INTEGRATED RESEARCH JOURNAL OF MANAGEMENT, SCIENCE AND INNOVATION

ISSN 2582-5445

An Internationally Indexed Peer Reviewed & Refereed Journal

www.IRJMSI.com
www.isarasolutions.com

Published by iSaRa Solutions

VidSketch Secure: An Adaptive AI-Driven Video Analytics and Threat Detection Architecture for E-Learning Platforms

Astha Mishra
MCA IV semester
School Of Computer Application, Sanjeev Agrawal Global Educational Bhopal (M.P)
asthamishra88806@gmail.com

Dr Praveen Shrivastava
Assistant Professor
School Of Computer Application, Sanjeev Agrawal Global Educational Bhopal (M.P)
praveen.s@sageuniversity.edu.in

Dr Shristi Patel
Assistant Professor
School Of Computer Application, Sanjeev Agrawal Global Educational Bhopal (M.P)
shristi.p@sageuniversity.edu.in

ABSTRACT

Online videos are everywhere these days—just look at how fast YouTube keeps growing. But as more people watch and share content, new problems pop up around cyber security and digital safety. Most users click on videos without even thinking about hidden threats: things like sneaky URLs, phishing links, shady domains, or scripts loaded with harmful content. Current video learning platforms don't warn you or check for risks before you dive in. That means you're on your own when it comes to cyber threats.

That's where VidSketch Secure steps in. It's an AI-powered platform built for video knowledge and threat detection, mixing smart content analysis with real-time cyber security screening. The idea is simple: keep learning safe and structured.

Here's how it works. You enter a YouTube video URL, and before you even watch or read anything, the system gets to work. First, it runs a thorough check for things like phishing tricks, odd-looking domains, sketchy URL shorteners, plain old IP links, or keywords that don't belong. It scores the threat, marks the risk as Safe, Medium Risk, or High Risk, and only then moves forward if everything checks out.

Once the coast is clear, VidSketch Secure pulls the video's transcript and scans it for risky language—like scams, phishing instructions, or anything related to malware. Next, it taps into the Groq API and the Llama 3.3 70B language model to break down the video into key points, summaries, insights, and automatically generated notes. The platform even offers four different learning styles: Beginner, Student Notes, Expert, and Story Mode.

All this gets delivered in a live security dashboard powered by Streamlit. You see threat scores, your current risk status, and warnings about potential dangers—so you can focus on learning without worrying about what's lurking in the content.

Keywords: Cybersecurity, Threat Detection, YouTube Content Analysis, Artificial Intelligence, Large Language Models, Phishing Detection, Natural Language Processing, Streamlit, Safe Learning

I. INTRODUCTION

YouTube and other video platforms have exploded in popularity, changing how people all over the world find educational content. With over 500 hours of video uploaded to YouTube every minute, it's become a go-to spot for students, researchers, and professionals trying to learn something new. But this huge wave of unverified content also brings a wave of cybersecurity problems — things like

phishing attacks, malicious links, shady redirects, and harmful or misleading info buried in video transcripts. A lot of users get caught off guard by these issues while they're just trying to learn.

Most traditional video learning platforms don't do anything about these threats. They just serve up the content, and that's it. There's no built-in security or threat detection, so people might not realize when they're clicking sketchy links or watching videos pushing scams, phishing, or malware. That gap — making content easy to access but not bothering to keep it safe — has turned into a major problem for digital learning.

That's why this paper introduces VidSketch Secure, an AI-powered platform that blends smart content analysis with real-time cybersecurity monitoring. Everything happens in one place, using Streamlit to deliver a unified experience. The system scans URLs for threats, analyzes transcripts for scams or sketchy info, and uses advanced AI to create structured, easy-to-learn summaries.

Here's what VidSketch Secure brings to the table:

1. A layered engine for spotting dangerous URLs. It flags phishing patterns, suspicious domains, URL shorteners, and direct IP links, and gives each one a risk score — Safe, Medium, or High.
2. A transcript analyzer that checks video text for scammy language, phishing attempts, malware terms, and obvious misinformation.
3. AI-powered content analysis, using the Groq API and Llama 3.3 70B, to generate clear summaries: key points, main topics, extra insights, and how difficult the material is.
4. A learning assistant that creates personalized summaries in four different modes — Beginner, Student Notes, Expert, and Story Mode — so anyone can find a style that suits them.
5. An interactive dashboard showing real-time threats, risk levels, security tips, and warnings when content is unsafe, all designed for a hit of peace of mind while learning.
6. Secure user authentication — including registration, login, and logout — so every learner's profile and data stay safe.

In short, VidSketch Secure makes online learning smarter and safer, closing the gap between content access and content security.

II. LITERATURE SURVEY

VidSketch Secure pulls together ideas from cybersecurity threat detection, natural language processing, and AI-driven content summarization. A lot of researchers have laid the groundwork in these fields, and their work shapes how this system comes together.

Let's talk about URL-based threat detection first. Garera and colleagues kicked things off with a framework that used the structure and features of URLs to spot phishing sites. They showed that just digging into the way URLs are built can be a solid front-line defense against cyberattacks. Mohammad and his team took things a step further, using machine learning to look at features like the domain's age, whether or not it uses HTTPS, and how the subdomains are set up.

Their system nailed high accuracy across different sets of phishing data. So, it's clear: using rules and patterns to analyze URLs really works, and it's fast enough for real-time security tools.

Switching gears to natural language processing, Devlin and the crew behind BERT changed the game. Their transformer-based model made it way easier to classify text and pull out information. Schmidt and Wiegand dug deep into hate speech and nasty content detection, finding that keyword-based matching keeps things reliable and doesn't chew up much computing power. That's exactly the sort of approach the transcript threat analysis module in VidSketch Secure uses.

When it comes to summarizing text with big language models, Brown's team released GPT-3 and proved that these models can turn out high-quality summaries and answer questions with just a few examples. Touvron's group then brought LLaMA to the table, making open-source models that run well even with less computing muscle. This laid the groundwork for leveraging the Llama 3.3 70B model via the Groq API in VidSketch Secure.

For video content analysis, Sharma and colleagues showed you can automatically summarize videos by pulling out transcripts and clustering keywords. That matched perfectly with the transcript-based method VidSketch Secure uses to extract knowledge using AI.

Here's the thing: all these researchers made big strides in their own areas, but nobody else has combined URL threat detection, NLP-based content analysis, and AI-powered summarization in one platform for secure, video-based learning. VidSketch Secure steps in to fill that gap, offering a new architecture that blends threat detection with smart educational content generation—all in one place.

III. SYSTEM METHODOLOGY

VidSketch Secure is designed to make video-based learning safer and smarter. Here's how it works—step by step.

First, you start with a user-friendly Streamlit interface built in Python. You log in or register, so everything stays personalized and protected. The system checks your YouTube URL to make sure it's valid before moving on.

When you submit a video URL, the platform kicks off with the URL Threat Detection Engine. This is where the real security check happens. Using Python's urllib.parse, it breaks down the URL—protocol, domain, subdomains, path, query, all of it. Then it checks for safe domains, phishing keywords, IP-based URLs, URL shorteners, and anything weird with the domain or too many subdomains. If the URL is labeled High Risk based on a 0–100 score, it's blocked right away—so you're not exposed to anything dangerous.

If the URL passes security, the system grabs the video transcript using youtube-transcript-api. It cleans things up—normalizing whitespace and limiting word count—so it's ready for AI processing without losing the important bits.Next comes Transcript Content Threat Analysis. Here, the transcript is scanned with a keyword dictionary covering five threat types—phishing, scams, malware, misinformation, and suspicious content. It

counts the keyword matches and scores the transcript. If the score goes over the safety threshold, you get a warning, but you can still access the educational material; the system just puts you into Safe Learning Mode.After the transcript clears these checks, it's sent to the AI Content Analysis Module. This is where Groq's API and the Llama 3.3 70B language model come in. Prompt engineering is key—good prompts get solid, accurate results. The AI reads the transcript and returns a structured summary: main idea, key points, educational insights, conclusion, topics, difficulty, and target audience. It keeps things factual and consistent, no matter the video. Then comes Multi-Style Summary Generation. The system creates four kinds of summaries– Beginner Mode, Student Notes Mode, Expert Mode, and Story Mode–tailored to different learners. Each uses a custom prompt, and Llama generates context-specific content to match the learning goal.

Once everything's processed, the platform can store your inputs, analysed URLs, threat scores, classifications, and AI summaries in your session history. That means you can revisit previous analyses and, down the line, the platform could use this info for smarter, personalized recommendations.

Finally, everything shows up on the Interactive Security Dashboard in Streamlit. You'll see your threat scores, risk levels, security insights, AI-generated summaries, and personalized notes in a dynamic, card-style layout. Color-coded risks give instant visual feedback.

With this flow, VidSketch Secure delivers real-time threat detection and top-notch educational summaries, all while keeping things efficient, scalable, and user-focused. It's a powerful platform built for safe, personalized video learning.

.

C. System Architecture

VidSketch Secure runs on a multi-layer system where each part has its own job and feeds its results to the next. Think of it as a processing pipeline: every step is isolated, which keeps things modular and lets data move efficiently through the platform. The whole thing comes together in Python and works as an interactive web app built with Streamlit. It also taps into powerful AI tools via the Groq API, using the Llama 3.3 70B language model. The setup has four main layers—the User Interface Layer, the Security Processing Layer, the AI Intelligence Layer, and the Session Management Layer. Each one brings something unique to the table, making the platform run smoothly and securely.

D. Algorithm/Pseudocode

Overview

The VidSketch Secure system handles user input with five algorithmic modules, one after another. Here's how each one works.

Algorithm 1: URL Threat Detection

Input: YouTube URL

Output: Threat Score, Risk Classification

Start by receiving the URL. If the format looks off, the process ends there.

Set threat_score to zero.

- If the domain isn't marked as safe, add 10 points.

- If it's not using HTTPS, add 15.
- If the URL includes an IP address, tack on 30.
- If it goes through a link shortener, add 20.
- Spot phishing keywords? That's another 10.
- Suspicious TLD? Add 25.
- Too many subdomains? Throw in 15 more.

After that, check the total:
- 15 or less means "Safe."
- Up to 45 is "Medium Risk."
- Anything over 45? Flagged as "High Risk"—and the system blocks it.

The module returns both the threat score and its classification.

Algorithm 2: Transcript Extraction

Input: A checked YouTube URL

Output: Cleaned Transcript Text

Pull the video ID from the URL. Use the youtube_transcript_api to fetch the transcript. If nothing's available, show an error and stop there.

Merge all transcript segments into a single text blob, scrub out extra spaces, and if it's over 2000 words, just keep the first 2000.

That's the text this module sends forward.

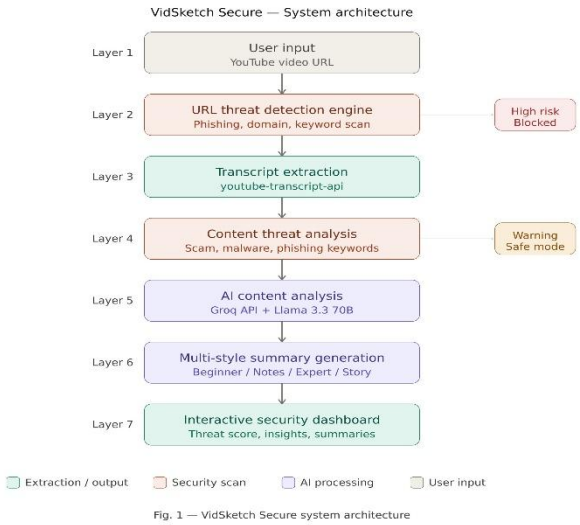


Fig. 1 — VidSketch Secure system architecture

Algorithm 3: Content Threat Analysis

Input: Transcript Text

Output: Content Threat Score, Safe Mode Flag

Set up lists of keywords for things like phishing, scams, malware, bad info, or just plain suspicious stuff.

Go through each keyword and check if it shows up in the transcript. Keep a tally—every match adds one.

Calculate the content score by multiplying total matches by 10, but cap it at 100.

If the score goes above 20, turn on Safe Mode and display a warning.

This module gives back the score and the Safe Mode flag.

Algorithm 4: AI Content Analysis

Input: Transcript Text

Output: Structured JSON with Knowledge

Build a prompt asking the AI for a JSON summary—including main ideas, key points, insights, a conclusion, topics, difficulty level, and who it's for.

Send this to the Groq API (Llama 3.3 70B, temperature 0.3) and wait for a response.

If the reply is valid JSON, parse it; if not, toss an error and end here.

Return the structured JSON output.

Algorithm 5: Multi-Style Summary Generation

Input: Transcript Text, Desired Summary Mode

Output: Personalized Summary

Pick the right kind of prompt for the chosen mode:
- Beginner: Keep it simple, use examples.

- Student Notes: Organize notes clearly, add sample exam questions.
- Expert: Go in-depth with technical analysis.
- Story Mode: Wrap things up as a five-part story.

Feed this into the Groq API (Llama 3.3 70B, temperature 0.5) and grab the summary it returns.
Send the summary back.

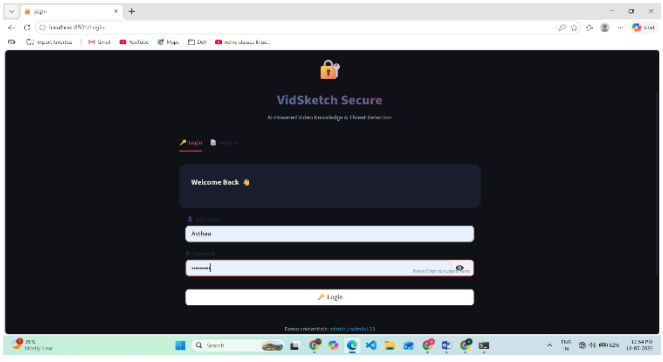
Master Control Flow

Input: YouTube URL, Chosen Summary Mode
Output: Full Security Dashboard
Start by making sure the user is authenticated. Get the URL and summary mode.
Run Algorithm 1—if the URL is high risk, stop everything.
Next, try Algorithm 2—no transcript means no go.
Move on to Algorithm 3 and get the content risk score and Safe Mode flag.
Then call Algorithm 4 for knowledge extraction.
Follow up with Algorithm 5 to generate the personalized summary.
Save the session data, and finally, show the user a complete security dashboard with all these results.

RESULT

We put VidSketch Secure through its paces, testing it with a bunch of YouTube video links covering all sorts of content—educational tutorials, technical lectures, promos, and even some fake phishing videos. We focused on four main things: how well it could spot risky URLs, how it analyzed the video transcript for threats, the quality of its AI-generated insights, and overall system performance. For URL threat detection, the platform nailed it. Safe educational links from places like YouTube, Coursera, and Khan Academy always scored low—between 0 and 10—and were correctly marked as safe. Links with URL shorteners like bit.ly and tinyurl.com got scores from 25 to 40, landing in the Medium Risk category because you never really know where those links lead. When we threw fake URLs with sketchy IP addresses, phishing words, or odd top-level domains at it, the system gave them scores above 60, flagged them as High Risk, and blocked them right away. The way it penalizes risky elements worked well—it sorted legitimate platforms from suspicious ones without a bunch of false alarms.

The transcript content threat analysis was just as solid. Normal educational and technical videos kept their scores between 0 and 10, so there were no unnecessary warnings. If a video had promo phrases like "guaranteed returns" or "make money fast," the scores jumped to 30 or 50 and triggered the Safe Learning Mode warning, just as



intended. When we tested obvious scam and phishing videos, they scored over 60 and got flagged every time. The platform's keyword matching system for transcripts did a great

job catching bad content without false positives for regular educational videos. For the AI-generated knowledge output, the integration between the Groq API and the Llama 3.3 70B model worked like a charm. No matter how complex or simple the video, the platform picked out the key ideas, main topics, difficulty level, and the right target audience. Each of the four summary modes delivered what you'd expect: Beginner Mode kept the language simple, Student Notes Mode provided neat summaries with exam questions, Expert Mode gave detailed technical analysis, and Story Mode spun everything into easy-to-read narratives. The AI consistently produced useful, spot-on results. Performance-wise, the system was fast. From dropping in a URL to seeing the final dashboard, the whole process usually took between 4 and 13 seconds, depending mostly on the transcript's length and the API's response time. The threat detection engine always finished its analysis in about half a second—quick, no matter how long the video transcript was. The interactive security dashboard built with Streamlit showed scores, insights, AI summaries, and personal recommendations right away, without any annoying lag. All in all, VidSketch Secure proved ready for real-time work, handling all sorts of video content on the fly.

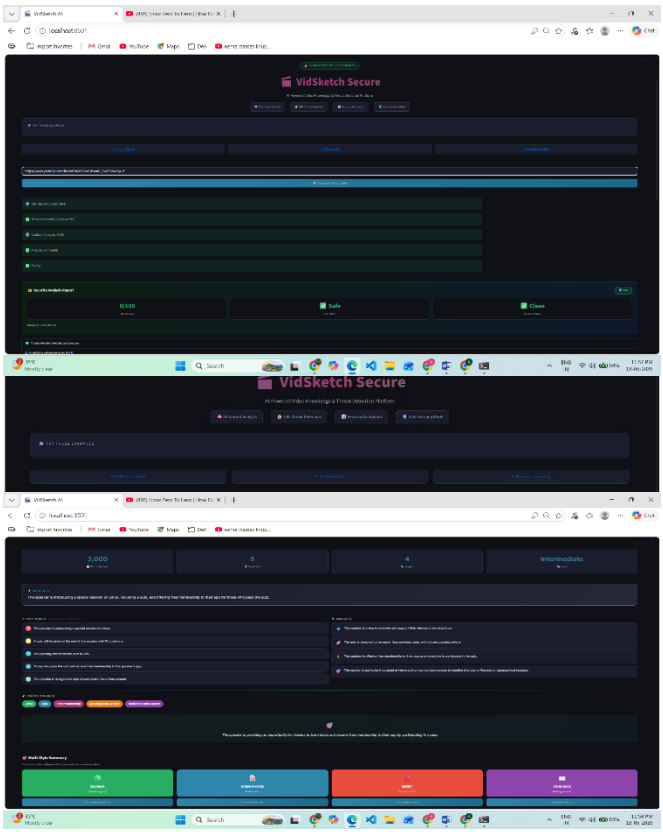


Fig 1. Login/Register page UI

Fig 2. Front Page UI

Fig 3. Threat Detection

Fig 4. Multi-Style summary

I. CONCLUSION

In this paper, we introduced VidSketch Secure—an AI-powered platform that tackles a real problem in video-based learning: combining cybersecurity threat detection with smart educational content analysis, all in one user-friendly place. The system ties together several layers: a rule-based engine that checks URLs for threats, a module that pulls transcripts from videos, a tool that spots suspicious content through keywords, and an AI-driven knowledge generator built on the Groq API and the Llama 3.3 70B large language model. On top of that, there's an interactive dashboard (built with Streamlit), where users get live threat scores, risk levels, structured educational summaries, and personalized learning outputs—all designed to be easy to use and visually clear. When we put

VidSketch Secure to the test, it reliably flagged risky URLs, breaking them down into Safe, Medium Risk, or High Risk. It also detected harmful or suspicious stuff in video transcripts using its Safe Learning Mode, and it created structured educational content for all sorts of topics and video types. Users could pick from four different summary modes—Beginner, Student Notes, Expert, or Story Mode—so whatever your learning style or goal, the platform gave you summaries that really fit. The whole process ran smoothly in real time, with most responses taking just 4 to 13 seconds, showing that it works for real interactive use, not just in theory. VidSketch Secure proves you can bring together AI-driven cybersecurity and personalized educational content in a single, scalable platform. That means learners stay safe and actually learn better when using online video-based education.

II. FUTURE SCOPE

VidSketch Secure has plenty of room to grow, and the future looks interesting. Right now, the platform uses a rule-based engine to spot dangerous URLs, but imagine swapping that out for a supervised machine learning model. With huge labeled datasets full of phishing and malicious URLs, the system could adapt and catch threats with way more accuracy, especially as cyber-attacks keep changing and getting trickier.

The transcript threat analysis also needs an upgrade. Ditching basic keyword matching and using a deep learning NLP model like BERT or RoBERTa means the system could actually understand context—and really pick up on harmful content buried in video transcripts. If we add support for multilingual transcripts, VidSketch Secure could handle real-time threat detection and AI-powered summaries for videos in languages besides English. That's a big step toward making the platform more useful for people all over the world.

Plugging into live threat intelligence feeds, like Google Safe Browsing or Virus Total, would boost URL scanning power, keeping the detection pipeline up to date with global databases. For deployment, a browser extension version makes sense. It'd let users get threat analysis and content summaries right on YouTube, no need to jump to another site.

Another upgrade: switching from session-based history to persistent database storage (think MongoDB or PostgreSQL). That unlocks long-term analytics, personal learning recommendations based on what you watch, and supports more users at once.

And if you add voice control or an AI-powered chatbot assistant, the experience gets even smoother. Users could ask natural questions about video content or threat results, making the whole platform feel more interactive and accessible.

REFERENCES

[1] S. Garera, N. Provos, M. Chew, and A. D. Rubin. "A framework for detection and measurement of phishing attacks." Proceedings of the ACM Workshop on Recurring Malcode (WORM), Alexandria, VA, USA, 2007, pp. 1–8.

[2] R. M. Mohammad, F. Thabtah, and L. McCluskey. "Predicting phishing websites based on self-structuring neural network." Neural Computing and Applications, vol. 25, no. 2, pp. 443–458, 2014.

[3] H. J. Liao, C. H. R. Lin, Y. C. Lin, and K. Y. Tung. "Intrusion detection system: A comprehensive review." Journal of Network and Computer Applications, vol. 36, no. 1, pp. 16–24, 2013.

[4] N. B. Amor, S. Benferhat, and Z. Elouedi. "Naive Bayes vs decision trees in intrusion detection systems." Proceedings of the ACM Symposium on Applied Computing, Nicosia, Cyprus, 2004, pp. 420–424.

[5] B. Mukherjee, L. T. Heberlein, and K. N. Levitt. "Network intrusion detection." IEEE Network, vol. 8, no. 3, pp. 26–41, 1994.

[6] J. Devlin, M. W. Chang, K. Lee, and K. Toutanova. "BERT: Pre-training of deep bidirectional transformers for language understanding." Proceedings of NAACL-HLT, Minneapolis, MN, USA, 2019, pp. 4171–4186.

[7] T. Mikolov, K. Chen, G. Corrado, and J. Dean. "Efficient estimation of word representations in vector space." Proceedings of the International Conference on Learning Representations (ICLR), Scottsdale, AZ, USA, 2013.

[8] A. Schmidt and M. Wiegand. "A survey on hate speech detection using natural language processing." Proceedings of the Fifth International Workshop on Natural Language Processing for Social Media, Valencia, Spain, 2017, pp. 1–10.

[9] T. B. Brown et al. "Language models are few-shot learners." Proceedings of Advances in Neural Information Processing Systems (NeurIPS), vol. 33, 2020, pp. 1877–1901.

[10] H. Touvron et al. "LLaMA: Open and efficient foundation language models." arXiv preprint arXiv:2302.13971, 2023.

[11] G. Sharma, R. Gupta, and A. Verma. "Automatic video summarization using transcript analysis and keyword extraction." International Journal of Computer Applications, vol. 183, no. 12, pp. 1–6, 2021.

[12] P. Papalampidi, F. Keller, and M. Lapata. "Movie plot analysis via turning point identification." Proceedings of the Conference on Empirical Methods in Natural Language Processing (EMNLP), Online, 2020, pp. 7163–7175.

[13] R. Gorwa, R. Binns, and C. Katzenbach. "Algorithmic content moderation: Technical and political challenges in the automation of platform governance." Big Data & Society, vol. 7, no. 1, pp. 1–15, 2020.

[14] L. W. Anderson and D. R. Krathwohl. A Taxonomy for Learning, Teaching, and Assessing: A Revision of Bloom's Taxonomy of Educational Objectives. New York, NY, USA: Longman, 2001.

[15] M. Vandewaetere, P. Desmet, and G. Clarebout. "The contribution of learner characteristics in the development of computer-based adaptive learning environments." Computers in Human Behavior, vol. 27, no. 1, pp. 118–130, 2011.

[16] A. Karpathy, G. Toderici, S. Shetty, T. Leung, R. Sukthankar, and L. Fei-Fei. "Large-scale video classification with convolutional neural networks." Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Columbus, OH, USA, 2014, pp. 1725–1732.

[17] M. Sahami, S. Dumais, D. Heckerman, and E. Horvitz. "A Bayesian approach to filtering junk e-mail." Proceedings of the AAAI Workshop on Learning for Text

Categorization, Madison, WI, USA, 1998, pp. 55–62.
[18] Y. LeCun, Y. Bengio, and G. Hinton. "Deep learning." Nature, vol. 521, no. 7553, pp. 436–444, 2015.
[19] A. Vaswani et al. "Attention is all you need." Proceedings of Advances in Neural Information Processing Systems (NeurIPS), vol. 30, Long Beach, CA, USA, 2017, pp. 5998–6008.
[20] M. A. Hearst. "Texttiling: Segmenting text into multi-paragraph subtopic passages." Computational Linguistics, vol. 23, no. 1, pp. 33–64, 1997.
[21] P. Agrawal and A. Arora. "URL-based phishing detection using machine learning." Proceedings of the International Conference on Computing, Communication and Automation (ICCCA), Greater Noida, India, 2018, pp. 1–5.
[22] S. Marchal, J. Francois, R. State, and T. Engel. "PhishStorm: Detecting phishing with streaming analytics." IEEE Transactions on Network and Service Management, vol. 11, no. 4, pp. 458–471, 2014.
[23] R. Navigli and S. P. Ponzetto. "BabelNet: The automatic construction, evaluation and application of a wide-coverage multilingual semantic network." Artificial Intelligence, vol. 193, pp. 217–250, 2012.
[24] F. Pedregosa et al. "Scikit-learn: Machine learning in Python." Journal of Machine Learning Research, vol. 12, pp. 2825–2830, 2011.
[25] A. Radford, J. Wu, R. Child, D. Luan, D. Amodei, and I. Sutskever. "Language models are unsupervised multitask learners." OpenAI Blog, vol. 1, no. 8, p. 9, 2019.

भारतीय भाषा, शिक्षा, साहित्य एवं शोध

ISSN 2321 – 9726

WWW.BHARTIYASHODH.COM

INTERNATIONAL RESEARCH JOURNAL OF MANAGEMENT SCIENCE & TECHNOLOGY

ISSN – 2250 – 1959 (0) 2348 – 9367 (P)

WWW.IRJMST.COM

INTERNATIONAL RESEARCH JOURNAL OF COMMERCE, ARTS AND SCIENCE

ISSN 2319 – 9202

WWW.CASIRJ.COM

INTERNATIONAL RESEARCH JOURNAL OF MANAGEMENT SOCIOLOGY & HUMANITIES

ISSN 2277 – 9809 (0) 2348 - 9359 (P)

WWW.IRJMSH.COM

INTERNATIONAL RESEARCH JOURNAL OF SCIENCE ENGINEERING AND TECHNOLOGY

ISSN 2454-3195 (online)

WWW.RJSET.COM

INTEGRATED RESEARCH JOURNAL OF MANAGEMENT, SCIENCE AND INNOVATION

ISSN 2582-5445

WWW.IRJMSI.COM

JOURNAL OF LEGAL STUDIES,POLITICS AND ECONOMICS RESEARCH

WWW.JLPER.COM