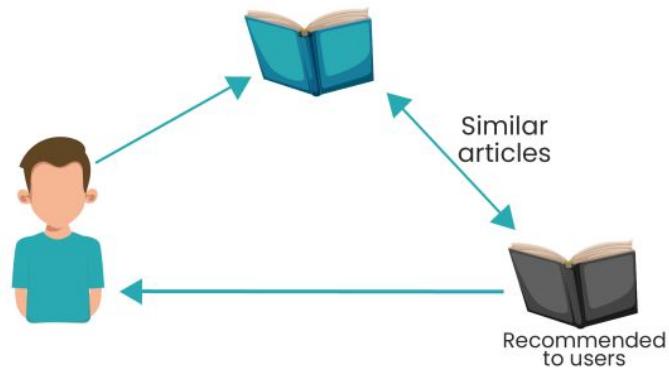
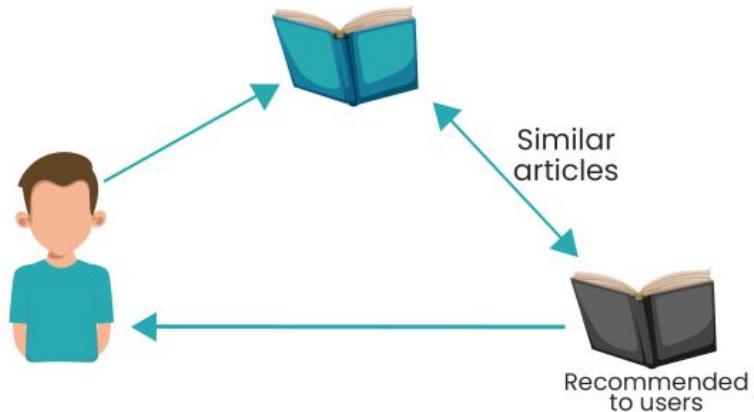


Text Similarity in Related Articles



What is Text Similarity?



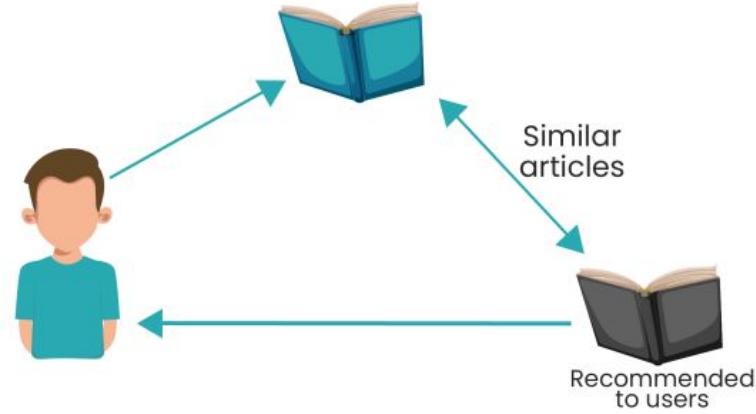
Introduction



Text similarity is comparing texts and determining how close they are semantically.

There are some techniques in text similarity:

- **Term Frequency-Inverse Document Frequency (TF-IDF)**
- **Word Embeddings - word2vector**
- **Bag-of-Words (BoW)**



Algorithm 1: TF-IDF

$$\text{TF-IDF: } w_{i,j} = tf_{i,j} \times \log \left(\frac{N}{df_i} \right)$$

1. Document

Index	Type	Size	Sentence
0	str	15	he is good boy.
1	str	17	she is good girl.
2	str	22	boy and girl are good.

2. Text Pre-processing

Index	Type	Size	after pre-processing
0	str	8	good boy
1	str	9	good girl
2	str	13	boy girl good

3. Term Frequency (TF)

$$TF = \left(\frac{\# \text{ rep of word in a sentence}}{\# \text{ words in a sentence}} \right)$$

	Boy	Girl	good
sen 0	1/2	0	1/2
sen 1	0	1/2	1/2
sen 2	1/3	1/3	1/3

4. Inverse Document Frequency (IDF)

$$IDF = \log\left(\frac{\# \text{ sentence}}{\# \text{ sentences containing word}}\right)$$

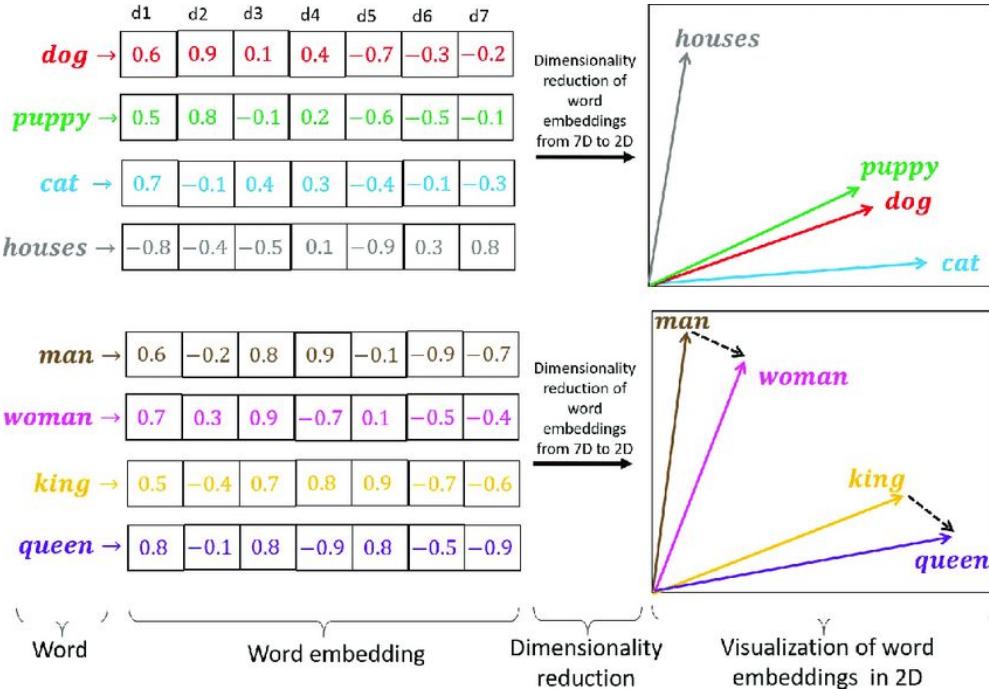
	Boy	Girl	Good
IDF	$\log(3/2)=0.176$	$\log(3/2)=0.176$	$\log(3/3)=1$

5. TF-IDF

	Boy	Girl	good
sen 0	0.08	0	0
sen 1	0	0.08	0
sen 2	0.06	0.06	0

Algorithm 2: word2vector

Word2Vec is a model to form / create word Embeddings. Its a modern way of representing Word , wherein each word is represented by a vector (Array of numbers based on Embedding Size).



Algorithm 3: Bag of word

The bag-of-words model is a model of text which uses a representation of text that is based on an unordered collection (or "bag") of words. It is used in natural language processing and information retrieval (IR).

I love this movie! It's sweet, but with satirical humor. The dialogue is great and the adventure scenes are fun... It manages to be whimsical and romantic while laughing at the conventions of the fairy tale genre. I would recommend it to just about anyone. I've seen it several times, and I'm always happy to see it again whenever I have a friend who hasn't seen it yet!



it	6
I	5
the	4
to	3
and	3
seen	2
yet	1
would	1
whimsical	1
times	1
sweet	1
satirical	1
adventure	1
genre	1
fairy	1
humor	1
have	1
great	1
...	...

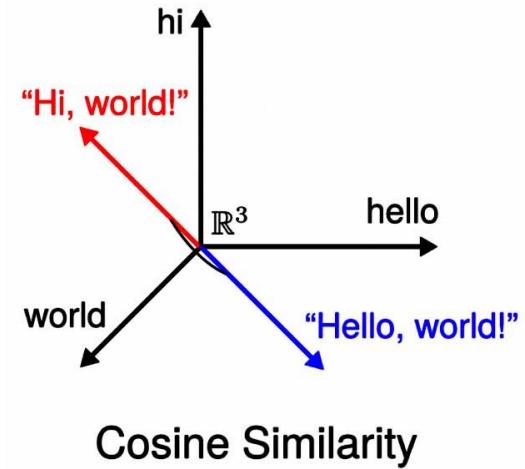
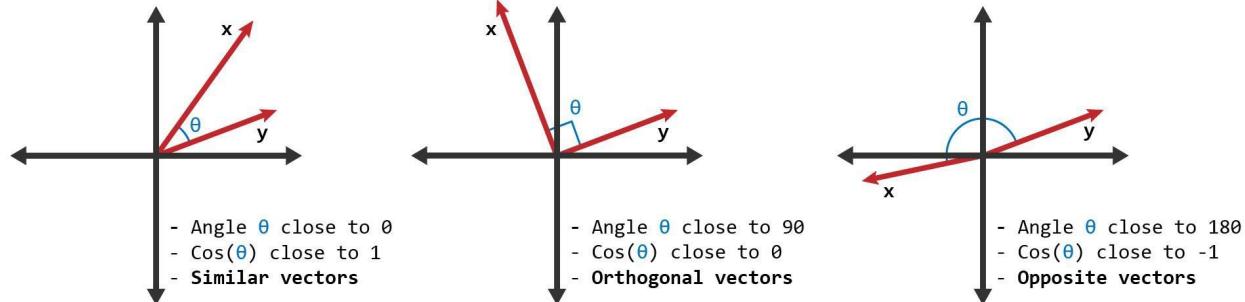
Document D1	<i>The child makes the dog happy</i> the: 2, dog: 1, makes: 1, child: 1, happy: 1
Document D2	<i>The dog makes the child happy</i> the: 2, child: 1, makes: 1, dog: 1, happy: 1

	child	dog	happy	makes	the	BoW Vector representations
D1	1	1	1	1	2	[1,1,1,1,2]
D2	1	1	1	1	2	[1,1,1,1,2]

Cosine Similarity

Cosine Similarity is a measure used to determine the similarity between two vectors in a multi-dimensional space. It calculates the cosine of the angle between the vectors, where a **value of 1** indicates that **similar vectors**, **0** indicates **orthogonality** (no similarity), and **-1** indicates **opposite vectors**.

$$\cos(\theta) = \frac{\mathbf{A} \cdot \mathbf{B}}{\|\mathbf{A}\| \|\mathbf{B}\|} = \frac{\sum_{i=1}^n A_i B_i}{\sqrt{\sum_{i=1}^n A_i^2} \sqrt{\sum_{i=1}^n B_i^2}}$$

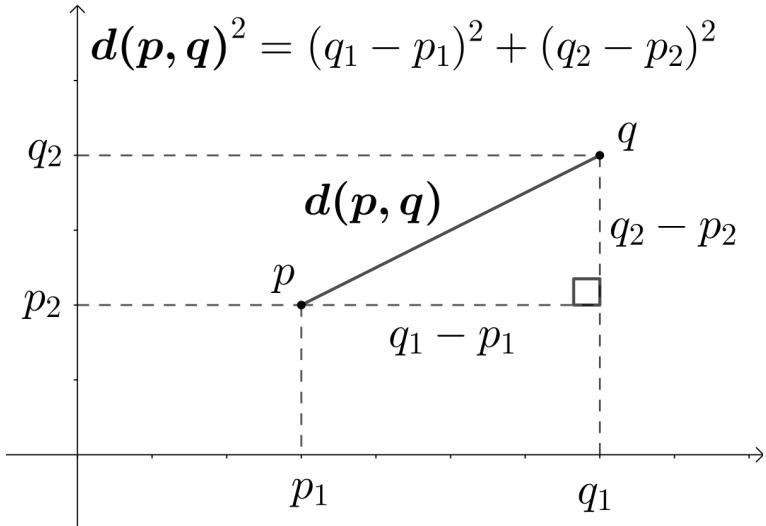


Euclidean Distance

The Euclidean distance formula is used to find the distance between two points on a plane. A straight-line distance between two points in n-dimensional space.

$$d(x, y) = \sqrt{\sum_{i=1}^n (y_i - x_i)^2}$$

$$\begin{aligned} d(p, q) &= d(q, p) \\ &= \sqrt{(q_1 - p_1)^2 + (q_2 - p_2)^2 + \dots + (q_n - p_n)^2} \end{aligned}$$



$$= \sqrt{\sum_{i=1}^n (q_i - p_i)^2}$$

Jaccard similarity

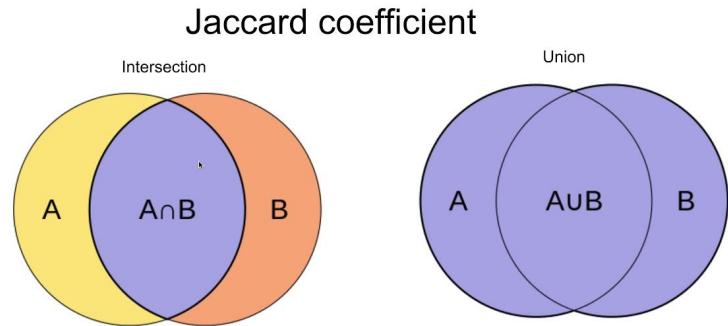
Jaccard similarity coefficient, is a statistic used for gauging the similarity and diversity of sample sets.

Doc1: **data is the new oil** of digital economy.
Doc2: **data is a new oil**.

$$J(doc_1, doc_2) = \frac{\{data, is, the, new, oil, of, digital, economy\} \cap \{data, is, a, new, oil\}}{\{data, is, the, new, oil, of, digital, economy\} \cup \{data, is, a, new, oil\}}$$

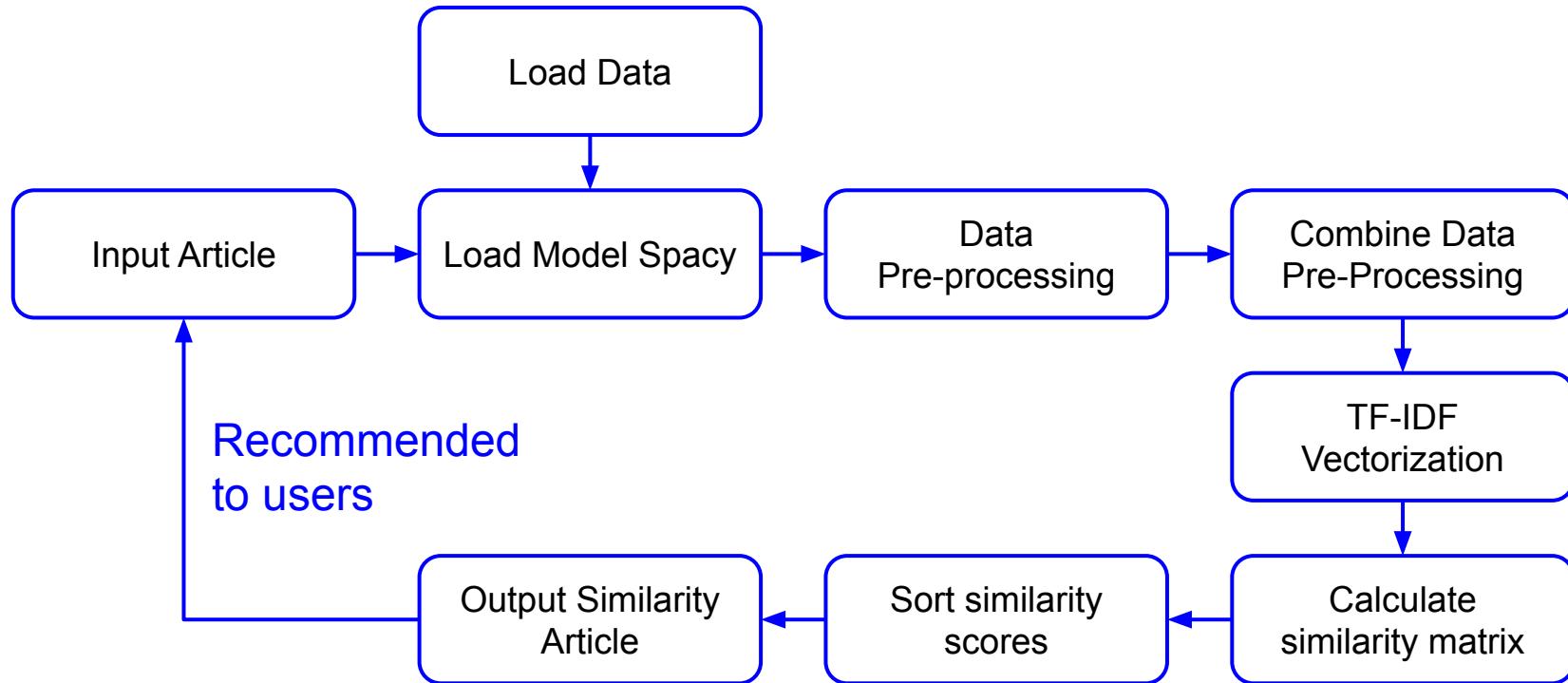
$$= \frac{\{data, is, new, oil\}}{\{data, a, of, is, economy, the, new, digital, oil\}}$$

$$= \frac{4}{9} = 0.444$$



$$J(A, B) = \frac{|A \cap B|}{|A \cup B|}$$

1. Flowchart of IF-IDF Algorithm



Result 1.1 of TF-IDF with Cosine similarity

Recommended

Algorithm 1: TF-IDF with Consine similarity							Vs	Kiripost API		
Categories	Input_title	No	Related_Article_Output	date	Score	Assumption		Date	Assumption	
Business	Annual HEINEKEN Sustainathon Invites University Students to Innovate for a Circular Cambodia	1	Winners Crowned for Sustainable Innovations in Heineken Sustainathon	2023-05-10	0.62143	High Accurate	Currently	EU and Cambodia Launch \$55.3M Initiative to Promote Green and Digital Jobs		Medium Accurate
		2	Heineken Partners with Grab to Drive Home Responsible Consumption Message in Cambodia	2024-01-28	0.51782	High Accurate	PSE Empowering Cambodian Students	PSE Empowering Cambodian Students		Medium Accurate
		3	Heineken's Worlds Together Campaign Brings together Community Gatherings	2023-10-02	0.498169	High Accurate	Transitioning to Greener Growth Model Key to Boosting Employment	Transitioning to Greener Growth Model Key to Boosting Employment		Low Accurate
		4	Heineken Cambodia Appoints Andy Hewson as New Managing Director	2023-09-07	0.489743	High Accurate	EU and Cambodia Unveil \$96.7M Investment in Four Initiatives	EU and Cambodia Unveil \$96.7M Investment in Four Initiatives		Low Accurate
		5	Glass Bottle Return Movement Gaining Ground in Cambodia's Hospitality Scene	2024-03-14	0.445468	Low Accurate	National Bank of Canada Refutes Sale of ABA	National Bank of Canada Refutes Sale of ABA		Low Accurate
Tech	For Social Platforms, the Outage Was Short. But People's Stories Vanished, and That's No Small Thing	1	BREAKING Meta platforms including Facebook Instagram and Messenger are Experiencing a Global Outage	2024-03-05	0.183855	High Accurate	Airbnb is Banning the Use of Indoor Security Cameras in the Platform's Listings	Airbnb is Banning the Use of Indoor Security Cameras in the Platform's Listings		Low Accurate
		2	Metas Meltdown Resolved After Hours of Outage	2024-03-06	0.139652	High Accurate	EU and Cambodia Launch \$55.3M Initiative to Promote Green and Digital Jobs	EU and Cambodia Launch \$55.3M Initiative to Promote Green and Digital Jobs		Low Accurate
		3	Capturing Cambodia's Stories of Change	2023-11-28	0.115229	Medium Accurate	PSE Empowering Cambodian Students	PSE Empowering Cambodian Students		Low Accurate
		4	Chaktomuk Short Film Festival Welcomes Rise in Entries	2022-11-23	0.094762	Medium Accurate	Transitioning to Greener Growth Model Key to Boosting Employment	Transitioning to Greener Growth Model Key to Boosting Employment		Low Accurate
		5	Microsoft Teams Recovering After Outage	2022-07-21	0.08586	Low Accurate	EU and Cambodia Unveil \$96.7M Investment in Four Initiatives	EU and Cambodia Unveil \$96.7M Investment in Four Initiatives		Low Accurate
Lifestyle	Taylor Swift Struck a Deal With Singapore Not to Perform in Any Other Southeast Asian Country	1	Singapore Prime Minister Defends Exclusive Deal With Taylor Swift That Riles Some Neighbors	2024-03-05	0.685372	High Accurate	Airbnb is Banning the Use of Indoor Security Cameras in the Platform's Listings	Airbnb is Banning the Use of Indoor Security Cameras in the Platform's Listings		Low Accurate
		2	Taylor Swift Eras Tour Concert Film Coming to Movie Theaters in October	2023-09-02	0.316985	High Accurate	EU and Cambodia Launch \$55.3M Initiative to Promote Green and Digital Jobs	EU and Cambodia Launch \$55.3M Initiative to Promote Green and Digital Jobs		Low Accurate
		3	Taylor Swift The Eras Tour Will be a Blockbuster and Might Shake up the Movie Business	2023-10-11	0.272954	High Accurate	PSE Empowering Cambodian Students	PSE Empowering Cambodian Students		Low Accurate
		4	Taylor Swift Is Named Time Magazines Person of the Year	2023-12-07	0.2394	Medium Accurate	Transitioning to Greener Growth Model Key to Boosting Employment	Transitioning to Greener Growth Model Key to Boosting Employment		Low Accurate
		5	Australia and Laos Elevate Bilateral Relations At Southeast Asian Summit	2024-03-06	0.196307	Low Accurate	EU and Cambodia Unveil \$96.7M Investment in Four Initiatives	EU and Cambodia Unveil \$96.7M Investment in Four Initiatives		Low Accurate
Education	UK University Students to Empower Cambodian Girls with STEM Skills	1	A STEM Classroom Is Creating Cambodia's Future Scientists	2022-10-28	0.307654	High Accurate	Airbnb is Banning the Use of Indoor Security Cameras in the Platform's Listings	Airbnb is Banning the Use of Indoor Security Cameras in the Platform's Listings		Low Accurate
		2	How STEM Is Modernizing Cambodia's Education System	2022-01-14	0.290512	High Accurate	EU and Cambodia Launch \$55.3M Initiative to Promote Green and Digital Jobs	EU and Cambodia Launch \$55.3M Initiative to Promote Green and Digital Jobs		Low Accurate
		3	Using Photography to Explore STEAM Skills	2023-11-18	0.284598	High Accurate	PSE Empowering Cambodian Students	PSE Empowering Cambodian Students		Low Accurate
		4	Strengthening STEM Education Is a Priority	2022-08-31	0.265646	Medium Accurate	Transitioning to Greener Growth Model Key to Boosting Employment	Transitioning to Greener Growth Model Key to Boosting Employment		Low Accurate
		5	Cambodia's Female Coders	2022-01-10	0.237902	Medium Accurate	EU and Cambodia Unveil \$96.7M Investment in Four Initiatives	EU and Cambodia Unveil \$96.7M Investment in Four Initiatives		Low Accurate
Health	Court Detains Lamborghini Driver Behind Fatal Collision	1	Lamborghini Smashes Into Motorbike Killing Driver	2024-02-23	0.459315	High Accurate	Airbnb is Banning the Use of Indoor Security Cameras in the Platform's Listings	Airbnb is Banning the Use of Indoor Security Cameras in the Platform's Listings		Low Accurate
		2	Court Summons for Deadly HitandRun Driver	2023-12-20	0.269611	High Accurate	EU and Cambodia Launch \$55.3M Initiative to Promote Green and Digital Jobs	EU and Cambodia Launch \$55.3M Initiative to Promote Green and Digital Jobs		Low Accurate
		3	Strong Action Key to Curb Surge in Fatal HitandRuns	2023-12-29	0.262686	High Accurate	PSE Empowering Cambodian Students	PSE Empowering Cambodian Students		Low Accurate
		4	PM Orders Police Chief to Tighten Alcohol Road Rules	2024-01-24	0.217533	Medium Accurate	Transitioning to Greener Growth Model Key to Boosting Employment	Transitioning to Greener Growth Model Key to Boosting Employment		Low Accurate
		5	Role of Licences in Driving Down Road Crashes	2024-02-02	0.213389	Medium Accurate	EU and Cambodia Unveil \$96.7M Investment in Four Initiatives	EU and Cambodia Unveil \$96.7M Investment in Four Initiatives		Low Accurate
Travel	AirAsia Cambodia: Enhance Intra-ASEAN Travel, Cambodian Income With Relaxed Visa Fees	1	AirAsia Cambodia Enhances Intra-ASEAN Travel Cambodian Income With Relaxed Visa Fees	2024-02-26	0.443662	High Accurate	Airbnb is Banning the Use of Indoor Security Cameras in the Platform's Listings	Airbnb is Banning the Use of Indoor Security Cameras in the Platform's Listings		Low Accurate
		2	AirAsia Cambodia Maiden Flight TakeOff in Q4	2023-06-21	0.354796	High Accurate	EU and Cambodia Launch \$55.3M Initiative to Promote Green and Digital Jobs	EU and Cambodia Launch \$55.3M Initiative to Promote Green and Digital Jobs		Low Accurate
		3	AirAsia Cambodia To Take Off to Expand ASEAN Market Share	2022-12-09	0.273101	High Accurate	PSE Empowering Cambodian Students	PSE Empowering Cambodian Students		Low Accurate
		4	AirAsia Cambodia Seeks Collaboration To Address High Aviation Costs	2024-03-19	0.244773	Medium Accurate	Transitioning to Greener Growth Model Key to Boosting Employment	Transitioning to Greener Growth Model Key to Boosting Employment		Low Accurate
		5	AirAsia Cambodia To Take Off in May Amid Aviation Costs Setback	2024-03-18	0.215015	Medium Accurate	EU and Cambodia Unveil \$96.7M Investment in Four Initiatives	EU and Cambodia Unveil \$96.7M Investment in Four Initiatives		Low Accurate

High Accurate: 18, Medium Accurate: 9, Low Accuracy: 3

Result 1.2 of TF-IDF with Jaccard similarity

Algorithm 1: TF-IDF with Jaccard similarity						Vs	Kiripost API		
Categories	Input_title	No	Related_Article_Output	Date	Score	Assumption	Currently	Date	Assumption
Business	Annual HEINEKEN Sustaination Invites University Students to Innovate for a Circular Cambodia	1	Glass Bottle Return Movement Gaining Ground in Cambodias Hospitality Scene	2024-03-14	0.246654	Medium Accurate	EU and Cambodia Launch \$55.3M Initiative to Promote Green and Digital Jobs	Medium Accurate	
		2	Heineken Partners with Grab to Drive Home Responsible Consumption Message in Cambodia	2024-01-28	0.246465	Medium Accurate	PSE Empowering Cambodian Students		Medium Accurate
		3	Heinekens Worlds Together Campaign Brings together Community Gatherings	2023-10-02	0.234127	Medium Accurate	Transitioning to Greener Growth Model Key to Boosting Employment		Low Accurate
		4	Heineken Cambodia Appoints Andy Hewson as New Managing Director	2023-09-07	0.197531	Medium Accurate	EU and Cambodia Unveil \$96.7M Investment in Four Initiatives		Low Accurate
		5	ABC Teams up with Cambodias Top Fashion Designers to Custom Make Exclusive Wedding for L	2023-12-08	0.188596	Medium Accurate	National Bank of Canada Refutes Sale of ABA		Low Accurate
Tech	For Social Platforms, the Outage Was Short. But People's Stories Vanished, and That's No Small Thing	1	National Project Supports People with Disabilities	2023-10-25	0.020134	Low Accurate	Airbnb is Banning the Use of Indoor Security Cameras in the Platform's Listings	Low Accurate	
		2	Airbnb Is Banning the Use of Indoor Security Cameras in the Platforms Listings Worldwide	2024-03-13	0.019417	Low Accurate	EU and Cambodia Launch \$55.3M Initiative to Promote Green and Digital Jobs	Low Accurate	
		3	Six Million People Celebrate Water Festival	2023-11-30	0.017699	Low Accurate	PSE Empowering Cambodian Students		Low Accurate
		4	Social Media App for Creating Community for Cambodians	2023-10-22	0.017699	Low Accurate	Transitioning to Greener Growth Model Key to Boosting Employment		Low Accurate
		5	Prahok Season Starts in Cambodia	2022-12-30	0.017241	Low Accurate	EU and Cambodia Unveil \$96.7M Investment in Four Initiatives		Low Accurate
Lifestyle	Taylor Swift Struck a Deal With Singapore Not to Perform in Any Other Southeast Asian Country	1	Singapore Prime Minister Defends Exclusive Deal With Taylor Swift That Riles Some Neighbors	2024-03-05	0.313531	High Accurate	Airbnb is Banning the Use of Indoor Security Cameras in the Platform's Listings	Low Accurate	
		2	Hong Kong Eyes Stronger Economic and Trade Ties with Thailand to Expand its Role in Southeast Asia	2023-10-11	0.114206	Low Accurate	EU and Cambodia Launch \$55.3M Initiative to Promote Green and Digital Jobs	Low Accurate	
		3	Australia and Laos Elevate Bilateral Relations At Southeast Asian Summit	2024-03-06	0.112069	Low Accurate	PSE Empowering Cambodian Students		Low Accurate
		4	Japan and ASEAN Bolster Ties at a Summit Focused on Security and Economy amid Tensions with China	2023-12-18	0.104167	Low Accurate	Transitioning to Greener Growth Model Key to Boosting Employment		Low Accurate
		5	Southeast Asian Leaders Are Besieged by Thorny Issues as They Hold an ASEAN Summit without China	2023-09-04	0.102686	Low Accurate	EU and Cambodia Unveil \$96.7M Investment in Four Initiatives		Low Accurate
Education	UK University Students to Empower Cambodian Girls with STEM Skills	1	Career Center Opens in Phnom Penh	2022-12-23	0.139676	Low Accurate	Airbnb is Banning the Use of Indoor Security Cameras in the Platform's Listings	Low Accurate	
		2	Khoding Hero Program Trains Svay Rieng Students to Join Digital Workforce	2022-08-19	0.131902	Low Accurate	EU and Cambodia Launch \$55.3M Initiative to Promote Green and Digital Jobs	Low Accurate	
		3	Cambodias Female Coder	2022-01-10	0.126649	Low Accurate	PSE Empowering Cambodian Students		Low Accurate
		4	Snooperd Empowers STEM Education	2023-12-17	0.126437	Low Accurate	Transitioning to Greener Growth Model Key to Boosting Employment		Low Accurate
		5	Role Model Raises Cambodias Female Tech Startup Scene	2022-04-21	0.12381	Low Accurate	EU and Cambodia Unveil \$96.7M Investment in Four Initiatives		Low Accurate
Health	Court Detains Lamborghini Driver Behind Fatal Collision	1	Lamborghini Smashes Into Motorbike Killing Driver	2024-02-23	0.244444	High Accurate	Airbnb is Banning the Use of Indoor Security Cameras in the Platform's Listings	Low Accurate	
		2	Role of Licences in Driving Down Road Crashes	2024-02-02	0.152542	Medium Accurate	EU and Cambodia Launch \$55.3M Initiative to Promote Green and Digital Jobs	Low Accurate	
		3	Court Summons for Deadly HitandRun Driver	2023-12-20	0.11828	Medium Accurate	PSE Empowering Cambodian Students		Low Accurate
		4	PM Orders Police Chief to Tighten Alcohol Road Rules	2024-01-24	0.114428	Medium Accurate	Transitioning to Greener Growth Model Key to Boosting Employment		Low Accurate
		5	PM Orders Arrest of Fatal HitandRun Suspect	2023-12-22	0.111111	Medium Accurate	EU and Cambodia Unveil \$96.7M Investment in Four Initiatives		Low Accurate
Travel	AirAsia Cambodia: Enhance Intra-ASEAN Travel, Cambodian Income With Relaxed Visa Fees	1	More than 2000 Expats Apply for Free Angkor Pass	2022-09-24	0.029851	Low Accurate	Airbnb is Banning the Use of Indoor Security Cameras in the Platform's Listings	Low Accurate	
		2	Cambodia Smashes World Record for Most Paper Hearts	2023-04-12	0.029126	Low Accurate	EU and Cambodia Launch \$55.3M Initiative to Promote Green and Digital Jobs	Low Accurate	
		3	Rice Exports Total 637004 Tons in 2022	2023-01-06	0.028169	Low Accurate	PSE Empowering Cambodian Students		Low Accurate
		4	Cambodias Expats to Enjoy Free Access to Angkor	2022-08-04	0.027397	Low Accurate	Transitioning to Greener Growth Model Key to Boosting Employment		Low Accurate
		5	Expats Can Enjoy Free Entry to Angkor for One Year	2022-09-02	0.027211	Low Accurate	EU and Cambodia Unveil \$96.7M Investment in Four Initiatives		Low Accurate

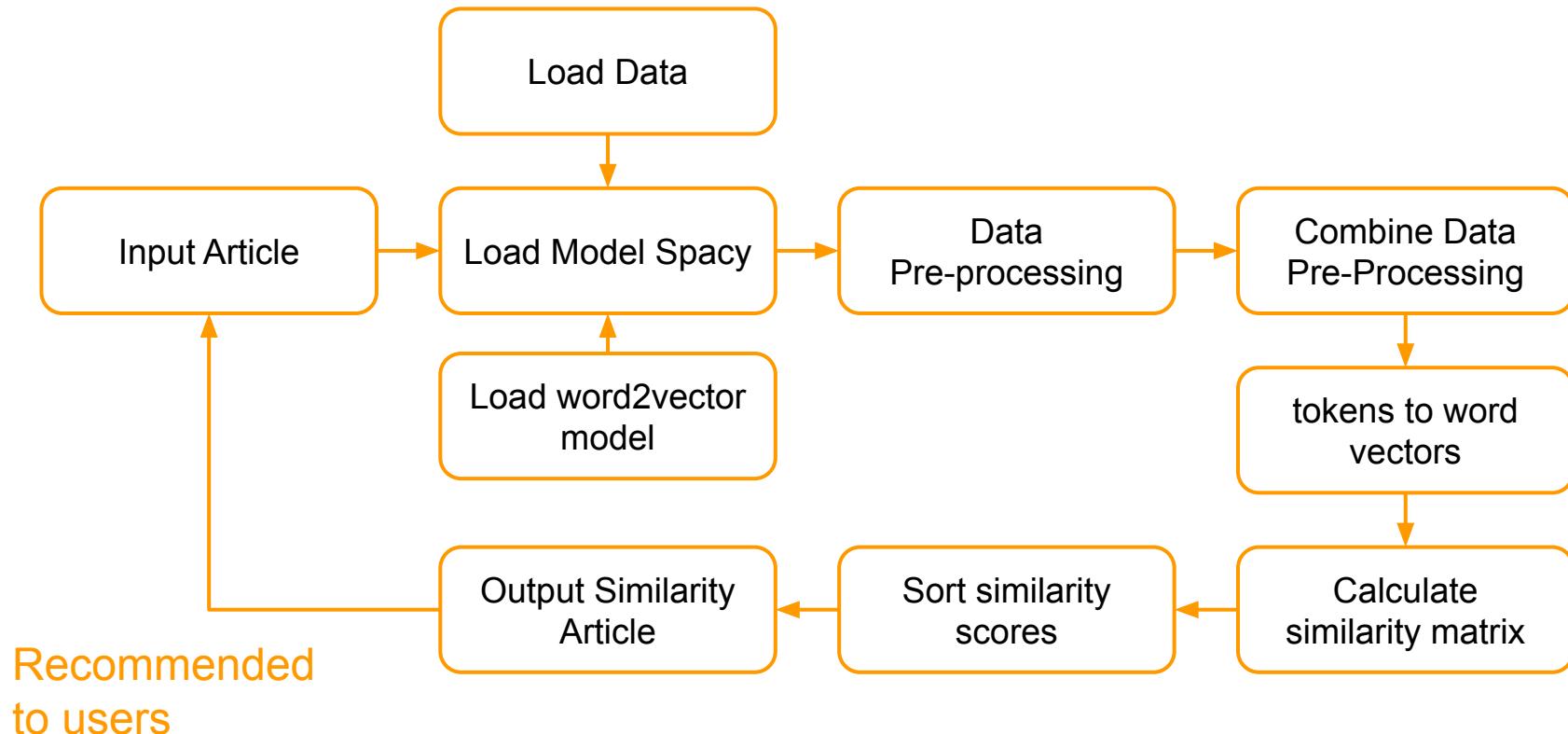
High Accurate: 2, Medium Accurate: 9, Low Accuracy: 19

Result 1.3 of TF-IDF with Euclidean distance

Algorithm 1: TF-IDF with Euclidean distance							Vs	Kiripost API		
Categories	Input_title	No	Related_Article_Output	date	Score	Assumption	Currently	Date	Assumption	
Business	Annual HEINEKEN Sustainathon Invites University Students to Innovate for a Circular Cambodia	1	Winners Crowned for Sustainable Innovations in Heineken Sustainathon	2023-05-10	[[1.11142]	High Accurate	EU and Cambodia Launch \$55.3M Initiative to Promote Green and Digital Jobs	Medium	Accurate	
		2	Heineken Cambodia Appoints Andy Hewson as New Managing Director	2023-09-07	[[1.22530]	High Accurate	PSE Empowering Cambodian Students	Medium	Accurate	
		3	Heineken Partners with Grab to Drive Home Responsible Consumption Message in Cambodia	2024-01-28	[[1.23100]	High Accurate	Transitioning to Greener Growth Model Key to Boosting Employment	Low	Accurate	
		4	Heineken's Worlds Together Campaign Brings together Community Gatherings	2023-10-02	[[1.24432]	High Accurate	EU and Cambodia Unveil \$96.7M Investment in Four Initiatives	Low	Accurate	
		5	Heineken Commemorates 150 Years of Good Times	2023-06-28	[[1.25744]	High Accurate	National Bank of Canada Refutes Sale of ABA	Low	Accurate	
Tech	For Social Platforms, the Outage Was Short. But People's Stories Vanished, and That's No Small Thing	1	BREAKING Meta platforms Including Facebook Instagram and Messenger are Experiencing a Glitch	2024-03-05	[[1.27761]	High Accurate	Airbnb is Banning the Use of Indoor Security Cameras in the Platform's Listings	Low	Accurate	
		2	Metas Meltdown Resolved After Hours of Outage	2024-03-06	[[1.31175]	High Accurate	EU and Cambodia Launch \$55.3M Initiative to Promote Green and Digital Jobs	Low	Accurate	
		3	Capturing Cambodias Stories of Change	2023-11-28	[[1.33024]	Low Accurate	PSE Empowering Cambodian Students	Low	Accurate	
		4	Chaktomuk Short Film Festival Welcomes Rise in Entries	2022-11-23	[[1.34553]	Low Accurate	Transitioning to Greener Growth Model Key to Boosting Employment	Low	Accurate	
		5	Microsoft Teams Recovering After Outage	2022-07-21	[[1.35213]	Low Accurate	EU and Cambodia Unveil \$96.7M Investment in Four Initiatives	Low	Accurate	
Lifestyle	Taylor Swift Struck a Deal With Singapore Not to Perform in Any Other Southeast Asian Country	1	Singapore Prime Minister Defends Exclusive Deal With Taylor Swift That Riles Some Neighbors	2024-03-05	[[1.01775]	High Accurate	Airbnb is Banning the Use of Indoor Security Cameras in the Platform's Listings	Low	Accurate	
		2	Taylor Swift Is Named Time Magazines Person of the Year	2023-12-07	[[1.19942]	Medium Accurate	EU and Cambodia Launch \$55.3M Initiative to Promote Green and Digital Jobs	Low	Accurate	
		3	Taylor Swift Eras Tour Concert Film Coming to Movie Theaters in October	2023-09-02	[[1.22225]	Medium Accurate	PSE Empowering Cambodian Students	Low	Accurate	
		4	Taylor Swift The Eras Tour Will be a Blockbuster and Might Shake up the Movie Business	2023-10-11	[[1.25216]	Medium Accurate	Transitioning to Greener Growth Model Key to Boosting Employment	Low	Accurate	
		5	Taylor Swift Kisses Travis Kelce on the Field After His Chiefs Beat the 49ers in the Super Bowl	2024-02-12	[[1.27752]	Medium Accurate	EU and Cambodia Unveil \$96.7M Investment in Four Initiatives	Low	Accurate	
Education	UK University Students to Empower Cambodian Girls with STEM Skills	1	Call for UK to Invest in Cambodias Agriculture Sector	2023-01-28	[[1.15421]	High Accurate	Airbnb is Banning the Use of Indoor Security Cameras in the Platform's Listings	Low	Accurate	
		2	UK and Cambodian Officials Meet to Strengthen Alliance	2024-03-06	[[1.16280]	High Accurate	EU and Cambodia Launch \$55.3M Initiative to Promote Green and Digital Jobs	Low	Accurate	
		3	Strengthening STEM Education Is A Priority	2022-08-31	[[1.21548]	High Accurate	PSE Empowering Cambodian Students	Low	Accurate	
		4	How STEM Is Modernizing Cambodias Education System	2022-01-14	[[1.21801]	Medium Accurate	Transitioning to Greener Growth Model Key to Boosting Employment	Low	Accurate	
		5	A STEM Classroom Is Creating Cambodias Future Scientists	2022-10-28	[[1.22191]	Medium Accurate	EU and Cambodia Unveil \$96.7M Investment in Four Initiatives	Low	Accurate	
Health	Court Detains Lamborghini Driver Behind Fatal Collision	1	Lamborghini Smashes Into Motorbike Killing Driver	2024-02-23	[[1.19915]	High Accurate	Airbnb is Banning the Use of Indoor Security Cameras in the Platform's Listings	Low	Accurate	
		2	PM Orders Police Chief to Tighten Alcohol Road Rules	2024-01-24	[[1.28709]	High Accurate	EU and Cambodia Launch \$55.3M Initiative to Promote Green and Digital Jobs	Low	Accurate	
		3	Court Summons for Deadly HitandRun Driver	2023-12-20	[[1.29997]	High Accurate	PSE Empowering Cambodian Students	Low	Accurate	
		4	Strong Action Key to Curb Surge in Fatal HitandRuns	2023-12-29	[[1.31236]	Medium Accurate	Transitioning to Greener Growth Model Key to Boosting Employment	Low	Accurate	
		5	Nutrition Advocacy Group Calls for NutriLatt Court Case Rethink	2024-01-10	[[1.31699]	Low Accurate	EU and Cambodia Unveil \$96.7M Investment in Four Initiatives	Low	Accurate	
Travel	AirAsia Cambodia: Enhance Intra-ASEAN Travel, Cambodian Income With Relaxed Visa Fees	1	AirAsia Cambodia Maiden Flight TakeOff in Q4	2023-06-21	[[1.13596]	High Accurate	Airbnb is Banning the Use of Indoor Security Cameras in the Platform's Listings	Low	Accurate	
		2	AirAsia Cambodia to Take Off to Expand ASEAN Market Share	2022-12-09	[[1.20573]	High Accurate	EU and Cambodia Launch \$55.3M Initiative to Promote Green and Digital Jobs	Low	Accurate	
		3	AirAsia Cambodia Seeks Collaboration To Address High Aviation Costs	2024-03-19	[[1.22900]	High Accurate	PSE Empowering Cambodian Students	Low	Accurate	
		4	AirAsia Cambodia To Take Off in May Amid Aviation Costs Setback	2024-03-18	[[1.25298]	High Accurate	Transitioning to Greener Growth Model Key to Boosting Employment	Low	Accurate	
		5	Thailand Reduces ReEntry Visa Fees for Migrant Workers	2024-02-10	[[1.25512]	High Accurate	EU and Cambodia Unveil \$96.7M Investment in Four Initiatives	Low	Accurate	

High Accurate: 19, Medium Accurate: 7, Low Accuracy: 4

2. Flowchart of word2vector Algorithm



Result 2.1 of Word2vector with Cosine similarity

Algorithm 2: Word2vector with Consine similarity							Vs	Kiripost API		
Categories	Input_title	No	Related_Article_Output	date	Score	Assumption		Currently	Date	Assumption
Business	Annual HEINEKEN Sustainathon Invites University Students to Innovate for a Circular Cambodia	1	Asian Cup Everything You Need to Know as World Cup Host Qatar Stages Latest Soccer Tournar	2024-01-11	0.992569	Low Accurate	EU and Cambodia Launch \$55.3M Initiative to Promote Green and Digital Jobs	Medium Accurate		
		2	Bamboo Airways Connects Siem Reap with Hanoi	2022-08-08	0.992378	Low Accurate	PSE Empowering Cambodian Students	Medium Accurate		
		3	Restaurant Week to Throw Spotlight on Quality US Produce in Cambodia	2023-11-14	0.992313	Low Accurate	Transitioning to Greener Growth Model Key to Boosting Employment		Low Accurate	
		4	Cambodia and Kuwait to Strengthen Bilateral Ties	2022-12-01	0.99227	Low Accurate	EU and Cambodia Unveil \$96.7M Investment in Four Initiatives		Low Accurate	
		5	DMC Students Showcase their Film Skills	2023-08-29	0.992211	Low Accurate	National Bank of Canada Refutes Sale of ABA		Low Accurate	
Tech	For Social Platforms, the Outage Was Short. But People's Stories Vanished, and That's No Small Thing	1	Cambodia to Host Mekong Tourism Forum 2023	2022-10-14	0.986834	Low Accurate	Airbnb is Banning the Use of Indoor Security Cameras in the Platform's Listings	Low Accurate		
		2	AirAsia Resumes Flights Between Sihanoukville and KL and Bangkok	2023-06-13	0.986595	Low Accurate	EU and Cambodia Launch \$55.3M Initiative to Promote Green and Digital Jobs	Low Accurate		
		3	Two Tourism Training Schools to Open	2022-11-18	0.98649	Low Accurate	PSE Empowering Cambodian Students		Low Accurate	
		4	Kun Lbokator Inscribed on UNESCO World Heritage List	2022-11-30	0.986443	Low Accurate	Transitioning to Greener Growth Model Key to Boosting Employment		Low Accurate	
		5	Cambodia Aims to Position Itself as A Solo Destination for Tourists	2023-08-14	0.986399	Low Accurate	EU and Cambodia Unveil \$96.7M Investment in Four Initiatives		Low Accurate	
Lifestyle	Taylor Swift Struck a Deal With Singapore Not to Perform in Any Other Southeast Asian Country	1	Taylor Swift Struck a Deal With Singapore Not to Perform in Any Other Southeast Asian Country	2024-03-06	0.993242	High Accurate	Airbnb is Banning the Use of Indoor Security Cameras in the Platform's Listings	Low Accurate		
		2	Cambodia Gears up to Host Mekong Tourism Forum	2023-02-28	0.99325	Low Accurate	EU and Cambodia Launch \$55.3M Initiative to Promote Green and Digital Jobs	Low Accurate		
		3	Wat Ounalom's Heritage Buildings Totally Razed	2022-12-22	0.99324	Low Accurate	PSE Empowering Cambodian Students		Low Accurate	
		4	Cambodia Successfully Hosts Mekong Tourism Forum	2023-04-28	0.993195	Low Accurate	Transitioning to Greener Growth Model Key to Boosting Employment		Low Accurate	
		5	First Destination Siem Reap Day Aims to Boost Tourism	2023-03-29	0.993155	Low Accurate	EU and Cambodia Unveil \$96.7M Investment in Four Initiatives		Low Accurate	
Education	UK University Students to Empower Cambodian Girls with STEM Skills	1	Niche in Auto Green Energy Sectors Can Lure EU Business	2023-03-26	0.995033	Low Accurate	Airbnb is Banning the Use of Indoor Security Cameras in the Platform's Listings	Low Accurate		
		2	Youth Future Fair Shapes Sustainable Future	2023-10-23	0.994884	Low Accurate	EU and Cambodia Launch \$55.3M Initiative to Promote Green and Digital Jobs	Low Accurate		
		3	Talking Business with Entrepreneurship Guru Stephen Paterson	2023-01-06	0.994795	Low Accurate	PSE Empowering Cambodian Students		Low Accurate	
		4	Sastraa Film cofounders high hopes for Cambodia's burgeoning film industry	2022-02-07	0.994766	Low Accurate	Transitioning to Greener Growth Model Key to Boosting Employment		Low Accurate	
		5	Richard Yim How to Secure Quality Investment	2022-12-14	0.994735	Low Accurate	EU and Cambodia Unveil \$96.7M Investment in Four Initiatives		Low Accurate	
Health	Court Detains Lamborghini Driver Behind Fatal Collision	1	Innovative Pilot Project Produces Rice Husk Briquettes	2022-12-09	0.994508	Low Accurate	Airbnb is Banning the Use of Indoor Security Cameras in the Platform's Listings	Low Accurate		
		2	Launching Cambodia's Credit Bureau	2022-12-07	0.994491	Low Accurate	EU and Cambodia Launch \$55.3M Initiative to Promote Green and Digital Jobs	Low Accurate		
		3	Ezeecom and Huawei Partner to Drive Cloud Adoption	2023-06-06	0.994427	Low Accurate	PSE Empowering Cambodian Students		Low Accurate	
		4	Mark Zuckerberg's Long Apology Tour: A Brief History	2024-02-01	0.994395	Low Accurate	Transitioning to Greener Growth Model Key to Boosting Employment		Low Accurate	
		5	Improving Cambodia's Agriculture Sector	2023-04-24	0.994323	Low Accurate	EU and Cambodia Unveil \$96.7M Investment in Four Initiatives		Low Accurate	
Travel	AirAsia Cambodia: Enhance Intra-ASEAN Travel, Cambodian Income With Relaxed Visa Fees	1	PropertyGuru Awards Celebrate Cambodia's Real Estate Success	2023-08-21	0.970795	Low Accurate	Airbnb is Banning the Use of Indoor Security Cameras in the Platform's Listings	Low Accurate		
		2	ACLEDA Crowned SME Financier of the Year	2022-09-23	0.969854	Low Accurate	EU and Cambodia Launch \$55.3M Initiative to Promote Green and Digital Jobs	Low Accurate		
		3	Head of FFC Reverses Decision to Resign	2023-05-09	0.969307	Low Accurate	PSE Empowering Cambodian Students		Low Accurate	
		4	ACLEDA Scoops Award for Aiding Cambodian SMEs	2022-09-22	0.969275	Low Accurate	Transitioning to Greener Growth Model Key to Boosting Employment		Low Accurate	
		5	Officials Arrive in Phnom Penh for ASEAN Summit	2022-11-07	0.969122	Low Accurate	EU and Cambodia Unveil \$96.7M Investment in Four Initiatives		Low Accurate	

High Accurate: 1, Medium Accurate: 0, Low Accuracy: 29

Result 2.2 of Word2vector with Jaccard similarity

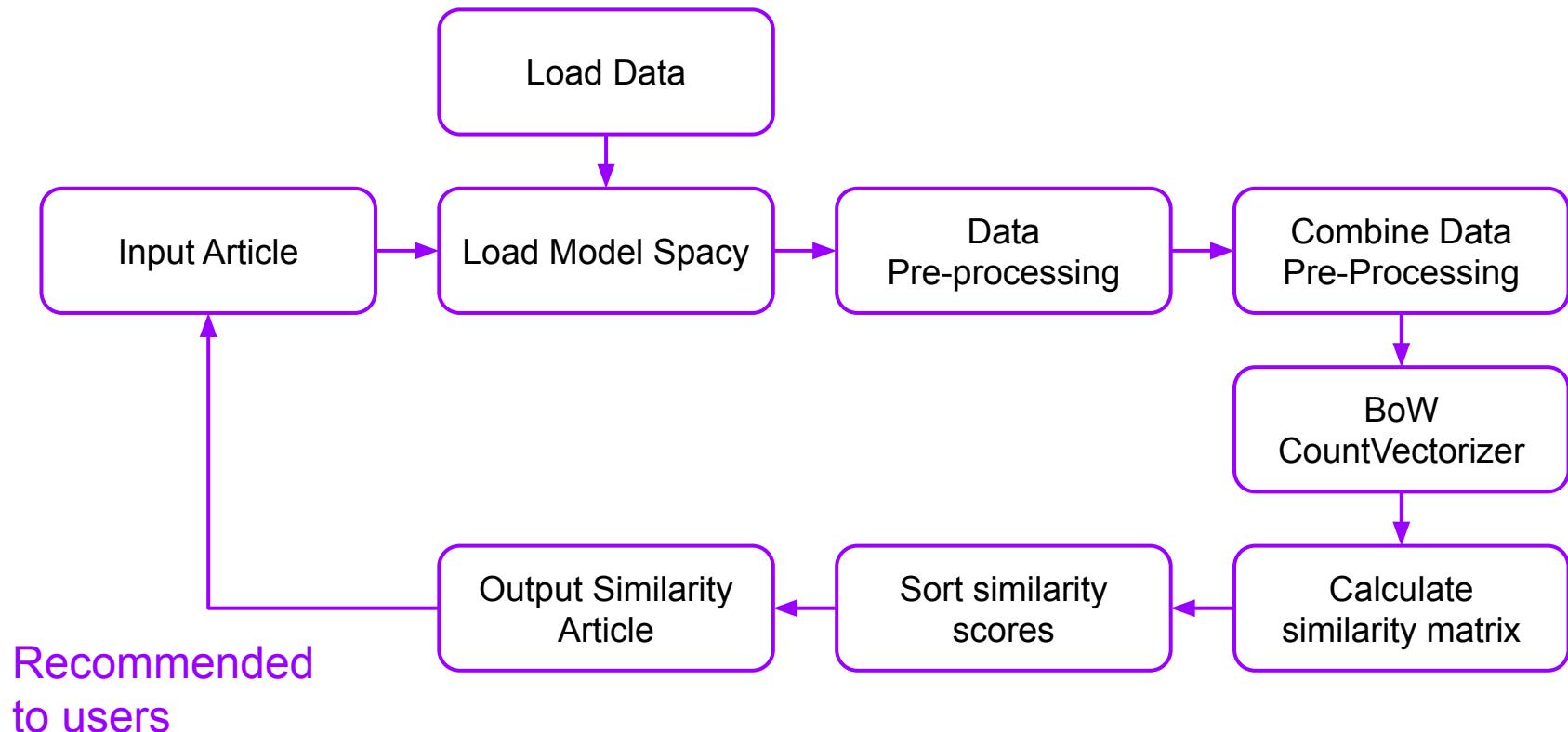
Algorithm 2: Word2vector with Jaccard similarity							VS	Kripost API		
Categories	Input_title	No	Related_Article_Output	Date	Score	Assumption		Date	Assumption	
Business	Annual HEINEKEN Sustaination Invites University Students to Innovate for a Circular Cambodia	1	Asian Cup Everything You Need to Know as World Cup Host Qatar Stages Latest Soccer Tourna	2024-01-11	0.203081	Low Accurate	Currently			
		2	Cambodia to Host Mekong Tourism Forum 2023	2022-10-14	0.208485	Low Accurate	EU and Cambodia Launch \$55.3M Initiative to Promote Green and Digital Jobs	Medium Accurate		
		3	AIESEC Cambodia Organizes Youth to Business Event	2023-10-26	0.208675	Low Accurate	PSE Empowering Cambodian Students	Medium Accurate		
		4	Cambodian Egamers Gear Up to Battle for the Kingdom	2023-02-27	0.209166	Low Accurate	Transitioning to Greener Growth Model Key to Boosting Employment	Low Accurate		
		5	Kun Lbokator Inscribed on UNESCO World Heritage List	2022-11-30	0.210048	Low Accurate	EU and Cambodia Unveil \$96.7M Investment in Four Initiatives	Low Accurate		
Tech	For Social Platforms, the Outage Was Short. But People's Stories Vanished, and That's No Small Thing	1	Cambodia to Host Mekong Tourism Forum 2023	2022-10-14	0.269386	Low Accurate	National Bank of Canada Refutes Sale of ABA	Low Accurate		
		2	Kun Lbokator Inscribed on UNESCO World Heritage List	2022-11-30	0.273387	Low Accurate	Airbnb is Banning the Use of Indoor Security Cameras in the Platform's Listings	Low Accurate		
		3	Cambodia Gears up to Host Mekong Tourism Forum	2023-02-28	0.279206	Low Accurate	EU and Cambodia Launch \$55.3M Initiative to Promote Green and Digital Jobs	Low Accurate		
		4	Cambodias Esporters Prepare for Phnom Penh Hosted Asia Cup	2023-06-06	0.279485	Low Accurate	PSE Empowering Cambodian Students	Low Accurate		
		5	Two Tourism Training Schools to Open	2022-11-18	0.279634	Low Accurate	Transitioning to Greener Growth Model Key to Boosting Employment	Low Accurate		
Lifestyle	Taylor Swift Struck a Deal With Singapore Not to Perform in Any Other Southeast Asian Country	1	Cambodia Gears up to Host Mekong Tourism Forum	2023-02-28	0.195325	Low Accurate	EU and Cambodia Unveil \$96.7M Investment in Four Initiatives	Low Accurate		
		2	Wat Ounaloms Heritage Buildings Totally Razed	2022-12-22	0.196614	Low Accurate	Airbnb is Banning the Use of Indoor Security Cameras in the Platform's Listings	Low Accurate		
		3	First Destination Siem Reap Day Aims to Boost Tourism	2023-03-29	0.197555	Low Accurate	EU and Cambodia Launch \$55.3M Initiative to Promote Green and Digital Jobs	Low Accurate		
		4	Cambodia Successfully Hosts Mekong Tourism Forum	2023-04-28	0.198173	Low Accurate	PSE Empowering Cambodian Students	Low Accurate		
		5	No More Discrimination Against Humanity Urges LGBT Group	2023-10-31	0.19851	Low Accurate	Transitioning to Greener Growth Model Key to Boosting Employment	Low Accurate		
Education	UK University Students to Empower Cambodian Girls with STEM Skills	1	Niche in Auto Green Energy Sectors Can Lure EU Business	2023-03-26	0.171742	Low Accurate	EU and Cambodia Unveil \$96.7M Investment in Four Initiatives	Low Accurate		
		2	Youth Future Fair Shapes Sustainable Future	2023-10-23	0.1736	Low Accurate	Airbnb is Banning the Use of Indoor Security Cameras in the Platform's Listings	Low Accurate		
		3	No More Discrimination Against Humanity Urges LGBT Group	2023-10-31	0.174164	Low Accurate	EU and Cambodia Launch \$55.3M Initiative to Promote Green and Digital Jobs	Low Accurate		
		4	Talking Business with Entrepreneurship Guru Stephen Paterson	2023-01-06	0.174186	Low Accurate	PSE Empowering Cambodian Students	Low Accurate		
		5	2023 in Other Words AI Might Be the Term of the Year but Consider These Farflung Contenders	2023-12-13	0.174365	Low Accurate	Transitioning to Greener Growth Model Key to Boosting Employment	Low Accurate		
Health	Court Detains Lamborghini Driver Behind Fatal Collision	1	Launching Cambodias Credit Bureau	2022-12-07	0.178907	Low Accurate	EU and Cambodia Unveil \$96.7M Investment in Four Initiatives	Low Accurate		
		2	Mark Zuckerbergs Long Apology Tour A Brief History	2024-02-01	0.180604	Low Accurate	Airbnb is Banning the Use of Indoor Security Cameras in the Platform's Listings	Low Accurate		
		3	Ezecom and Huawei Partner to Drive Cloud Adoption	2023-06-06	0.180615	Low Accurate	EU and Cambodia Launch \$55.3M Initiative to Promote Green and Digital Jobs	Low Accurate		
		4	Innovative Pilot Project Produces Rice Husk Briquettes	2022-12-09	0.18138	Low Accurate	PSE Empowering Cambodian Students	Low Accurate		
		5	Cultural Club Defies Closure and Financial Struggles	2024-02-27	0.18224	Low Accurate	Transitioning to Greener Growth Model Key to Boosting Employment	Low Accurate		
Travel	AirAsia Cambodia: Enhance Intra-ASEAN Travel, Cambodian Income With Relaxed Visa Fees	1	ACLEDA Crowned SME Financier of the Year	2022-09-23	0.393206	Low Accurate	EU and Cambodia Unveil \$96.7M Investment in Four Initiatives	Low Accurate		
		2	ACLEDA Scoops Award for Aiding Cambodian SMEs	2022-09-22	0.409243	Low Accurate	Airbnb is Banning the Use of Indoor Security Cameras in the Platform's Listings	Low Accurate		
		3	Hun Manet Youngest Leader at ASEAN Summit	2023-09-05	0.410797	Low Accurate	EU and Cambodia Launch \$55.3M Initiative to Promote Green and Digital Jobs	Low Accurate		
		4	ACLEDA Securities Most Traded Value on Cambodian Stock Exchange	2022-07-11	0.416501	Low Accurate	PSE Empowering Cambodian Students	Low Accurate		
		5	MJQE Net Profit Skyrockets to 23m in FY22	2023-07-11	0.418062	Low Accurate	Transitioning to Greener Growth Model Key to Boosting Employment	Low Accurate		
High Accurate: 0, Medium Accurate: 0, Low Accuracy: 30										

Result 2.3 of Word2vector with Euclidean distance

Algorithm 2: Word2vector with Euclidean_distance						Vs	Kiripost API		
Categories	Input_title	No	Related_Article_Output	date	Score	Assumption	Currently	Date	Assumption
Business	Annual HEINEKEN Sustanathon Invites University Students to Innovate for a Circular Cambodia	1	Asian Cup Everything You Need to Know as World Cup Host Qatar Stages Latest Soccer Tournar	2024-01-11	0.203081	Low Accurate	EU and Cambodia Launch \$55.3M Initiative to Promote Green and Digital Jobs	Medium Accurate	
		2	Cambodia to Host Mekong Tourism Forum 2023	2022-10-14	0.208485	Low Accurate	PSE Empowering Cambodian Students		Medium Accurate
		3	AIESEC Cambodia Organizes Youth to Business Event	2023-10-26	0.208675	Low Accurate	Transitioning to Greener Growth Model Key to Boosting Employment		Low Accurate
		4	Cambodian Egamers Gear Up to Battle for the Kingdom	2023-02-27	0.209166	Low Accurate	EU and Cambodia Unveil \$96.7M Investment in Four Initiatives		Low Accurate
		5	Kun Lbokator Inscribed on UNESCO World Heritage List	2022-11-30	0.210048	Low Accurate	National Bank of Canada Refutes Sale of ABA		Low Accurate
Tech	For Social Platforms, the Outage Was Short. But People's Stories Vanished, and That's No Small Thing	1	Cambodia to Host Mekong Tourism Forum 2023	2022-10-14	0.269386	Low Accurate	Airbnb is Banning the Use of Indoor Security Cameras in the Platform's Listings	Low Accurate	
		2	Kun Lbokator Inscribed on UNESCO World Heritage List	2022-11-30	0.273387	Low Accurate	EU and Cambodia Launch \$55.3M Initiative to Promote Green and Digital Jobs	Low Accurate	
		3	Cambodia Gears up to Host Mekong Tourism Forum	2023-02-28	0.279206	Low Accurate	PSE Empowering Cambodian Students		Low Accurate
		4	Cambodias Esporters Prepare for Phnom Penh Hosted Asia Cup	2023-06-06	0.279485	Low Accurate	Transitioning to Greener Growth Model Key to Boosting Employment	Low Accurate	
		5	Two Tourism Training Schools to Open	2022-11-18	0.279634	Low Accurate	EU and Cambodia Unveil \$96.7M Investment in Four Initiatives		Low Accurate
Lifestyle	Taylor Swift Struck a Deal With Singapore Not to Perform in Any Other Southeast Asian Country	1	Cambodia Gears up to Host Mekong Tourism Forum	2023-02-28	0.195325	Low Accurate	Airbnb is Banning the Use of Indoor Security Cameras in the Platform's Listings	Low Accurate	
		2	Wat Ounaloms Heritage Buildings Totally Razed	2022-12-22	0.196614	Low Accurate	EU and Cambodia Launch \$55.3M Initiative to Promote Green and Digital Jobs	Low Accurate	
		3	First Destination Siem Reap Day Aims to Boost Tourism	2023-03-29	0.197555	Low Accurate	PSE Empowering Cambodian Students		Low Accurate
		4	Cambodia Successfully Hosts Mekong Tourism Forum	2023-04-28	0.198173	Low Accurate	Transitioning to Greener Growth Model Key to Boosting Employment	Low Accurate	
		5	No More Discrimination Against Humanity Urges LGBT Group	2023-10-31	0.19851	Low Accurate	EU and Cambodia Unveil \$96.7M Investment in Four Initiatives		Low Accurate
Education	UK University Students to Empower Cambodian Girls with STEM Skills	1	Niche in Auto Green Energy Sectors Can Lure EU Business	2023-03-26	0.171742	Low Accurate	Airbnb is Banning the Use of Indoor Security Cameras in the Platform's Listings	Low Accurate	
		2	Youth Future Fair Shapes Sustainable Future	2023-10-23	0.1736	Low Accurate	EU and Cambodia Launch \$55.3M Initiative to Promote Green and Digital Jobs	Low Accurate	
		3	No More Discrimination Against Humanity Urges LGBT Group	2023-10-31	0.174164	Low Accurate	PSE Empowering Cambodian Students		Low Accurate
		4	Talking Business with Entrepreneurship Guru Stephen Paterson	2023-01-06	0.174186	Low Accurate	Transitioning to Greener Growth Model Key to Boosting Employment	Low Accurate	
		5	2023 in Other Words AI Might Be the Term of the Year but Consider These Farflung Contenders	2023-12-13	0.174365	Low Accurate	EU and Cambodia Unveil \$96.7M Investment in Four Initiatives		Low Accurate
Health	Court Detains Lamborghini Driver Behind Fatal Collision	1	Launching Cambodias Credit Bureau	2022-12-07	0.178907	Low Accurate	Airbnb is Banning the Use of Indoor Security Cameras in the Platform's Listings	Low Accurate	
		2	Mark Zuckerbergs Long Apology Tour A Brief History	2024-02-01	0.180604	Low Accurate	EU and Cambodia Launch \$55.3M Initiative to Promote Green and Digital Jobs	Low Accurate	
		3	Ezecom and Huawei Partner to Drive Cloud Adoption	2023-06-06	0.180615	Low Accurate	PSE Empowering Cambodian Students		Low Accurate
		4	Innovative Pilot Project Produces Rice Husk Briquettes	2022-12-09	0.18138	Low Accurate	Transitioning to Greener Growth Model Key to Boosting Employment	Low Accurate	
		5	Cultural Club Defies Closure and Financial Struggles	2024-02-27	0.18224	Low Accurate	EU and Cambodia Unveil \$96.7M Investment in Four Initiatives		Low Accurate
Travel	AirAsia Cambodia: Enhance Intra-ASEAN Travel, Cambodian Income With Relaxed Visa Fees	1	ACLEDA Crowned SME Financier of the Year	2022-09-23	0.393206	Low Accurate	Airbnb is Banning the Use of Indoor Security Cameras in the Platform's Listings	Low Accurate	
		2	ACLEDA Scoops Award for Aiding Cambodian SMEs	2022-09-22	0.409243	Low Accurate	EU and Cambodia Launch \$55.3M Initiative to Promote Green and Digital Jobs	Low Accurate	
		3	Hun Manet Youngest Leader at ASEAN Summit	2023-09-05	0.410797	Low Accurate	PSE Empowering Cambodian Students		Low Accurate
		4	ACLEDA Securities Most Traded Value on Cambodian Stock Exchange	2022-07-11	0.416501	Low Accurate	Transitioning to Greener Growth Model Key to Boosting Employment	Low Accurate	
		5	MJQE Net Profit Skyrockets to 23m in FY22	2023-07-11	0.418062	Low Accurate	EU and Cambodia Unveil \$96.7M Investment in Four Initiatives		Low Accurate

High Accurate: 0, Medium Accurate: 0, Low Accuracy: 30

3. Flowchart of BoW Algorithm



Result 3.1 of Bag of word with Cosine similarity

Algorithm 3: Bag of word with Cosine similarity							Vs	Kiripost API		
Categories	Input_title	No	Related_Article_Output	date	Score	Assumption	Currently	Date	Assumption	
Business	Annual HEINEKEN Sustainathon Invites University Students to Innovate for a Circular Cambodia	1	Heineken Partners with Grab to Drive Home Responsible Consumption Message in Cambodia	2024-01-28	0.588042	High Accurate	EU and Cambodia Launch \$55.3M Initiative to Promote Green and Digital Jobs PSE Empowering Cambodian Students Transitioning to Greener Growth Model Key to Boosting Employment EU and Cambodia Unveil \$96.7M Investment in Four Initiatives National Bank of Canada Refutes Sale of ABA	Medium Accurate Medium Accurate Low Accurate Low Accurate Low Accurate		
		2	Winners Crowned for Sustainable Innovations in Heineken Sustainathon	2023-05-10	0.583091	High Accurate				
		3	Heineken's Worlds Together Campaign Brings together Community Gatherings	2023-10-02	0.573357	High Accurate				
		4	Heineken Cambodia Appoints Andy Hewson as New Managing Director	2023-09-07	0.553354	High Accurate				
		5	Glass Bottle Return Movement Gaining Ground in Cambodia's Hospitality Scene	2024-03-14	0.523219	Low Accurate				
Tech	For Social Platforms, the Outage Was Short. But People's Stories Vanished, and That's No Small Thing	1	PM Allocates 450M to Social Protection System	2024-02-06	0.246598	Low Accurate	Airbnb is Banning the Use of Indoor Security Cameras in the Platform's Listings EU and Cambodia Launch \$55.3M Initiative to Promote Green and Digital Jobs PSE Empowering Cambodian Students Transitioning to Greener Growth Model Key to Boosting Employment EU and Cambodia Unveil \$96.7M Investment in Four Initiatives	Low Accurate Low Accurate Low Accurate Low Accurate Low Accurate		
		2	Cambodia's Rising Elderly Population Needs More Priority Policies	2023-10-30	0.240498	Low Accurate				
		3	Social Security Week to Tackle Efficiency and Transparency	2023-02-10	0.217674	Low Accurate				
		4	Fighting for the Rights of Cambodians with Disabilities	2023-08-31	0.198775	Low Accurate				
		5	Call for More Awareness for People with Disabilities	2023-10-31	0.193798	Low Accurate				
Lifestyle	Taylor Swift Struck a Deal With Singapore Not to Perform in Any Other Southeast Asian Country	1	Singapore Prime Minister Defends Exclusive Deal With Taylor Swift That Riles Some Neighbors	2024-03-05	0.758516	High Accurate	Airbnb is Banning the Use of Indoor Security Cameras in the Platform's Listings EU and Cambodia Launch \$55.3M Initiative to Promote Green and Digital Jobs PSE Empowering Cambodian Students Transitioning to Greener Growth Model Key to Boosting Employment EU and Cambodia Unveil \$96.7M Investment in Four Initiatives	Low Accurate Low Accurate Low Accurate Low Accurate Low Accurate		
		2	Southeast Asian Leaders Are Besieged by Thorny Issues as They Hold an ASEAN Summit without Taylor Swift	2023-09-04	0.313462	Low Accurate				
		3	Taylor Swift The Eras Tour Will be a Blockbuster and Might Shake up the Movie Business	2023-10-11	0.283043	Medium Accurate				
		4	Taylor Swift Eras Tour Concert Film Coming to Movie Theaters in October	2023-09-02	0.271411	Medium Accurate				
		5	Australia and Laos Elevate Bilateral Relations At Southeast Asian Summit	2024-03-06	0.268896	Low Accurate				
Education	UK University Students to Empower Cambodian Girls with STEM Skills	1	How STEM Is Modernizing Cambodia's Education System	2022-01-14	0.461497	High Accurate	Airbnb is Banning the Use of Indoor Security Cameras in the Platform's Listings EU and Cambodia Launch \$55.3M Initiative to Promote Green and Digital Jobs PSE Empowering Cambodian Students Transitioning to Greener Growth Model Key to Boosting Employment EU and Cambodia Unveil \$96.7M Investment in Four Initiatives	Low Accurate Low Accurate Low Accurate Low Accurate Low Accurate		
		2	A STEM Classroom Is Creating Cambodia's Future Scientists	2022-10-28	0.461476	High Accurate				
		3	Khodong Hero Program Trains Svay Rieng Students to Join Digital Workforce	2022-08-19	0.427643	Low Accurate				
		4	Career Center Opens in Phnom Penh	2022-12-23	0.416502	Low Accurate				
		5	Using Photography to Explore STEAM Skills	2023-11-18	0.373	Medium Accurate				
Health	Court Detains Lamborghini Driver Behind Fatal Collision	1	Lamborghini Smashes Into Motorbike Killing Driver	2024-02-23	0.48745	High Accurate	Airbnb is Banning the Use of Indoor Security Cameras in the Platform's Listings EU and Cambodia Launch \$55.3M Initiative to Promote Green and Digital Jobs PSE Empowering Cambodian Students Transitioning to Greener Growth Model Key to Boosting Employment EU and Cambodia Unveil \$96.7M Investment in Four Initiatives	Low Accurate Low Accurate Low Accurate Low Accurate Low Accurate		
		2	Strong Action Key to Curb Surge in Fatal Hit-and-Runs	2023-12-29	0.340535	High Accurate				
		3	Court Summons for Deadly Hit-and-Run Driver	2023-12-20	0.333826	High Accurate				
		4	Role of Licences in Driving Down Road Crashes	2024-02-02	0.329167	Medium Accurate				
		5	Calls to Urgently Review Draft Alcohol Law	2024-01-28	0.294192	Medium Accurate				
Travel	AirAsia Cambodia: Enhance Intra-ASEAN Travel, Cambodian Income With Relaxed Visa Fees	1	AirAsia Cambodia to Take Off to Expand ASEAN Market Share	2022-12-09	0.369897	High Accurate	Airbnb is Banning the Use of Indoor Security Cameras in the Platform's Listings EU and Cambodia Launch \$55.3M Initiative to Promote Green and Digital Jobs PSE Empowering Cambodian Students Transitioning to Greener Growth Model Key to Boosting Employment EU and Cambodia Unveil \$96.7M Investment in Four Initiatives	Low Accurate Low Accurate Low Accurate Low Accurate Low Accurate		
		2	AirAsia Cambodia Seeks Collaboration To Address High Aviation Costs	2024-03-19	0.271572	High Accurate				
		3	AirAsia Cambodia To Take Off in May Amid Aviation Costs Setback	2024-03-18	0.264002	High Accurate				
		4	CP Cambodia Becomes Official Sponsor of ASEAN Para Games	2023-05-31	0.253024	Low Accurate				
		5	Thailand Reduces ReEntry Visa Fees for Migrant Workers	2024-02-10	0.252357	Medium Accurate				

High Accurate: 13, Medium Accurate: 6, Low Accuracy: 11

Result 3.2 of Bag of word with Jaccard similarity

Algorithm 3: Bag of word with Jaccard similarity							Vs	Kiripost API		
Categories	Input_title	No	Related_Article_Output	date	Score	Assumption	Currently	Date	Assumption	
Business	Annual HEINEKEN Sustination Invites University Students to Innovate for a Circular Cambodia	1	Schools Off During SEA and ASEAN Para Games	2023-03-23	0.037037	Low Accurate	EU and Cambodia Launch \$55.3M Initiative to Promote Green and Digital Jobs	Medium Accurate		
		2	Business Model Competition Opens to Uni Students	2022-12-06	0.037037	Low Accurate	PSE Empowering Cambodian Students		Medium Accurate	
		3	US and Cambodia Discuss Nuclear Security	2024-02-23	0.035714	Low Accurate	Transitioning to Greener Growth Model Key to Boosting Employment		Low Accurate	
		4	Google Cloud Brings Event to Cambodia to Boost Tech Scene	2024-03-01	0.033708	Low Accurate	EU and Cambodia Unveil \$96.7M Investment in Four Initiatives		Low Accurate	
		5	AIESEC Cambodia Organizes Youth to Business Event	2023-10-26	0.032609	Low Accurate	National Bank of Canada Refutes Sale of ABA		Low Accurate	
Tech	For Social Platforms, the Outage Was Short. But People's Stories Vanished, and That's No Small Thing	1	Prahok Season Starts in Cambodia	2022-12-30	0.037037	Low Accurate	Airbnb is Banning the Use of Indoor Security Cameras in the Platform's Listings	Low Accurate		
		2	BREAKING Meta platforms including Facebook Instagram and Messenger are Experiencing a Glitch	2024-03-05	0.031746	High Accurate	EU and Cambodia Launch \$55.3M Initiative to Promote Green and Digital Jobs	Low Accurate		
		3	Metas Meltdown Resolved After Hours of Outage	2024-03-06	0.03125	High Accurate	PSE Empowering Cambodian Students		Low Accurate	
		4	Digital Menu Scanning Startup Published on Android OS	2023-04-05	0.029412	Low Accurate	Transitioning to Greener Growth Model Key to Boosting Employment		Low Accurate	
		5	Social Media App for Creating Community for Cambodians	2023-10-22	0.028571	Low Accurate	EU and Cambodia Unveil \$96.7M Investment in Four Initiatives		Low Accurate	
Lifestyle	Taylor Swift Struck a Deal With Singapore Not to Perform in Any Other Southeast Asian Country	1	Singapore Prime Minister Defends Exclusive Deal With Taylor Swift That Riles Some Neighbors	2024-03-05	0.069231	High Accurate	Airbnb is Banning the Use of Indoor Security Cameras in the Platform's Listings	Low Accurate		
		2	Cambodia Defeats Singapore 30 in Mens Volleyball at SEA Games	2023-05-04	0.04918	Low Accurate	EU and Cambodia Launch \$55.3M Initiative to Promote Green and Digital Jobs	Low Accurate		
		3	Cambodian Swimmer Misses Out on Medal at 50m Backstroke	2023-05-08	0.044776	Low Accurate	PSE Empowering Cambodian Students		Low Accurate	
		4	Cambodian Womens U22 Make History Securing a Spot in the Semifinals	2023-05-07	0.042553	Low Accurate	Transitioning to Greener Growth Model Key to Boosting Employment		Low Accurate	
		5	Cambodian King of Chess Wins Gold Medal in 5Minute Rapid Chess	2023-05-06	0.037975	Low Accurate	EU and Cambodia Unveil \$96.7M Investment in Four Initiatives		Low Accurate	
Education	UK University Students to Empower Cambodian Girls with STEM Skills	1	Secondary School Students to Compete in Firstever Mechatronic Design Competition	2023-10-30	0.038835	Low Accurate	Airbnb is Banning the Use of Indoor Security Cameras in the Platform's Listings	Low Accurate		
		2	MoU Strengthens Technical and Vocational Internships	2023-11-18	0.034483	Low Accurate	EU and Cambodia Launch \$55.3M Initiative to Promote Green and Digital Jobs	Low Accurate		
		3	Lighthouse Club Cambodia Holds Fundraiser	2023-03-27	0.033898	Low Accurate	PSE Empowering Cambodian Students		Low Accurate	
		4	AIESEC Cambodia Organizes Youth to Business Event	2023-10-26	0.032967	Low Accurate	Transitioning to Greener Growth Model Key to Boosting Employment		Low Accurate	
		5	PM Calls for Calm After Collapse of Two US Banks	2023-03-16	0.03125	Low Accurate	EU and Cambodia Unveil \$96.7M Investment in Four Initiatives		Low Accurate	
Health	Court Detains Lamborghini Driver Behind Fatal Collision	1	Lamborghini Smashes Into Motorbike Killing Driver	2024-02-23	0.03125	High Accurate	Airbnb is Banning the Use of Indoor Security Cameras in the Platform's Listings	Low Accurate		
		2	PM Orders Police Chief to Tighten Alcohol Road Rules	2024-01-24	0.030928	Medium Accurate	EU and Cambodia Launch \$55.3M Initiative to Promote Green and Digital Jobs	Low Accurate		
		3	1709 Killed on Cambodias Roads in 2022	2023-01-30	0.021739	Low Accurate	PSE Empowering Cambodian Students		Low Accurate	
		4	PM Orders Arrest of Fatal HitandRun Suspect	2023-12-22	0.021505	Low Accurate	Transitioning to Greener Growth Model Key to Boosting Employment		Low Accurate	
		5	BREAKING Tycoon Leng Channa Arrested for Fraud Breach of Trust and Bad Cheques Court	2024-03-12	0.02	Low Accurate	EU and Cambodia Unveil \$96.7M Investment in Four Initiatives		Low Accurate	
Travel	AirAsia Cambodia: Enhance Intra-ASEAN Travel, Cambodian Income With Relaxed Visa Fees	1	Sar Kheng Tests Positive for Covid19	2023-05-12	0.040404	Low Accurate	Airbnb is Banning the Use of Indoor Security Cameras in the Platform's Listings	Low Accurate		
		2	More than 2000 Expats Apply for Free Angkor Pass	2022-09-24	0.040323	Low Accurate	EU and Cambodia Launch \$55.3M Initiative to Promote Green and Digital Jobs	Low Accurate		
		3	Cambodia and Kuwait to Strengthen Bilateral Ties	2022-12-01	0.035714	Low Accurate	PSE Empowering Cambodian Students		Low Accurate	
		4	PMs Personal Call for China to Build HighSpeed Railway	2023-01-16	0.034884	Low Accurate	Transitioning to Greener Growth Model Key to Boosting Employment		Low Accurate	
		5	Ministry of Post and Telecommunications Signs Strategic Cooperation Agreement with Viettel Car	2022-11-09	0.034091	Low Accurate	EU and Cambodia Unveil \$96.7M Investment in Four Initiatives		Low Accurate	

High Accurate: 4, Medium Accurate: 1, Low Accuracy: 25

Result 3.3 of Bag of word with Euclidean distance

Algorithm 3: Bag of word with Euclidean_distance							Vs	Kiripost API		
Categories	Input_title	No	Related_Article_Output	date	Score	Assumption		Date	Assumption	
Business	Annual HEINEKEN Sustainathon Invites University Students to Innovate for a Circular Cambodia	1	Winners Crowned for Sustainable Innovations in Heineken Sustainathon	2023-05-10	[[1.11142]	High Accurate	Currently	EU and Cambodia Launch \$55.3M Initiative to Promote Green and Digital Jobs	Medium Accurate	
		2	Heineken Cambodia Appoints Andy Hewson as New Managing Director	2023-09-07	[[1.22530]	High Accurate	PSE Empowering Cambodian Students		Medium Accurate	
		3	Heineken Partners with Grab to Drive Home Responsible Consumption Message in Cambodia	2024-01-28	[[1.23100]	High Accurate	Transitioning to Greener Growth Model Key to Boosting Employment		Low Accurate	
		4	Heineken's Worlds Together Campaign Brings together Community Gatherings	2023-10-02	[[1.24432]	High Accurate	EU and Cambodia Unveil \$96.7M Investment in Four Initiatives		Low Accurate	
		5	Heineken Commemorates 150 Years of Good Times	2023-06-28	[[1.25744]	High Accurate	National Bank of Canada Refutes Sale of ABA		Low Accurate	
Tech	For Social Platforms, the Outage Was Short. But People's Stories Vanished, and That's No Small Thing	1	BREAKING Meta platforms including Facebook Instagram and Messenger are Experiencing a Glitch	2024-03-05	[[1.27761]	High Accurate	Airbnb is Banning the Use of Indoor Security Cameras in the Platform's Listings	Low Accurate		
		2	Metas Meltdown Resolved After Hours of Outage	2024-03-06	[[1.31175]	High Accurate	EU and Cambodia Launch \$55.3M Initiative to Promote Green and Digital Jobs	Low Accurate		
		3	Capturing Cambodias Stories of Change	2023-11-28	[[1.33024]	Low Accurate	PSE Empowering Cambodian Students		Low Accurate	
		4	Chaktomuk Short Film Festival Welcomes Rise in Entries	2022-11-23	[[1.34553]	Low Accurate	Transitioning to Greener Growth Model Key to Boosting Employment		Low Accurate	
		5	Microsoft Teams Recovering After Outage	2022-07-21	[[1.35213]	Low Accurate	EU and Cambodia Unveil \$96.7M Investment in Four Initiatives		Low Accurate	
Lifestyle	Taylor Swift Struck a Deal With Singapore Not to Perform in Any Other Southeast Asian Country	1	Singapore Prime Minister Defends Exclusive Deal With Taylor Swift That Riles Some Neighbors	2024-03-05	04:25:50	[[1.25500]	High Accurate	Airbnb is Banning the Use of Indoor Security Cameras in the Platform's Listings	Low Accurate	
		2	Taylor Swift Is Named Time Magazines Person of the Year	2023-12-07	08:06:15	[[1.16250]	High Accurate	EU and Cambodia Launch \$55.3M Initiative to Promote Green and Digital Jobs	Low Accurate	
		3	Taylor Swift Eras Tour Concert Film Coming to Movie Theaters in October	2023-09-02	06:50:52	[[1.16254]	High Accurate	PSE Empowering Cambodian Students		Low Accurate
		4	Taylor Swift The Eras Tour Will be a Blockbuster and Might Shake up the Movie Business	2023-10-11	00:53:29	[[1.16254]	High Accurate	Transitioning to Greener Growth Model Key to Boosting Employment		Low Accurate
		5	Taylor Swift Kisses Travis Kelce on the Field After His Chiefs Beat the 49ers in the Super Bowl	2024-02-12	08:01:02	[[1.20252]	High Accurate	EU and Cambodia Unveil \$96.7M Investment in Four Initiatives		Low Accurate
Education	UK University Students to Empower Cambodian Girls with STEM Skills	1	Call for UK to Invest in Cambodias Agriculture Sector	2023-01-28	[[1.15421]	High Accurate	Airbnb is Banning the Use of Indoor Security Cameras in the Platform's Listings	Low Accurate		
		2	UK and Cambodian Officials Meet to Strengthen Alliance	2024-03-06	[[1.16280]	High Accurate	EU and Cambodia Launch \$55.3M Initiative to Promote Green and Digital Jobs	Low Accurate		
		3	Strengthening STEM Education Is A Priority	2022-08-31	[[1.21548]	High Accurate	PSE Empowering Cambodian Students		Low Accurate	
		4	How STEM Is Modernizing Cambodias Education System	2022-01-14	[[1.21801]	High Accurate	Transitioning to Greener Growth Model Key to Boosting Employment		Low Accurate	
		5	A STEM Classroom Is Creating Cambodias Future Scientists	2022-10-28	[[1.22191]	High Accurate	EU and Cambodia Unveil \$96.7M Investment in Four Initiatives		Low Accurate	
Health	Court Detains Lamborghini Driver Behind Fatal Collision	1	Lamborghini Smashes Into Motorbike Killing Driver	2024-02-23	[[1.19915]	High Accurate	Airbnb is Banning the Use of Indoor Security Cameras in the Platform's Listings	Low Accurate		
		2	PM Orders Police Chief to Tighten Alcohol Road Rules	2024-01-24	[[1.28709]	High Accurate	EU and Cambodia Launch \$55.3M Initiative to Promote Green and Digital Jobs	Low Accurate		
		3	Court Summons for Deadly HitandRun Driver	2023-12-20	[[1.29997]	Medium Accurate	PSE Empowering Cambodian Students		Low Accurate	
		4	Strong Action Key to Curb Surge in Fatal HitandRuns	2023-12-29	[[1.31236]	Low Accurate	Transitioning to Greener Growth Model Key to Boosting Employment		Low Accurate	
		5	Nutrition Advocacy Group Calls for Nutrilat Court Case Rethink	2024-01-10	[[1.31699]	Low Accurate	EU and Cambodia Unveil \$96.7M Investment in Four Initiatives		Low Accurate	
Travel	AirAsia Cambodia: Enhance Intra-ASEAN Travel, Cambodian Income With Relaxed Visa Fees	1	AirAsia Cambodia Maiden Flight TakeOff in Q4	2023-06-21	[[1.13596]	High Accurate	Airbnb is Banning the Use of Indoor Security Cameras in the Platform's Listings	Low Accurate		
		2	AirAsia Cambodia to Take Off to Expand ASEAN Market Share	2022-12-09	[[1.20573]	High Accurate	EU and Cambodia Launch \$55.3M Initiative to Promote Green and Digital Jobs	Low Accurate		
		3	AirAsia Cambodia Seeks Collaboration To Address High Aviation Costs	2024-03-19	[[1.22900]	High Accurate	PSE Empowering Cambodian Students		Low Accurate	
		4	AirAsia Cambodia To Take Off in May Amid Aviation Costs Setback	2024-03-18	[[1.25298]	Medium Accurate	Transitioning to Greener Growth Model Key to Boosting Employment		Low Accurate	
		5	Thailand Reduces ReEntry Visa Fees for Migrant Workers	2024-02-10	[[1.25512]	Medium Accurate	EU and Cambodia Unveil \$96.7M Investment in Four Initiatives		Low Accurate	

High Accurate: 22, Medium Accurate: 3, Low Accuracy: 5

Word2vector on Kiripost Related News Model

TECH

For Social Platforms, the Outage Was Short. But People's Stories Vanished, and That's No Small Thing

Once upon a time, there was a brief outage on some social media platforms. It got fixed. The end. On the face of it, kind of a boring story.



BUSINESS

Annual HEINEKEN Sustainathon Invites University Students to Innovate for a Circular Cambodia

In partnership with GIZ Cambodia, the second annual HEINEKEN Sustainathon returns this year under the theme of 'Circularity', with applications opening this March.



RELATED NEWS



Meta's Meltdown Resolved After Hours of Outage



BUSINESS
Shinhan Bank (Cambodia) Plc., Signs Long-Term Strategic...

RELATED NEWS



EDUCATION
Reading Campaign Promotes Literacy in Remote Areas



BUSINESS
Shinhan Bank (Cambodia) Plc., Signs Long-Term Strategic...

LIFESTYLE

Taylor Swift Struck a Deal With Singapore Not to Perform in Any Other Southeast Asian Country

Taylor Swift stole the show at an Asian summit Tuesday when Singapore's leader was prompted to defend his tiny country's exclusive concert deal with the singer that risks bad blood in the region by preventing her from performing in neighboring nations.



RELATED NEWS



EDUCATION
Reading Campaign Promotes Literacy in Remote Areas



BUSINESS
Shinhan Bank (Cambodia) Plc., Signs Long-Term Strategic...

TF-IDF on Kiripost Related News Model

BUSINESS

BREAKING - Meta platforms, including Facebook, Instagram, and Messenger, are Experiencing a Global Outage

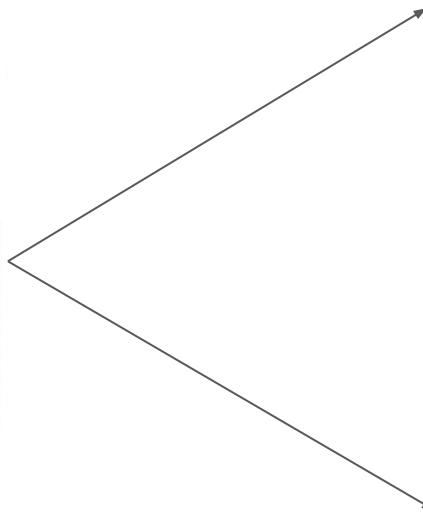
According to Downdetector.com, Facebook, Instagram, and Facebook Messenger experienced outages within the last 24 hours.



TECH

For Social Platforms, the Outage Was Short. But People's Stories Vanished, and That's No Small Thing

Once upon a time, there was a brief outage on some social media platforms. It got fixed. The end. On the face of it, kind of a boring story.



TECH

Meta's Meltdown Resolved After Hours of Outage

Meta's meltdown has been resolved after hours of being unable to access Facebook, Facebook Messenger and Instagram worldwide



Comparison Algorithm

Characteristic	TF-IDF	Word2Vec	Bag of words
Semantic Accuracy	Moderate	High	Low
Context Sensitivity	Low	High	Low
Dimensionality Reduction	Not Applied	Applied	Not Applied
Vocabulary Sensitivity	High	Medium-High	Low-Medium
Pre-trained Models	No	Yes	No
Out-of-Vocabulary Handling	No	Yes	No
Interpretability	High	Low	High
Computational Complexity	Low	Medium-High	Low-Medium
Training Data Size	N/A	Large (Depends on Corpus)	N/A
Memory Usage	Low	High (Depends on Model Size)	Low
Suitable For	Small to Medium-Sized Corpora	Large Corpora	Small Corpora

Summarize Algorithm

Algorithm	similarity metrics	High Accurate	Medium Accurate	Low Accurate	Others
Word2Vector	Cosine Similarity	0	2	28	Current Algorithm on kiripost
TF-IDF	Cosine Similarity	18	9	3	Detect Importation word and transform word to TF-IDF vector and making similarity score, finally output similarity article
	Jaccard similarity	2	9	19	
	Euclidean Distance	19	7	4	
Word2Vector	Cosine Similarity	1	0	29	Tokens word to vectors and making Similarity
	Jaccard similarity	0	0	30	
	Euclidean Distance	0	0	30	
Bag of word	Cosine Similarity	13	6	11	Count word and sort then convert to vector and making Similarity
	Jaccard similarity	4	1	25	
	Euclidean Distance	22	3	5	

Matrix Accurate performance

Algorithm/similarity metrics	Cosine Similarity	Jaccard similarity	Euclidean Distance
TF-IDF	High Accurate	Low Accurate	High Accurate
Word2Vector	Low Accurate	Low Accurate	Low Accurate
Bag of word	High Accurate	Low Accurate	High Accurate

Comparison Computational

Algorithm	Accurate	Computational Time	Performance Search Related Article
TF-IDF	High Accurate	4m 8s	~1.7s
Word2Vector	Low Accurate	19m 37s	~1.8s
Bag of word	High Accurate	3m 55s	~2.5s

Conclusion

Kiripost leverages content recommendation systems:

- optimize user experience
- increase engagement
- facilitate content discovery
- drive revenue for online media platforms
- benefiting both users and platform operators.

Our approach, utilizing algorithms like TF-IDF and Bag-of-Words (BoW) coupled with similarity metrics such as Cosine Similarity and Euclidean Distance, demonstrates higher accuracy compared to Word2Vec.

Thank
You!

Q&A