A Systematic Study of Hindi–Dogri Text Translation Using Neural Machine Translation

[1] **Joginder Kumar**
[1] Department of Computer Applications, Govt. Degree College Paloura
[1] joginder1981@gmail.com

ABSTRACT

While there have been notable steps taken towards Neural Machine Translation (NMT) for Indic Languages, the gap in precise translation output is still considerable. Unlike any other language, the primary reason for this gap is the lack of in-depth parallel corpora which causes refrains in its development. As a outcome, most Indian languages are still categorized as recourse depleted or untouched languages. To alleviate these barriers, there are multiple state of the art NMT frameworks that have been designed, with the pivot-based approach emerging as the most popular one. This approach, which is based on redistribution strategy, uses a mid-sentence language to ease the burden of translation. For the problem of translating between Dogri and English, a pivot based NMT framework is proposed which uses Hindi as a primary language. The pivot strategy is especially useful because Dogri is a low-resource language and has significant linguistic connections with Hindi. Furthermore, availability of systems that perform accurately during Hindi to English transitions solves the issue efficiently. The model provides a BLEU score of 20.97, a Sacre BLEU score of 21.54 and a chrF score of 50.01, proving that it is efficient for low resourced language translation.

Keywords - Neural Machine Translation, Low-Resource Languages, Pivot-Based Translation, Dogri-to-English Translation, Hindi Pivot, Transfer Learning, BLEU Score, SacreBLEU, chrF Score.

I INTRODUCTION

Multilingualism is a social problem and barrier to communication, considering that language functions as an important channel for interaction. To alleviate this problem, Machine Translation (MT) was created. It is a technology that automatically translates [1] texts, using computers to perform linguistically based changes from one narrative to another. Machine Translation has a long-standing history, and for decades has relied on so called “rule-based systems" and statistical models. Translation was usually done using predefined linguistic rules and lexicons, or by means of bilingual corpora statistical analysis. Neither of them is without pitfalls. Neural Machine Translation (NMT), however, does away with these limitations. Under deep neural networks, machine translation shifts to text stories expressed in the target language, where primary probabilistic estimates set out a complicated mapping between the source and target languages. Such enormous datasets allow for iterative control of translation quality through 'supervised training', where the model actively updates itself to approach the ideal function. This function measures how disparate the predicted translations are from the established ones, whiles controlling for parameters to lessen the negative results.

Across the landscape of translation using AI algorithms, the most important development is Neural Machine Translation (NMT) systems, which significantly automate and strengthen MT performance. However, the situation continues to be challenging, especially with languages that do not have enough annotated resources for training. So-called low-resource NMT deals with the creation of translation models for languages with limited linguistic resources such as low parallel corpora, insufficient computing facilities, and little NLP scholarship. One of the common ways to counter this problem is through the use of pivot language that enables translation from one language to another via an intermediate language that does not have align text with the two paraphrased languages. The performing of this method can be done without possessing the complex grammatical structures by doing this in a two-step process. The initial stage involves the translation of the main language to the vocabulary of the main language before, in the second step, the target language is produced.

The Dogri language is a member of the Indo-Aryan branch of the Indo-European languages. It is spoken chiefly in the Jammu region of India, and makes its way across the border to some parts of Pakistan. As most languages, Dogri also has a range of dialects which exhibit the differences in grammar and vocabulary of the language. On the other hand, English is a global lingua franca used in virtually all fields of human activities as a primary language. The construction of Dogri-English linguistic translation NMT systems should be facilitated due to the higher linguistic mobility and cross-linguistic interaction. But the main problem is the absence of sufficient quality parallel corpora, which is one of the main obstacles that slow down the development of these systems accurate NMT models for this language pair [2-5].

The primary goal of this study is to execute translation from Dogri to English while combining the use of Hindi as an intermediary language. In specific terms, the study aims to accomplish the following goals:

1. Create an NMT based framework that can translate low resource Indic language Dogri to English.

2. Solve the problems regarding the data sparsity of the semantic and syntactic representations that exist due to unavailability of parallel corpora sparseness for Dogri-English translation. This issue is solved through the pivot-based transfer learning where Hindi is used as a pivot language because it has a lot of resources and shares many features with Dogri [5-9].

This research is organized as follows: First, Section II illustrates the gaps and challenges that is vital to discuss with respect to the influence of machine translation methodologies on lesser known languages. Then, Section III discusses the study's models architecture, the chosen datasets, and the metrics used for evaluating performance. After that, Section IV analyses the results, providing the models' comparative analysis. Finally, Section V share overall conclusions, the new proposals or inventions in refinement of the translation systems for lesser-known languages, as well as the other pieces of information that came to light during the research. It is obvious that this approach attempts to close outstanding issues, and at the same time contribute to the development of NMT technology for poorly resourced languages.

II RELATED WORK

The field of automated language processing has been improved greatly by NMT. NMT surpassed previous systems based on statistical models of language processing. One of the crucial technological steps was the introduction of the transformer architectures [9-11] that utilize self-attention for enhanced contextual understanding. Regardless of the progress in NMT, there is still a persistent representation gap within low-resource languages caused by the deficit in parallel corpora. To address this problem, researchers have suggested parent-child transfer learning where a low resource NMT model is created from a high resource language model through transferring weights. This makes it possible to use transfer learning in a more powerful way and leads to much higher translation quality for languages that do not have enough linguistic resources.

More recent studies have focused on the application of UNMT for improving English-to-Telugu translation in the context of low-resource constraints. This approach has led to novel combinatorial frameworks that employ cross-lingual embeddings together with pseudo-parallel corpus reordering techniques. While these approaches lead to better main body translation, the problem of rare word translation in Indic languages is still apparent [11-14]. Also, researchers, with the aid of pre-trained word embeddings and custom architectures, have worked to translate Kashmiri into English and Hindi, which has significantly improved the BLEU scores from baseline models. This indicates that games may also be won with a lack of linguistic data through architectural changes.

In another step forward, a pivot-based NMT framework was created to improve the translation quality of non-English language pairs. A notable study reported best results employed English as medium for translation between Russian and Chinese, greatly reducing error propagation and optimizing BLEU scores. This study showed how the Russian language resource can be used extensively while minimizing the English required. In UNMT, other regions of the world - and especially India - have worked with languages like Kannada, achieving good results in English-Kannada translation using a pre-trained Cross-Lingual Language Model. While promising, the accuracy is limited due to the distinct structural differences between these languages, the available hardware, and data [14-19].

Prior research has explored the impact that the choice of the pivot language has on translation accuracy. While back-translating an English text to Arabic, Chinese, and paraphrasing monolingually, researchers uncovered variances in the assessments provided by the BLEU automated evaluation system and the human evaluator. A multi-pivot technique was also tested with English and Thai serving as pivots for Lao from Chinese. The results emphasized the importance [19-25] of the language quality and language resources for translation productivity. Separately, Mainpuri to Hindi translation was done with English as the pivot, providing more support as to the usefulness of pivot NMT with poor language resources.

The broadness literature provides a discussion of the challenges in translating less-resourced languages. Different strategies have been employed to solve the data scarcity issue, such as parent–child transfer learning, unsupervised learning, and pretrained word embeddings. Frameworks that are pivot based have also been useful for solving problems of linguistic resource deficiencies. However, many of the Indian languages like Dogri, Bodo, and Khasi have not yet been studied in the context of NMT. Follow up studies need to focus on these in order to improve and build more comprehensive translation systems for under-resourced languages [25-29].

III PROPOSED METHODOLOGY

This part examines transfer learning-based techniques and its application with NMT involving other than English language pairs. To efficiently tackle these data deficiencies, the methodology utilizes a sequential training pipeline: a source-to-pivot model is pre-trained first, then a pivot-to-target model is trained with parallel corpora for each language pair. This model incorporates progressive knowledge transfer by using the pre-trained source-to-pivot network as the encoder for the source language, and the pre-trained pivot-to-target decoder as the target language decoder. This creates a smooth initialization path for adapting from source to target.

Table 1. Description of dataset

Language Pair	*Data Source*	*Sentence Count*
Hindi-English	CCMatrix Dataset	300,000

Dogri-Hindi	In-House Dataset (Derived from AIR's 'Mann Ki Baat')	7,458
Dogri-Hindi	Optical Character Recognition (OCR) Extracted Data	1,235
Dogri-Hindi	Machine-Generated Corpus (Google Translate)	291,307

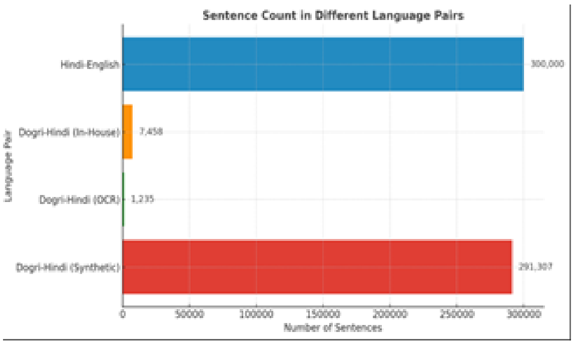

Fig. 1. To illustration of Graphical of Dataset Description

Table 2. Description dataset splitting into test & train

Dataset Partition	Sample Count	Purpose
Training Set	300,000	Used for model learning and parameter optimization
Validation Set	2,500	Utilized for hyperparameter tuning and overfitting prevention
Testing Set	1,500	Reserved for evaluating generalization and performance metrics

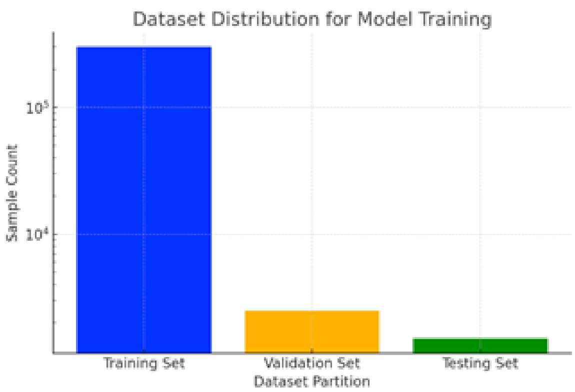

Fig. 2. Illustration of Data Splitting into Testing and Training Dataset

A. Data Collection and Description

A structured set of methodologies is prepared to synthesize and compile a dataset that facilitates the use of the zero-source Dogri language. This part explains how the data was collected as well as the reference materials that were used for the corpus building. Mann Ki Baat is an Indian radio program hosted by the Prime Minister Narendra Modi where he shares his thoughts on various issues via All India Radio and online platforms. The program started in the year 2014 and comes with multiple episodes. The program covers topics such as education, public health, national security, policy talk and socio-economic issues, and everything else in between chatting directly with the citizens of India. Because there are little resources available concerning the Dogri language, documents provided by the broadcasting institutions are made accessible in the form of digitized PDFs and scanned manuscripts. OCR technology makes it possible to convert scanned documents and images of text into machine-readable text. The use of such technologies aids in the ease of obtaining operational transcripts and other vital radio-supplementary structured linguistic data. It is especially useful for increasing the efficiency and ease with which Mann Ki Baat archives can be accessed and processed.

Jammu and Kashmir is a region marked by high linguistic diversity with Urdu, Dogri, and Kashmiri as the major languages. While translation for Urdu is readily available on the web, including Google's Hindi to Urdu and English to Urdu translation services, Kashmiri still does not have adequate resources and tools available, although the University of Kashmir is working towards developing machine translation systems for the language. There are, however, no major efforts in machine translation for Dogri. Software localization has been aided by DIT, India for Dogri, which has resulted in the development of several localized software like Open Office, Firefox, Thunderbird, Pidgin Messenger, Sunbird Calendar and Scribus. Still, there is no working system that allows for machine translation of Dogri. This work is the first step in developing an automated translation system from Hindi to Dogri and is a major leap in the field of computational linguistics for low-resource languages.

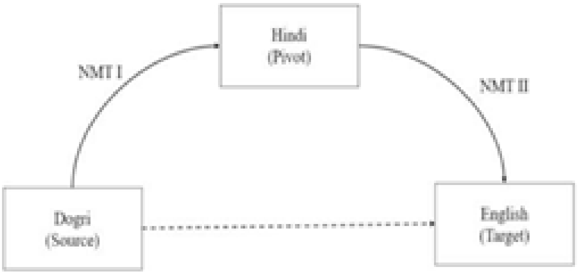

Fig. 3. Illustration of Translation of Dogri-Hindi-English approaches

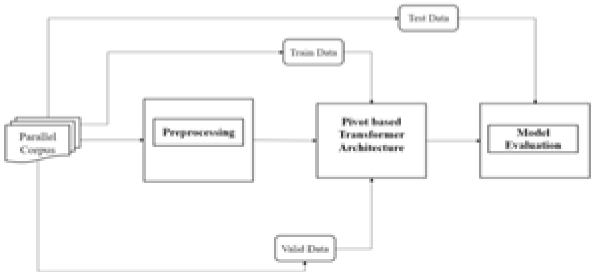

Fig. 4. llustration of NMT Approaches Dogri – Hindi using Pivot Method

Synthetic data creation was used as an additional strategy for collecting data. For this task, Google Translate was utilized to create the datasets. An overview of the assembled training data that was obtained from various sources is given in Table I, while Table II reports on the obtained statistical parameters pertaining to the training, validation, and testing datasets.

B. Data Preprocessing

In the framework of quantitative analysis, data cleansing is the preparation of the data that has to be analysed, which includes transliteration, tokenization, Byte Pair Encoding (BPE), and embedding. Transliteration aids in enabling written communication between people who speak different languages, by changing words from one language into another phonemically which is especially useful for language with different writing systems. For Dogri, which is phonemically similar to Punjabi, transliteration was utilized to fetch the resourceful language model and embeddings trained on Punjabi texts.

Tokenization the process of breaking any information into smaller units referred to as tokens which in case can be a word, sub words, or even punctuation marks for more detailed language processing. A method that uses tokenized sub words is the Byte Pair Encoding (BPE) which was implemented over 64k times for representation optimization. Furthermore, embedding features of less resourced languages are very well enriched by trained models. When refining our corpus we used Fast Text embedding trained in Punjab, Hindi, and English. This combination provides sufficient coverage for non-monolingual contexts.

C. Design of System

We focus on the proposed pivot-based machine translation model for Dogri to English using Hindi as the pivot language. The investigation revolves around the intricacies of neural architecture, following a systematic approach found in Figs. 1 and 2. The pivot-based translation framework is heuristic in nature, starting with building an NMT model for Dogri to Hindi translation, which is labelled NMT-I for this study. This phase uses an encoder that preserves the meaning of source sentences. The next step processes this Hindi representation through a decoder to translate it to English. This step paves the way for the pivot-based mechanism that leads to the creation of the Hindi to English NMT model, NMT-II for this study.

Table 3. Translation division analysis

Model ID	*Translation Type*	*Source → Target*	*Architecture*	*Performance Metrics*

NMT I	Encoder Model	Dogri → Hindi	6L Encoder, 6L Decoder, 2048D, 50 Epochs	BLEU (Train/Valid): 5.61 / 25.67, BLEU (Test): 27.60, SacreBLEU : 54.00
NMT II	Decoder Model	Hindi → English	6L Encoder, 6L Decoder, 2048D, 50 Epochs	BLEU (Train/Valid): 2.36 / 7.75, BLEU (Test): 13.50, SacreBLEU : 50.80

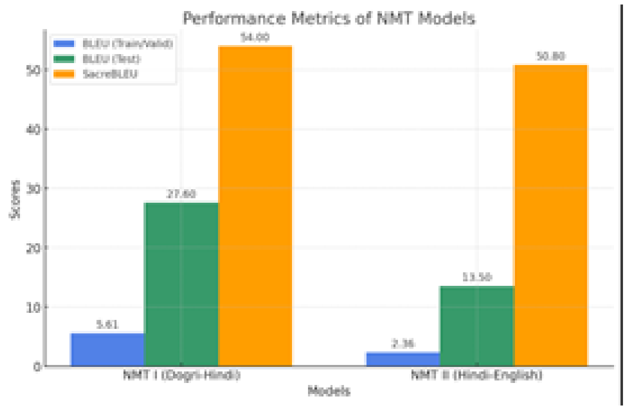

Fig. 5. Illustration NMT Models Performance Metrics

D. Presentation of Performance Metrics

The Bilingual Evaluation Understudy (BLEU) metric is commonly used for measuring the quality of machine translations by comparing the output of machine translation systems with human reference translations. The values of the BLEU scores is given on a scale of 0 to 100, where 100 reflects a perfect match of the translation with the reference text.

Furthermore, the CHaRacter-level F-score (chrF) is another such measure of the quality of translations, but it does so by comparing n-grams at the character level for the system output and a reference translation. This approach allows the evaluator to take a more detailed look into languages that are morphologically complex as it has a score range starting from 0 up to 100. To some extent, SacreBLEU would be received more positively as it is a reworked version of the original BLEU metric that aims towards higher accuracy and dependable results when evaluating a translation model with multiple datasets.

IV RESULTS AND DISCUSSIONS

A total of nine experimental models were developed systematically, distinguishing each model with respect to unique architectural features such as a defined configuration of encoder and decoder layers, and latent dimensionality. These models were created with a purpose to test the degree of their impact on the translation accuracy. The assessment covered key evaluation criteria like the highest BLEU test score achieved, SacreBLEU, and chrF scoring. Structural configurations of the encoder-decoder models are provided in Table III, while the corresponding performance metrics are illustrated in Figures 4 and 5. Moreover, Table IV presents a detailed overview of the nine outlined experimental models by hyperparameter setting and evaluation results. Table III has listed two core models: NMT I with Dogri as the source language and Hindi as the target language (encoder model) and NMT II with Hindi as source and English as target language (decoder model). These base models were the basis for the pivot models which were created by modifying the four most impactful hyperparameters to translation performance.

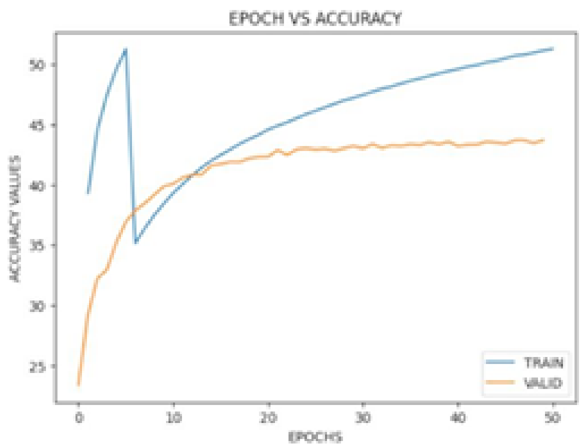

Fig. 6. Illustration NMT I Model Epoch vs Accuracy

It is written in the Devanagari script which is used for Hindi and several other North Indian languages. These phonemes are further classified into vowels and consonants. Altogether, this language has 10 distinct vowel phonemes and 28 consonant phonemes.

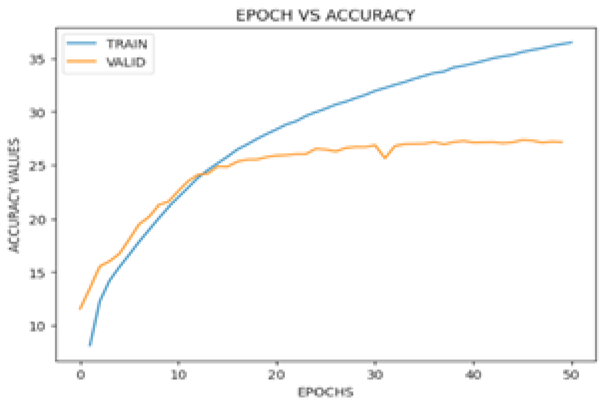

Fig. 7. Illustration NMT II Model Epoch vs Accuracy

The most optimal experimental model is defined by an architecture that contains 4 encoder layers and 4 decoder layers with ten training spans, a latent space dimension of 1024, and a total training span of 50 epochs. The configuration for this model resulted in the highest evaluation metrics, earned BLEU4 score of 20.97, SacreBLEU score of 21.54 and chrF score of 50.01. These scores are presented further in Table IV. It is reasonable to assume that some hyperparameter tuning was performed to try and get the optimal score. In addition, Fig 6 enables one to compare and analyze the BLEU scores across different parameter combinations for both the validation and test datasets. Besides, Table V allows comparison between the model-generated output for the input sentences and the reference translations through which evaluation is made. The translated sequences demonstrate one pattern of structure and outline a few changes in meaning, which shows the enduring complications related to accurate and context sensitive translations, especially to less resourced languages like Dogri.

Table 4. Comparative analysis of experimental models

Model Parameters Varied While Training				Model Evaluation Metrics		
Encoder Layers	Decoder Layers	Latent Dimension	Epochs Trained	BLEU	SacreBLEU	chrF
2	2	256	50	12.01	12.98	42.03
2	2	512	50	13.57	14.13	43.30
2	2	1024	50	13.25	14.57	43.87
4	4	256	50	19.21	20.03	48.97
4	4	512	50	20.10	20.98	49.05
4	**4**	**1024**	**50**	**20.97**	**21.54**	**50.01**
6	6	256	50	19.56	20.20	49.23
6	6	512	50	19.19	19.90	49.54
6	6	1024	50	19.06	19.50	48.67

Even though there are noticeable gaps between the reference translations and outputs from the machine translation system, these gaps also indicate the level of understanding and meaning thereof. These gaps emphasize the advantages and shortcomings of machine translation models. Most modern models have challenges dealing with differences in syntax, context, and linguistics.

The gaps noted explain the intricacy of natural language processing because 100% accuracy is almost impossible due to the ambiguous nature of language, its cultural content, and in the case of low resource languages. However, the movement towards improvement and progress of machine translation works will always be focused on making languages more understandable so that the message, even though changed with some minor issues, is understood to an acceptable amount.

Source Sentence	Reference Translation	System Translation
[illegible]	The police officer said a policeman visited the village the next day after the complaint was received.	The police have been asked to go to the spot, said a senior police official.
[illegible]	Congress president Rahul Gandhi released the party' s manifesto for the seven phased Lok Sabha elections on Tuesday.	Congress president Rahul Gandhi on Tuesday released seven phases for the Lok Sabha elections
[illegible]	Gurugram police officials were designated outside the metro station to ensure traffic flow remains unhampered from the drill while the DMRC officials regulated passenger queues to ensure there is no disruption to Metro services	The DMRC was appointed that the officials appointed as the traffic police officials to ensure that passengers coming out of the Metro stations would be taken out to ensure that passengers will remain closed.
[illegible]	Rainbows caused by sunlight always appear in the section of sky directly opposite the Sun.	Because of sunlight, rainbows always appear in the part of the sky directly in front of the sun.

Fig. 8. Illustration BLEU Score Validation Train and Test Dataset

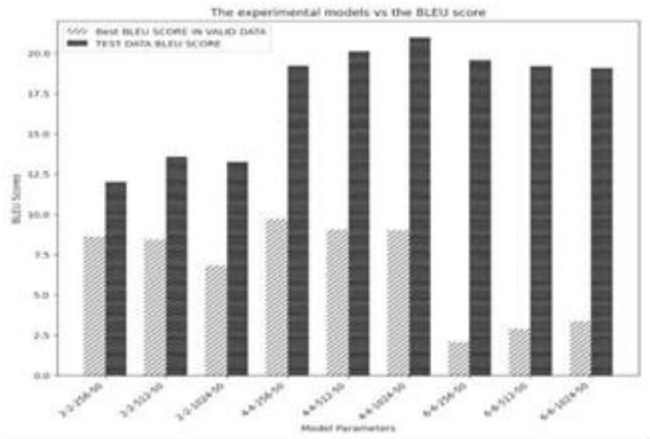

Fig. 9. Illustration Translation of Reference to System Translation

V CONCLUSION

An effort is being made to improve the quality of research in the field of Indic language processing by pursuing strategies that automatically translate between Dogri and other languages through linguistic transfer that is systematic in nature. At the heart of this initiative lies the design and implementation of a pivot neural machine translation system where Hindi is the pivot language. The obtained scores, BLEU 20.97, SacreBLEU 21.54, and chrF 50.01, indicate that this technique does work by boosting the structural and morphological similarities of Dogri and Hindi. Nonetheless, even with these results, the scarcity of quality parallel corpora for the Indic languages continues to be an issue.

Finally, to aid in the creation of these corpora, the initial criteria for the machine translation framework are met through incorporating concepts of automated translation and NMT. Additionally, using more developed methodologies of knowledge transfer and zero-shot learning enables greater efficiency in corpus and translation system in the inflating scope. The combination of more sophisticated pivot-based techniques and the greater use of intermediate languages enables the construction of a more simple yet powerful translation model while optimizing for language.

REFERENCES

[1] Nair, Aarathi Rajagopalan, et al. "Neural Machine Translation Using a Pivot Approach: Dogri to English." 2024 15th International Conference on Computing Communication and Networking Technologies (ICCCNT). IEEE, 2024.

[2] Mujadia, Vandan, and Dipti Misra Sharma. "BhashaVerse: Translation Ecosystem for Indian Subcontinent Languages." arXiv preprint arXiv:2412.04351 (2024).

[3] Bisht, Akhilesh, and Deepa Gupta. "Neural machine translation for low resource indian language: Hindi-Kangri." Journal of Intelligent & Fuzzy Systems Preprint (2024): 1-13.

[4] Sarkar, Tiyas, et al. "Review Paper of Performance Analysis in Wireless Sensor Networks." Kilby 100 (2023): 7th.

[5] Singh, Dalwinder, et al. "Predictions on the Future of Agriculture and Recent Developments in Agricultural Technology." International Conference On Artificial Intelligence Of Things For Smart Societies. Cham: Springer Nature Switzerland, 2024.

[6] Singh, Dalwinder, et al. "of Agriculture and Recent Developments in Agricultural." The Smart IoT Blueprint: Engineering a Connected Future: Guiding Principles and Practical Strategies for Seamless Integration 297 (2024).

[7] T. Sarkar, M. Rakhra, V. Sharma, A. Maan, A. Singh and G. S. Cheema, "Comparative study on Loan Approval Prediction Utilizing Machine Learning Approaches," 2024 7th International Conference on Contemporary Computing and Informatics (IC3I), Greater Noida, India, 2024, pp. 72-78, doi: 10.1109/IC3I61595.2024.10829014.

[8] H. Kaur, T. Sarkar, B. Kaur, A. Singh, G. S. Cheema and S. Das, "Using Machine Learning Techniques in Business Analytics: Lessons from Agile Project Management," 2024 7th International Conference on Contemporary Computing and Informatics (IC3I), Greater Noida, India, 2024, pp. 214-220.

[9] T. Sarkar, M. Rakhra, V. Sharma, A. Singh, K. Jairath and A. Maan, "Comparing Traditional vs Agile Methods for Software Development Projects: A Case Study," 2024 7th International Conference on Contemporary Computing and Informatics (IC3I), Greater Noida, India, 2024, pp. 221-227, doi: 10.1109/IC3I61595.2024.10829321.

[10] S. Takkar, A. S. Rayat, V. Sharma, T. Sarkar, K. Jairath and A. Singh, "Problems and Optimal Methodologies Agile Project Management Solution in Publicly accessible Sector: A Comprehensive Evaluation of the Literature," 2024 7th International Conference on Contemporary Computing and Informatics (IC3I), Greater Noida, India, 2024, pp. 207-213, doi: 10.1109/IC3I61595.2024.10828832.

[11] S. Singh, T. Sarkar, M. Mangla, M. Rakhra, A. Singh and K. Jairath, "FinTech Edge: Utility Computing & Artificial Intelligence Technologies for Smart Financial Acquisition & Blockchain in the Financial Industries," 2024 7th International Conference on Contemporary Computing and Informatics (IC3I), Greater Noida, India, 2024, pp. 228-234, doi: 10.1109/IC3I61595.2024.10828908.

[12] V. K. Singh, B. Moharana, T. Sarkar, A. Maan, A. Bhattacherjee and M. Rakhra, "Practical Assessment of Large Language Models in Cloud Computing Using Real-World Data Applications," 2024 International Conference on Cybernation and Computation (CYBERCOM), Dehradun, India, 2024, pp. 356-362.

[13] B. Moharana, V. K. Singh, T. Sarkar, A. Maan, M. Rakhra and P. Chakrabarty, "A Comprehensive Review on Performance Analysis of Enterprise Cloud Computing Architecture," 2024 International Conference on Cybernation and Computation (CYBERCOM), Dehradun, India, 2024, pp. 184-191, doi: 10.1109/CYBERCOM63683.2024.10803178.

[14] A. K. Verma, M. Rakhra, A. Bhattacherjee, A. Maan, T. Sarkar and V. K. Pandey, "A Suggested Model for Using Multi Factor Authentication Framework in Cloud Computing for SME," 2024 International Conference on Cybernation and Computation (CYBERCOM), Dehradun, India, 2024, pp. 192-197, doi: 10.1109/CYBERCOM63683.2024.10803146.

[15] B. Moharana, V. K. Singh, T. Sarkar, D. Singh, M. Rakhra and V. K. Pandey, "Automated Questions Answering Generation System Adopting NLP and T5," 2024 International Conference on Cybernation and Computation (CYBERCOM), Dehradun, India, 2024, pp. 363-369, doi: 10.1109/CYBERCOM63683.2024.10803238.

[16] V. Sharma, M. Rakhra, V. Sharma, T. Sarkar, A. Singh and K. Jairath, "Hybrid Approaches for Stocks Price Prediction: A New Institutive Way of Neural Architecture Design," 2024 International Conference on Electrical Electronics and Computing Technologies (ICEECT), Greater Noida, India, 2024, pp. 1-6, doi: 10.1109/ICEECT61758.2024.10739249.

[17] M. Rakhra, T. Sarkar, V. Sharma, P. K. Angra, A. Kumar and P. Singh, "An In-Depth Analysis Of Local Mechanic Support For Immediately Apparent Repairs And Scheduled Maintenance Services," 2024 International Conference on Electrical Electronics and Computing Technologies (ICEECT), Greater Noida, India, 2024, pp. 1-7, doi: 10.1109/ICEECT61758.2024.10739024.

[18] T. Sarkar, A. Singh and R. Kakkar, "Enhancing Deepfake Video Detection Performance with a Hybrid CNN Deep Learning Model," 2024 International Conference on Electrical Electronics and Computing Technologies (ICEECT), Greater Noida, India, 2024, pp. 1-7.

[19] T. Sarkar, M. Rakhra, V. Sharma, S. Takkar and K. Jairath, "Comparative Study of Object

Recognition Utilizing Machine Learning Techniques," 2024 International Conference on Communication, Computer Sciences and Engineering (IC3SE), Gautam Buddha Nagar, India, 2024, pp. 726-731, doi: 10.1109/IC3SE62002.2024.10593475.

[20] A. K. Verma, Shruti and T. Sarkar, "Utilizing Imaging Steganographic Improvement using LSB & Image Decoder," 2024 International Conference on Communication, Computer Sciences and Engineering (IC3SE), Gautam Buddha Nagar, India, 2024, pp. 144-150, doi: 10.1109/IC3SE62002.2024.10593623.

[21] P. Chakrabarty, T. Sarkar, M. Rakhra, K. Jairath and V. Sharma, "Enhanced Data Security Framework Using Lightweight Cryptography and Multi-Level Encryption," 2024 International Conference on Communication, Computer Sciences and Engineering (IC3SE), Gautam Buddha Nagar, India, 2024, pp. 720-725.

[22] T. Sarkar, M. Rakhra, V. Sharma and A. Singh, "An Empirical Comparison of Machine Learning Techniques for Bank Loan Approval Prediction," 2024 International Conference on Communication, Computer Sciences and Engineering (IC3SE), Gautam Buddha Nagar, India, 2024, pp. 137-143, doi: 10.1109/IC3SE62002.2024.10593355.

[23] S. Talwani, M. Rakhra and T. Sarkar, "Perspectives on the Future of Agriculture and Recent Developments in Agricultural Technology," 2024 11th International Conference on Reliability, Infocom Technologies and Optimization (Trends and Future Directions) (ICRITO), Noida, India, 2024, pp. 1-6.

[24] T. Sarkar, B. Moharana, M. Rakhra and G. S. Cheema, "Comparative Analysis of Empirical Research on Agile Software Development Approaches," 2024 11th International Conference on Reliability, Infocom Technologies and Optimization (Trends and Future Directions) (ICRITO), Noida, India, 2024, pp. 1-6, doi: 10.1109/ICRITO61523.2024.10522134.

[25] A. K. Verma, T. Sarkar, M. Rakhra, V. K. Pandey and P. Chakrabarty, "Studying the Impact of Voice-Activated Chatbot's with Artificial Intelligence on Consumer Interaction in India," 2024 4th International Conference on Technological Advancements in Computational Sciences (ICTACS), Tashkent, Uzbekistan, 2024, pp. 1708-1713, doi: 10.1109/ICTACS62700.2024.10841074.

[26] B. Moharana, V. K. Singh, T. Sarkar, A. Hassan and V. K. Pandey, "Predicting Indian Master's and PhD Enrollment using Data Mining & Machine Learning Approaches," 2024 4th International Conference on Technological Advancements in Computational Sciences (ICTACS), Tashkent, Uzbekistan, 2024, pp. 1714-1720.

[27] T. Sarkar, M. Rakhra, P. Singh, P. K. Angra, G. S. Cheema and N. Mathur, "Comprehensive Review on Fog Computing: Obstacles, Feature Traits, Limitations, and Potential Leverages on Application," 2024 4th International Conference on Technological Advancements in Computational Sciences (ICTACS), Tashkent, Uzbekistan, 2024, pp. 404-410, doi: 10.1109/ICTACS62700.2024.10840690.

[28] T. Sarkar, A. Bhattacherjee and M. Mukhia, "Large Language Model for Practical Evaluation in Cloud Computing with Applications of Real World Data," 2024 4th International Conference on Technological Advancements in Computational Sciences (ICTACS), Tashkent, Uzbekistan, 2024, pp. 1721-1729.