Breaking Barriers: Exploring the Diagnostic Potential of Speech Narratives in Hindi for Alzheimer's Disease

Kritesh Rauniyar[1], Shuvam Shiwakoti[1], Sweta Poudel[2], Surendrabikram Thapa[3], Usman Naseem[4], Mehwish Nasim[5,6]
[1]Delhi Technological University, India
[2]Kathmandu Engineering College, Tribhuvan University, Nepal
[3]Virginia Tech, USA [4]University of Sydney, Australia
[5]University of Western Australia [6]Flinders University, Australia

Abstract

Alzheimer's Disease (AD) is a neurodegenerative disorder that affects cognitive abilities and memory, especially in older adults. One of the challenges of AD is that it can be difficult to diagnose in its early stages. However, recent research has shown that changes in language, including speech decline and difficulty in processing information, can be important indicators of AD and may help with early detection. Hence, the speech narratives of the patients can be useful in diagnosing the early stages of Alzheimer's disease. While the previous works have presented the potential of using speech narratives to diagnose AD in high-resource languages, this work explores the possibility of using a low-resourced language, i.e., Hindi language, to diagnose AD. In this paper, we present a dataset specifically for analyzing AD in the Hindi language, along with experimental results using various state-of-the-art algorithms to assess the diagnostic potential of speech narratives in Hindi. Our analysis suggests that speech narratives in the Hindi language have the potential to aid in the diagnosis of AD. Our dataset and code are made publicly available at `https://github.com/rkritesh210/DementiaBankHindi`.

1 Introduction

Alzheimer's Disease (AD) is the most typical kind of dementia, characterized by a specific pattern of cognitive and functional deterioration brought on by aging that may eventually lead to death (Soria Lopez et al., 2019). This condition is mostly seen in adults over 60. Hampel et al. (2011) predicted that by 2040, more than 80 million people would be affected by dementia globally, up from an estimated 24 million in 2001.

In the early stages of AD, it is common to experience subtle language impairments such as problems with word finding and comprehension, the use of incorrect words, ambiguous referents, loss of verbal fluency, speaking too much or too loudly, repeating ideas, straying from the topic, which worsens in the moderate and severe stages (Meghanani et al., 2021). This shows that the temporal aspects of spontaneous speech are impacted by this disease. With the advancement of technology, machine learning approaches have been widely applied in the early diagnosis of AD utilizing neuroimaging scans such as Magnetic Resonance Imaging (MRI) and Positron Emission Tomography (PET) (Thapa et al., 2020b). However, this technique for identifying AD patients from Control Normal (CN) is limited to medical personnel (Thapa et al., 2020b). Szatloczki et al. (2015) showed that linguistic analysis could be used to identify AD more accurately than other types of cognitive testing. The temporal features of spontaneous speech, such as speech pace, frequency, and length of pauses, are sensitive detectors of the early stage of the illness, allowing an early and straightforward linguistic screening for AD. Thus, speech might be a straightforward but crucial characteristic that can be utilized to create potent AI models for AD diagnosis.

Through groundbreaking advancements in NLP, machines are now able to comprehend human language with unprecedented accuracy, unlocking a new realm of possibilities for data analysis and knowledge extraction (Naseem et al., 2021). Due to its ability to analyze language patterns and detect small alterations that may signal cognitive deterioration, NLP has grown in prominence in identifying AD (Thapa et al., 2022). NLP has been used to diagnose AD largely in high-resource languages like English. However, there is potential for this approach to be adapted to the low-resource languages in developing countries, including those spoken in India. Therefore the motivation behind this work is to promote the use of automated NLP-based tools for detecting AD in a low-resource language. Such a method will result in an accurate, quick, and economical AD diagnosis. Our contri-

Proceedings of the 5th Clinical Natural Language Processing Workshop, pages 24–30
July 14, 2023

butions are as follows:

- A new dataset for the low-resource language, *Hindi*, is created. The dataset (DementiaBankHindi) includes transcripts from 168 patients with Alzheimer's disease (AD) and 98 healthy control normal (CN) participants. The scientific community is expected to benefit from this original dataset.

- Using Hindi transcripts, an NLP-based methodology is given for the early diagnosis of AD patients. We have also explored how various machine translation-based systems perform in diagnosing AD using Hindi transcripts.

2 Related Works

In recent times, there has been significant research in the field of detecting AD using data-driven approaches (Adhikari et al., 2022). The rapid growth of NLP techniques has led to increased utilization of speech and linguistic features for detecting AD. Consequently, machine learning (ML) techniques are extensively employed in this domain. Classical ML-based methods require manual feature engineering. Such feature extraction methods can vary widely for different languages and can get outdated easily for evolving languages (Thapa et al., 2020a). Driven by the limitation of manual feature engineering of classical ML methods for such a diverse and complex task, in more recent times researchers have leveraged deep learning methods for the detection of AD. Karlekar et al. (2018) applied three neural models Convolutional Neural Network (CNN), Long Short-Term Memory (LSTM), and a stronger CNN-LSTM for detection of AD against CN on transcripts of Dementia bank's dataset of cookie theft picture description. With their initial approach, they were able to achieve an accuracy of 82.8%, 83.7%, and 84.9% for the CNN, LSTM, and CNN-LSTM architecture respectively. But when fed with POS-tagged data, their best-performing model CNN-LSTM model achieved an accuracy of 91.1%. Wang et al. (2021) took a multimodal approach where they also leveraged acoustic features for dementia detection. They used a CNN-attention Network and explored audio-based, text-based, and multimodal approaches using both audio and text-based features. They reported that a multimodal approach (C-Attention-Unified model) using Linguistic features and X-vector (acoustic) features performed best and could detect AD with an accuracy of 77.2% and an F1 score of 0.763.

Work has also been done for the detection of AD in languages other than English. Guo et al. (2020) proposed an autoencoder-based method to augment the Mandarin corpus (Liu et al., 2019) with a larger English dataset from DementiaBank and used a contrastive learning method based on BERT embeddings. With the data augmentation method, they achieved an accuracy of 81.6% in AD prediction. Rentoumi et al. (2017) used a dataset of transcripts of Boston cookie theft picture descriptions in the Greek language. The samples were obtained from native Greek Speakers diagnosed with Alzheimer's and normal controls. They extracted a total of 10 features based on Lexical and Syntactic measures and employed Naive Bayes (NB) and SVM with SMO (Sequential Minimal Optimization) classifiers.

Though extensive work has been done in the field of dementia detection for widely used languages such as English and Mandarin (Chinese) as per WHO's report, 58% of dementia patients worldwide belong to low-income, middle-income countries (Chen et al., 2019). This highlights the importance of building NLP-based diagnostic tools for lesser-known and low-resource languages. India as one of the lower-middle-income (Review, 2023) countries, is estimated to have dementia prevalence in 7.4% of the population, for ages 60 and above, resulting in about 8.8 million Indians older than 60 years living with dementia (Lee et al., 2023). Hindi is the most spoken language in India and to the best of our knowledge, no work has been done in the Hindi language for the detection of Alzheimer's disease. Thus, we believe our annotated dataset could serve as a stepping stone toward the detection of AD in the Hindi language and would contribute to further research in this field.

3 Dataset

DementiaBank's Pitt Corpus is utilized in this study. DementiaBank results from an experiment conducted by Becker et al. (1994) that contains audio recordings and transcripts for the Boston Cookie Theft picture description task. The task required the participants to describe a scene, as shown in Figure 1.

The transcription for the recordings was done

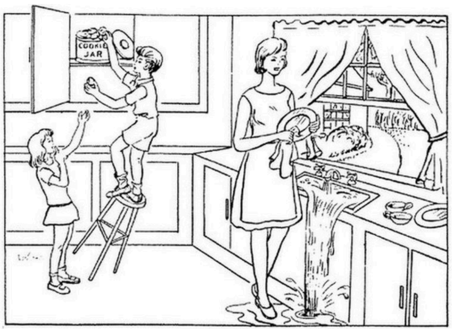

Figure 1: Boston cookie theft picture. This picture is widely used in the diagnosis of AD where patients are made to describe the scene.

manually using the CHAT (Codes for the Human Analysis of Transcripts) protocol (MacWhinney, 2017). The experiment consisted of 292 participants, with 194 having some sort of dementia. This resulted in 309 transcripts for the dementia category, as some participants had multiple recording sessions. This study deals with AD diagnosis. Thus, only the 255 transcripts from 168 AD patients and 244 transcripts from 98 CN participants were used in the study. Table 1 gives the demographics of the participants of the experiment.

Attributes	AD	CN
No. of subjects	168	98
Sex	113F / 55M	67F / 31M
Avg. Age	71.2	64.7
Avg. MMSE	19.9	29.1

Table 1: Demographics of the participants.

The transcriptions of recordings are accessible in English, which were translated into Hindi by three fluent Hindi speakers. We decided on manual translation since it is more likely to capture the social subtleties needed for the language.

We also created four more datasets using neural machine translation. BLEU (Bilingual Evaluation Understudy) score (Papineni et al., 2002) was used to compare machine translation against manual translation. mBART-50 (Tang et al., 2020), Google Translate, M2M-100 (Fan et al., 2020), and OPUS-MT (Tiedemann and Thottingal, 2020) were used as translation models. The BLEU score for each translation model is shown in Table 2. The table reflects that the text translated through the neural models is inaccurate. This shows that although deep learning models have been very prominent for various tasks, they still lack human-level performance that requires capturing niche social subtleties. Thus, manual translation was used in the study.

Translation models	BLEU-score
mBART-50	0.342
Google Translate	0.503
M2M-100	0.350
OPUS-MT	0.267

Table 2: BLEU score of different translation models

3.1 Exploratory Data Analysis

The top 10 important words from the entire dataset, along with their corresponding translation and TF-IDF (Term Frequency - Inverse Document Frequency) scores, are displayed in Table 3. TF-IDF scores are used to determine which words are more important to the document. A high TF-IDF score denotes the high significance of the word in the document.

Word	Translation	TF-IDF
कुकी	Cookie	0.3109
पानी	Water	0.2969
बर्तन	Utensils	0.2630
सिंक	Sink	0.2416
स्टूल	Stool	0.2414
माँ	Mother	0.2362
जार	Jar	0.2294
लड़की	Girl	0.2124
लड़का	Boy	0.2109
लगता	Seems	0.1942

Table 3: Top-10 most frequent words in the dataset with corresponding TF-IDF scores

3.2 Data Preprocessing

A crucial phase of data preprocessing in English is to convert all the words to uppercase or lowercase. Unlike English, case insensitivity is a feature of the Hindi language. As a result, no such modification is necessary. The punctuation marks, such as commas, semicolons, etc. that do not contribute any substantive significance to the content are eliminated in this study. Eliminating stop words from classification jobs while using NLP is another common practice that often enhances the model's performance. However, stop words like “and", “therefore," and others were frequently repeated by AD patients so, for this reason, stop words were not removed as they maintain the language traits of AD people (Khodabakhsh et al., 2014; Adhikari et al., 2021). Furthermore, Khodabakhsh et al. (2014) also suggested that pause words such as ‘um,’ ‘uh,’ and ‘ah’ were more frequently used by AD patients; as a result, they were not removed in the preprocessing phase and were translated as is.

Model	Manual Translation			OPUS-MT			M2M-100			Google Translate			mBART-50		
	Acc↑	**MMAE↓**	**F1$_{macro}$↑**	**Acc↑**	**MMAE↓**	**F1$_{macro}$↑**	**Acc↑**	**MMAE↓**	**F1$_{macro}$↑**	**Acc↑**	**MMAE↓**	**F1$_{macro}$↑**	**Acc↑**	**MMAE↓**	**F1$_{macro}$↑**
RF	0.727	0.287	0.702	0.672	0.327	0.662	0.727	0.267	0.717	0.696	0.304	0.690	0.654	0.343	0.639
NB	0.732	0.237	0.718	0.593	0.386	0.527	0.684	0.277	0.651	0.660	0.310	0.624	0.551	0.460	0.456
LR	0.726	0.269	0.721	0.666	0.332	0.653	0.690	0.307	0.680	0.709	0.292	0.704	0.654	0.343	0.639
SVC	0.732	0.267	0.731	0.690	0.309	0.683	0.690	0.309	0.683	0.703	0.298	0.697	0.648	0.353	0.637
XGB	0.709	0.291	0.702	0.666	0.335	0.659	0.690	0.307	0.680	0.690	0.310	0.684	0.690	0.309	0.683
ADA	0.715	0.285	0.708	0.678	0.321	0.669	0.684	0.317	0.682	0.721	0.281	0.718	0.696	0.302	0.688
LSTM	0.836	0.146	0.836	0.727	0.256	0.726	0.727	0.274	0.723	0.781	0.212	0.781	0.727	0.270	0.725
Bi-LSTM	0.872	0.127	0.869	0.800	0.200	0.797	0.745	0.272	0.730	0.763	0.242	0.758	0.745	0.263	0.738
BERT (Hindi)	0.842	0.125	0.840	0.820	0.199	0.807	0.717	0.276	0.717	0.700	0.300	0.694	0.740	0.260	0.734
ALBERT	0.829	0.169	0.828	0.714	0.283	0.711	0.794	0.208	0.791	0.800	0.200	0.797	0.743	0.267	0.735
XLM-RoBERTa	**0.880**	**0.120**	**0.879**	0.800	0.200	0.798	0.820	0.180	0.819	0.760	0.240	0.753	0.820	0.180	0.816
RoBERTa	0.860	0.140	0.859	0.740	0.260	0.739	0.780	0.220	0.779	0.720	0.280	0.719	0.780	0.220	0.779

Table 4: Baseline results with different algorithms for multiple translation models. The DementiaBank was translated manually and also using various machine translation algorithms.

4 Experimental Results and Discussion

We developed benchmarks using a variety of approaches, including traditional machine learning methods, deep learning, and transformer-based models. To evaluate the results for each baseline, we used accuracy, macro-mean-squared-error (MMAE), and F1-score (macro) as assessment metrics. Accuracy is a trivial evaluation metric in classification tasks. However, we use macro MAE and macro F1 to account for imbalanced datasets. Using macro MAE and macro F1 score gives equal weight to each class, regardless of size.

4.1 Benchmark Algorithms

We performed benchmarks with various machine learning and deep learning algorithms.

Machine Learning Algorithms: We employed Random Forest (RF) (Svetnik et al., 2003), Naive Bayes (NB) (Rish et al., 2001), Logistic Regression (LR), Support Vector Classification (SVC) (Hsu et al., 2003), XGBoost (XGB) (Chen et al., 2015), and AdaBoost (ADA) (Schapire, 2013) as our classical machine learning techniques. The vectorization of the corpus was done using the TF-IDF vectorizer.

Deep Learning Algorithms: LSTM (Hochreiter and Schmidhuber, 1997) and bidirectional LSTM were used as deep learning algorithms. Word embedding was done using the TensorFlow tokenizer. For transformer-based models, we used FillMask models for RoBERTa and ALBERT. We also implemented BERT (Doiron, 2023), ALBERT (Joshi, 2022), XML-RoBERTa (Pandya et al., 2021), and RoBERTa (Huang et al., 2021) for the benchmark evaluations.

4.2 Results and Analysis

The comprehensive classification results for diagnosing AD and CN are shown in Table 4. With an F1-score and accuracy of 0.879 and 0.880, respectively, the XML-Roberta model performed the best among all algorithms. SVC and NB beat the other ML methods with an accuracy of 0.732. DL and ML models did not perform as good as transformer-based models. The requirement for more sophisticated and reliable algorithms for text identification is highlighted by the model's substantially lower F1 score of ML models compared to transformer-based models. Similarly, we run the benchmark evaluations for the translations using neural machine translation. The translation made using OPUS-MT showed a f1-score of 0.807. This shows that the machine translation-based transcripts can also capture the nuances in speech which are necessary to delineate AD patients from CN groups. The benchmark evaluations done on manual translation have shown remarkable performance compared to the translations done by machine translation algorithms. This shows that automated translations may not account for social subtleties.

5 Conclusion

Our work presents a novel dataset in the Hindi language that classifies the speech of AD patients against CN individuals. AD cannot be cured, so its detection and management become crucial. Speech impairment is one of the most common symptoms of AD. Hence, we have created this dataset that has the potential to significantly aid in the development of automated, speedy, and cost-effective systems for detecting AD. We also performed benchmarks on the created dataset and achieved the highest accuracy of 0.880 and an F1 score of 0.879 with the XLM-Roberta model. Such considerable benchmark results encourage further research in the field by extending the dataset and creating more sophisticated and domain-specific models. Results show there is room for improvement in constructing superior models.

Limitations

Some potential limitations of this work include a relatively small sample size, which may limit the generalizability of the results. Hindi is spoken differently across India, hence the translations made by the three translators may not be representative. Study did not examine the potential impact of regional dialects or variations in Hindi, on the accuracy of the diagnosis. Finally, the study focused solely on the use of speech narratives and did not explore the other types of data, e.g., imaging or genetic data, which could be important for the diagnosis of AD.

References

Surabhi Adhikari, Surendrabikram Thapa, Usman Naseem, Priyanka Singh, Huan Huo, Gnana Bharathy, and Mukesh Prasad. 2022. Exploiting linguistic information from Nepali transcripts for early detection of Alzheimer's Disease using natural language processing and machine learning techniques. *International Journal of Human-Computer Studies*, 160:102761.

Surabhi Adhikari, Surendrabikram Thapa, Priyanka Singh, Huan Huo, Gnana Bharathy, and Mukesh Prasad. 2021. A comparative study of machine learning and NLP techniques for uses of stop words by patients in diagnosis of Alzheimer's Disease. In *2021 International Joint Conference on Neural Networks (IJCNN)*, pages 1–8. IEEE.

James T Becker, François Boiler, Oscar L Lopez, Judith Saxton, and Karen L McGonigle. 1994. The natural history of Alzheimer's Disease: description of study cohort and accuracy of diagnosis. *Archives of Neurology*, 51(6):585–594.

Jia-Hung Chen, Chien-Tai Hong, Dean Wu, Wen-Chou Chi, Chia-Feng Yen, Hua-Fang Liao, Lung Chan, and Tsan-Hon Liou. 2019. Dementia-related functional disability in moderate to advanced parkinsons disease: assessment using the World Health Organization disability assessment schedule 2.0. *International Journal of Environmental Research and Public Health*, 16(12):2230.

Tianqi Chen, Tong He, Michael Benesty, Vadim Khotilovich, Yuan Tang, Hyunsu Cho, Kailong Chen, Rory Mitchell, Ignacio Cano, Tianyi Zhou, et al. 2015. XGBoost: extreme gradient boosting. *R package version 0.4-2*, 1(4):1–4.

N. Doiron. 2023. Hindi-Bert. `https://huggingface.co/monsoon-nlp/hindi-bert`. [Online; accessed 12-Apr-2023].

Angela Fan, Shruti Bhosale, Holger Schwenk, Zhiyi Ma, Ahmed El-Kishky, Siddharth Goyal, Mandeep Baines, Onur Celebi, Guillaume Wenzek, Vishrav Chaudhary, Naman Goyal, Tom Birch, Vitaliy Liptchinsky, Sergey Edunov, Edouard Grave, Michael Auli, and Armand Joulin. 2020. Beyond English-Centric Multilingual Machine Translation.

Zhiqiang Guo, Zhaoci Liu, Zhenhua Ling, Shijin Wang, Lingjing Jin, and Yunxia Li. 2020. Text classification by contrastive learning and cross-lingual data augmentation for Alzheimers Disease detection. In *Proceedings of the 28th International Conference on Computational Linguistics*, pages 6161–6171.

Harald Hampel, David Prvulovic, Stefan Teipel, Frank Jessen, Christian Luckhaus, Lutz Frölich, Matthias W. Riepe, Richard Dodel, Thomas Leyhe, Lars Bertram, Wolfgang Hoffmann, and Frank Faltraco. 2011. The future of Alzheimer's Disease: The next 10 years. *Progress in Neurobiology*, 95(4):718–728. Biological Markers for Neurodegenerative Diseases.

Sepp Hochreiter and Jürgen Schmidhuber. 1997. Long short-term memory. *Neural Computation*, 9(8):1735–1780.

Chih-Wei Hsu, Chih-Chung Chang, Chih-Jen Lin, et al. 2003. A practical guide to support vector classification.

Xixuan Huang, Nankai Lin, Kexin Li, Lianxi Wang, and Suifu Gan. 2021. Hinplms: Pre-trained language models for hindi. In *2021 International Conference on Asian Language Processing (IALP)*, pages 241–246. IEEE.

Raviraj Joshi. 2022. L3Cube-HindBERT and DevBERT: Pre-Trained BERT Transformer models for Devanagari based Hindi and Marathi Languages. *arXiv preprint arXiv:2211.11418*.

Sweta Karlekar, Tong Niu, and Mohit Bansal. 2018. Detecting linguistic characteristics of Alzheimer's Dementia by interpreting neural models. *arXiv preprint arXiv:1804.06440*.

Ali Khodabakhsh, Serhan Kuxuolu, and Cenk Demirolu. 2014. Natural language features for detection of Alzheimer's Disease in conversational speech. In *IEEE-EMBS International Conference on Biomedical and Health Informatics (BHI)*, pages 581–584.

Jinkook Lee, Erik Meijer, Kenneth M Langa, Mary Ganguli, Mathew Varghese, Joyita Banerjee, Pranali Khobragade, Marco Angrisani, Ravi Kurup, Sankha Shubhra Chakrabarti, et al. 2023. Prevalence of dementia in India: National and state estimates from a nationwide study. *Alzheimer's & Dementia*.

Zhaoci Liu, Zhiqiang Guo, Zhenhua Ling, Shijin Wang, Lingjing Jin, and Yunxia Li. 2019. Dementia detection by analyzing spontaneous Mandarin speech. In *2019 Asia-Pacific Signal and Information Processing Association Annual Summit and Conference (APSIPA ASC)*, pages 289–296. IEEE.

Brian MacWhinney. 2017. Tools for analyzing talk part 1: The chat transcription format. *Carnegie.[Google Scholar]*, 16.

Amit Meghanani, CS Anoop, and Angarai Ganesan Ramakrishnan. 2021. Recognition of Alzheimers Dementia from the transcriptions of spontaneous speech using fastText and CNN models. *Frontiers in Computer Science*, 3:624558.

Usman Naseem, Imran Razzak, Shah Khalid Khan, and Mukesh Prasad. 2021. A comprehensive survey on word representation models: From classical to state-of-the-art word representation language models. *Transactions on Asian and Low-Resource Language Information Processing*, 20(5):1–35.

Hariom A. Pandya, Bhavik Ardeshna, and Dr. Brijesh S. Bhatt. 2021. Cascading adaptors to leverage english data to improve performance of question answering for low-resource languages.

Kishore Papineni, Salim Roukos, Todd Ward, and Wei-Jing Zhu. 2002. BLEU: a method for automatic evaluation of machine translation. In *Proceedings of the 40th annual meeting of the Association for Computational Linguistics*, pages 311–318.

Vassiliki Rentoumi, George Paliouras, Eva Danasi, Dimitra Arfani, Katerina Fragkopoulou, Spyridoula Varlokosta, and Spyros Papadatos. 2017. Automatic detection of linguistic indicators as a means of early detection of Alzheimer's Disease and of related dementias: A computational linguistics analysis. In *2017 8th IEEE International Conference on Cognitive Infocommunications (CogInfoCom)*, pages 000033–000038. IEEE.

World Population Review. 2023. Middle-Income Countries 2023. `https://worldpopulationreview.com/country-rankings/middle-income-countries`. [Online; accessed 12-Apr-2023].

Irina Rish et al. 2001. An empirical study of the naive bayes classifier. In *IJCAI 2001 workshop on empirical methods in artificial intelligence*, volume 3, pages 41–46.

Robert E Schapire. 2013. Explaining Adaboost. *Empirical Inference: Festschrift in Honor of Vladimir N. Vapnik*, pages 37–52.

Jose A. Soria Lopez, Hector M. González, and Gabriel C. Léger. 2019. Chapter 13 - Alzheimer's Disease. In Steven T. Dekosky and Sanjay Asthana, editors, *Geriatric Neurology*, volume 167 of *Handbook of Clinical Neurology*, pages 231–255. Elsevier.

Vladimir Svetnik, Andy Liaw, Christopher Tong, J Christopher Culberson, Robert P Sheridan, and Bradley P Feuston. 2003. Random Forest: a classification and regression tool for compound classification and qsar modeling. *Journal of Chemical Information and Computer Sciences*, 43(6):1947–1958.

Greta Szatloczki, Ildiko Hoffmann, Veronika Vincze, Janos Kalman, and Magdolna Pakaski. 2015. Speaking in Alzheimers Disease, is that an early sign? importance of changes in language abilities in Alzheimers Disease. *Frontiers in aging neuroscience*, 7:195.

Yuqing Tang, Chau Tran, Xian Li, Peng-Jen Chen, Naman Goyal, Vishrav Chaudhary, Jiatao Gu, and Angela Fan. 2020. Multilingual translation with extensible multilingual pretraining and finetuning.

Surendrabikram Thapa, Surabhi Adhikari, Usman Naseem, Priyanka Singh, Gnana Bharathy, and Mukesh Prasad. 2020a. Detecting Alzheimers Disease by exploiting linguistic information from nepali transcript. In *Neural Information Processing: 27th International Conference, ICONIP 2020, Bangkok, Thailand, November 18–22, 2020, Proceedings, Part IV 27*, pages 176–184. Springer.

Surendrabikram Thapa, Awishkar Ghimire, Surabhi Adhikari, Akash Kumar Bhoi, and Paolo Barsocchi. 2022. Cognitive internet of things (IoT) and computational intelligence for mental well-being. In *Cognitive and Soft Computing Techniques for the Analysis of Healthcare Data*, pages 59–77. Elsevier.

Surendrabikram Thapa, Priyanka Singh, Deepak Kumar Jain, Neha Bharill, Akshansh Gupta, and Mukesh Prasad. 2020b. Data-driven approach based on feature selection technique for early diagnosis of Alzheimers Disease. In *2020 International Joint Conference on Neural Networks (IJCNN)*, pages 1–8. IEEE.

Jörg Tiedemann and Santhosh Thottingal. 2020. OPUS-MT Building open translation services for the World. In *Proceedings of the 22nd Annual Conferenec of the European Association for Machine Translation (EAMT)*, Lisbon, Portugal.

Ning Wang, Yupeng Cao, Shuai Hao, Zongru Shao, and KP Subbalakshmi. 2021. Modular multi-modal attention network for Alzheimer's Disease detection using patient audio and language data. In *Interspeech*, pages 3835–3839.

A Appendix

The example of translated transcripts for control normal (CN) participants and AD patients are shown in Table 5 and Table 6 respectively. The translations show that google translate was able to generate more similar translations as manual translation than other translation algorithms. As the evaluation of translation can be subjective, we use BLEU scores as mentioned in Table 2.

	CHAT ID: 002-2.cha
Transcript from DementiaBank	a boy and a girl are in the kitchen with their mother. and the little boy is getting a cookie for the little girl but he's on a stool it's about to fall down. the mother is washing dishes. and she's obviously thinking of something else because the water's pouring out over the sink. she's finished with some of the dishes. it seems to be summertime because there are bushes. the window is open. there seems to be some kind of a breeze because the curtains on the sill there blowing. must be fairly hot. the mother's in a sleeveless dress. the children are in short sleeve tops and have sandals. the little boy has tennis shoes on. the mother obviously is unaware of what the children are doing. she's going to be aware of it very shortly. I mean how much more do you want hm okay.
Manual Translation	एक लड़का और एक लड़की रसोई में अपनी माँ के साथ हैं। और छोटा लड़का छोटी लड़की के लिए एक कुकी ले रहा है लेकिन वह एक स्टूल पर है जो नीचे गिरने वाला है। माँ बर्तन धो रही है। और वह स्पष्ट रूप से कुछ और सोच रही है क्योंकि सिंक के ऊपर से पानी बह रहा है। उसने कुछ बर्तन समाप्त कर लिए हैं। ऐसा लगता है कि गर्मियों का समय है क्योंकि झाड़ियाँ हैं। खिड़की खुली है। ऐसा लगता है कि किसी तरह की हवा चल रही है क्योंकि वहाँ की सिल पर लगे पर्दे उड़ रहे हैं। काफी गर्म होना चाहिए। माँ एक बिना आस्तीन की पोशाक में है।
Google TM	एक लड़का और एक लड़की अपनी माँ के साथ रसोई में हैं। और छोटे लड़के को छोटी लड़की के लिए एक कुकी मिल रही है, लेकिन वह एक स्टूल पर है जो नीचे गिरने वाला है। माँ व्यंजन धो रही है। और वह स्पष्ट रूप से कुछ और के बारे में सोच रही है क्योंकि पानी सिंक पर बाहर निकल रहा है। वह कुछ व्यंजनों के साथ समाप्त हो गई है। यह गर्मियों में लगता है क्योंकि वहाँ झाड़ियों हैं। खिड़की खुली है । वहाँ कुछ प्रकार की हवा लगती है क्योंकि वहाँ पर पर्दे बहते हैं। काफी गर्म होना चाहिए। एक स्लीवलेस ड्रेस में माँ की। बच्चे छोटी आस्तीन में होते हैं और सैंडल होते हैं। छोटे लड़के के पास टेनिस के जूते हैं। मां स्पष्ट रूप से इस बात से अनजान है कि बच्चे क्या कर रहे हैं। वह बहुत जल्द ही इसके बारे में जागरूक हो रही है। Imean कितना अधिक आप hm ठीक चाहते हैं।
mBART-50 TM	एक लड़का और एक लड़की मां के साथ रसोईघर में हैं। और छोटा लड़का छोटे लड़की के लिए एक कुकी ले रहा है लेकिन वह एक कुर्सी पर है यह नीचे गिरने के लिए जा रहा है। मां बर्तन धो रही है। और वह स्पष्ट रूप से कुछ और के बारे में सोच रहा है क्योंकि पानी Sink के ऊपर बह रहा है। वह कुछ बर्तनों के साथ समाप्त हो गया है। लगता है यह गर्मी का समय है क्योंकि वहाँ झाड़ियाँ हैं। विंडो खुला है। ऐसा लगता है कि हवा कुछ तरह से चल रही है क्योंकि सील पर परदे उड़ रहे हैं। काफी गर्म होना चाहिए। मां बेहिश्त पोशाक में है। बच्चों को कमज़ोर सूती के टोप पहने जाते हैं और सैंडल पहने जाते हैं। छोटे लड़के पर टेनिस जूते हैं। मां स्पष्ट रूप से नहीं जानती कि बच्चे क्या कर रहे हैं। वह बहुत जल्दी यह जानने के लिए जा रहा है। imean कितना अधिक आप चाहते हैं hm ठीक है।
M2M-100 TM	एक लड़का और एक लड़की अपनी मां के साथ रसोई में हैं। और छोटा लड़का छोटी लड़की के लिए एक कुकी मिल रहा है लेकिन वह एक मल पर है यह गिरने के लिए तैयार है। मां ने खाना धोया है। और वह स्पष्ट रूप से कुछ और के बारे में सोच रहा है क्योंकि पानी स्नान पर बह रहा है। वह कुछ डिश के साथ खत्म हो गया है। ऐसा लगता है कि यह गर्मियों का समय है क्योंकि वहाँ झाड़ियाँ हैं। खिड़की खुली है। ऐसा लगता है कि वहाँ कुछ तरह का एक बर्फ है क्योंकि वहां ब्लेड पर पर्दे फेंक रहे हैं। काफी गर्म होना चाहिए। मम्मी के कपड़े बेकार हैं। बच्चों के पास छोटी–छोटी बूंदें हैं और सैंडल हैं। बच्चे के पास टेनिस जूते हैं। माता–पिता स्पष्ट रूप से यह नहीं जानते हैं कि बच्चे क्या कर रहे हैं। वह जल्द ही इसके बारे में जान लेगी। इमेन कितना अधिक आप चाहते हैं ठीक है।
OPUS-MT TM	एक लड़का और लड़की रसोई में अपनी माँ के साथ हैं. और छोटा लड़का छोटी लड़की के लिए एक कुकी हो रहा है...... लेकिन वह यह नीचे गिर करने के बारे में है. माँ बरतन धो रही है. और वह स्पष्ट रूप से कुछ और के बारे में सोच रहा है क्योंकि पानी डूबने पर बाहर बहा रहा है. वह कुछ व्यंजनों के साथ समाप्त हो गया है. ऐसा लगता है कि गर्मियों का समय है क्योंकि वहाँ झाड़ी हैं. विंडो खुला है. ऐसा लगता है कि एक प्रकार की हवा है क्योंकि वहाँ के परदे वहाँ बाढ़ की बाढ़ से गुज़र रहे हैं । काफी गर्म होना चाहिए. माँ एक बेकार पोशाक में है. बच्चे छोटी सीटों पर हैं और जूते हैं. छोटे लड़के के जूते. स्पष्ट रूप से माँ को पता नहीं कि बच्चे क्या कर रहे हैं । वह इसके बारे में बहुत जल्द पता करने जा रहा है. Dmm ठीक है चाहते हैं कितना अधिक.

Table 5: Example of the original and translated transcripts of control normal (CN) participants

	CHAT ID: 051-1.cha
Transcript from DementiaBank	the boy reaching for the cookies is gonna fall down on the um what uh shes saying shho i think shho but give me a cookie too um mother is wiping the dish sink is running over the uh im to tell everything i see all the action yeah splashing the water gram well xxx and the girl saying shho give me a cookie too thats all i see
Manual Translation	लड़का कुकीज़ के लिए पहुँच रहा है उम पर गिरने वाला है वह क्या कह रही है उह क्या कह रही है शो मुझे लगता है शो लेकिन मुझे एक कुकी भी दे उम चने के पानी के छींटे अच्छी तरह से और लड़की कह रही है कि मुझे एक कुकी भी दो वह सब मैं देख रहा हूँ।
Google TM	कुकीज़ के लिए पहुंचने वाला लड़का उम पर गिरने वाला है जो उह शेस कह रहा है कि मुझे लगता है कि मुझे लगता है कि मुझे एक कुकी भी है, उम माँ डिश सिंक को पोंछ रही है, उह इम के ऊपर चल रही है सब कुछ बताने के लिए मैं सभी एक्शन को देख रहा हूं पानी चने की अच्छी तरह से xxx और शॉ कहती है कि मुझे एक कुकी भी है जो मैं देख रहा हूं।
mBART-50 TM	लड़का कुकी के लिए पहुंचने के लिए नीचे गिरने के लिए है um क्या uh वह कह रही है shho मैं सोचता हूँ shho लेकिन मुझे एक कुकी भी दे रहा हूँ um माँ बर्तन धो रहा है धोने के लिए कूल्हे पर चल रहा है uh im सब कुछ बताने के लिए मैं देख रहा हूँ सभी क्रिया हाँ छिड़का हुआ पानी ग्राम अच्छी तरह से xxx और लड़की कह रही है shho मुझे एक कुकी भी दे रहा है कि सब मैं देख रहा हूँ।
M2M-100 TM	कुकीज़ के लिए पहुंचने वाला लड़का ओम पर गिरने जा रहा है कि ओह क्या कहता है shho मैं सोचता हूं shho लेकिन मुझे एक कुकी दें ओम माँ डिश स्नान पर चल रही है ओह im सब कुछ बताने के लिए मैं देखता हूं सभी कार्रवाई हाँ पानी gram अच्छी xxx और लड़की कहती है shho मुझे एक कुकी देना भी यही सब मैं देखता हूं।
OPUS-MT TM	कुकीज़ के लिए प्रवेश करने वाला लड़का उमर पर गिर रहा है क्या उह वह कह रहा है hhhh मैं लगता है पर नीचे गिर रहा है लेकिन मुझे एक कुकी भी एक उमर पानी सिंक बंद कर रहा है सब कुछ बताने के लिए उह, मैं सब कुछ अच्छी तरह से देखने के लिए जा रहा हूँ और लड़की कहते हैं कि मुझे एक कुकी है कि मैं भी देख रहा हूँ कि मैं देख रहा हूँ

Table 6: Example of the original and translated transcripts of patients with AD