TamilATIS: Dataset for Task-Oriented Dialog in Tamil

Ramaneswaran S and **Sanchit Vijay** and **Kathiravan Srinivasan**
Vellore Institute Of Technology, Vellore
kathiravan.srinivasan@vit.ac.in
{s.ramaneswaran2018, sanchit.vijay2018}@vitstudent.ac.in

Abstract

Task-Oriented Dialogue (TOD) systems allow users to accomplish tasks by giving directions to the system using natural language utterances. With the widespread adoption of conversational agents and chat platforms, TOD has become mainstream in NLP research today. However, developing TOD systems require massive amounts of data, and there has been limited work done for TOD in low-resource languages like Tamil. Towards this objective, we introduce TamilATIS - a TOD dataset for Tamil which contains 4874 utterances. We present a detailed account of the entire data collection and data annotation process. We train state-of-the-art NLU models and report their performances. The Joint BERT model with XLM-Roberta as utterance encoder achieved the highest score with an intent accuracy of 96.26% and slot F1 of 94.01%.

1 Introduction

Task-oriented dialog (TOD) systems enable a user to use natural language directions to complete specific tasks. Recently, such systems have been successfully deployed in smart applications such as Amazon's Echo and Spotify's Car Thing.

There are several components that are critical to the performance of a TOD system. These components are Natural Language Understanding, Dialogue State Tracking (DST), and Response Selection. In this work, we focus on the NLU component. NLU aims to semantically parse an input utterance and typically has two tasks: intent classification and slot filling (refer Table 1 for example).

Intent classification deals with identifying the underlying motivation or the goal of the user query. It is modelled as a sequence classification task. The simplicity and conciseness of the utterances, paired with the necessity to scale to multiple domains, pose bottlenecks to intent detection. Slot filling deals with identifying entities present in an utterance that corresponds to certain slots in the user query. This is typically cast as a token classification or span identification task. Slot filling is a challenging task in NLU. The model needs to adapt to unseen domains and identify entities that it has not encountered in training before.

Find	morning	flights	to	Chennai
O	B-period	O	O	B-fromcity
Intent: find-flight				

Table 1: Example of user utterance with their corresponding BIO annotation and intent.

Intent classification and slot filling have been widely researched for the English language. The approaches presented in these works achieve excellent performance due to the availability of large amounts of high-quality and human-annotated datasets. However, such performance has not been achieved for several low-resource languages due to a lack of data. Developing TOD datasets for low-resource languages is essential to the proliferation of NLP technologies in these communities and contributes towards inclusivity and diversity of language resources.

To facilitate this, we present a dataset named TamilATIS, which contains 4874 utterances in Tamil and their corresponding slots and intent annotations. The following are some of the main contributions of this work:

- We present a TOD dataset for Tamil - TamilATIS with 4874 utterances.
- We perform initial experiments with state-of-the art slot filling and intent detection models to establish the baselines.

The full dataset and the source code of the baseline models are available at https://github.com/ramaneswaran/tamil_atis

Proceedings of the Second Workshop on Speech and Language Technologies for Dravidian Languages, pages 25 - 32

	Utterance	Intent
Ex. 1	(Morning flights from Vadodara to Vijayawada on April 20) ஏப்ரல் **20** அன்று வதோதராவில் இருந்து விஜயவாடாவிற்கு காலை விமானங்கள்	atis_flight
Slots	B.MN B.DN O B.FCN O B.TCN B.POD	
Ex. 2	(What is the fare for a taxi to Agartala?) அகர்தலாவிற்கு ஒரு டாக்ஸிக்கு என்ன கட்டணம்	atis_ground_fare
Slots	B.CN O B.TPT O O	
Ex. 3	(What is the seat capacity of the 733?) **733** இன் இருக்கை திறன் என்ன?	atis_capacity
Slots	B.AC O O O O	
Ex. 4	(I would like to take a flight from Kochi to Tiruchirappalli on Saturday morning in Vistara) விஸ்தாராவில் சனிக்கிழமை காலை கொச்சியிலிருந்து திருச்சிராப்பள்ளிக்கு விமானம் செல்ல விரும்புகிறேன்	atis_flight
Slots	B.AN B.DDN B.DPD B.FCN B.TCN O O O	

Figure 1: Examples of utterances and their annotations from the Tamil ATIS dataset. The words in blue are the slot values

2 Related Work

Intent classification and slot filling are two key challenges modelled separately or jointly for Natural Language Understanding (NLU). Joint modelling approaches have attained state-of-the-art performance, and have demonstrated that there exists a significant correlation between the two tasks. Prior works have implemented CNN-CRF (Xu and Sarikaya, 2013), RecNN (Guo et al., 2014), joint RNN-LSTM (Hakkani-Tür et al., 2016), attention-based BiRNN (Liu and Lane, 2016), and slot-gated attention-based model (Goo et al., 2018) and more recently have used BERT (Chen et al., 2019a) and BiLSTM based (Haihong et al., 2019) approaches.

Intent classification and slot filling functions are core modules for NLU in Task-Oriented Dialogue (TOD) systems (Chen et al., 2016; Takanobu et al., 2019; Kummerfeld et al., 2019; Liang et al., 2019; Campagna et al., 2020; Hosseini-Asl et al., 2020; Ham et al., 2020). Since these tasks are characterized as sequence classification and token tagging tasks, sentence encoder models have been utilized to solve them. The two extensively utilized large scale datasets in English (high resource language) for this purpose in NLU are: ATIS (Price, 1990), which features audio recordings of individuals booking flight reservations, and SNIPS (Coucke et al., 2018), which is gathered from Snips' personal voice assistant. Dao et al. (2021) introduced a low-resource language dataset in Vietnamese. Apart from these monolingual English corpora, Schuster et al. (2019) presented a new dataset of 57k annotated utterances in English (43k), including low-resource Spanish (8.6k), and Thai (5k), spanning the topics of weather, alarm, and reminder.

In Indian languages, there have been works to synthesize training data by using Google Translate and proposed CNN+LSTM based architecture (Gupta et al., 2020). Malviya et al. (2021) released a Hindi Dialogue Restaurant Search (HDRS) corpus consisting of 1.4k human-to-human typed dialogues collected using the Wizard-of-Oz paradigm and compared various state-of-the-art DST models. Small sized datasets were constructed manually and used with Indic and Code-Switched TOD systems (Jayarao and Srivastava, 2018). (Kanakagiri and Radhakrishnan, 2021) used mBERT based semantic tracking to associate the slot tokens to the respective tokens in the utterance and employed Google Translate API, morphological characteristics and semantics based heuristic slot aligner to publish a dataset for dravidian languages like Kannada and Tamil.

3 Tamil ATIS

The earliest Old Tamil documents are small inscriptions in Adichanallur dating from 905 BC to 696

Name	Language	Intent	Slot	Description
HDRS 2021	hi	No	Yes	H2H dialogue corpus for restaurant domain in Hindi
TaskMaster-1 2018	hi,mr,bn,gj	Yes	No	Google's Taskmaster-1 dataset for intent classification automatically translated to 4 indian languages
CoMTIC 2021	hi-en	Yes	No	Hindi-english code-mixed dataset for intent classification
Codemix-DSTC2 2018	hi,bn,gj,ta	Yes	Yes	DSTC2 dataset manually converted to codemix and slot labels manually annotated
Codemix-SNIPS 2020	hi-en	Yes	No	SNIPs dataset manually converted to hindi-english code-mixed form.
TOD-Dravidian 2021	kn, ta	Yes	Yes	MTOD dataset automatically translated and slots automatically annotated
Ours	ta	Yes	Yes	ATIS dataset automatically translated to Tamil and slot labels manually annotated

Table 2: Comparison of various datasets for TOD in Indian languages.

Figure 2: Translation results from Google Translate API (in blue) and IndicTrans (in red).

BC. Tamil has the oldest ancient non-Sanskritic Indian literature of any Indian language. Tamil uses agglutinative grammar, which uses suffixes to indicate noun class, number, case, verb tense, and other grammatical categories. Tamil's standard metalinguistic terminology and scholarly vocabulary is itself Tamil, as opposed to the Sanskrit that is standard for most Aryan languages. Tamil has many forms, in addition to dialects: a classical literary style based on the ancient language (cankattami), a modern literary and formal style (centami), and a current colloquial form (kotuntami) (Sakuntharaj and Mahesan, 2021, 2017, 2016; Thavareesan and Mahesan, 2019, 2020a,b, 2021). These styles blend into one another, creating a stylistic continuity. It is conceivable, for example, to write centami using cankattami vocabulary, or to utilize forms connected with one of the other varieties while speaking kotuntami (Subalalitha, 2019; Srinivasan and Subalalitha, 2019; Narasimhan et al., 2018). Tamil words are made up of a lexical root and one or more affixes. The majority of Tamil affixes are suffixes. Tamil suffixes are either derivational suffixes, which modify the part of speech or meaning of the word, or inflectional suffixes, which designate categories like as person, number, mood, tense, and so on. There is no ultimate limit to the length and scope of agglutination, which might result in large words with several suffixes, requiring many words or a sentence in English (Anita and Subalalitha, 2019b,a; Subalalitha and Poovammal, 2018).

The TamilATIS corpus is collected to promote research and development in the field of task-oriented dialogue systems for the Tamil language. It contains 4874 utterances related to airline-related enquiries.

Table 2 gives an overview of various TOD datasets available for Indian languages. We observe that only two of them, TOD-Dravidian and Codemix-DSTC2 contain Tamil utterances. Our dataset is different from these two. While TOD-Dravidian is annotated using automatic methods, TamilATIS is curated by hand. Codemix-DSTC2 contains Tamil-english code-mixed utterances, unlike ours which does not focus on code-mixing and rather contains utterances in pure Tamil.

Below, we describe the data collection and data annotation processes and give detailed statistics about the TamilATIS dataset.

3.1 Data Collection

We derive the Tamil ATIS dataset by automatically translating a modified version of the ATIS dataset (Hemphill et al., 1990) to Tamil and then manually annotating the slot labels. The ATIS dataset is a standard benchmark dataset for intent classification and slot filling. It consists of audio recordings and manual transcripts of humans enquiring about

Annotator Identity	Educational Background	Native Proficiency
1	Bachelors	✓
2	Bachelors	✓
3	Masters	✓

Table 3: Annotators and their details

Annotators	κ
α_1 α_2	0.94
α_1 α_3	0.95
α_2 α_3	0.97

Table 4: Cohen's κ agreement obtained during the annotation dry run. α_1 α_2 α_3 are the three annotators

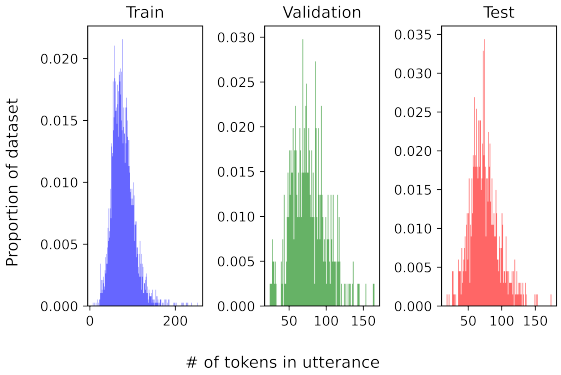

Figure 3: Distribution of number of tokens in each utterance

flight-related information on an automated airline travel inquiry system.

We experiment with two methods for translation: IndicTrans (Ramesh et al., 2021) and Google Translate API. We randomly sampled 50 utterances from the ATIS dataset and translated them using IndicTrans and the Google Translate API and manually inspected the translation quality. We noticed that translated utterances obtained from Google Translate were of much higher quality. Figure 2 shows some examples where IndicTrans did not give correct translations. For eg. in the first example, IndicTrans is not able to identify Vistara as an airline and combines it with Delhi, and in the second example, airline information is lost in translation. However, in both these cases, Google translate was able to provide proper translations.

3.2 Annotation Setup

For annotation, we follow earlier work on TOD (Malviya et al., 2021) where each utterance is annotated by one annotator. Since we derive utterances from ATIS, we already have a list of slot labels expected in each utterance and the annotator has to correctly map the slot label to the correct value in the utterance.

To aid with the annotation, we designed an interface that provides the annotators with an easy-to-use platform for annotation. Each annotator was assigned random batches of utterances and they worked independently in their own schedule.

3.3 Annotators

For the annotation process, we had 3 annotators. Two of the annotators are bachelor's student and one of them is a master's student. All three of the annotators have native language proficiency in Tamil. The details of the annotators are summarized in Table 3.

3.4 Annotation Process

Before the start of the annotation process, we briefed the annotators about TOD and trained them to identify intent and slots in an utterance from the TamilATIS dataset using examples covering all the different intent and slot labels. We conducted the annotation in two phases the dry run and a final annotation.

Dry run. We conducted a dry run on a subset of 200 utterances. We ensure that we uniformly sample the utterances from the dataset to ensure we have all the intents and slot labels in this subset. We then asked each of the three annotators to independently annotate each utterance. After this annotation, we computed the Cohen's κ for each pair of annotators. Table 4 shows the κ scores obtained. The high κ scores indicate that the annotators got a good grasp of the annotation process.

Final annotation. After the dry run, we started the final annotation process. In this stage, we also asked the annotators to reject an utterance if the translation was not correct. A total of 78 utterances were rejected in this phase.

3.5 Corpus Statistics

The corpus statistics for the TamilATIS dataset are given in Table 5. The minimum and maximum utterance lengths are 7 and 252 respectively, while the minimum and the maximum number of tokens in the utterance are 2 and 29. However, these are edge cases and the average utterance length and number of tokens are 76 and 8 respectively. We

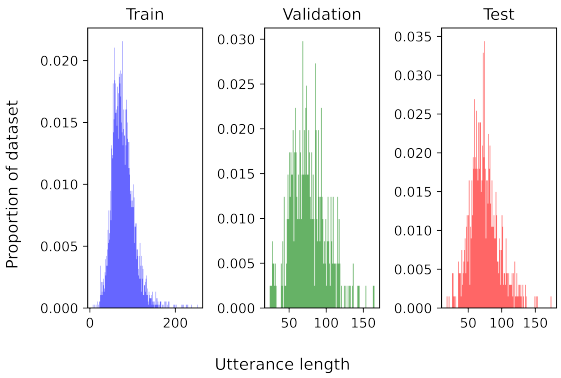

Figure 4: Distribution of utterance length

Vocabulary Size	1819
Total utterances	4874
# of intents	23
# of slot labels	45
# of unique slot values	885
Average length of utterances	76
Average # of tokens in utterance	8
Min & Max # tokens	2, 29
Min & Max utterance length	7, 252

Table 5: Statistics of the TamilATIS dataset

can observe in Figure 4 and 3 that the length and number of tokens in the utterances are consistent across the train, validation and test split.

4 Experimental Settings

We benchmark the TamilATIS dataset on eight state-of-the-art NLU models. In this section, we describe the models used and present the baseline results obtained. The problem of intent detection and slot filling can be cast as a generation task or a classification task, and in our baseline models, we include both of these types of architectures.

- **Seq2seq**:(Liu and Lane, 2016) propose an attention-based encoder-decoder model for joint intent detection and slot filling. Due to the explicit alignment requirement in the slot-filling task, the authors use an attention mechanism to incorporate alignment information into the encoder-decoder framework.
- **Slot-Gated:**(Goo et al., 2018) propose a slot-gated joint model that explicitly models the relationship between the slot and the intent attention vectors.
- **Capsule NLU:**(Zhang et al., 2019) propose hierarchical capsule nets to model the semantic hierarchy present among words, slots, and the intent of the utterance. They use context-aware word representations and dynamic routing to perform intent detection and slot-filling.
- **SF-ID:**(E et al., 2019) propose a bi-directional interrelated model for joint intent detection and slot filling. An SF-ID network is used to establish connections between the two tasks to help them promote each other mutually.
- **Stack-Propagation:**(Qin et al., 2019) propose a stack-propagation framework to incorporate the intent information during slot tagging. This allows the model to capture the intent of semantic knowledge. Moreover, to avoid error propagation in the model, token-level intent detection is performed.
- **SlotRefine:**(Wu et al., 2020) cast the task of joint intent detection and slot filling as a tag generation task and propose a non-autoregressive model for it. They use a two-pass mechanism to explicitly predict the slot boundary.
- **GL-GIN:**(Qin et al., 2021) propose a non-autoregressive model for joint intent detection and slot filling. It employs graph interaction networks to model slot dependency to model the interaction between intents and all the slots in the utterance.
- **JointBERT:**(Chen et al., 2019b) propose a joint model for intent detection and slot filling using the BERT model. The intent detection task is modelled as sequence classification while the slot filling task is modelled as token classification and the losses from the two models are jointly optimized.

5 Result and Analysis

In this section, we report the results obtained by the baseline models. We evaluate the NLU performance for slot filling using the F1(Micro) score and intent prediction using accuracy. The score obtained by each of the baselines is shown in Table 6.

Since the focus of these experiments is to just establish baselines and provide a starting point for further exploration, we restrict ourselves from in-depth error analysis.

Model	Intent (Acc)	Slot (F1)
Seq2seq ♣	83.11	56.99
Slot-Gated	93.87	91.31
Capsule NLU	89.33	88.48
SF-ID	91.92	92.47
Stack-Propagation ♣	93.27	91.84
SlotRefine ♣	94.30	92.10
GL-GIN ♣	91.33	91.94
JointBERT	96.26	94.01

Table 6: Performance of various baselines on the TamilATIS dataset. ♣ indicates that the model uses a generative approach and the rest of the models use a classification approach.

Encoder	Intent (Acc)	Slot (F1)
mBERT	95.21	93.64
indicBERT	93.57	89.28
Muril	87.44	81.74
XLM-Roberta	96.26	94.01

Table 7: Performance obtained from the JointBERT architecture by using different multilingual models as utterance encoders.

The baselines can be broadly classified into two approaches, generative approaches and classification-based approaches. The lowest scoring model is Seq2seq, which uses a simple encoder-decoder architecture to generate intent and slot details. This approach scores a decent accuracy of 83.11 for intent detection but a low F1 score of 56.99 for slot filling. Other generative approaches like Stack-Propagation, SlotRefine and GL-GIN achieve much better performances. These approaches have components in their architecture (like Stack-propagation framework, graph interaction layers etc) that help them explicitly model the relationship between the slots and intent leading to superior performance.

Classification based approaches like SF-ID, Capsule NLU and Slot-Gated achieve performance similar to the three generative architectures mentioned before. All of these approaches we discussed use sophisticated techniques to better model the interaction between intent and slot information and yield noticeable improvements over a simple architecture like Seq2Seq.

However, the best score is obtained by JointBERT. It obtains an intent accuracy of 96.26% and slot F1 score of 94.01%. This is an absolute improvement of 1.96% and 1.91% in intent accuracy and slot F1 over the second-best performing model (SlotRefine). JointBERT shows the effectiveness of transformer architectures pre-trained on large datasets when applied to downstream tasks like intent detection and slot filling.

We further investigate the JointBERT architecture by using different multilingual models as utterance encoders. Table 7 gives an overview of the score obtained by using different utterance encoders. XLM-Roberta and mBERT perform similarly, with XLM-Roberta getting slightly scores. IndicBERT and Muril, two transformer models pretrained on Indian languages however fail to produce scores as good as mBERT and XLM-Roberta. The superior performance of these two encoders could be attributed to the robust architecture and training strategy of XLM-Roberta and the large amount of data used to pretrain XLM-Roberta and mBERT.

6 Conclusion

In this paper, we presented TamilATIS, a TOD dataset in Tamil with 4874 utterances. We benchmarked the dataset with eight state-of-the-art NLU models and reported their intent accuracy and slot F1. Both generative and classification-based approaches perform similarly and achieve high intent accuracy and slot F1 score. We also highlighted the importance of modelling the relation between intent detection and slot labelling to yield performance improvement. The Joint BERT model with XLM-Roberta as utterance encoder achieved the highest score with an intent accuracy of 96.26% and slot F1 of 94.01%.

In future work, we plan to extend this dataset to other low-resource Dravidian languages like Malayalam, Kannada and Telugu. This would contribute towards the proliferation of TOD technology in the communities that speak these languages and also promote the development of multi-lingual TOD models for Dravidian languages. Having multi-domain utterances is another important research direction.

References

R Anita and CN Subalalitha. 2019a. An approach to cluster Tamil literatures using discourse connectives. In *2019 IEEE 1st International Conference on Energy, Systems and Information Processing (ICESIP)*, pages 1–4. IEEE.

R Anita and CN Subalalitha. 2019b. Building discourse parser for Thirukkural. In *Proceedings of the 16th*

International Conference on Natural Language Processing, pages 18–25.

Suman Banerjee, Nikita Moghe, Siddharth Arora, and Mitesh M. Khapra. 2018. A dataset for building code-mixed goal oriented conversation systems. In *COLING*.

Giovanni Campagna, Agata Foryciarz, Mehrad Moradshahi, and Monica Lam. 2020. Zero-shot transfer learning with synthesized data for multi-domain dialogue state tracking. pages 122–132.

Qian Chen, Zhu Zhuo, and Wen Wang. 2019a. Bert for joint intent classification and slot filling.

Qian Chen, Zhu Zhuo, and Wen Wang. 2019b. Bert for joint intent classification and slot filling.

Yun-Nung (Vivian) Chen, Dilek Z. Hakkani-Tür, Gökhan Tür, Jianfeng Gao, and Li Deng. 2016. End-to-end memory networks with knowledge carryover for multi-turn spoken language understanding. In *INTERSPEECH*.

Alice Coucke, Alaa Saade, Adrien Ball, Théodore Bluche, Alexandre Caulier, David Leroy, Clément Doumouro, Thibault Gisselbrecht, Francesco Caltagirone, Thibaut Lavril, Maël Primet, and Joseph Dureau. 2018. Snips voice platform: an embedded spoken language understanding system for private-by-design voice interfaces. *ArXiv*, abs/1805.10190.

Mai Hoang Dao, Thinh Hung Truong, and Dat Quoc Nguyen. 2021. Intent detection and slot filling for vietnamese.

Haihong E, Peiqing Niu, Zhongfu Chen, and Meina Song. 2019. A novel bi-directional interrelated model for joint intent detection and slot filling.

Chih-Wen Goo, Guang Gao, Yun-Kai Hsu, Chih-Li Huo, Tsung-Chieh Chen, Keng-Wei Hsu, and Yun-Nung Chen. 2018. Slot-gated modeling for joint slot filling and intent prediction. In *Proceedings of the 2018 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 2 (Short Papers)*, pages 753–757, New Orleans, Louisiana. Association for Computational Linguistics.

Daniel Guo, Gokhan Tur, Wen-tau Yih, and Geoffrey Zweig. 2014. Joint semantic utterance classification and slot filling with recursive neural networks. In *2014 IEEE Spoken Language Technology Workshop (SLT)*, pages 554–559.

Akshat Gupta, Xinjian Li, Sai Rallabandi, and Alan Black. 2020. Acoustics based intent recognition using discovered phonetic units for low resource languages.

E. Haihong, Peiqing Niu, Zhongfu Chen, and Meina Song. 2019. A novel bi-directional interrelated model for joint intent detection and slot filling. In *ACL*.

Dilek Hakkani-Tür, Gokhan Tur, Asli Celikyilmaz, Yun-Nung Vivian Chen, Jianfeng Gao, Li Deng, and Ye-Yi Wang. 2016. Multi-domain joint semantic frame parsing using bi-directional rnn-lstm. In *Proceedings of The 17th Annual Meeting of the International Speech Communication Association (INTERSPEECH 2016)*. ISCA.

Donghoon Ham, Jeong-Gwan Lee, Youngsoo Jang, and Kee-Eung Kim. 2020. End-to-end neural pipeline for goal-oriented dialogue systems using GPT-2. In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, pages 583–592, Online. Association for Computational Linguistics.

Charles T. Hemphill, John J. Godfrey, and George R. Doddington. 1990. The atis spoken language systems pilot corpus. HLT '90, page 96–101, USA. Association for Computational Linguistics.

Ehsan Hosseini-Asl, Bryan McCann, Chien-Sheng Wu, Semih Yavuz, and Richard Socher. 2020. A simple language model for task-oriented dialogue. In *Advances in Neural Information Processing Systems*, volume 33, pages 20179–20191. Curran Associates, Inc.

Pratik Jayarao and Aman Srivastava. 2018. Intent detection for code-mix utterances in task oriented dialogue systems. *2018 International Conference on Electrical, Electronics, Communication, Computer, and Optimization Techniques (ICEECCOT)*, pages 583–587.

Tushar Kanakagiri and Karthik Radhakrishnan. 2021. Task-oriented dialog systems for Dravidian languages. In *Proceedings of the First Workshop on Speech and Language Technologies for Dravidian Languages*, pages 85–93, Kyiv. Association for Computational Linguistics.

Jonathan K. Kummerfeld, Sai R. Gouravajhala, Joseph J. Peper, Vignesh Athreya, Chulaka Gunasekara, Jatin Ganhotra, Siva Sankalp Patel, Lazaros C Polymenakos, and Walter Lasecki. 2019. A large-scale corpus for conversation disentanglement. In *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics*, pages 3846–3856, Florence, Italy. Association for Computational Linguistics.

Weixin Liang, Youzhi Tian, Chengcai Chen, and Zhou Yu. 2019. Moss: End-to-end dialog system framework with modular supervision.

Bing Liu and Ian Lane. 2016. Attention-based recurrent neural network models for joint intent detection and slot filling.

Shrikant Malviya, Rohit Mishra, Santosh Barnwal, and Uma Shanker Tiwary. 2021. Hdrs: Hindi dialogue restaurant search corpus for dialogue state tracking in task-oriented environment. *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, PP:1–1.

Siddhartha Mukherjee, Anish Nediyanchath, Abhishek Singh, Vinuthkumar Prasan, Divya Verma Gogoi, and Surya Pratap Singh Parmar. 2021. Intent classification from code mixed input for virtual assistants. In *2021 IEEE 15th International Conference on Semantic Computing (ICSC)*, pages 108–111.

Anitha Narasimhan, Aarthy Anandan, Madhan Karky, and CN Subalalitha. 2018. Porul: Option generation and selection and scoring algorithms for a tamil flash card game. *International Journal of Cognitive and Language Sciences*, 12(2):225–228.

P. J. Price. 1990. Evaluation of spoken language systems: the ATIS domain. In *Speech and Natural Language: Proceedings of a Workshop Held at Hidden Valley, Pennsylvania, June 24-27,1990*.

Libo Qin, Wanxiang Che, Yangming Li, Haoyang Wen, and Ting Liu. 2019. A stack-propagation framework with token-level intent detection for spoken language understanding.

Libo Qin, Fuxuan Wei, Tianbao Xie, Xiao Xu, Wanxiang Che, and Ting Liu. 2021. Gl-gin: Fast and accurate non-autoregressive model for joint multiple intent detection and slot filling.

Gowtham Ramesh, Sumanth Doddapaneni, Aravinth Bheemaraj, Mayank Jobanputra, Raghavan AK, Ajitesh Sharma, Sujit Sahoo, Harshita Diddee, Mahalakshmi J, Divyanshu Kakwani, Navneet Kumar, Aswin Pradeep, Kumar Deepak, Vivek Raghavan, Anoop Kunchukuttan, Pratyush Kumar, and Mitesh Shantadevi Khapra. 2021. Samanantar: The largest publicly available parallel corpora collection for 11 indic languages.

Ratnasingam Sakuntharaj and Sinnathamby Mahesan. 2016. A novel hybrid approach to detect and correct spelling in Tamil text. In *2016 IEEE International Conference on Information and Automation for Sustainability (ICIAfS)*, pages 1–6.

Ratnasingam Sakuntharaj and Sinnathamby Mahesan. 2017. Use of a novel hash-table for speeding-up suggestions for misspelt Tamil words. In *2017 IEEE International Conference on Industrial and Information Systems (ICIIS)*, pages 1–5.

Ratnasingam Sakuntharaj and Sinnathamby Mahesan. 2021. Missing word detection and correction based on context of Tamil sentences using n-grams. In *2021 10th International Conference on Information and Automation for Sustainability (ICIAfS)*, pages 42–47.

Sebastian Schuster, S. Gupta, Rushin Shah, and Mike Lewis. 2019. Cross-lingual transfer learning for multilingual task oriented dialog. In *NAACL*.

R Srinivasan and CN Subalalitha. 2019. Automated named entity recognition from tamil documents. In *2019 IEEE 1st International Conference on Energy, Systems and Information Processing (ICESIP)*, pages 1–5. IEEE.

C. N. Subalalitha. 2019. Information extraction framework for Kurunthogai. *Sādhanā*, 44(7):156.

CN Subalalitha and E Poovammal. 2018. Automatic bilingual dictionary construction for Tirukural. *Applied Artificial Intelligence*, 32(6):558–567.

Ryuichi Takanobu, Hanlin Zhu, and Minlie Huang. 2019. Guided dialog policy learning: Reward estimation for multi-domain task-oriented dialog. In *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP)*, pages 100–110, Hong Kong, China. Association for Computational Linguistics.

Sajeetha Thavareesan and Sinnathamby Mahesan. 2019. Sentiment analysis in Tamil texts: A study on machine learning techniques and feature representation. In *2019 14th Conference on Industrial and Information Systems (ICIIS)*, pages 320–325.

Sajeetha Thavareesan and Sinnathamby Mahesan. 2020a. Sentiment lexicon expansion using Word2vec and fastText for sentiment prediction in Tamil texts. In *2020 Moratuwa Engineering Research Conference (MERCon)*, pages 272–276.

Sajeetha Thavareesan and Sinnathamby Mahesan. 2020b. Word embedding-based part of speech tagging in Tamil texts. In *2020 IEEE 15th International Conference on Industrial and Information Systems (ICIIS)*, pages 478–482.

Sajeetha Thavareesan and Sinnathamby Mahesan. 2021. Sentiment analysis in Tamil texts using k-means and k-nearest neighbour. In *2021 10th International Conference on Information and Automation for Sustainability (ICIAfS)*, pages 48–53.

Di Wu, Liang Ding, Fan Lu, and Jian Xie. 2020. Slotrefine: A fast non-autoregressive model for joint intent detection and slot filling.

Puyang Xu and Ruhi Sarikaya. 2013. Convolutional neural network based triangular crf for joint intent detection and slot filling. IEEE - Institute of Electrical and Electronics Engineers. Best Paper Award.

Chenwei Zhang, Yaliang Li, Nan Du, Wei Fan, and Philip S. Yu. 2019. Joint slot filling and intent detection via capsule neural networks.