Intent Detection in Urdu Queries using Fine-tuned BERT models

Sana Shams
Department of CS
University of Engineering and Technology, Lahore
Pakistan
sana.shams@kics.edu.pk

Bareera Sadia
CLE- KICS
University of Engineering and Technology, Lahore
Pakistan
bareera.sadia@kics.edu.pk

Muhammad Aslam
Department of CS
University of Engineering and Technology, Lahore
Pakistan
maslam@uet.edu.pk

***Abstract*— User's intent detection provides essential cues in query understanding and accurate information retrieval through search engines and task-oriented dialogue systems. Intent detection from user queries is challenging due to short query length and lack of sufficient context. Further, limited prior research in query intent detection has been conducted for Urdu, an under-resourced language. With the recent success of Bidirectional Encoder Representation from Transformers (BERT), that provides pre-trained language models, we propose to develop intent detection model for Urdu by fine-tuning BERT variants for intent detection task. We conduct rigorous experimentation on mono and cross-lingual transfer learning approaches by using pre-trained BERT models i.e. mBERT, ArBERT, and roBERTa-urdu-small and two query datasets. Experimental evaluation reveal that the fine-tuned models of mBERT and roBERTa-urdu-small achieve 96.38% and 93.30% accuracy respectively on datasets I and II outperforming strong statistical and neural network baselines.**

Keywords—BERT, fine-tuning, intent detection, search queries, Urdu

I. INTRODUCTION

Intent detection is a critical component of natural language understanding that classifies text into predefined intent labels. User intent is defined as "the expression of an effective, cognitive, or situational goal" during users' interaction with web systems such as search engine or dialogue systems [1]. Intent detection in search queries is challenging as queries are semantically ambiguous due to their loose grammatical structure, and limited context in short length. This is further complicated in Urdu due to its morphological richness. This is illustrated in the query examples below:

ہم|*PRP* سٹائل|*NN* ایوارڈز|*NN* *(1)*

خطوط|*NN* مرزا|*NNP* غالب|*JJ* *(2)*

Due to limited context, named entity "ہم" in (1) is mislabeled as PRP (Pronoun) when it is NNP (Proper Noun, TV channel's name). In addition all three units are individually tagged whereas, it is a single named entity "Hum Style Awards". In (2) due to loose grammatical structure, phrases NN خطوط and مرزا|NNP غالب|JJ lack syntactic connection of a possessive pronoun i.e. "کے" (s') between them though they are semantically related. Analyzing further, "Hum Style Awards" in (1) is a complete indivisible unit representing an intention to visit the website of this show directly, while decomposition into multiple units may imply a different intent to seek information about each units.

Traditionally, intent detection models used features extracted from query texts and search logs e.g. word n-grams, past user-engagements, clicked URLs etc. to statistically classify queries using machine learning (ML) based classifiers e.g. decision trees, naïve bayes or support vector machine (SVM). Recently, in deep neural network (DNN) based architectures encoded input representation of query texts, e.g. word embeddings are extensively used [2]. Word embeddings e.g. Word2Vec [3], GloVe [4], ELMo [5] and fastText [6] etc. are vector representations learnt from large corpora that map input words into a high dimensional semantic space. Pre-trained Bidirectional Encoder Representations from Transformers (BERT) [7] models are attention based contextual language representation models that encode positional information and word relations within the vector space. There are two popular approaches to train the word embedding; i) train the weights of the embedding layer from scratch ii) use pre-trained word embedding developed using a large corpus. Models built from scratch consume more time and resources therefore leveraging pre-trained models to develop optimized word or aggregate sentence representation is more cost-effective.

Pre-trained models can be fine-tuned using task specific annotated datasets. Fine-tuning is a specialized case of transfer learning, in which the pre-trained model is used for feature extraction and the last output layer is tuned according to the downstream task. In majority of the scenarios, fine-tuned models achieve comparable or better results than pre-trained equivalents, even if the labelled data for the downstream task is small. Further, fine-tuning is efficient as it does not require immense compute and memory resources as compared to the models built from scratch.

In this paper we fine-tune three variants of pre-trained BERT models; mBERT, ArBERT[1] and roBERTa-urdu-small[2] for downstream task of intent detection. We report the results of each intent detection model using four evaluation metrics; precision, recall, accuracy and F1. The main contributions of this study are.

- Development of two datasets of Urdu web queries namely Urdu Web Queries (UWQ-22) and Urdu ATIS (U-ATIS) having 3 and 16 intents respectively.

1 https://huggingface.co/asafaya/bert-base-arabic

2 https://huggingface.co/urduhack/roberta-urdu-small

978-1-6654-6477-2/22/$31.00 ©2022 IEEE

- Comparative analysis of models developed by fine-tuning BERT variants for specialized downstream task of intent detection.

II. RELATED WORK

User intent detection has been largely analyzed as a classification problem from written and spoken queries. Datasets for intent detection have been largely annotated using taxonomy proposed in [8]. Later studies have reported newer taxonomies by improving or extending upon the Broders' classification [9-11]. These datasets have been extracted from search logs of popular search engines. Some of the logs are available in public domain for research e.g. MSN Search Query log[3], Yahoo[4] and AOL Search query logs [12] in English as well as Russian[5], Chinese[6], Chilean [13] and Vietnamese [14].

Machine learning (ML) methods use supervised and unsupervised learning techniques for intent detection. For unsupervised learning, heuristic [15], graph [16] and rules based approaches, modelling linguistic properties of query intents have been proposed. SVM has been extensively used for detecting user intent [17-19] in the supervised learning paradigm. These models use feature engineering and domain knowledge to devise discriminative features extracted from corpora, query's text, or from the query logs for intent detection. Intent detection architectures based on neural network (NN) have shown better performance than ML based techniques [20, 21]. Recently, DNN based models and architectures have been excessively applied in natural language processing (NLP) tasks, specifically intent detection, for improved performance compared to techniques involving feature engineering. Convolutional neural networks (CNN) and recurrent neural networks (RNN) architectures [22, 23] have considerably outperformed earlier approaches. Long short term memory (LSTM) networks (both single and bi-directional)[24, 25], with strength to model long distance dependencies, have been significantly applied for intent detection in queries.

The transformer-based pre-trained BERT models have further improved the state-of-the-art performance on many NLP tasks, specifically intent detection. In [26] a hybrid BERT-Cap based model was proposed for intent detection reporting results on four datasets; two English and two Chinese Q&A datasets. The model achieved 0.967 accuracy on one of the Chinese datasets. In [13] BETO model (a Spanish pre-trained BERT model from the Universiedad de Chile) was fine-tuned using domain specific vocabulary extension and achieved 79.48% accuracy on downstream named entity recognition task. Another study presented in [27] reported a fine-tuned model for Urdu sentiment analysis using pre-trained multilingual BERT. The fine-tuned model achieved F1 score 81.49 outperforming the pre-trained mBert as well as statistical and neural baselines. Pre-trained BERT for Arabic has also been utilized for sentence continuation inference in Urdu text. The study highlighted the effectiveness of fine-tuning based on cross-lingual transfer learning models for enhanced task specific performance.

Studies presented above highlight the effectiveness of mono and cross-lingual transfer learning approaches in multiple NLP tasks, thus motivating its applicability for intent detection in Urdu text. To demonstrate the effectiveness of mono-lingual transfer learning, we experimented with roBERTa-urdu-small. For analyzing impact of cross-lingual learning, pre-trained Arabic BERT is used. As currently, no publically available large Urdu BERT model is available, therefore pre-trained BERT multilingual model (mBERT) trained on 103 languages including Urdu, is used, as a hybrid learning scenario between mono and cross-lingual learning. These models are fine-tuned for intent detection task to evaluate their performance with statistical and neural baselines.

III. METHODOLOGY

Fine-tuning entails training a dataset specific to a task. BERT models pre-trained on massive data, can be rapidly fine-tuned for specific NLP tasks by adding an additional untrained layer of neurons on top of the final BERT layer, and re-training the model for 2-4 extra epochs. In our research we develop intent detection models by fine-tuning three pre-trained variants of BERT models; mBERT, ArBERT and roBERTa-urdu-small and comparatively analyze their intent detection performance on two datasets of Urdu queries.

Selection of the pre-trained models was based on the following:

- Pre-trained mBERT incorporates Urdu in addition to other languages
- roBERTa-urdu-small is the only publically available BERT variant for Urdu
- ArBERT includes Arabic which is an ancestor language of Urdu, and shares minimal vocabulary with Urdu

The following section describes the datasets and fine-tuned models in further detail.

A. Dataset

Two datasets were used in the experiments. The first dataset, includes indigenous Urdu search queries extracted from Humkinar,[7] an Urdu search engine. The second dataset is a publically available dataset, ATIS benchmark evaluation corpus for natural language understanding that has been translated into Urdu. Both datasets are described in the following sections.

1) Urdu Web Queries (UWQ-22) Dataset

The Urdu web queries dataset [31] is the first dataset comprising of native web queries extracted from search records of Humkinar [28]. The dataset comprised of 11,751 search records from 165 users. The total queries in the dataset comprised of 42,214 terms of which 38,789 were unique. The mean length of web queries is 3.63 terms. After extraction, dataset has been preprocessed before intent annotation. During pre-processing queries were trimmed for trailing spaces at the beginning and end. The resulting dataset was normalized through diacritics removal, systematic whitespace normalization, removal of punctuations or accents and removal of English alphabets. After normalization, duplicate queries and records resulting in null queries were removed again. The final dataset after preprocessing was 8518 queries.

3 http://www.sobigdata.eu/content/query-log-msn-rfp-2006
4 https://webscope.sandbox.yahoo.com/
5 http://switchdetect.yandex.ru/en/datasets
6 http://www.sogou.com/labs/
7 https://www.humkinar.com.pk/

The Urdu web queries dataset was annotated with three intents: Informational (INFO), Navigational (NAV) and Transactional (TRAN) following the intent annotation taxonomy proposed in [8]. The distribution of the three intents across the Urdu web queries dataset is presented in Table 1.

TABLE I. INTENT DISTRIBUTION WITH MEAN QUERY LENGTH IN URDU WEB QUERIES DATASET

Intent	Frequency	Coverage	Mean Query Length
INFO	6495	76.25 %	4.0
NAV	845	9.92 %	3.2
TRAN	1178	13.83 %	3.7
Total	8518	100%	3.63

It is evident from Table I that the largest number of queries are INFO followed by TRAN and NAV queries. The dataset included a balanced coverage across AOL query classification domains presented in [29]. The domain wise distribution of the queries across 11 query domains is presented in Table II.

TABLE II. DOMAIN-WISE INTENT DISTRIBUTION IN URDU WEB QUERIES DATASET

Domain	INFO (%)	TRAN (%)	NAV (%)
Books	7%	2%	28%
Business	8%	15%	1%
Entertainment	9%	10%	36%
Health	9%	1%	5%
Travel	7%	2%	2%
Technology	12%	9%	3%
News	10%	7%	7%
Fact-Info	7%	20%	4%
Shopping	10%	7%	4%
Geography	11%	19%	1%
Sports	10%	8%	9%

Table II shows that technology related queries are in majority in informational queries; Fact-Info queries are in majority in transactional queries, and entertainment queries are in majority in navigational queries [31]. For experimentation, the dataset was divided into train and test sets using a standard ratio of 80:20. The training dataset contained 6818 queries and the testing set had 1700 queries.

2) *Urdu ATIS (U-ATIS)*

In this study, the ATIS corpus annotated with 17 intents has been used to develop the translated Urdu ATIS (U-ATIS) for intent detection experiments. One intent label having multi-intents (two) in the original dataset was removed. The remaining dataset was translated from English to Urdu using automatic machine translations by Google's toolkit. To evaluate the correctness and completeness of the automatically generated translations the dataset was completely reviewed by two native Urdu speakers, verifying that the translated phrases were logical and complete. The intent labels were used as they were and were not translated. The final statistics of Urdu ATIS dataset (U-ATIS) are given in Table III presenting the intent labels and statistics related to the train and test splits.

TABLE III. DATASET STATISTICS OF URDU ATIS (U-ATIS)

Urdu –ATIS Dataset	Counts
Intents	16
Train set	4965
Test set	828

B. *Fine-Tuned Model for Urdu Intent Detection*

Pre-trained BERT models are extensively used language modeling architectures. For fine-tuning, BERT is initialized with the pre-trained parameters, followed by updating model weights using task specific labelled data. Figure 1 shows the high-level architecture of fine-tuning BERT (multilingual) for intent detection.

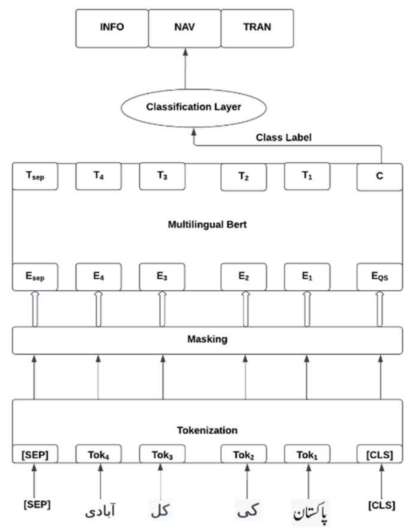


Fig. 1. Fine-tuning BERT (multilingual) for downstream task of intent detection

As shown in Fig. 1, the input representation consists of sequence of queries retrieved from the dataset. As a first step the input sequence is tokenized. [CLS] and [SEP] are special classification tokens that represent the initial and final tokens of the input query. For classification tasks, the final hidden state corresponding to [CLS] is used as the aggregate representation of the complete sequence. The encoder assigns a unique input representation to each token, constructed by aggregating the token, segment, and positional embeddings [7].

The second stage called masked language modelling entail randomly hiding/masking a certain number of tokens and then predicting the masked tokens. Typically 15% of tokens are masked at random in a given sequence with a [MASK] token. For example as in Fig. 1, the word "abadi (آبادی)" could be masked and replaced with [MASK] token. By iterating the input query with masked tokens in the model, it can make precise predictions of the original masked tokens, by learning bidirectional contextual representations of the sequences.

At the output, the representation of the special token, [CLS] is fed into softmax classifier, an output layer, for specific predictions such as sequence tagging or question answering. These predictions are used as the model's fine-tuning output. For intent detection task, a classification layer is added on top of the transformer output for [CLS] token from pre-trained BERT model. It forms a dense, fully-connected layer having K X H dimensions (K= no. of classes and H = size of the hidden state). A loss function such as *cross entropy* is used to minimize the loss.

TABLE IV. COMPARATIVE EVALUATION OF BERT VARIANTS VS. FINE-TUNED BERT MODELS FOR INTENT DETECTION ON U-ATIS AND UWQ-22 DATASETS

	Dataset I: U-ATIS				Dataset II: UWQ-22			
Model	Prec.	Recall	F1	Accuracy	Prec.	Recall	F1	Accuracy
ArBERT	0.5525	0.05917	0.05917	0.07149	0.7973	0.10229	0.10229	0.03414
Intent_ArBERT	***0.9357***	***0.942***	***0.942***	***0.9356***	***0.8885***	***0.8912***	***0.8912***	***0.8891***
roBERTa-small-Ur	0.5119	0.1521	0.1521	0.2303	0.5831	0.7636	0.7636	0.6613
Intent_roBERTa-sm-ur	***0.952***	***0.948***	***0.948***	***0.9412***	***0.942***	***0.9514***	***0.9514***	***0.933***
mBERT	0.325	0.042	0.033	0.0423	0.633	0.673	0.609	0.714
Intent_mBERT	***0.9605***	***0.9638***	***0.9592***	***0.9638***	***0.8908***	***0.8936***	***0.8917***	***0.8936***

TABLE V. COMPARATIVE EVALUATION OF FINE-TUNED BERT MODELS WITH ML AND DL BASED MODELS FOR INTENT DETECTION ON U-ATIS AND UWQ-22 DATASETS

	Dataset I: U-ATIS				Dataset II: UWQ-22			
Model	Prec.	Recall	F1	Accuracy	Prec.	Recall	F1	Accuracy
Baseline I: SVM	0.86	0.88	0.85	0.88	0.83	0.82	0.82	0.82
Baseline II: NB	0.84	0.80	0.81	0.80	0.80	0.84	0.80	0.81
Baseline III: CNN	0.859	0.859	0.859	0.859	0.875	0.7	0.764	0.88
Baseline IV: BLSTM	0.862	0.860	0.860	0.860	0.747	0.788	0.765	0.859
Intent_ArBERT	0.9357	0.942	0.942	0.9356	0.8885	0.8912	0.8912	0.8891
Intent_roBERTa-small-Ur	0.952	0.948	0.948	0.9412	**0.9420**	**0.9514**	**0.9514**	**0.9330**
Intent_mBERT	**0.9605**	**0.9638**	**0.9592**	**0.9638**	0.8908	0.8936	0.8917	0.8936

IV. EXPERIMENTAL SETUP

In all the experiments, the models were trained on a core i7 machine, 2.0 GHz, with 64GB RAM and 2X12GB GeForce RTX 3060 graphical processing units.

A. Baselines

Intent detection results using fine-tuned BERT models are compared across the following four baselines.

SVM/NB (Baseline I-II): Baselines I and II are curated with TF-IDF features to represent the query, and support vector machine (SVM) with linear kernel and multinomial naïve bayes (NB) as classifiers.

Convolutional Neural Network (CNN) (Baseline III): This baseline is setup using the architecture proposed in [30] having n-gram convolutional layer and pooling operation for text classification.

BLSTM (Baseline IV): In this baseline, bi-directional long short term memory (BLSTM) network, [24] having bi-directional forward layer is used and the last hidden state is used for classification.

B. Fine-Tuning Set-up

The following pre-trained models are used to fine-tune for intent detection task.

1) BERT multilingual

The mBERT model is based on single language base BERT, which consists of 12-layers, 768-hidden, 12-heads, 110M parameter. mBERT model is trained using the top 104 languages (including Urdu) having the largest Wikipedia pages.

2) Arabic_BERT

The "arabic-bertbase" model was pre-trained on approximately 8.2 billion words. This model is trained using 3M training steps with batch size of 128.

3) roBERTa-small-urdu

roBERTa-small-urdu is the only publically available pre-trained model using roBERTa architecture for Urdu. This model is trained on a web corpus of Urdu newspapers.

The hyper parameters used for BERT fine tuning experiments are provided in Table VI. The three pre-trained models were fine-tuned using respective training sets with annotated intents and the separately held out test sets have been used for evaluation.

TABLE VI. HYPERPARAMETERS USED TO FINE TUNE BERT MODELS ON UWQ-22 AND U-ATIS DATASETS

Hyper parameter	Value
Learning Rate	5e-5
No. of Epochs	5
Max sentence length	512

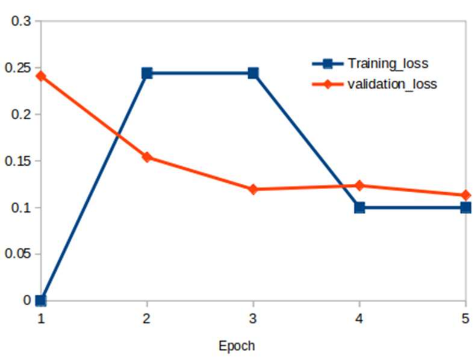
Hyper parameter	Value
Batch size	16
Optimizer	Adam, 5e-8
Weight Decay	0.05
Loss function	CE Loss

V. Evaluation Results

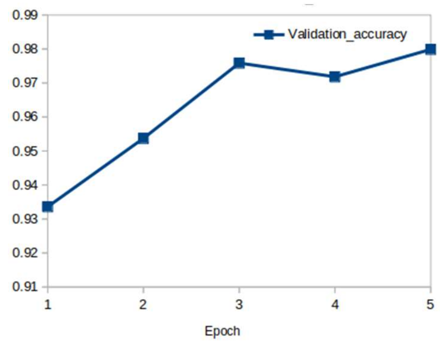
Fig. 2 present the training and validation accuracies and loss during fine-tuning of BERT multilingual on U-ATIS dataset. The optimal number of epochs, 5 can be seen from the training accuracy and loss curves, as at epoch 5 the training set converges to the highest value and the loss is the minimum. Table IV presents the performance of pre-trained BERT multilingual, Arabic and roBERTa-urdu-small vs. the respective fine-tuned models for intent detection tasks. To evaluate performance. precision, recall, F1-scores and accuracy was calculated. It is evident from the results that models fine-tuned on limited task specific dataset have performed exceptionally better than their pre-trained models.

On U-ATIS dataset, the pre-trained models have performed poorly, however BERT models fine-tuned on annotated training datasets have shown more than 90% improvement in accuracy compared to the pre-trained models. One reason for this stark improvement could be the domain specificity of the dataset that has been adequately learnt by the fine-tuned models. Intent_mBERT, has performed the best with 96.35% accuracy and 0.9592 F1-score. The training and validation loss vs. epoch representation during fine-tuning BERT multilingual on U-ATIS dataset is presented in Fig.2.

On UWQ-22 dataset the pre-trained models have shown satisfactory results compared with their performance on U-ATIS. One reason for the improved performance could be similarity of syntactic structures in the datasets (since majority of web query dataset content is also related to web corpus) on which the models are trained. Intent_roBERTa_sm_ur performs the best with 93.3% accuracy and 0.9514 F1 score.



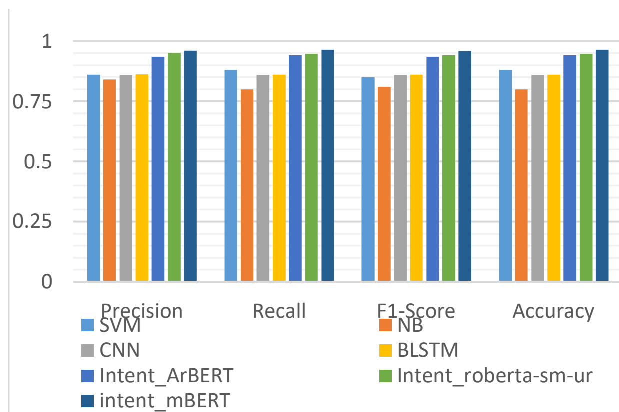
(a)



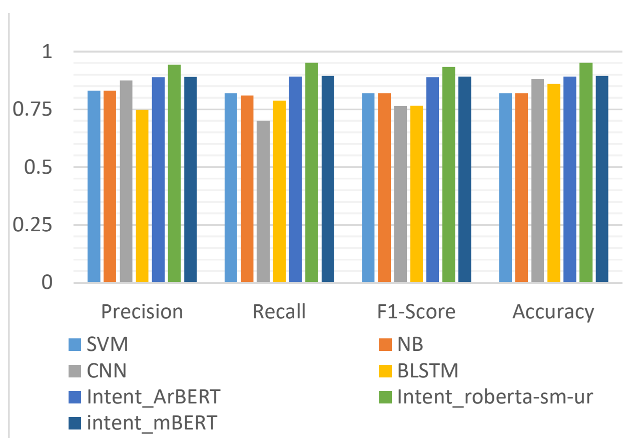
(b)

Fig. 2. Graphical representtaion of (a) loss and (b) accuracy curves during fine-tuning of BERT (multilingual) for Intent detection on U-ATIS dataset.

Table V compares the results of fine-tuned models with machine learning based and deep learning based baselines. The results show that on U-ATIS dataset, fine-tuned model Intent_mBERT outperform all other baselines with precision, recall, F1 score and accuracy of 0.9605, 0.9638, 0.9592, and 96.38% respectively. On UWQ-22 dataset, fine-tuned model Intent_roBERTa-sm-ur outperform all other baselines with precision, recall, F1 score and accuracy of 0.9420, 0.9514, 0.9514, and 93.30% respectively. The comparative results are also presented in Fig 3.



(a)



(b)

Fig. 3. Comparative precision, recall, accuracy and f1-score of baseline models and fine-tuned BERT models on intent detection for (a) U-ATIS dataset and (b) UWQ-22 dataset

VI. Conclusions

Transfer learning has achieved state of the art results with minimal task specific training. In this paper we tackle the problem of intent detection in Urdu search queries by fine tuning pre-trained BERT models using two intent annotated Urdu queries datasets. Through rigorous experimental evaluation, we compare the performance of fine-tuned models with machine learning based and deep learning based intent detection baselines and alternate classification techniques. We demonstrate that fine-tuned BERT multilingual (Intent_mBERT) attained 96.38% accuracy on U-ATIS dataset, while fine-tuned roBERTa-small-urdu (Intent_roBERTa-sm-ur) attained 95.15% accuracy on the UWQ-22 datasets. In the future, further research could be conducted to boost the systems' performance by using hybrid models leveraging input embedding from the fine-tuned BERT models and using neural networks for intent detection.

References

1. B. J. Jansen, D.L. Booth, and A. Spink, "Determining the user intent of web search engine queries," in 16th international conference on World Wide Web. 2007, Association for Computing Machinery: Banff, Alberta, Canada. p. 1149–1150.
2. J.K. Kim, G. Tur, A. Celikyilmaz, B. Cao, and Y. Wang, "Intent detection using semantically enriched word embeddings," in 2016 IEEE Spoken Language Technology Workshop (SLT). 2016, San Diego, CA. p. 414-419.
3. T. Mikolov,I. Sutskever, K. Chen, G. Corrado, and J. Dean, "Distributed representations of words and phrases and their compositionality," in 26th International Conference on Neural Information Processing Systems - Volume 2. 2013, Curran Associates Inc., Lake Tahoe, Nevada. p. 3111–3119.
4. J. Pennington, R. Socher, and C. Manning.," GloVe: Global vectors for word representation," in 2014 Conference on Empirical Methods in Natural Language Processing, EMNLP, 2014. Doha, Qatar.
5. M.E. Peters, M. Neumann, M. Iyyer, M. Gardner, C. Clark, K.Lee and L. Zettlemoyer, "Deep contextualized word representations," in Conference of the North American Chapter of the Association for Computational Linguistics (NAACL): Human Language Technologies (HLT) 2018. New Orleans, Louisiana:
6. E. Grave, P. Bojanowski, P. Gupta, A. Joulin, and T. Mikolov, "Learning word vectors for 157 languages," in 11th International Conference on Language Resources and Evaluation (LREC 2018). 2018. Miyazaki, Japan.
7. J. Devlin, M. Chang, K. Lee, and K. Toutanova, " BERT: Pre-training of deep bidirectional transformers for language understanding," in Conference of the North American Chapter of the Association for Computational Linguistics (NAACL): Human Language Technologies (HLT), Volume 1, 2019. Minneapolis, Minnesota.
8. A. Broder,"A taxonomy of web search," SIGIR Forum, 2002. 36(2): p. 3-10.
9. B.J. Jansen, D.L. Booth, and A. Spink, "Determining the informational, navigational, and transactional intent of Web queries," in Information. Processing. Management. 2008, 44(3): p. 1251–1266.
10. D.E. Rose, and D. Levinson, "Understanding user goals in web search," in 13th international conference on World Wide Web. 2004, ACM: New York, NY, USA. p. 13-19.
11. R. Baeza-Yates, L. Calderón-Benavides, and C. González-Caro, "The intention behind web queries," In: F. Crestani, P. Ferragina, M. Sanderson, (eds) String Processing and Information Retrieval. SPIRE 2006. Lecture Notes in Computer Science, vol 4209. Springer, Berlin, Heidelberg 2006.
12. G. Pass, A. Chowdhury, and C. Torgeson, "A picture of search," in 1st International Conference on Scalable Information Systems. 2006, ACM: Hong Kong. p. 1.
13. F. Fernández-Martínez, C. Luna-Jiménez, R. Kleinlein, D. Griol, Z. Callejas and J.M. Montero, "Fine-tuning BERT models for intent recognition using a frequency cut-off strategy for domain-specific vocabulary extension," Applied Sciences, 2022. 12(3): p. 1610.
14. N.T.T. Trang, D.T.D. Anh, V.Q. Viet and P. Woomyoung, P.," Advanced joint model for Vietnamese intent detection and slot tagging," in Vo, NS., Vien, QT., Ha, DB. (eds) Industrial Networks and Intelligent Systems. INISCOM 2022. Lecture Notes of the Institute for Computer Sciences, Social Informatics and Telecommunications Engineering, vol 444. Springer, Cham.
15. I. -H. Kang, and G. Kim, "Query type classification for web document retrieval," in 26th annual international ACM SIGIR conference on Research and development in informaion retrieval. 2003, ACM: Toronto, Canada. p. 64-71.
16. U. Lee, Z. Liu, and J. Cho, "Automatic identification of user goals in web search," in 14th international conference on World Wide Web. 2005, ACM: Chiba, Japan. p. 391-400.
17. D. Wu, Y. Zhang, S. Zhao and T. Liu, "Identification of web query intent based on query text and web knowledge," in 1st International Conference on Pervasive Computing, Signal Processing and Applications, 2010, pp. 128-131
18. M. Marcelo, and J. Zamora., "Identifying the iIntent of a user query using support vector machines," in String Processing and Information Retrieval. 2009. Finland: Springer-Verlag Berlin Heidelberg.
19. P. Ren, Z. Chen, X. Song, B. Li, H. Yang, and J. Ma," Understanding temporal intent of user query based on time-based query classification," in: G. Zhou, J. Li, D. Zhao, and Y. Feng, (eds) Natural Language Processing and Chinese Computing. NLPCC 2013. Communications in Computer and Information Science, vol 400. Springer, Berlin, Heidelberg
20. Song, I. and J. Diederich, "Intent extraction from text messages," Neural Information Processing. Theory and Algorithms, ICONIP 2010. Lecture Notes in Computer Science, ed. M.B.S.U.B.A. Wong K.W. Vol. 6443. 2010, Berlin Heidelberg: Springer-Verlag.
21. Bhardwaj, A.R.A.I., "User intent classification using memory networks: A comparative analysis for a limited data scenario,"arXiv, 2017. abs/1706.06160.
22. Ravuri, S.V. and A. Stolcke. "Recurrent neural network and LSTM models for lexical utterance classification," in INTERSPEECH. 2015.
23. M.A. Ali, B. Zaity, P. Drobintsev, H. Wannous, I. Chernoruckiy, and A. Filchenkov, "Joint slot filling and intent detection in spoken language understanding by hybrid CNN-LSTM model," in 2020 1st International Conference on Control, Robotics and Intelligent System (CCRIS '20). Association for Computing Machinery, New York, NY, USA, 112–117.
24. K. Sreelakshmi, P. C. Rafeeque, S. Sreetha and E. S. Gayathri,, "Deep bi-directional LSTM network for query intent detection," Procedia Computer Science, 2018. 143: p. 939-946.
25. Y.Y.K.T.L. Shi, and D. Jiang, "Deep LSTM based feature mapping for query classification," in North American Chapter of the Association for Computational Linguistics (NAACL): Human Language technologies. (HLT) 2016, Association of Computational Linguistics: San Diego, California. p. 1501-1511.
26. H. Liu, Y. Liu, L. -P. Wong, L.-K. Lee, T. Hao, and Z. Yang, "A hybrid neural network BERT-Cap based on pre-trained language model and capsule network for user intent classification," Complex., 2020. 2020: p. 11.
27. L. Khan, A. Amjad, N. Ashraf and H.-T. Chang, "Multi-class sentiment analysis of urdu text using multilingual BERT," Scientific Reports, 2022. 12(1): p. 5436.
28. H. M. Shafiq, B. Tahir, M. A. Mehmood, D. Pinto, V. Singh, and F. Perez. "Towards building an Urdu language corpus using Common Crawl," J. Intell. Fuzzy Syst., 2020. 39(2): p. 2445–2455.
29. B.J. Jansen, and D. Booth, "Classifying web queries by topic and user intent," in CHI '10 Extended Abstracts on Human Factors in Computing Systems. 2010, ACM: Atlanta, Georgia, USA. p. 4285-4290.
30. Y. Kim, "Convolutional neural networks for sentence classification," in Empirical Methods in Natural Language Processing (EMNLP) 2014. Doha, Qatar.
31. S. Shams, and M. Aslam, "Improving user intent detection in Urdu web queries with capsule net architectures," Applied Sciences. 2022, in press.