INDIAN JOURNAL OF SCIENCE AND TECHNOLOGY

SYSTEMATIC REVIEW

HiTEK Pre-processing for Speech and Text: NLP

Naveenkumar T Rudrappa[1]*, Mallamma V Reddy[2], M Hanumanthappa[3]

1 Research Scholar, Department of Computer Science, Rani Channamma University, Belagavi, Karnataka, India
2 Assistant Professor, Department of Computer Science, Rani Channamma University, Belagavi, Karnataka, India
3 Professor of Computer Science and Applications, Jnanabharati Campus, Bangalore University, Bangalore, Karnataka, India

OPEN ACCESS

Received: 14-02-2023

Accepted: 22-04-2023

Published: 12-05-2023

Citation: Rudrappa NT, Reddy MV, Hanumanthappa M (2023) HiTEK Pre-processing for Speech and Text: NLP. Indian Journal of Science and Technology 16(19): 1413-1421. https://doi.org/10.17485/IJST/v16i19.296

* **Corresponding author**.

naveenkumartr@rcub.ac.in

Funding: None

Competing Interests: None

Published By Indian Society for Education and Environment (iSee)

ISSN
Print: 0974-6846
Electronic: 0974-5645

Abstract

Objective: To develop a system that accepts a sentence consisting of two and/or four languages and convert it to a target language text, termed as Cross Language Speech Identification and Text Translation System. **Methods:** A combinatorial model consisting of Hidden Markov Model, Artificial Neural Networks, Deep Neural Networks and Gaussian Mixture Model are utilized for direct and indirect speech mapping. Trained dataset consisting of thousand phonemes for each of the Hindi, Telugu, English and Kannada languages, initially for bank, hospital domains, later the grammatical phonemes of each language were added and wave files consisting of cross lingual spoken sentence were created which incurred a six months period to build from scratch, as cross lingual vocal data-set is not available. Hindi language dataset Shabdanjali was also referred. The basic parameters considered for creation of structured dataset are loudness, pause, pitch, tone, noise cancellation, sampling frequency, threshold etc. **Findings:** Comparative analysis of various techniques, target languages and features are tabulated. Research idea emerged from the comparative analysis of Monolingual Systems where there was a gap for cross lingual speech to text translation. The architecture can be enhanced in future for other regional languages of India. **Novelty:** A new bench mark for Cross Language dataset was created. This work presents CLSITT tool applicable in transforming public speeches spoken in multiple languages to a selected target language and the tool is helpful for a regional news editor, rural and agricultural activities, medical applications, defence and so on.

Keywords: Artificial Intelligence (AI); Deep Learning (DL); Machine Learning (ML); Natural Language Processing (NLP)

1 Introduction

Natural Language Processing involves the processing of speech, text or both. CLSITT plays a major role for the users who are not skilled in multiple languages but need to carry the research and understand the concepts present in other languages. Spoken

languages have a specific set of rules, symbols and grammar that form a meaningful conversation. Previous research focussed on Monolingual, Bilingual, Cross Lingual[(1)] Text Translation and Monolingual Speech[(2)] Translation but their exists a gap in the previous research that Multilingual Speech and Cross Lingual Speech is not addressed. Hence, in the current research we have addressed the issue of Multilingual and Cross Lingual Speech.

Social networks like Facebook, LinkedIn[(3)] and others generate a huge amount of data on the world wide web but nowadays due to the user-friendly technologies users can create, share new opinions, ideas and content either through spoken words or typed text, resulting in massive data. This unstructured data cannot be processed by the machines directly and hence needs machine learning algorithms. Websites are algorithm dependent designed to retrieve, split, process, check spelling, count the words and summarize text. These systems had limited capabilities for meaning extraction from the sentences. NLP mandates elevated symbolic[(4)] capabilities for dynamic binding, treatment of repeated composite structures, attainment of syntactical, semantical and lexical phases, routing control of information across several modules, preparation of basic language constructs and text summarization. Human beings have an inbuilt capability of processing both visible(through eyes) and audible information[(5)]. Speech is an alternate for textual communication but with added qualitative characteristics which help to understand human emotions.

Major advancements in machine translation[(6)] involves translating a text or speech from one specific language[(7)] to target language, the research has focused on translating cross lingual speech to target language text. The objective of the research is to develop a system that accepts multilingual oral sentence consisting of two, three or four languages and translate it to a specific target language, the system named as Cross Lingual Speech to Text Translation System. For example transforming public speeches spoken in multiple languages to a selected target language that reduces the overhead of manual intervention for a regional news editor. This paper focuses on monolingual and multilingual (Cross) speech to text translation for Hindi, Telugu, English and Kannada languages.

1.1 Terminologies and notations

Natural Language Processing is a subset of Linguistics, and consists of various fields like Artificial Intelligence (AI), Machine Learning (ML) and Deep Learning (DL) as shown in Figure 1 .

1. Linguistics is a study of language[(8)], its structure, grammar, syntax, and phonetics.
2. Natural Language Processing (NLP) is subset of Linguistics a process of extracting meaning and learning from textual data or speech.
3. Artificial Intelligence (AI) is a subset of NLP which transforms human intelligence[(9)] into machines based on rules, is used to solve problems and demonstrate intelligent behavior.
4. Machine Learning (ML)[(10)] is a subset of AI which enables self learning by machines using some input data, make accurate predictions, have both supervised and unsupervised methods.
5. Deep Learning[(11)] (DL) is a subset of ML that consists of algorithms, trained to mimic human brain and accomplish tasks like identifying patterns. Processing of spoken language involves making computers to perform a task. The NLP system may convert speech to text, text to speech, one language text to another language text or one language speech to another language speech. It may also accept speech or text as input by an electronic robot to perform some action like lifting a pen or playing a music file. Basic notations used in NLP are shown in Figure 2.

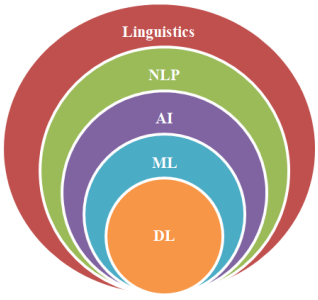

Fig 1. Combinatorial Fields of Linguistics

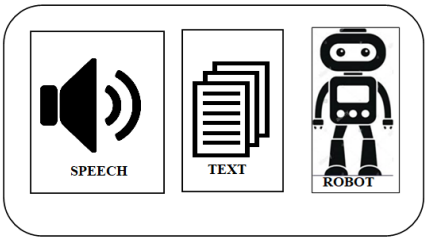

Fig 2. Basic Notations in NLP

1.2 NLP Components

There are 2 main components of any NLP system, Understanding of Natural Language and Generation of Natural Language:[12]

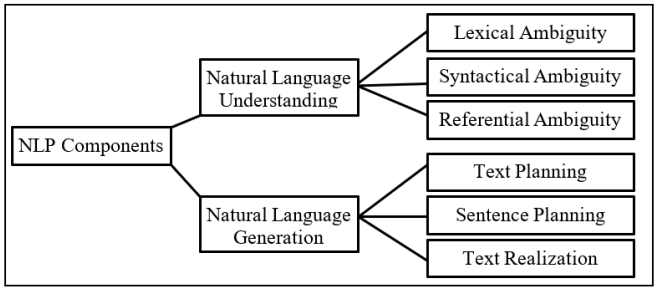

Fig 3. Components of an NLP System

1. Natural Language Understanding (NLU) uses computer software to understand a dialect by accepting input text or speech. It involves i. Lexical Ambiguity (LA) that is it refers to the multiple meanings of a word and its context. It can be resolved using parts of speech tagging. Ex: The Fish is ready to eat. Is the fish ready to eat its food or is it ready for someone else to be eaten. ii. Syntactical Ambiguity (SA) means that the word refers to a different meaning in the same context. Ex: John met Ajay and Vijay. They went to a restaurant. They refer to Ajay or Vijay or all the three. iii. Referential Ambiguity (RA) occurs when a spoken word or written text, could imply to more than one person or things in other sentences or in the same sentence. Ex: The girl explained her mother about the theft. She was shocked. She is referentially ambiguous as it refers to both the girl and mother.
2. Natural Language Generation (NLG): It is the production of meaningful words, phrases and statements by a computer machine. It consists of i. Text planning (TP): It is the process of retrieval of related data from a repository. ii. Sentence planning (SP): It is the process of word selection to form a meaningful statement. iii. Text Realization (TR): It is the process of matching a sentence plan into a context.

2 Methodology

2.1 Dataset

HiTEK Hybrid Cross Lingual Dataset is constructed consisting of thousand phonemes for each language. Some of the other datasets which are referred in our research are Shabdajnali, https://www.manythings.org/anki/. The dataset consists of tab delimited Bilingual Sentence Pairs of the parallel corpus of source as well as the target language files having one sentence per line. The dataset is collected from several sources and combined.

2.2 Preprocessing and cleaning dataset

The proposed system stores both raw, cleaned speech and text corpus for English, Hindi, Kannada and Telugu language sentences in separate speech (.wav) and text (.txt) files. An instance of the below mentioned documents will be placed in one

data directory. The data is then classified as follows:

Table 1. Speech File Description

Speech Files	Description
Source	This is a training file having HiTEK languages within a speech sentence. (Source Languages)
Target	This is a training file having both speech and textual sentences in all the four languages. (Target Language)
Source Validation	Validation data consisting of Hindi, Telugu, English and Kannada language words. (Source Languages)
Target Validation	Validation data consisting of Hindi, Telugu, English and Kannada language words. (Target Language)
Source Test	Initial stage where human speaker checks for proper functioning of hardware and corresponding software.

This section presents a workflow of an NLP system. Input to the system may be speech or text. The detail diagram is depicted in Figure 5

1. Speech Repository serves as the central storage for raw input, supervised, intermediate speech [13] files obtained after each processing phase, phase wise log file and the final speech output obtained.
2. Text Repository serves as the central storage for raw input, supervised, intermediate text [14] files obtained after each processing phase, phase wise log file and the final text output obtained.
3. Speech Segmentation [15] is the process of identifying boundaries between words, syllables, phones or phonemes in spoken languages to give a proper meaning for a word or a sentence in silent and noisy situations. Ex: CAN=3 Segments. Speech specific phonological [16] rules are productive in nature. Ex: Usage of the word "THE" to construct a sentence.
4. Text Tokenization [17] is a task of dividing a textual sentence into a predefined set of tokens like space, tab, or punctuation. Ex: A sentence can be divided into words and a paragraph can be divided into sentences, here words and sentences act as tokens.
5. Morphological Analysis [18] is the process of formation of words from the smallest primitive chunks or constituents called Morphemes. Ex LOVE cannot be divided, hence it is a morpheme in itself. Ex: Word UNBREAKABLE= UN + BREAK + ABLE, [3 Morphemes].
6. Lexical analysis is a procedure [19] of breaking huge sized content into constituent words, statements and paragraphs. It is used to understand the relationships between and meaning of words in a sentence but within a context by utilizing dictionary. A cricket stadium is validated as a location in the computer machine.
7. Syntactic Analysis [20] checks for words, grammar in a sentence and their arrangement among the words and validates whether the natural language aligns with grammatical rules. It consists of Grammatical rules (Techniques are Stemming, Dependency Parsing and Word Separation) and Parts of Speech tagging (Technique is word annotation). Ex: A word may be annotated as a noun in the preceding statement and an adjective or verb in a succeeding statement. Table 2 visualizes a set of POS [21] tags and Figure 4 illustrates an example sentence [22].

Table 2. Parts of Speech Tag Set

Key	POS	Key	POS
CC	Coordination of a Conjunction	NN	Noun Singular
CD	Cardinality of a Number	NNS	Plural of a Noun
DT	Determiner	NNP	Singular of a Proper Noun
SYM	Symbol	NNPS	Plural of a Proper Noun
UH	Interjection	VB	Verb
IN	Preposition	VBD	Past form of verb
JJ	Adjective	VBG	Present participle of a verb
JJR	Comparative of an Adjective	VBN	Past Participle of a Verb
JJS	Superlative of an Adjective	PRP	Pronoun Personal
PDT	Pre-determiner	RB	Adverb
POS	Possessive End	RBR	Adverb, Comparative
FW	Foreign Word	RBS	Adverb, Superlative

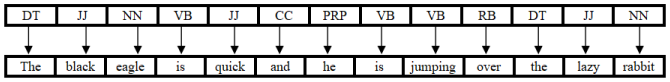

Fig 4. POS tagging for a sentence

8. Semantic Analysis: is an algorithmic activity to understand [23], merge and analyze the meaning of words, integrate words to form phrases and unite phrases to form a well-structured sentence. Ex: Rat eats Cat is syntactically correct but invalidated and Cat eats Rat is both syntactically correct and validated by the semantic analyzer.

9. Discourse Analysis [24] is the analysis of the meaning of a sentence that is dependent on other sentences. It can be applied to a complete vocal conversation or a collection of written text or non-verbal communication like gestures as explained in Table 3. Ex: She needed it depends on previous discourse [25] context.

Table 3. Discourse Analysis for various types of text

	Level of communication	Analysis
1	Conversational codes	Communication among individuals like user response and interrupts
2	Non-verbal Communication	Speaker hand gestures, pauses, tone to illustrate emotions and intentions
3	Genre	A text category to analyze the relation and communication goals. Ex: Political speeches, Newspaper articles
4	Structure	Analysis of a text structure to focus on the construction of a narration
5	Grammar	Construction of sentences. Ex: Active or passive voice, verb tenses, question building can depict the meaning and intention
6	Vocabulary	Formation of words, clauses and statements for detailed context analysis

10. Pragmatic Analysis helps people to identify the intentions of a situational conversation by the application of rules. It mandates the inclusion of some extra textual meaning, goals, plans and world knowledge into a conversational context. Ex: Close the door may be an order or request. To find out the difference inference modules and world knowledge is to be retrieved from the data stores.

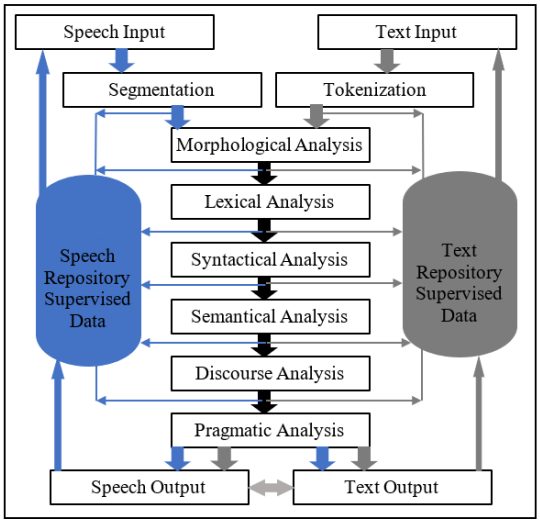

Fig 5. Workflow of a Natural Language Processing System

2.3 Output: Processing of NLP leads to four different deliverables such as

- Input Speech to Output Text,
- Input Speech to Output Speech,
- Input Text to Output Speech and
- Input Text to Output Text.

Resultant NLP Processing System may be helpful in the following applications and described in Figure 6.

1. Automatic Text Summarization is the process of intelligently shortening long pieces of text such that the meaning is not altered. Ex: Grammarly, Microsoft-word, Google docs etc.
2. Named Entity Recognition is an activity of locating and classifying an entity into predefined text categories like organizations, person names, location etc.
3. Sentiment Analysis is the process of identifying sentiments which may be positive, negative, neutral opinion from speech or text reviews in Social Media, Online Survey Responses and attributes like clothing size that fit or not fit.
4. Speech Recognition enables computers to recognize and transform spoken language into text or action. Ex: Alexa, Siri and Cortona.
5. Topic Segmentation divides written text speech or recordings into shorter coherent segments. Ex: Classification of news articles into Sports, Politics, Entertainment and so on.
6. Intent Classification is the process of classifying text according to a goal. Ex: Request for help, urgent problem, emergency situation and dangerous water level.
7. Identify Customer types like variety of products which are followed by customers. Ex: Lifetime value, Product Preferences, Privileged user, Value added customer.
8. Text Classification is used to categorize, understand, analyze sentiment and process unstructured text. 9. Auto Search and Auto Correct helps us to improve our writing vocabulary by checking for spelling mistake, detecting grammatical issues, or a statement structural error. Ex: Search Engines like Duck DuckGo, Google.
9. Translation Systems help to transform text in one language to a text in another language. Ex: Google Translate.
10. NLP Tool Kits in India and other countries for language understanding and generation like open sources, commercial and research oriented software's useful for societal benefits.

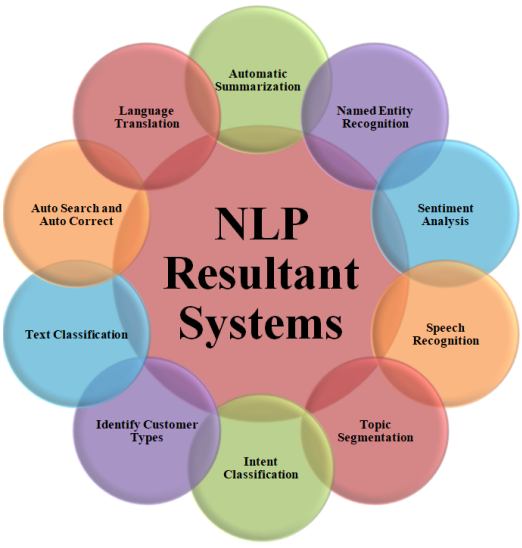

Fig 6. NLP Resultant Systems

3 Results and Discussion

Table 4. Comparison of Monolingual Translation Systems Made in India

Prototype Name	Technique	Target Language	Origin	Features incorporated
Anusaaraka	Direct	Bengali, Kannada, Marathi, Punjabi	1995, IIT- Kanpur, IIIT-Hyderabad	1. Based on rich linguistic rule 2. Bilingual dictionary, 3. Word by word translation
Hindi to Punjabi MTS	Direct	Hindi to Punjabi	2007 Punjab University, Patiala	
Mantra 1995	Transfer Based	English to Hindi, Gujarati, Telegu. Hindi to English, Bengali, Marathi	2004, CDAC, Pune	1. Uses Tree Adjoining Grammar Formalism 2. Multiple Bilingual dictionary 3. Grammatical rules 4. Intermediate source language representation
Matra	Transfer Based	English to Hindi	2004 CDAC, Pune	1. Human assisted translation project uses rule bases and heuristics. 2. multiple bilingual dictionary, 3. grammatical rules 4. Intermediate source language representation
Shakti	Transfer Based	English to Hindi, Marathi, Telegu	2004 IISc Bangalore and IIIT, Hyderabad	1. Working by combining rule based and statistical approach. 2. multiple bilingual dictionary, 3. grammatical rules 4. Intermediate source language representation
Anubaad	Transfer Based	English to Bengali	2005 CDAC, Kolkata	1. Hybrid system which uses n-gram approach for POS tagging at sentence level 2. multiple bilingual dictionary, 3. grammatical rules 4. Intermediate source language representation
Anglabharti	Interlingual	English to Hindi, Tamil	1991 IIT, Kanpur	1. Uses intermediate structure Pseudo Lingua for Indian Languages 2. based on Interlingua 3. contains transfer approach
AnglaHindi	Interlingual	English to Hindi	1991 IIT, Kanpur	1. Uses rule-bases, example-base and statistics to obtain translation for frequently encountered noun and verb phrases 2. based on Interlingua 3. contains transfer approach
UNL English Hindi MTS	Interlingual	English to Hindi	2001 IIT, Mumbai	1. Uses Universal Natural Language as interlingua, based on Interlingua 2. contains transfer approach
English Hindi MTS	Statistical Machine Translation	English to Hindi	2010 IIIT, Hyderabad	1. Combines Rule Based Machine Translation and phrase based SMT 2. Contain statistical methods and rules with huge language corpus
English Malyalam MTS	Statistical Machine Translation	English to Malyalam	2010 Cochin University ,Cochin	1. Uses SMT by using monolingual Malayalam corpus and a bilingual English/Malayalam corpus in the training phase 2. contain statistical methods and rules with huge language corpus
Vaasaanubaada	Example Based Machine Translation	Bengali Assamese	2002 Pondichery University	1. Pre-processing and post processing task, longer sentences fragmented at punctuation, backtracking for unmatched results. 2. based on previous examples from bilingual corpora
Anubharti	Example Based Machine Translation	Hindi-English	2004, IIT, Kanpur	1. Hybrid Example based system which combines pattern based and example based approach 2. Based on previous examples from bilingual corpora

Continued on next page

Table 4 continued

Shiva	Example Based Machine Translation	Hindi-English	2004, IISc Bangalore, IIIT Hyderabad	1. Uses linguistic rules and statistical approach to infer linguistic information. 2. based on previous examples from bilingual corpora
Samantar	Neural Machine Translation	Assamese, Bengali, Gujarati, Hindi, Kannada, Malayalam, Marathi, Oriya, Punjabi, Tamil, and Telugu.	As per the report "Machine Translation :The National Scenario" published in April 2021 by Dr. S. K. Srivaastava	1. Uses neural networks

All the above mentioned systems provide single source to single target language translation but our research has developed Cross Lingual Speech to Text Translation System that has addressed Trilingual and Quadlingual Speech to Text translation, output is as shown below:

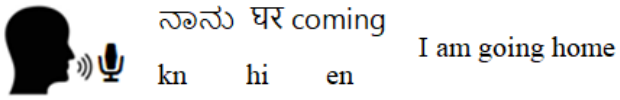

Fig 7. Trilingual Speech to Text where kn is Kannada, hi is Hindi and en is English

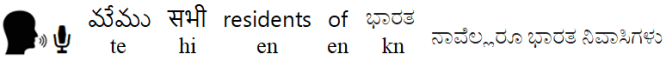

Fig 8. Quadlingual Speech to Text where te is telugu, hi is Hindi, en is English and kn is Kannada.

Processing of natural language enables computer to perform activities of reading, understanding and retrieving meaning from spoken languages. Some applicable fields where NLP can enhance benefits are in Health Care, Marketing, Hiring and Recruitment, Advertising, Commerce, E-Governance and Education.

4 Conclusion

The review has resulted in the comparative analysis of monolingual systems where there was a gap for cross lingual speech to text translation, produced a generic system that can accept oral sentences having two to four languages to be translated to a selected single target language. Ex Transforming public speeches spoken in multiple languages to a selected regional language helpful for a regional news editor. Existing systems are monolingual but the research has resulted in the development of two to four languages spoken within a single sentence for the languages Hindi, Telugu, Kannada, and English. The architecture can be enhanced in future to incorporate other regional languages of India.

References

1) Patel J, Makvana K, Shah P. Cross-lingual Information Retrieval: application and Challenges for Indian Languages. In: 2019 IEEE 5th International Conference for Convergence in Technology (I2CT). IEEE. 2019;p. 1–4. Available from: https://doi.org/10.1109/I2CT45611.2019.9033563.
2) Shah S, Sitaram S. Using Monolingual Speech Recognition for Spoken Term Detection in Code-switched Hindi-English Speech. In: 2019 International Conference on Data Mining Workshops (ICDMW). IEEE. 2019;p. 1–5. Available from: https://doi.org/10.1109/ICDMW48858.2019.9024755.
3) Nagarhalli TP, Mhatre S, Patil S, Patil P. The Review of Natural Language Processing Applications with Emphasis on Machine Learning Implementations. *2022 International Conference on Electronics and Renewable Systems (ICEARS)*. 2022;p. 1353–1358. Available from: https://doi.org/10.1109/ICEARS53579.2022.9752326.
4) Khurana D, Koli A, Khatter K, Singh S. Natural language processing: state of the art, current trends and challenges. *Multimedia Tools and Applications*. 2023;82(3):3713–3744. Available from: https://doi.org/10.1007/s11042-022-13428-4.
5) Banane M, Erraissi A. A comprehensive study of Natural Language processing techniques Based on Big Data. *2022 International Conference on Decision Aid Sciences and Applications (DASA)*. 2022;p. 1492–1497. Available from: https://doi.org/10.1109/DASA54658.2022.9765270.
6) Saini S, Sahula V. A Survey of Machine Translation Techniques and Systems for Indian Languages. In: 2015 IEEE International Conference on Computational Intelligence & Communication Technology. IEEE. 2015;p. 676–681. Available from: https://doi.org/10.1109/CICT.2015.123.

7) Vyas R, Joshi K, Sutar H, Nagarhalli TP. Real Time Machine Translation System for English to Indian language. In: 2020 6th International Conference on Advanced Computing and Communication Systems (ICACCS). IEEE. 2020;p. 838–842. Available from: https://doi.org/10.1109/ICACCS48705.2020.9074265.
8) Satpathy S, Mishra SP, Nayak AK. Analysis of Learning Approaches for Machine Translation Systems. In: 2019 International Conference on Applied Machine Learning (ICAML). IEEE. 2019;p. 160–164. Available from: https://doi.org/10.1109/ICAML48257.2019.00038.
9) Sohail AS, Sameen M, Ahmed Q. Formal Notations of Linguistic Analysis for Monetary Policy. In: 2019 International Conference on Green and Human Information Technology (ICGHIT). IEEE. 2019;p. 119–121. Available from: https://doi.org/10.1109/ICGHIT.2019.00035.
10) Hanif A, Zhang X, Wood S. A Survey on Explainable Artificial Intelligence Techniques and Challenges. In: 2021 IEEE 25th International Enterprise Distributed Object Computing Workshop (EDOCW). IEEE. 2021;p. 81–89. Available from: https://doi.org/10.1109/EDOCW52865.2021.00036.
11) Durga S, Nag R, Daniel E. Survey on Machine Learning and Deep Learning Algorithms used in Internet of Things (IoT) Healthcare. In: 2019 3rd International Conference on Computing Methodologies and Communication (ICCMC). IEEE. 2019;p. 1018–1022. Available from: https://doi.org/10.1109/ICCMC.2019.8819806.
12) Jiang K, Lu X. Natural Language Processing and Its Applications in Machine Translation: A Diachronic Review. In: 2020 IEEE 3rd International Conference of Safe Production and Informatization (IICSPI). IEEE. 2020;p. 210–214. Available from: https://doi.org/10.1109/IICSPI51290.2020.9332458.
13) Zhang X, Peng Y, Xu X. An Overview of Speech Recognition Technology. In: 2019 4th International Conference on Control, Robotics and Cybernetics (CRC). IEEE. 2019;p. 81–85. Available from: https://doi.org/10.1109/CRC.2019.00025.
14) Jaiswal RK, Dubey RK. Concatenative Text-to-Speech Synthesis System for Communication Recognition. In: 2021 5th International Conference on Electronics, Communication and Aerospace Technology (ICECA). IEEE. 2021;p. 867–872. Available from: https://doi.org/10.1109/ICECA52323.2021.9675855.
15) Patel P, Singh NC, Torppa M. Understanding the role of cross-language transfer of phonological awareness in emergent Hindi–English biliteracy acquisition. *Reading and Writing*. 2022. Available from: https://doi.org/10.1007/s11145-022-10253-x.
16) Kanan T, Sadaqa O, Aldajeh A, Alshwabka H, Al-Dolime W, Alzu'bi S, et al. A Review of Natural Language Processing and Machine Learning Tools Used to Analyze Arabic Social Media. In: 2019 IEEE Jordan International Joint Conference on Electrical Engineering and Information Technology (JEEIT). IEEE. 2019;p. 622–628. Available from: https://doi.org/10.1109/JEEIT.2019.8717369.
17) Azkan C, Moller F, Meisel L, Otto B. Service Dominant Logic Perspective on Data Ecosystems - A Case Study based Morphology. 2020. Available from: https://www.researchgate.net/publication/341205004_Service_Dominant_Logic_Perspective_on_Data_Ecosystems_-_A_Case_Study_based_Morphology.
18) Luthra S, Goel V, Kumar R. Exploration of Text Mining and Lexical Analysis Using a Novel Approach. In: 2019 International Conference on Intelligent Computing and Control Systems (ICCS). IEEE. 2019;p. 187–192. Available from: https://doi.org/10.1109/ICCS45141.2019.9065853.
19) Nandwani P, Verma R. A review on sentiment analysis and emotion detection from text. *Social Network Analysis and Mining*. 2021;11(1):81. Available from: https://doi.org/10.1007/s13278-021-00776-6.
20) Dwaraki A, Freedman R, Zilberstein S, Wolf T. Using Natural Language Constructs and Concepts to Aid Network Management. In: 2019 International Conference on Computing, Networking and Communications (ICNC). 2019;p. 802–808. Available from: https://doi.org/10.1109/ICCNC.2019.8685639.
21) Advaith V, Shivkumar A, Lakshmi BSS. Parts of Speech Tagging for Kannada and Hindi Languages using ML and DL models. *2022 IEEE International Conference on Electronics, Computing and Communication Technologies (CONECCT)*. 2022. Available from: https://doi.org/10.1109/CONECCT55679.2022.9865745.
22) De Luca G, Chen Y. Semantic Analysis of Concurrent Computing in Decentralized IoT and Robotics Applications. In: 2019 IEEE 14th International Symposium on Autonomous Decentralized System (ISADS). IEEE. 2019;p. 1–8. Available from: https://doi.org/10.1109/ISADS45777.2019.9155627.
23) Janda HK, Pawar A, Du S, Mago V. Syntactic, Semantic and Sentiment Analysis: The Joint Effect on Automated Essay Evaluation. *IEEE Access*. 2019;7:108486–108503. Available from: https://doi.org/10.1109/ACCESS.2019.2933354.
24) Liu M. Discourse Analysis Model for Literary Works Appreciation and Evaluation by Regressive Analysis. In: 2020 International Conference on Artificial Intelligence and Education (ICAIE). IEEE. 2020;p. 503–505. Available from: https://doi.org/10.1109/ICAIE50891.2020.00121.
25) Matyakubovna KM, Eshimovna QN. The Role Of Text Pragmatics In The Study Of Computer Linguistics. In: 2021 International Conference on Information Science and Communications Technologies (ICISCT). IEEE. 2021;p. 1–3. Available from: https://doi.org/10.1109/ICISCT52966.2021.9670319.