|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | | | | | | | | |
|  | | SEARCH ENGINE AND DOCUMENT CLUSTERING | | | | |  | |
|  |  | | | | | | |  |
|  | | | | Faculty of Engineering, Environment and computing, Coventry University, Coventry, United Kingdom.  MSc Data Science and Computational Intelligence  ============================= |  | | | |
|  | | | | UCHECHUKWU FREDRICK OKONKWO STID: 11738697  *okonkwou2@uni.coventry.ac.uk* |  | | | |
|  | | | | 11th APRIL 2021—7071CEM INFORMATION RETRIEVAL—DR SEYED MOUSAVI |  | | | |
|  | | |  | | |  | | |

Table of Contents

1. Introduction to Search Engines
   1. The web crawler**…………………………………………………………………………………….. 3**
   2. The Indexer**………………………………………………………………………………………….. 6**
   3. Query processor**……………………………………………………………….…………………….. 7**
2. Introduction to document clustering**……………………………………………………………………...13**
   1. Clustering**…………………………………………………………………………………………….13**
   2. Query result**…………………………………………………………………………………………...14**
   3. Topic modelling using latent dirichlet allocation (LDA)**……………………………………….…….17**
3. References**…………………………………………………………………………………………………22**
4. Appendix**………………………………………………………………………………………………….. 22**
   1. Code snippet………………………………………………………………………………………….. 23
   2. Document gotten from CNN news website used for clustering task………………………………….30

**TASK ONE**

**1.0 Introduction To Search Engine**

A search engine is a software system that is designed to carry outweb searches. They search the web in a systematic manner for specific information based on a textual web search query. The search results are generally presented in a line of outputs which are often referred to as the search engine results pages (SERPs). The output information retrieved could be links to web pages, images, videos ,articles, infographics, research papers and other relevant search outputs. Web search engines get their information by crawling the web from url sites to another. First the restrictions on the type of information or amount of information is checked before the crawling task is commenced using the “robots.txt” to be on the safe side. The robots.txt basically contains directives for crawlers telling which pages to crawl and which ones not to. After checking for the robots.txt, the crawler then sends the crawled information back to be indexed in a limited period of time spent on the web sites. A query from a user could be a single word or a sentence with specific words. The index helps find information relating to the query as quickly as possible.

The search engine maintains the following processes in real time:

* Web crawler
* Indexing
* Query processor

In this task I will be carrying out this steps in developing a search engine. The data will be obtained from the coventry university pureportal web pages. This vertical search engine specialises in retrieving only publications strictly by members of the School of Humanities (SH) at Coventry University: <https://pureportal.coventry.ac.uk/en/organisations/school-of-humanities>.

* 1. **Web Crawler**

A total number of 10 pages were crawled. As shown in figure 1 below, a total of 600 publications and 339 authors where crawled from the school of humanities publication pages. This 600 publications includes those publications by authors who are not staffs of coventry university. So I restricted my crawler to retrieve only publications that were published by coventry university staffs/authors which narrows the result to 339 publications and 339 authors as seen in figure 2 below. All information crawled are the publication link, author link (coventry person profile link), authors name, publication title and they are all 339 in number because I fit this data in a structured dataframe.

Screenshot shows the number of publications and the pages crawled

Graphical user interface, text, application

Description automatically generated

Screenshot shows the new number of publication links that matches the number of author links

Graphical user interface, text, application, email

Description automatically generated

Web crawlers can easily build multiple levels of indexes by employing the bredth-first search and this is implemented by starting from the source web page and then it visits all the links from that source which was executed in this crawling task. Figure 2 shows how I implemented the breadth-first as a baseline crawler. I started my crawling from the root node which is the publication links and traversed all the connected nodes to the root node subsequently.



As seen in screenshot 2, I inserted the if statement in the for loop when extracting the publication authors so I can get only authors who are staffs of coventry university.

  if "https://pureportal.coventry.ac.uk/en/persons/" not in links2:

        continue

All extracted contents were put into a dataframe with heads: title, author, paperlink, authorslink and id as seen in the screenshot 3 below.

Screenshot 3 showing the data frame of the extracted information

Table

Description automatically generated

I also did some preprocessing on the data crawled before parsing it through the indexing stage. First , I eliminated punctuations, stop words and converted all words to lowercase. Next I used the POS Tag to create a function to classify words to its correct part of speech. I also did the natural language processing canonicalisation which involves the process of grouping together the inflected forms of a word so they can be analysed as a single item as recognised by the word’s lemma or dictionary root form.

This is seen in the screenshot 4 below.

Screenshot 4 shows the pre-process stage

Graphical user interface, text, application

Description automatically generated

* 1. **Indexer**

In implemented the indexer using the inverted index structure which is constructed from scratch every time the crawler is runned. This type of inverted index is a word-level inverted index also called the “full inverted index” or “inverted list”. It contains the positions of each word that exist within a document. This form of indexing provides more functionality in phrase searching. The output is shown in screenshot 5 below.

Screenshot 5 shows the implementation of inverted index and the indexed output

Graphical user interface, text, application

Description automatically generated

Next I downloaded the google news vector words to vector the model which also helps with ranking the result. This step is necessary to encode text data as words or sentences into high-dimensional vectors in an organized manner. The word2vec model is a famous machine learning model used in natural language processing(NLP) used to extract the notion of relatedness across words and it finally encodes this relatedness into vector similarities. After computing the vector of words, I proceeded to constructing a database to stored the vectored output in the created database.as seen in the screenshot 6 below. This database has 85 rows and 301 columns of vector documents.

Screenshot 6 shows the word vectorization implementation and the vectorized output.

A picture containing text

Description automatically generated

**1.3 Query Processor**

Firstly I created a test split and then converted the test split into a function. I used the data structure index (IDX) defined in columns in the database to significantly speed up data retrieval operations. this function helps to retrieve publication ids that is related to the search word. Then I mapped the ids to relevant publication titles then printed out the information related to the query. Since the data is in a dataframe, I wrote some lines of code to print out the entire row of the retrieved query and return the entire row as the search output. As shown in screenshot 7 below.

Next, I computed the cosine similarities between the retrieved documents and the queries then finally converted the query processor, the indexer and its relevancy to return a valid output as ranked results. The sample output is shown in screenshot 8 below.

Screenshot 7

Text

Description automatically generated

Screenshot 8 shows a sample of the final result after entering the query

Graphical user interface, text, application

Description automatically generated

Screen shot 9 to 20 below are the search result showing the related publication tittle, author name, publication link and author’s link of various string inputs in the query.

Screenshot 9

Graphical user interface, text, application, email

Description automatically generated

Screenshot 10

Graphical user interface, text, application, email

Description automatically generated

Screenshot 11

Graphical user interface, text, application

Description automatically generated

Screenshot 12

Graphical user interface, text, application, email

Description automatically generated

Screenshot 3

Graphical user interface, text, application, email

Description automatically generated

Screenshot 14

Graphical user interface, text, application, email

Description automatically generated

Screenshot 15

Text

Description automatically generated

Screenshot 16

Graphical user interface, text, application, email

Description automatically generated

Screenshot 17

A picture containing text

Description automatically generated

Screenshot 18

Graphical user interface, text, application, Word

Description automatically generated

Screenshot 19

A picture containing text

Description automatically generated

Screenshot 0

Graphical user interface, text, application

Description automatically generated

**TASK TWO**

**2.0 Introduction to Document Clustering**

In this task I performed document clustering using the KMeans clustering algorithm. My document was generated manually from the cnn news rss feed. It was generated from five different news sections which are US politics, health, sports, entertainment and business rspectively. A total of 133 different documents was collected from this five various news sections. This is seen in visual 1 below.

VISUAL shows the details of the dataset collected

Text

Description automatically generated

I saved it as a csv file (encoder utf-8). It consist of two columns and 133 rows. Visual 2 belows show the first 20 rows of the dataset.

VISUAL shows the first 20 rows of the dataset

Graphical user interface, application

Description automatically generated

I started by importing the necessary libraries for this tasks as seen in visual 3 below. Libraries like pandas for dataframe visualization, the natural language processing toolkit for document processing like removing stop words, tokenizing and stemming as well as the matplotlib.pyplot to plot the elbow curve.

VISUAL

Text

Description automatically generated

I also performed feature extraction by vectorizing the documents using the tf-idfvectorizer from the sklearn featiure extraction dependency library.

VISUAL showing the vectorization process

A picture containing table

Description automatically generated

**2.2 Query result**

I used the elbow method to determine the optimal number of clusters by fitting the model with k ranging from 2-12 plotted against the sum of square errors (SSE) or inertia plot or within-cluster SSE plot on the Y-axis and the number of clusters on the X-axis. The curve is show in visual 5 below.

Fit 2 clusters

Fit 4 clusters

Fit 6 clusters

Fit 8 clusters

Fit 10 clusters

Fit 12 clusters

cluster no. of input documents, in the order they received:

[ 4 11 2 2 1 1 8 9 9 9 9 9 11 1 2 1 2 0 11 11 8 1 1 1

1 1 11 1 9 11 1 8 10 10 1 1 9 10 9 10 1 8 6 1 5 5 1 8

6 4 4 5 5 5 5 6 1 1 1 5 5 5 6 7 1 1 1 1 1 4 4 4

1 1 1 9 1 1 1 4 1 4 1 3 0 1 1 1 3 7 1 7 1 1 7 1

8 7 3 8 1 1 9 5 9 1 1 3 1 3 3 11 3 8 3 1 8 7 0 1

2 1 10 1 0 6 10 1 1 7 1 6 1]

VISUAL shows the elbow curve of SSE vs number of document clusters

Chart, line chart

Description automatically generated

I tried various queries to determine the cluster each enterd query belonged to. Below are the various queries tried and the assigned clusters as shown in visual 6 below. From the query output we notice that the queries belongs mostly to the clusters 1 and 9 rspectively.

VISUAL shows the assigned clusters for various query input

Text

Description automatically generated

Below are the queries and the clusters the queries are assigned to :

The move, which Twitter said would affect more than 300 accounts belonging to Russia, reflects a new policy the company is unveiling to ensure the free flow of information. The policy states that Twitter will not amplify or recommend accounts on its platform run by governments that are "engaged in armed interstate conflict" and that are also simultaneously restricting "access to free information.

The move, which Twitter said would affect more than 300 accounts belonging to Russia, reflects a new policy the company is unveiling to ensure the free flow of information. The policy states that Twitter will not amplify or recommend accounts on its platform run by governments that are "engaged in armed interstate conflict" and that are also simultaneously restricting "access to free information. (0, 17) 0.43111481248990974

(0, 16) 0.6893555964251167

(0, 14) 0.5821759872488401

[9]

your query belongs to cluster: 9

A SpaceX capsule carrying three paying customers and a former NASA astronaut has arrived at the International Space Station, finishing the first leg of this first-of-its-kind mission that will last about 10 days

A SpaceX capsule carrying three paying customers and a former NASA astronaut has arrived at the International Space Station, finishing the first leg of this first-of-its-kind mission that will last about 10 days

[1]

your query belongs to cluster: 1

The song features vocals by Andriy Khlyvnyuk from the Ukrainian band Boombox. The band used audio of Khlyvnyuk singing in central Kyiv, the Ukrainian capital, where he performed "a rousing Ukrainian protest song written during the first world war which has been taken up across the world over the past month in protest" against the Russian invasion.

The song features vocals by Andriy Khlyvnyuk from the Ukrainian band Boombox. The band used audio of Khlyvnyuk singing in central Kyiv, the Ukrainian capital, where he performed "a rousing Ukrainian protest song written during the first world war which has been taken up across the world over the past month in protest" against the Russian invasion. (0, 31) 1.0

[1]

your query belongs to cluster: 1

Afterwards, Woods confirmed he will play at the 150th Open Championship at St. Andrews, Scotland in July. Woods said he was still undecided about playing in golf's next major, the PGA Championship at Southern Hills Country Club in Tulsa, Oklahoma, in May.

Afterwards, Woods confirmed he will play at the 150th Open Championship at St. Andrews, Scotland in July. Woods said he was still undecided about playing in golf's next major, the PGA Championship at Southern Hills Country Club in Tulsa, Oklahoma, in May. (0, 17) 1.0

[9]

your query belongs to cluster: 9

Former President Donald Trump endorsed Dr. Mehmet Oz for US Senate on Saturday, just weeks before the contested Republican primary in Pennsylvania. "This is all about winning elections in order to stop the Radical Left maniacs from destroying our Country," Trump said in the statement endorsing the celebrity heart surgeon and television personality.

Former President Donald Trump endorsed Dr. Mehmet Oz for US Senate on Saturday, just weeks before the contested Republican primary in Pennsylvania. "This is all about winning elections in order to stop the Radical Left maniacs from destroying our Country," Trump said in the statement endorsing the celebrity heart surgeon and television personality. (0, 17) 1.0

[9]

your query belongs to cluster: 9

The midfielder was surrounded by defenders just outside the box but with three quick moves he left three opponents on the ground before finally finding a gap and rifling a precise shot past goalkeeper Yassine Bounou and inside the far post.

The midfielder was surrounded by defenders just outside the box but with three quick moves he left three opponents on the ground before finally finding a gap and rifling a precise shot past goalkeeper Yassine Bounou and inside the far post.

[1]

your query belongs to cluster: 1

In his musical plea, Davidson lamented the runtimes of films like "The Batman," which was nearly three hours, as was recent best picture nominee "Drive My Car." "Spider-Man: No Way Home" was almost two-and-a-half hours, and the 2021 remake of "West Side Story" ran a few minutes longer. When long runtimes plague both prestige films and superhero flicks, what gives?

In his musical plea, Davidson lamented the runtimes of films like "The Batman," which was nearly three hours, as was recent best picture nominee "Drive My Car." "Spider-Man: No Way Home" was almost two-and-a-half hours, and the 2021 remake of "West Side Story" ran a few minutes longer. When long runtimes plague both prestige films and superhero flicks, what gives? (0, 10) 1.0

[1]

your query belongs to cluster: 1

Social anxiety is one of the most common mental disorders around the world, Fallon Goodman, an assistant professor of psychology at the University of South Florida, told CNN's Chasing Life podcast: "About 4% of the world population will have social anxiety disorder at some point in their lives, which is roughly 300 million people."

Social anxiety is one of the most common mental disorders around the world, Fallon Goodman, an assistant professor of psychology at the University of South Florida, told CNN's Chasing Life podcast: "About 4% of the world population will have social anxiety disorder at some point in their lives, which is roughly 300 million people." (0, 31) 0.676817243228065

(0, 25) 0.37574595331539656

(0, 13) 0.36504805985278466

(0, 9) 0.37574595331539656

(0, 1) 0.3553706380683241

[1]

your query belongs to cluster: 1

Cardiovascular disease is a leading killer worldwide, taking nearly 18 million lives every year, according to the World Health Organization. In the United States alone, the US Centers for Disease Control and Prevention says heart disease takes a life every 36 seconds.

Cardiovascular disease is a leading killer worldwide, taking nearly 18 million lives every year, according to the World Health Organization. In the United States alone, the US Centers for Disease Control and Prevention says heart disease takes a life every 36 seconds. (0, 32) 0.3742732354172195

(0, 31) 0.4371794409254968

(0, 13) 0.4715940922433049

(0, 9) 0.48541436390411197

(0, 6) 0.45909213580628466

[1]

your query belongs to cluster: 1

Polls have closed in California's 22nd Congressional District where voters cast their ballots in the special election to replace Republican Rep. Devin Nunes, who resigned in January to head the Trump Media & Technology Group.

Polls have closed in California's 22nd Congressional District where voters cast their ballots in the special election to replace Republican Rep. Devin Nunes, who resigned in January to head the Trump Media & Technology Group.

[1]

your query belongs to cluster: 1

Illinois Republican Rep. Adam Kinzinger is sick and tired of what he is hearing -- and not hearing -- from some of his Republican colleagues when it comes to Russia and authoritarian ruler Vladimir Putin.

Illinois Republican Rep. Adam Kinzinger is sick and tired of what he is hearing -- and not hearing -- from some of his Republican colleagues when it comes to Russia and authoritarian ruler Vladimir Putin. (0, 16) 1.0

[9]

your query belongs to cluster: 9

**2.3 Topic Modelling**

I decided to carryout one extra text clustering/classifaction analysis on the dataset to compare the result using the latent dirichlet allocation (LDA) which a topic modelling algorithm. I want to determine the hidden topics in this document. The core idea is to take a matrix of documents and terms and try to decompose it into two separate matrices which are the document-topic matrix and the topic term matrix by way of dimension reduction or noise reducing technique.

First on this task is the pre-processing stage to eliminate the stop words, punctuation marks and finaly lemmatizing then saving in a new column the processed document. This is shown in the visual 7 below.

VISUAL showing the pre-processing and the processed document

Text

Description automatically generated

Next I plotted the most occuring words on a word cloud. This word cloud is show in visual 8 below

VISUAL

Text

Description automatically generated

My next step was to prepare the data for the topic modelling before it is passed into the LSA algorithm. I tokenized and stemmed the processed document and plotted the number of words in each document. As seen in visual 9 below.

VISUAL

Chart

Description automatically generated

After decomposing the documents into tokens I created a dictionary to keep track of the vocabularies and finally created a corpus of vectorized documents which will be passed into the LDA algorithm to determine the number of topics in the corpus. This shown in visual 10 below. After this pre-processing, the unique number of words before filtering process was 1437 and after filtering the unique number of words amount to 216.

VISUAL

Text

Description automatically generated

Next stage is fitting the model to the corpus to determine the topics. The best number of topics will be determined using the coherence score. The topic ranging from 3 – 10 with the highest cohernece score will be selected as the best number of topics in this modelling analysis. The coherence score (c\_v) plotted against the number of topics is show in visual 12 below.

VISUAL shows the coherence scores for topics 3-10

Graphical user interface, text, application

Description automatically generated

VISUAL show the coherence score vs the number of topics

Chart, line chart

Description automatically generated

As seen the screenshot above we see that topic 9 has the highest number of coherence score amongst other number of topics. This helps measure how well topics are extracted.

VISUAL shows the intertopic distance map and the top most salient words for each topic.

Chart

Description automatically generated

Visual 14 below shows the predicted topics for each document in the corpus and the length of the document in words.

VISUAL shows the dominant topics from the trained data

Graphical user interface, text, application

Description automatically generated

VISUAL shows the top words in each topic

A picture containing text

Description automatically generated

As seen in visual 15 above, we see the top dominant words in each topic. Topic one seems to be about sport as we see words like FIFA and league. Topic 2 has to do with health since it has words like maternal, death, live and death. Topic 8 also has some words that pertains to health as we see words like anxiety, social, life an so on.

**3.0 References**

**"https://www.bbc.com/news/business",**

**"https://www.bbc.com/news/technology",**

**"https://www.bbc.com/news/science\_and\_environment",**

**"https://www.bbc.com/news/entertainment\_and\_arts",**

**"https://www.bbc.com/news/health"**

**4 Appendix**

**4.1 Code snippet Task one (Web crawler)**







****

****



****

**4.1 DOCUMENT USED FOR CLUSTERING**

|  |
| --- |
| **news\_text** |
| A White House official said Biden, Vice President Kamala Harris and Obama "will deliver remarks celebrating the success of the Affordable Care Act and Medicaid in extending affordable health insurance to millions of Americans as a part of the President's agenda to cut costs for American families." |
| Obama made his first post-presidential trip to the White House Tuesday and showed he'd not lost a political step, telling demoralized Democrats -- worried about a possible midterm election shellacking -- to buck up and remember what political power is for. |
| The special primary election in California's 22nd Congressional District will head to a runoff, CNN projects. |
| Republican Connie Conway is leading the field and CNN projects she will advance, but the second spot in the runoff is still too early to call. |
| As of about 2 a.m. ET, Conway had about 35% of the vote, while Democrat Lourin Hubbard was in second with about 20% and Republican Matt Stoll and Democrat Eric Garcia were behind him with about 15% each. |
| Polls have closed in California's 22nd Congressional District where voters cast their ballots in the special election to replace Republican Rep. Devin Nunes, who resigned in January to head the Trump Media & Technology Group. |
| Six contenders were vying to fill the vacant seat in a district covering portions of Fresno and Tulare counties from Clovis to areas south of Visalia. But the winner may only end up serving in Congress for a few months after the district was sliced into pieces by the state's non-partisan redistricting commission. It will cease to exist in its current form when the new Congress is sworn in next year. |
| The official said the sweeping package "will impose significant costs on Russia and send it further down the road of economic, financial, and technological isolation." |
| The new sanctions package will ban all new investment in Russia, increase sanctions on financial institutions and state-owned enterprises in Russia, and sanction Russian government officials and their family members. |
| A Western official familiar with the plans said the US could apply sanctions on Russian President Vladimir Putin's adult children -- he has acknowledged two daughters -- as early as Wednesday. The Biden administration is also eyeing an expansion of sanctions on Sberbank, Russia's largest financial institution, and Alfa Bank, another large lender, that official said. |
| The new sanctions package will mark the latest escalation in efforts by the US and its allies to impose costs on Russia for its invasion and, over time, cut off critical economic sectors the country utilizes to wage the ongoing war. They also follow new revelations of further atrocities committed by Russian forces in northern Ukraine, with the images of the atrocities committed in Bucha serving as an accelerator to ongoing discussions between the US and its European allies to ramp up the economic costs, officials said. |
| Illinois Republican Rep. Adam Kinzinger is sick and tired of what he is hearing -- and not hearing -- from some of his Republican colleagues when it comes to Russia and authoritarian ruler Vladimir Putin. |
| In a two-minute video posted to Twitter on Tuesday -- which had already garnered more than 400,000 views by midafternoon -- Kinzinger sharply criticizes members of his party for what he sees as their unwillingness to speak out against the crisis in Ukraine. |
| "We are being governed by a bunch of children," says Kinzinger. "By a bunch of people that are not serious about running the United States of America and truly don't understand the threat that's out there from Vladimir Putin, from China and from some of these actors in the world that want to destroy our place here." |
| Washington (CNN)Yet another false claim about CNN's coverage of the war in Ukraine is swirling on social media -- this one an absurd conspiracy theory that CNN faked its footage from the scene of a fire caused by a Russian attack on a fuel storage site in the Ukrainian city of Lviv. |
| The conspiracy theory, which is being spread in multiple languages by pro-Russia accounts and others, is that CNN video from the Lviv fire on Saturday was actually filmed in the Canadian city of Edmonton. |
| Why would anyone say such a thing? Because, as CNN anchor Don Lemon was reporting live from the Lviv fire, a firefighter seen behind him was wearing a coat that said "EDMONTON" on the back. |
| Former House Speaker Newt Gingrich says he is backing trucking company owner Mike Collins for a US House seat in Georgia, hours after a video had been released where Gingrich endorsed another candidate in the race. |
| Gingrich announced his support for Republican candidate Vernon Jones in a video released earlier Tuesday by the Jones campaign -- even though Gingrich had endorsed Collins, whom he had long known, months ago. |
| New York (CNN Business)The Federal Reserve's fight against inflation will spark a recession in the United States that begins late next year, Deutsche Bank warned on Tuesday. |
| The recession call -- the first from a major bank -- reflects growing concern that the Fed will hit the brakes on the economy so hard that it will inadvertently end the recovery that began just two years ago. |
| "We no longer see the Fed achieving a soft landing. Instead, we anticipate that a more aggressive tightening of monetary policy will push the economy |
| Gogoro, a Taiwanese electric scooter startup that's backed by Al Gore and one of Apple's biggest suppliers, is about to have its day on Wall Street. |
| The company will debut on the Nasdaq on Tuesday after merging with a special purpose acquisition company, Poema Global (PPGH). The deal values the vehicle maker at about $2.35 billion. That's more than triple its last disclosed valuation of $800 million in September 2017. |
| The Taipei-based firm has raised at least $335 million in cash through the deal, which closed Monday, and will begin trading under ticker symbol "GGR." |
| Gogoro was founded in 2011 by CEO Horace Luke, a self-described "mad scientist" who previously led product strategy for XBox at Microsoft (MSFT) and innovation at electronics giant HTC. The company, which is also backed by Foxconn and Temasek, sells its own electric scooters and subscriptions to its battery swapping network. |
| Twitter (TWTR) said Tuesday it will no longer amplify or recommend Russian government accounts on its platform, bringing its stance on Kremlin-linked accounts closer in line with its approach to Russian state media. |
| Twitter accounts operated by the Russian government will no longer "be amplified or recommended to people on Twitter, including across the Home Timeline, Explore, Search, and other places on the service," the company said in a blog post. |
| The move, which Twitter said would affect more than 300 accounts belonging to Russia, reflects a new policy the company is unveiling to ensure the free flow of information. The policy states that Twitter will not amplify or recommend accounts on its platform run by governments that are "engaged in armed interstate conflict" and that are also simultaneously restricting "access to free information. |
| The World Bank said on Tuesday that countries in Asia may have to deal with three major economic shocks this year — the war in Ukraine, a sharp slowdown in China, and rate hikes from the US Federal Reserve. |
| The development bank cut its 2022 growth forecast for East Asia and the Pacific (EAP) region to 5% from 5.4%, and warned that growth could slump to 4% if conditions weaken further, trapping 6 million more people in poverty. |
| It also lowered China's growth forecast, estimating that the world's second largest economy will now grow at just 5% this year, sharply down on last year's 8.1%. That's also lower than China's official target of about 5.5%. |
| Shocks emanating from the war in Ukraine could affect the region "most concretely" by disrupting the supply of commodities and increasing financial stress, the organization added. |
| "The war and sanctions are likely to increase international prices of food and fuel, hurting consumers and growth," it said, adding that the number of poor in the Philippines, for example, could increase by 1.1 million if cereal prices rise 10% over the year. |
| Are upper middle class and affluent shoppers starting to feel the pinch from higher prices and geopolitical concerns? If so, what does that mean for the rest of American consumers? |
| Stock market volatility, inflation and concerns about rising interest rates could soon slow all parts of the economy — even demand for luxury goods. |
| Gary Friedman, the CEO of high end furniture retailer RH (RH), said in its earnings call last week that "we have experienced softening demand in the first quarter that coincided with Russia's invasion of Ukraine in late February and the market volatility that followed." |
| "I don't think anybody really understands how high prices are going to go everywhere, in restaurants, in cars and everything," Friedman said, adding that those increases will hurt all consumers and that companies like his are going to be in a "tricky space." |
| McDonald's first unveiled the spicy chicken nuggets in 2020, making them the first new McNugget flavor in the United States since the item was introduced in 1983. The nuggets have have a a breaded tempura coating of cayenne and chili peppers, packing a more flavorful punch compared to its regular chicken nuggets. |
| The chain credits the nuggets for boosting sales two years ago during the initial onset of the pandemic. McDonald's US sales jumped 7.5% in the fourth quarter of 2021 thanks in part to menu price increases, powerful digital sales and popular items like its crispy chicken sandwiches and the limited-time McRib sandwiches. |
| Limited-time items are an important marketing tool for the food industry to draw foot traffic and interest. These items offer consumers some variety — a key for fast food, as it's a notoriously thin-margin business with little customer loyalty. Offering something new gives people a reason to come back. |
| Chuck Skypeck, the technical brewing projects director at the Brewers Association, has owned and operated craft breweries for 21 years. He said the cause of skunking wasn't really understood until around the 1960s. |
| When hops in beer are exposed to strong light, a photooxidation reaction takes place, creating the compound 3-methyl-2-butene-1-thiol. To prevent the process of skunking from occurring, beer brewers have opted for darkly tinted glass. It's why you see so many beers in brown glass bottles today. |
| "It's a simple reaction that creates what most people see as an undesirable flavor," Skypeck said. "So, anything that shields the beer from that will preserve its flavor, so it is served as the brewer intends." |
| Hence the popular choice among brewers: brown. But that's not the only color seen in beer bottles; some come in green glass, too. What's behind that choice? Given that green is not as protective from light, the reason for its use is mainly marketing, according to Skypeck. |
| "If you look to see what brands are in green glass, you're going to probably find mostly heritage brands that have been around for a while," he said. "A fair amount of European brewers use green glass. Their green bottles are their image. And again, we're talking decades ago, there was a certain association of quality and uniqueness with green glass." |
| The US also had the highest maternal mortality rate: 23.8 deaths per 100,000 live births, more than triple the rate of any other country studied. When looking only at Black maternal mortality, the rate jumped to 55.3 deaths per 100,000 live births. |
| By comparison, in Norway in 2019, the last year for which data was available, there were zero maternal deaths. |
| "A high rate of cesarean sections, inadequate prenatal care, and elevated rates of chronic illnesses like obesity, diabetes, and heart disease may be factors contributing to the high U.S. maternal mortality rate. Many maternal deaths result from missed or delayed opportunities for treatment," the researchers wrote. |
| Researchers with the Commonwealth Fund pulled data from the fund's 2020 International Health Policy Survey and the Organisation for Economic Co-operation and Development. Statistics on health care were compared among Australia, Canada, France, Germany, the Netherlands, New Zealand, Norway, Sweden, Switzerland, the UK and the US. |
| "There's this misconception that cannabis is safe," said study author Brianna Moore, an assistant professor at the Colorado School of Public Health in Aurora, Colorado. |
| "So some women may use it in pregnancy, thinking that it's a safe alternative to other medicines, even prescribed medications," Moore said. "Yet studies show connections between marijuana use during pregnancy and low-birth weight in babies and behavioral problems later in childhood, and there may be links to glucose and weight issues as well." |
| Use of cannabis by pregnant women has been growing in the United States and other countries such as Canada in recent decades. A 2019 analysis of over 450,000 American women by the National Institute on Drug Abuse found cannabis use in pregnancy more than doubled between 2002 and 2017. |
| Past studies have shown the use of cannabis during pregnancy is linked to abnormal neurological development, autism and hyperactivity, attention issues, and other cognitive and behavioral issues in children, according to the US Centers for Disease Control and Prevention. |
| "What's interesting, too, is that we measured this at midpregnancy," Moore said. "These women very much knew they were pregnant, and they were past that first trimester where cannabis is often used for morning sickness." |
| Levels of CBD (cannabidiol) were also found. The chemical CBD is marketed over the counter and online as being "nonpsychoactive" and a safe option for anxiety, depression, sleep, pain, nausea and more. Many CBD products are made from industrial hemp, a type of cannabis plant that has little THC. |
| "They'll grow so fast they may overcompensate, which actually puts them at risk for obesity, metabolic syndrome and Type 2 diabetes later in life," she said. "That's why our study was designed to go back when the children were older to see how they fared." |
| At age 5, the children in the study were measured for fat and lean mass and were given a fasting glucose test to determine their blood sugar levels. |
| "The concern is that relative to their body size, they have more fat mass, which over the long term could put them at risk for metabolic syndrome, Type 2 diabetes and other conditions later in life," she said. |
| Based on all that is known, "women should refrain from using when they're pregnant or breastfeeding," Moore said. "And if no one's told them marijuana isn't actually safe to use in pregnancy, that messaging really needs to get out there. " |
| People with prescriptions for those sleep aids, which include zolpidem and temazepam, were over four times as likely to die from accidents and health conditions stemming from use compared with those who did not use such drugs, the study said. Even people who took fewer than two sleeping pills a month were three times more likely to die than those who didn't. |
| If you are using, or thinking of using a prescription or over-the-counter sleep aid, what do you need to know to do so safely? We asked Dr. Jing Wang, an assistant professor of pulmonary, critical care and sleep medicine at the Icahn School of Medicine at Mount Sinai in New York City. |
| Since my job as a CNN writer requires being in constant communication with various kinds of people, many are shocked when I reveal that I've struggled with social anxiety for years. |
| Ever since elementary school, I've experienced intense fear during social interactions both virtual and in person, whether it was as small as figuring out a response to a text or as big as presenting at a work meeting. These are hallmarks of social anxiety, a phobia of being judged, negatively evaluated or rejected in a social or performance situation, according to the Anxiety & Depression Association of America. |
| Social anxiety is one of the most common mental disorders around the world, Fallon Goodman, an assistant professor of psychology at the University of South Florida, told CNN's Chasing Life podcast: "About 4% of the world population will have social anxiety disorder at some point in their lives, which is roughly 300 million people." |
| It can really limit a person's life. "Clinical social anxiety is that level of social anxiety that keeps you from really living your life," Wendy Suzuki, a professor of neural science at New York University, told Chasing Life. "It keeps you inside because the fear of having these social interactions is so high." |
| "The treatment of choice (is) cognitive behavioral therapy, as it is with most anxiety disorders," clinical psychologist Ramani Durvasula, a professor of psychology at California State University, Los Angeles, told Chasing Life. "If you can help a person change how they think or perceive about a situation, you'll likely change the reactions and the behaviors." |
| Cardiovascular disease is a leading killer worldwide, taking nearly 18 million lives every year, according to the World Health Organization. In the United States alone, the US Centers for Disease Control and Prevention says heart disease takes a life every 36 seconds. |
| Eating at least two servings of avocado a week reduced the risk of having a heart attack by 21% when compared to avoiding or rarely eating avocados. However, there was not an equivalent benefit in reducing the risk for stroke, according to the study published Wednesday in the Journal of the American Heart Association. |
| A serving of avocado, which is a fruit, was defined as "½ avocado or ½ cup of avocado, which roughly weighs 80 grams," said study author Lorena Pacheco, a postdoctoral research fellow in the department of nutrition at Harvard T.H. Chan School of Public Health in Boston. |
| "Although no one food is the solution to routinely eating a healthy diet, this study is evidence that avocados have possible health benefits," said Cheryl Anderson, chair of the American Heart Association's Council on Epidemiology and Prevention, in a statement. Anderson was not involved in the study. |
| "We desperately need strategies to improve intake of AHA-recommended healthy diets — such as the Mediterranean diet — that are rich in vegetables and fruits," said Anderson, who is also professor and dean of the Herbert Wertheim School of Public Health and Human Longevity Science at University of California San Diego. |
| After the "Saturday Night Live" star bemoaned the lack of "short-ass movies" in a digital short of the same name, Netflix introduced a new category for films that fit Davidson's definition of a film that runs no more than an hour and 40 minutes long. |
| Yep, the streamer has assorted some "short-ass movies" that won't make your legs fall asleep. Among them: the fourth and fifth entries in the "Scary Movie" series, the comedy "Zoolander" and the acclaimed "Stand By Me," among some other random options. |
| In his musical plea, Davidson lamented the runtimes of films like "The Batman," which was nearly three hours, as was recent best picture nominee "Drive My Car." "Spider-Man: No Way Home" was almost two-and-a-half hours, and the 2021 remake of "West Side Story" ran a few minutes longer. When long runtimes plague both prestige films and superhero flicks, what gives? |
| "I've gotten Ryan Reynolds, [Chris] Pratt, [Chris] Hemsworth, Matt Damon," said Pine, making his bid for "best dressed Chris" in a blush, embroidered button-down, striped maroon-and-white pants and smoking slippers. |
| When Pine finally told the party patron that he was Captain America (again, not true; the only heroic captain Pine has played is Kirk in the "Star Trek" reboot films), the attendee said, "'That's my man!'" according to Pine. |
| The Hollywood Chrises -- all action stars (three of them from Marvel films), White and mostly blond -- are confused for each other fairly often. Even Arnold Schwarzenegger, father-in-law of Pratt ("Guardians of the Galaxy," "Jurassic Park," "Parks and Recreation"), called Pratt by Evans' last name, according to People. Some celebs just can't catch a break, minus the millions of dollars they make from their mega-successful franchises. |
| "I am happier, I am more present, I connect more with people," Gomez said. "I understand how powerful the internet is, and in so many ways it's done the best things for the world. But for me, I get to my news that is actually important I get through people in my life." |
| The singer/actress/producer, who currently has more than 300 million followers on Instagram, is promoting Wondermind, a mental health platform she helped found. |
| Gomez, who has been outspoken about her challenges including, having bipolar disorder, told "GMA" the goal is to make resources available to those in need. |
| "There are places where people go when they need help, and it's unfortunate that they cost ridiculous amounts of money," Gomez said. "But [as with] Planned Parenthood, there's a place for women to feel okay and to feel understood, and I want that for mental health." |
| Denzel Washington spoke publicly about his interaction with Will Smith on Oscar night and what happened after Smith slapped Chris Rock on stage. |
| Washington was a guest and speaker at Bishop T.D. Jakes' leadership summit on Saturday. During the two-hour discussion about faith and his award-winning acting career, Washington was asked about the infamous slap and what happened afterward. |
| "Well, there's a saying when the devil ignores you, then you know you're doing something wrong," Washington said in video of the exchange posted on Jakes' verified Twitter account. |
| "You know, the devil goes, 'Oh, no, leave him alone, he's my favorite. Don't bother him.' Conversely, when the devil comes at you, maybe it's because you're trying to do something right. And for whatever reason, the devil got a hold of that circumstance that night." |
| As Marvel movies go, "Morbius" is more a sip than a gulp, a relatively small-boned Jekyll-and-Hyde tale that moves another Spider-Man villain into the spotlight. Significantly better than "Venom" but still somewhat lacking in bite, this origin story perhaps inevitably grows more pallid toward the end but until then proves just tasty enough to merit the giving it a shot. |
| Having sat on the shelf for 21 months since its originally scheduled release, the timing of the film's arrival merely underscores Jared Leto's range, from this living vampire to an entrepreneurial vampire in "WeCrashed" entrepreneurial vampire to his scenery-chewing vampire in "House of Gucci." |
| The character's core, however, is a classic literary one, with Dr. Michael Morbius having grown up hobbled by a rare blood disorder, winning a Nobel prize for his development of artificial blood before seeking to turn his scientific genius loose on a cure for his condition. |
| Working with another scientist, Martine (Adria Arjona), Morbius has concocted a serum that draws upon the blood of vampire bats. But while the human trial gives him extraordinary strength and vitality, it also creates an unquenchable thirst for blood, one that his artificial creation only goes so far in satisfying. |
| Directed by Daniel Espinosa (whose credits include the sci-fi thriller "Life") from a screenplay by Matt Sazama and Burk Sharpless, "Morbius" has ample company in the world of tortured antiheroes and villains whose well-intentioned scientific goals go terribly wrong. (Another Spider-Man foe, the Lizard, shares virtually the exact same back story.) |
| While the modest scale actually works to the film's advantage, playing more like an old-fashioned monster movie than a superhero yarn, the plot is so simplistic and thin that the filmmakers labor to prolong the narrative until its climactic battle. |
| The special effects similarly alternate between slow-motion and lightning speed in demonstrating those vampiric attributes, although the central visual might be the way Morbius shifts from a mortal guise to his vampire one, which, parents should be forewarned, might induce sleepless nights among younger kids. |
| Nicole Farley readily admits her industry is "cut-throat," but that does little to dampen her ambitions of reaching the top. |
| A football coach who has worked with youth and academy teams on both sides of the Atlantic -- Chelsea, Arsenal and Reading in England and LA Galaxy and New York Red Bulls in the United States -- Farley's goal is to gain her professional license. |
| "Coaching with the first team manager, somewhere on that grass," she tells CNN Sport. "It could be the male professional game or the female [...] The highest I can go." |
| Farley is currently based at Premier League side West Ham having arrived last year as the club's first ever female coach. |
| She says her coaching career to date has been driven resilience, self-belief, and a willingness to step out of her comfort zone -- qualities that are all the more necessary when you're a woman working in a male-dominated environment. |
| Barcelona's teenage sensation Pedri scored a magnificent goal in the second half to secure a 1-0 win over fading Sevilla on Sunday as they climbed above their opponents into second spot in LaLiga after extending their winning streak to six games. |
| The 19-year-old Pedri created and scored the winner in a similar style to the trademark goals Lionel Messi scored for Barcelona before leaving to join Paris St Germain last year. |
| The midfielder was surrounded by defenders just outside the box but with three quick moves he left three opponents on the ground before finally finding a gap and rifling a precise shot past goalkeeper Yassine Bounou and inside the far post. |
| Its maiden season will be staged in five remote locations which have already been damaged or affected by climate issues, including the Arctic and the Amazon, raising awareness of climate change. |
| Hamilton, 35, called the opportunity to have his own team an "exciting new project" and said the series appealed to him because of its "environmental focus." |
| "Every single one of us has the power to make a difference, and it means so much to me that I can use my love of racing, together with my love for our planet, to have a positive impact," he said in a statement. |
| "I'm looking forward to the team taking part in this new series and I think it's incredible that we can do so whilst raising awareness about the climate crisis." |
| The six-time world champion's team, X44, is named in reference to his F1 racing number and will enter season one of the Extreme E Championship. The first X Prix race is scheduled to take place in early 2021. |
| It had been a tight game in the early stages before the Suns stepped it up in the third quarter to establish a healthy lead. |
| Devin Booker scored 32 points and made six three-pointers to hand the Western Conference leaders a franchise-record 63rd win of the season. |
| Lakers guard Russell Westbrook scored 28 points and Anthony Davis registered 21 points, but it was all too little, too late for Los Angeles which has been hampered by a string of injuries this season. |
| "This year, we had more starting lineups than wins," Davis told reporters. |
| "Our goal was to win a championship ... I feel like we had the pieces, but injuries got in the way of that. And that was the difference in the season." |
| Rafael Nadal, the 21-time grand slam champion, has suffered a stress fracture in one of his ribs and will be out for four to six weeks, the Spaniard announced on Twitter Tuesday. |
| Nadal's perfect start to 2022 was ended on Sunday by American Taylor Fritz who secured a surprise victory in the final at Indian Wells, beating Nadal 6-3 7-6 to win his first Masters 1000 title. |
| The 35-year-old was struggling with injury coming into Sunday's final and was seen wincing throughout the match, regularly touching the left side of his torso. |
| Nadal came into the final on a 20-match winning streak -- his best ever start to a season -- that included his record-breaking 21st grand slam victory at the Australian Open. |
| "Hello everyone, I wanted to announce that I've come back to Spain and immediately went to see my medical team so they could run some tests after the Indian Wells final, which I played hurt," wrote Nadal. |
| "In the end, it turns out I have a stress fracture on one of my ribs and I will be out between 4 and 6 weeks. This isn't good news and I didn't expect this. I'm down and sad about this because of the incredible start to the year I had. |
| "A very important time of year was coming up with a lot of good feelings and good results. But well, I have always had that fighting spirit and of overcoming things and what I will do is have patience and work hard after my recovery. Once again thank you all for the support," added Nadal. |
| Golf, basketball, hockey and other sports have all suspended their seasons. Major league baseball says it may start up in late spring. But competitive video gaming is going strong as esports tournaments move their operations completely online. |
| Some companies, including the Electronic Sports League, Riot Games and Activision Blizzard (ATVI), resumed their competitions online after pulling events in the US, Europe and Asia with a brief hiatus or no delay. |
| "It's pretty incredible that both Overwatch League and Call of Duty League will be the only city-based leagues still fortunate enough to operate in this current environment," Pete Vlastelica, president and CEO of Activision Blizzard Esports, said in a statement to CNN Business. |
| "I miss traditional sports, like we all do. At the same time I'm proud our leagues can compete online," he said, noting that Call of Duty League is still delayed but will be back online "soon." Activision did not comment on the delay. |
| Esports experienced a 37% increase in the average number of viewers per tournament last quarter, compared to the same period last year, according to Stream Hatchet, a video game analytics company. |
| As the esports industry draws more viewership online, it's also attracting new interest in a world where traditional sports are on pause. |
| Esports is the fast-growing, major money-making world of competitive video gaming. This global phenomenon is in a league of its own, where athletes may not look like your traditional sports stars -- yet have huge earning potential, massive brand endorsements and even their own fans. |
| While many of the top players reside in Europe or the US, the African continent is looking to share in a global market that pulled in over $1 billion in 2019. |
| One consulting firm projects Africa's gaming industry will increase by 12% in the next five years, with Egypt and South Africa leading the industry in revenue. |
| Friday's World Cup draw will see Qatar become the center of the footballing world, but the week's build-up has been dominated by non-sporting issues. |
| At FIFA's 72nd annual Congress on Thursday, Norway's FA President Lise Klaveness delivered a stinging speech that labeled the decision to allow Qatar to host the World Cup as "unacceptable" and demanded that FIFA do more to defend its principles of human rights. |
| Klaveness' comments came just a day after worldwide players' union FIFPRO and a global trade union federation published an open letter that criticized Qatar for continued poor working conditions for migrants in the run-up to December's tournament, even as it noted some "encouraging signs of progress." |
| A former player for the Norwegian national team and the country's first ever female football president, Klaveness said the sport's "core interests" were disregarded when Qatar won the bid in 2010. |
| "Human rights, equality, democracy: the core interests of football were not in the starting 11 until many years later," Klaveness told the delegation in Doha, Qatar's capital. |
| "These basic rights were pressured onto the field as substitutes mainly by outside voices. FIFA has addressed these issues, but there is still a long way to go." |