Available Online at www.ijcsmc.com

International Journal of Computer Science and Mobile Computing

A Monthly Journal of Computer Science and Information Technology

ISSN 2320–088X
IMPACT FACTOR: 7.056

IJCSMC, Vol. 10, Issue. 11, November 2021, pg.12 – 18

Leveraging Cloud Computing and Microservices for Tweet Analysis

Manoj Kumar Malik; Harsh S Malkani; Prinshu Dahiya
Maharaja Surajmal Institute Technology, New Delhi, India
manojmalik@msit.in, malkaniharsh@outlook.com, prinshudahiya7@gmail.com
DOI: 10.47760/ijcsmc.2021.v10i11.003

Abstract: Internet has become a platform for online learning, exchanging ideas and sharing opinions. Social networking site like Twitter is a widely used platform where millions of tweets are tweeted every day and most of these tweets never reach their intended audiences and fail fulfill their purposes because they are lost in a huge sea of tweets that are often irrelevant. Analysis of a segment of tweets might not truly reflect the real sentiments of the overall tweets on a topic which is a challenge. To tackle this problem, we introduce efficient techniques with which tweets are extracted, translated and sentiment analysis is performed on both text and images. These results are shown graphically and tabularly with other useful and important data such as username and hashtags used in tweets. For controlled access and security, login and registration features are incorporated.
Keywords- Cloud Computing, Microservices, Sentiment analysis, API, Cognitive services.

---

1. Introduction

Twitter has about 192 million daily active users who tweet 500 million tweets per day [1]. These tweets are expressed in many languages with one report showing the number of languages to be 104[2]. With such a high volume of tweets and their behavior and languages being so different and unique every time, keeping tab on tweets and treating them with equal importance irrespective of language is a hard and mundane task. About 75 percent of B2B businesses choose to market their products on Twitter [2] therefore it becomes an important platform for them to analyze the data and retrieve the sentiments towards them or their products.

What makes it more challenging is the short lifespan of tweets, that is the same tweet can be tweeted at separate times and days to reach new audiences, which is about 15-20 minutes [3]. Within this time, the tweet needs to be noticed by the right audience to make an impact. Noticing these critical tweets holds equal importance for both the user and the businesses irrespective of the language.

Instead of creating inferior algorithms or techniques, this paper focuses on using existing widely available and far more superior cloud computing models to augment the capabilities of the project. These services are available for public use and are economic as well as affordable when used efficiently.

To connect these microservices, RESTful API is used to make requests and to collect responses. APIs enable modular programming, through information hiding, which allows users to use the interface without bothering with the implementation [4]. This enhances the security of the code by showing only the implementation and provides a level of abstraction. All the services used are being called using REST API calls and the responses are stored in database so that they can be directly extracted if the same input occurs again. This prevents data redundancy and ensures that the paid services are being used only for new input.

2. Study

Most of the research papers published on sentiment analysis do not give importance to translation of tweets. The process of translating tweets is one of the most important tasks. India alone has 22 different languages, and a critical tweet can be in any one of those languages. A business operating in India needs to have familiarity with all these languages which requires extremely skilled workers which are less in numbers. Leaving any of those tweets while analyzing can lead to improper results. Azure Translator can translate 90 languages and dialects [5], hence helping in rectifying the said problem.

A common technique used while developing software applications is the monolithic approach, that is a single application holds all the code and handles all the responsibilities. It is still good for small teams where the size of the code is small, but as the code keeps getting bigger, it becomes difficult to manage, scale and time to market the product. The main drawback is that a single error in the code and crash the entire application [6].

Microservices solves most of these drawbacks. Microservices arranges an application in a collection of loosely coupled services. These services are logically connected and offer more fault tolerance than a monolithic approach. If a microservice encounters an error, a backup microservice and take its place to show the output. Microservices are deployed using containers which are lightweight virtual machines and are easier to scale [7].

The entire algorithm of our approach can be applied using a single API. But when there are multiple concurrent users and if the processing power of the virtual machine reaches its peak, the entire virtual machine will scale instead of just the process that is taking the most processing power hence increasing the overall cost of application. This may also have perceivable impact on new users that are trying to use the API. Using independent microservices can decrease the overall load on the main API, and only those processes will scale that need more processing power and will scale out when they do not. This proves to be very cost efficient and has more fault tolerance as, if a microservice fails the API can still return a degraded output.

3. Methodology

In this method, we first contact the Twitter API that sends response in the form of json, which is illustrated using the figure below.

{"msg":["RT @BradSmi: We need laws that ensure that all people can vote legally and securely. There are key aspects of Georgia\u2019s election law that f\u2026","@jkline120 \u2764","@msWordExpert Yes, way.","@snydez It's a secret. \ud83d\udcce","@Okay_Themb *silence*\n\n\"Yes.\"","@ScottieN0 \ud83e\udd7a","@cd0880 We are aligned with this.","@brianfagioli \ud83d\ude02","@threddyrex Oops, too slow.","@AngelaTidwell There's always one.","@NanoPolymath Yes to silence and yes to the power move.","@bradnelson \ud83c\udfc3\u200d\u2642\ufe0f\ud83d\udca8","Van Gogh\u2019s Starry Night in Microsoft Paint. \ud83d\ude31 Wow. https://t.co/5w48XufqIP","RT @MSCloud: Join @BradSmi at #Envision, where he'll discuss the digital divide and how organizations can operate more securely and sustain\u2026","As the digital divide widens, Canada is taking action to provide in-demand technology skills to young residents.\n\n\u2026 https://t.co/onVsE9CkS8","You: Does anyone have any questions before we move on?\n\n*silence* \n\nYou: No? Okay, moving on.","@PrawnAGE_OG Relatable.","@GovINd_kadel We believe in you.","@StephenWuchia About how much math are we talking here?","@AndreaRichman29 You're welcome.","@shaakunthala A very strategic approach.","@PIXEL_INK There's always room for improvement.","@Arron2Smith \ud83d\ude02","@DaveStinton \ud83d\ude02","@Xbox \ud83e\uddf2","@ChloeCondon We support this.","@CrucialMemory Same tbh.","@MarsNeedsTacos \ud83e\uddb2","They say knowledge is power. \ud83d\udca1\n\nHere's how open datasets are tackling one of the world\u2019s greatest challenges: https://t.co/gwOx8M6bZB","Now let\u2019s go over this one more time. https://t.co/Sd8FHZnhSd","RT @MicrosoftLife: In creating environments where everyone can achieve their purpose, Chief Accessibility Officer Jenny Lay-Flurrie and US\u2026","RT @MicrosoftLife: General Manager Erin Rifkin reflects on family and career while sharing her childhood recipe for a Southern fried chicke\u2026","@threddyrex Coming right up, Threddy.","@TheXGSFansoft \ud83d\ude02","@Paint3Dedu Thank you for this. \ud83d\ude4c","@RicoSynth A vibe.","@justinmpec \ud83d\ude02","@jlnhlfan \u2764","@n_silva You have great taste.","@bdsams \ud83d\ude02","@ZachWigal \ud83e\udd18","RT @MSFTEnable: \"Much more than a compliance exercise or checkbox, Manish says that accessibility must be a core part of building a better\u2026","Ready to roll. \u26a1\ufe0f https://t.co/4jCxkZTyhq","@BlueDRIFT Got it! On the way.","@aliciajeanss How about both?","@abhimjoshi Here you go: \ud83c\udf75","@SQL_Gal We like your style, Krista. \u2615","@NulFlar The GIF we didn't know we needed.","@kyperdana \ud83d\ude02","@UryensW Too late? We don't understand.","@TrunksWD .@YorkshireTea \ud83d\udc40","@robaxm Exactly.","@SherwoodOADLZ \u2615\u2615\ud83c\udf75\ud83c\udf75\ud83c\udf75","@ifermimm It's that kind of day.","@coaster_katy Very important distinction.","@sbisson Noooooo \ud83d\ude31","@BlondiebutGeeky Relatable.","@ajohnsocyber You got it, Ann.","@Mercedent Underrated beverage choice.","@Antonuval Of course, hydration is key.","@RazvanIon Great choice.","@jaydestro Coming right up.","@PrawnAGE_OG If only.","@Kaitlin_Knapp1 Priorities!","RT @Xbox: Amazing work by amazing artists celebrating amazing women. \u200b\n\nCheck out 7 new gamerpics and the new profile themes we've released\u2026","@MiStudent41 That is the question.","@Lil_RedTaco Happy to help.","@AbSumanthSuma You're on mute.","@CenturyLinkHome Every time.","@jaydestro \ud83d\udc38","@JenicaNorris \"Does anybody really care? \"","We're partnering with academics and thought leaders from around the globe to provide perspectives on how organizati\u2026 https://t.co/2gHcDZuFxh","Would anyone like some coffee or tea?","Nearly one year ago, inclusive of LinkedIn, we announced our #DigitalSkills initiative. Thirty million learners lat\u2026 https://t.co/TVKBilQB9t","How it started \u27a1\ufe0f How it\u2019s going https://t.co/NZZOKY3mPY","*Joins meeting*\n\n*mutes mic*\n\n*unmutes mic*\n\n*Today is actually Tuesday*\n\n*mutes mic*","RT @BradSmi: Since last June, @Microsoft and @LinkedIn has helped over 30 million people worldwide gain access to digital skills. This year\u2026","Out with the old, in with the new. \u2601\ufe0f\n\nHelp your business stay future-ready and resilient with the connected cloud: https://t.co/9RknVnSuZ5","Make the most of your time online and achieve your goals with fresh updates to Microsoft Edge and Bing.\n\nCheck them\u2026 https://t.co/7EE12mOcEp","RT @MSCloud: Watch #Envision Episode 2 highlights and join us for Episode 3, where @trevornoah and @BradSmi will discuss how to expand acce\u2026","@verge \ud83d\udc40","RT @verge: This Microsoft Flight Simulator mod features the cargo ship stuck in the Suez Canal https://t.co/IlXAYIjSu9 https://t.co/0GNcmmU\u2026","#NewProfilePic https://t.co/3ig2YsjS12","@nunixtech \"I've come to talk with you again.\"","@houssemayadil Our thoughts exactly.","@flipsparky360uk Excellent GIF.","@joocarter Accurate.","@NanoPolymath \ud83d\ude02","@mightymarkuk \ud83d\ude0e","@JenMsft We knew he would.","@TechDArchive \ud83d\ude4c","@hibacodes It's only fair.","RT @XboxWire: Today, @ID_Xbox spotlighted over 60 games coming to Xbox, including over 20 of your favorite indie games that will launch day\u2026","@danielgilbert Incredible. \ud83d\udc4f","Like this tweet if you think we should update our profile picture to dark mode. https://t.co/kX9WZ0kgxk","@deltaonealpha All snacks, all the time.","@bodecolander Motion granted.","@CgGeagea Touch\u00e9.","@epzik8 A long snack break is also acceptable.","@stemong \ud83d\ude02","@PrawnAGE_OG \ud83d\ude4c","@CogKosmos Snacks >>>","@HayoAHernandez2 \u2764","@DajImoto That depends on how many snacks you wish to consume.","@jonnytrickz You're welcome.","@dreamiebelieve On the way.","@danielgilbert What would your tagline be?","@ZitahM Very helpful coworkers.","@ndcnn \ud83d\ude39","@EmmaThommo Jealous.","@maryyemac Sometimes it's like that.","@katewhalls Yes.","@EngalsCaptain \ud83d\ude02","@gingerhat Rude, honestly.","Storytelling is a powerful way to drive sustainable culture change. It\u2019s not just about inspiration and emotion; it\u2026 https://t.co/Gk552jUvXi","All those in favor of a short snack break, say \u201caye.\u201d","Don't miss your chance to catch @sandylocks discuss how she came to coin the term \"intersectionality\" and how overl\u2026 https://t.co/LumxjCfz3y","@JimmyRamokgopa \u2764","@oliverjones A bold choice.","@marinalikethec Following. Would love to know how this turns out.","@diya404 You're welcome.","@18hallibhai1 You're ahead of the game.","@ohsaunak \ud83d\ude4c","\"How students feel has a big impact on their ability to learn,\" says Barbara Holzapfel, Microsoft VP of education.\u2026 https://t.co/gk065XXH3d","@hellosaumil Facts.","@ConsultingHumor You might be onto something there.","@shownde \"Clickity-clackity\" \ud83d\ude02","@A_V_Fernandez \ud83d\ude02","RT @burtonsnowboard: Adaptive riding is the front line of snowboard progression. Riders like Holly Palmer are showing us that with the righ\u2026","Microsoft Teams: Your microphone is muted.\n\nYou: I\u2019m just eating these overstuffed vegetable samosas, okay?","Proudly partnering with @burtonsnowboard to design custom gear that empowers all riders to pursue their passion.\u2026 https://t.co/lgq7nGXSXE","@VoteOmarNarvaez \ud83d\ude02","@XRays_Insider \ud83d\ude02\ud83d\udc9a","@Bassam_KE \"5 minute break\" \ud83d\ude02","@TheKaraAnton @jenna_clause \ud83d\udc4f","@Mrhankey91 You have to think big, Robert.","@mlaw91 @MSPowerApps Consider Floofy hired. \ud83d\ude3b","@Finellskii We're so proud.","@monishn We try.","@JenMsft That GIF \ud83d\ude02","@BrittanyChaffee You woke up and chose chaos. \u2728","@MaxPollard92 This is too real, Max.","@seangovind \ud83d\ude02","@jaydestro Someone has to.","@Xbox This is art.","@TwitterNktg Thanks.","@maingear \ud83d\ude4c","@McDonalds No, we went with CTRL+C and CTRL+VVVVVV.","At #Include2021, NYU Professor @DollyChugh discussed how to be effective, inclusive, and self-aware allies in the w\u2026 https://t.co/ltQLk4NyP0","Your funny tweet: 4 likes\n\nYour funny Microsoft Teams chat message: 7 laugh reactions, 2 heart reactions, 3 likes","Academics, behavioral scientists, and others have examined the intentional actions that can help us to recalibrate\u2026 https://t.co/yALdCpQC3Y","@XboxGamePassPC Never would've guessed.","@XboxGamePass Never would've guessed.","@Xbox Never would've guessed.","@CPStansbury \ud83d\ude4c","@NoXGameDev \ud83d\ude02","@Writerboy11 How many are there?","@AlguemStar \ud83e\udd2f","@bandrefilipe Well it was worth a try.","@DailySparkTeam You're welcome.","@ricardodsanchez Whoops. \ud83d\ude02","@Baspower_ Proud of you.","@brennadactyl Then it looks like we're at an impasse.","@erlaxs You get a free pass, just this once.","@Oajimoto Yes.","@OlaseniAd We are honored.","@SpiritJuiceZee \ud83d\udc40","@lukemoloney Nice save.","@steinekin They're gonna know.","@ASTROTR No. \u2764","@G4TV Same.","@cypherofTyr @steinekin You gotta do what you gotta do.","Reminder: don\u2019t forget to do that thing you said you\u2019d do today.","@McDonalds What extra sauce?","@TwitterNktg Personally offended there isn't a tech bracket.","@Outlook Attaching an old email thread to a new email thread is such a flex.","Khulile Vilakazi-Ofosu created a line of diverse dolls that celebrate the beauty of all children. When women achiev\u2026 https://t.co/a0gx1qgGWs","@laaurenmartin *becomes more angry*\n\n\"WHO IS GOING TO LOOK AT ME NOW? \"","@LisaJJohns Always remember to take breaks. \u2764","@Lindsbinds621 This is a whole mood.","@PinsPintaku Yes.","@libbyjones715 Relatable.","@VersaVejo \ud83d\ude4c","@JeffJohnsonJr_ \ud83d\udc40","@OlaseniAd Living in the future.","@drmarkdewit The vibes are strong.","@crashwawa It's pretty magical. \u2728","@lostatcore We love the honesty. \ud83d\ude02","@mac_lomaire Only the best cubes.","@refiloeofficial We're proud of you, too.","@cpmcbirds They had something to say.","@MiStudent41 Smh.","@thegirlontumblr Yes, thank you for asking.","@virtualpenny @SesameWorkshop \u2764","@n_silva We'll take it.","@RettMiller_IT Oops. \ud83d\ude02","@codeswift27 \ud83d\udc40","@chargejuix Why does that always happen? \ud83d\ude02","@ponokohq This GIF will give us nightmares, thanks.","@HeatheadMilitia No one has to know. \ud83e\udd2b"]}

Figure 1. Data returned by Twitter API, before processing.

This raw data holds additional fields that needs to be cleaned. For example, tweet id, image URLs, etc. This json file is then manipulated to show the results graphically and to extract some other useful data such as usernames and hashtags found in tweets.

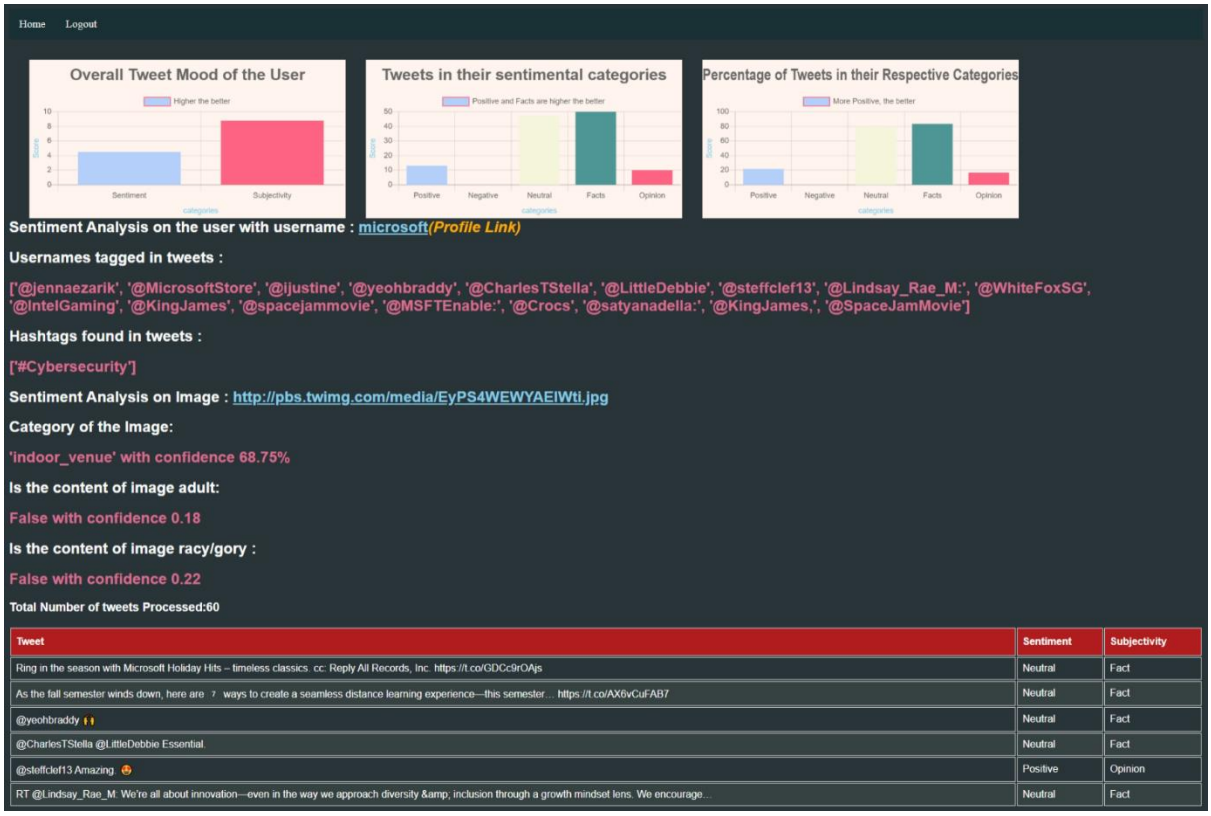

Figure 2. Result of analysis of tweets tweeted by @microsoft handle after processing.

Our approach analyses and processes tweets tweeted by a user, towards a user or having hashtag. To achieve our goal, we developed a RESTful API with login and registration that extracts and translates tweets for uniformity and returns the results in the form of graphs. The steps and tools used in our method are as follows:

3.1 Tweet Extraction:

For tweet extraction we used Tweepy to retrieve tweets from Twitter API. The tweets are returned in json format which are being transverse using Python.

3.2 Sentiment Analysis:

Using a natural language processing library called TextBlob, the sentiments and polarity are being returned between the scale of -1 to 1 which are then converted into tags such as positive, negative, and neutral.

3.3 Azure Translator:

Azure Translator is a cloud-based machine translation that translates text in real-time using REST API call.

To translate tweets, Azure Translator is being used which takes a tweet as an input and returns the translated tweet in json format which then can be extracted for further use.[5]

3.4 Azure Functions:

The entire application is divided into microservices instead of making it a single entity which can slow down when there are multiple concurrent users. Microservices using Azure functions tackles this problem by scaling automatically and are more cost effective [9].

3.5 Charts.js:

It is an open source HTML5 based JavaScript charts that can be integrated and used in websites [10]. It offers animated, interactive graphs on any website for free without the need to write code. It is being used in this paper to show sentiments, number of tweets in sentimental categories and overall mood of the user.

4. Results and Charts

In our approach the user interacts with the microservices through the API. This API calls these microservices and collects their responses to show the structured output to the user and stores it to the database. This can be summarized using the given chart.

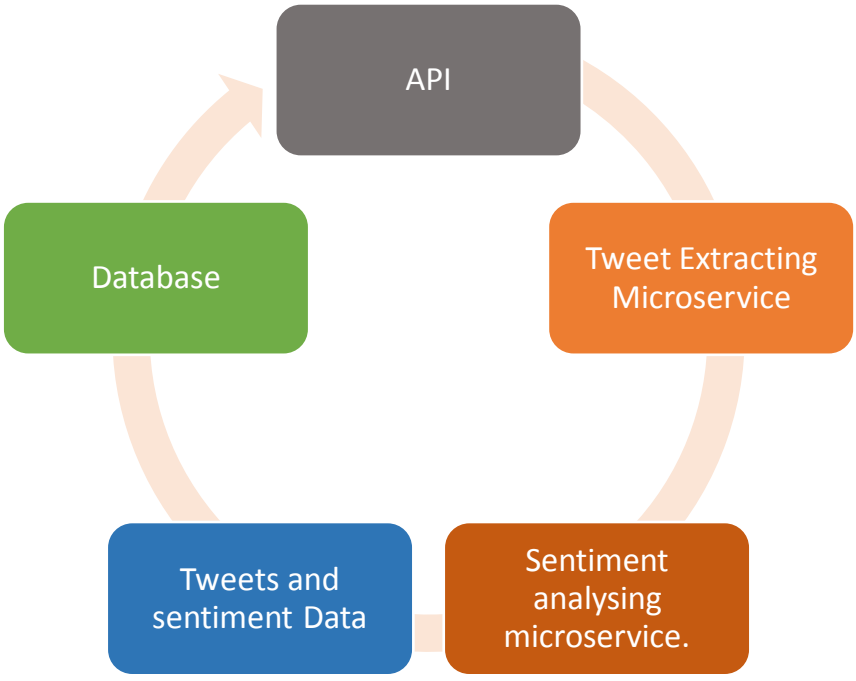

Figure 3. Flow chart of the process

It also performs analysis on tweeted images and provides useful information like the category of the image and checks if the content in the image is adult, racy, or gory. Below is an example of an image tweeted by Microsoft twitter handle of its indoor office space [11].

Figure 4. Microsoft indoor office space given as input.

The project returns its analysis with confidence score as follows:

```
Sentiment Analysis on Image : http://pbs.twimg.com/media/EyPS4WEWYAEIWti.jpg
Category of the Image:
'indoor_venue' with confidence 68.75%
Is the content of image adult:
False with confidence 0.18
Is the content of image racy/gory :
False with confidence 0.22
```

Figure 5: Analysis of image as output.

As shown above, this paper successfully performs sentiment analysis effectively and efficiently. All the cloud computing services used in this process are in their base plain which are free to try. Hence making this project economical. To increase the speed and quota limit of the entire algorithm the individual services can be scaled up by subscribing to a plan.

5. Conclusion and Scope

This project helps in analysis of enormous number of tweets in different languages and provides sentiment analysis on any topic. The resulting API has security features that provides controlled access to its services. Using this API, the users can be shown different data depending on their login status and their preferences. For ease of user, it also provides graphical representation of sentiment analysis for better understanding.

The API is very cost effective and can be easily replicated using any other cloud provider of choice. Analysing a large number of tweets can help in reflecting the real sentiments on a topic instead of a segment that may or not be reflecting the true sentiments.

This project can be further built to be applied to data from other social networking sites in file formats such as csv.

References

[1]. Oberlo, 10 Twitter Statistics Every Marketer should know in 2021[Infographic], https://www.oberlo.in/blog/twitter-statistics#:~:text=Summary%3A%20Twitter%20Statistics,-Here%27s%20a%20summary&text=There%20are%20192%20million%20daily,on%20Twitter%20is%203.39%20minutes (accessed 16 June 2021)

[2]. Parc, Languages and social network behaviors: Top 10 languages on Twitter, https://www.parc.com/blog/languages-and-social-network-behaviors-top-10-languages-on-twitter/ (accessed 16 June 2021)

[3]. MtoM Consulting, Updated: Lifespan of a Social Media Post, https://mtomconsulting.com/updated-lifespan-of-a-social-media-post/ (accessed 16 June 2021)

[4]. Wikipedia, API, https://en.wikipedia.org/wiki/API

[5]. Microsoft Azure, Translator | Microsoft Azure, https://azure.microsoft.com/en-us/services/cognitive-services/translator/ (accessed 16 June 2021)

[6]. David Jaramillo, Duy V Nguyen, Robert Smart, 2016, Leveraging microservices architecture by using Docker technology, IEEE Southeastcon , https://doi.org/10.1109/SECON.2016.7506647

[7]. Wikipedia, Microservices, https://en.wikipedia.org/wiki/Microservices (accessed 16 June 2021)

[8]. Microsoft Docs, Quickstart: Read client library or REST API – Azure Cognitive Services, https://docs.microsoft.com/en-us/azure/cognitive-services/computer-vision/quickstarts-sdk/client-library?pivots=programming-language-python&tabs=visual-studio (accessed 16 June 2021)

[9]. Microsoft Docs, Azure Functions documentation, https://docs.microsoft.com/en-in/azure/azure-functions/ (accessed 16 June 2021)

[10]. Charts.js , Charts.js, https://www.chartjs.org/docs/latest/ (accessed 16 June 2021)

[11]. Twitter, Microsoft on Twitter, https://twitter.com/Microsoft/status/1379177787362250756

