
The Minimal isiNdebele Explicit and Implicit Question-answering Text Datasets for Zero-shot Learning**Promise Tshepiso Malatji¹, Thihe Isaiah Modipa²**promise.malatji@ul.ac.za¹, thihe.modipa@ul.ac.za²^{1,2}Department of Computer Science, University of Limpopo, Polokwane, South Africa

Article Information

Received : 29 Mar 2026

Revised : 12 Jun 2026

Accepted : 30 Jun 2026

Keywords

data scarcity, low-resource language, question-answering, zero-shot, agglutinative

Abstract

The use of Large Language Models (LLMs) for task specific has gained more popularity for resourced languages like English due to availability of text data. There are no readily available datasets to fine tune for task specific in low resourced languages like isiNdebele. Data scarcity remains the main obstacle for many low resourced languages as it hinders on the development of language models. In this study we proposed the creation of two isiNdebele Question-answering (QA) text datasets, for explicit and implicit QA. Forty-five matrix question papers from the South African Department of Basic Education website were downloaded. A data verification process was performed using human-in-the-loop technique to validate the created datasets. Further data augmentation was performed to increase the implicit dataset from 3012 to 36560 context-question-answer triplets. The augmented implicit dataset was then used to fine-tune the mT5 model on a zero-shot. The two datasets were accepted with a 98.33% acceptance percentage by the participants. The mT5 model performed exceptionally with ROUGE-L and BERTScore F1 of 0,992 and 0,999 respectively. The model made 94.7% accurate predictions with a perplexity score of 1,157. The results indicate that the multilingual model and transfer learning have great potential of dealing with low resourced languages in such as QA.

A. Introduction

The use of Large Language Models (LLMs) for task specific has gained more popularity for resourced languages like English due to availability of text data [1], [2], [3]. Thai and English languages have made significant progress in the development of state-of-the-art QA models [4]. English QA datasets is readily available for use. There are two types questioning, the open-ended and close-ended question. And different close-ended questioning datasets are made available with the multiple-choice questions gaining more attention [5].

There are no readily available datasets to fine tune for task specific in low resourced languages [6]. Data scarcity remains the main obstacle for many low resourced languages as it hinders on the development of language models [1], [7]. Enormous amount of data is needed to effectively train and fine-tune a model, and it becomes impossible for low resourced languages like isiNdebele language [8]. In the study to develop the Arabic “why” QA model, only 3146 context-question-answer triplets were used to trained and validation the model [9]. A different dataset with 110 “why” questions was used for testing the developed model. The data size is very small compared to data sizes used in resourced languages.

IsiNdebele language is considered a low-resourced language with minimal online presence [10]. This is a morphological rich language with agglutinative structure which might present natural language processing challenges when fine-tuning [11]. Words are formed by concatenating two or more words together. A morpheme is regarded as the smaller meaningful unit of a language that is used to form a word [12]. There are two types of morphemes, free and bound morphemes. Each morpheme can be categorised as either free or bound morpheme. Free morphemes may exist independently while bound morphemes are dependent on other morphemes within a word, they cannot exist alone [13], [14]. The isiNdebele word *ezinganakukhubazeka* (those that do not have disability) is composed of these two words *ezingana* and *kukhubazeka* which are free morphemes.

In the study by [1] they used zero-shot and few-shot scenarios on multilingual models to develop story ending generation (SEG) models. The transfer learning technique is wildly used when developing models for low resourced languages [15], [16], [17]. This technique is beneficial in tranfering knowledge from well-resourced languages in the pretrained model to low resourced languages which share morphological structures. This method is adopted to address the data scarcity on low-resourced languages, and the consistent high accuracy rate results obtained indicate the potential of multilingual models to deal with complex learning challenges [17].

A QA model that answers the “why” question is developed for the Arabic language [9]. Developing the “why” QA model is complex due to the complexity of the responses. The purpose of this study is to create two isiNdebele QA datasets from the minimal data available online. The dataset created can be used to fine-tune pre-trained models for QA tasks. Future researchers can utilise the study as a benchmark for other low resourced languages.

B. Research Method

The flowchart in figure 1 depicts the process followed to create the two isiNdebele QA datasets. The source for the data is the South African Basic Education website. Forty-five matric question papers and their respective memorandums were downloaded. The dataset is created manually by copying the extract, question, and answer to form the Explicit QA dataset. The data quality verification process was then performed using the human-in-the-loop method.

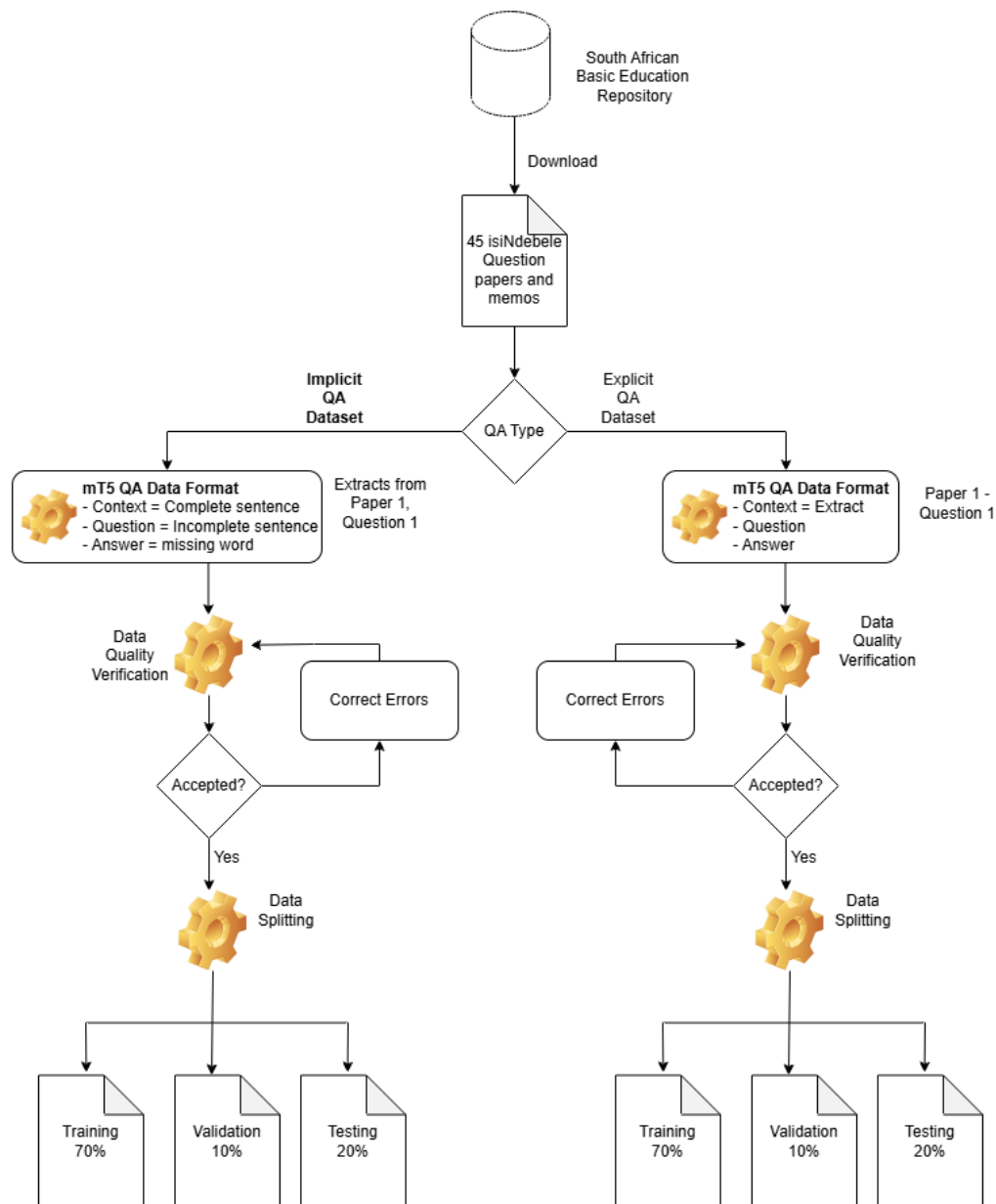


Figure 1. The flowchart depicting the process of isiNdebele explicit and implicit QA dataset creation.

Similar steps are followed in creating the Implicit QA dataset. For the Implicit QA dataset, the fill-in questioning style was used. The extracts from the 45

question papers are divided into sentences. And a single word in each sentence creating an incomplete sentence was randomly mask. To form the triplets, the complete sentence, the incomplete sentence, and the masked/missing word are the context, question and answer respectively. The same data quality verification is performed to determine if the data is acceptable or not. The last step in both datasets is to split them into training, validation and testing. Some studies use the ratio 70:15:15 while others partition the dataset into 70:10:20 for training: validation: testing respectively [18], [19].

- **Data Collection**

We downloaded grade 12 isiNdebele Home, First Additional and Second Additional language question papers and their respective memorandums. The papers are from the National Senior Certificate (NSC) examinations known as matric of the Curriculum Assessment Policy Statements (CAPS). The collected papers are for the examination that were written since the inception of CAPS in 2008. And the extracts are taken from the internet and translated to isiNdebele. The Department of Basic Education (DBE) remains the only source for the type of data we need for the task. The format for fine-tuning multilingual text-to-text transfer transformer (mT5) model for question-answering is a set of triplets: context, question and answer. This type of data is available only in paper 1, under section A, comprehension. Only paper 1 question papers were used to develop the dataset.

- **Data Analysis**

Figure 2 shows the word length distribution for the paragraphs dataset that is created by combining the extracts. All the 45 extracts from which the questions and answers were based on are combined to form a new plain text file named paragraphs. The word length distribution of the combined extracts is depicted on Figure 2. The combined extracts are normally distributed with a total of 16452 words and 7478 unique words. As shown in figure 2, eight is the mode of the length distribution for all the datasets in this study. The top three most common word lengths make up 39.26% of the whole dataset, with word length seven, eight and nine having 13.67%, 14.57% and 11.02%, respectively.

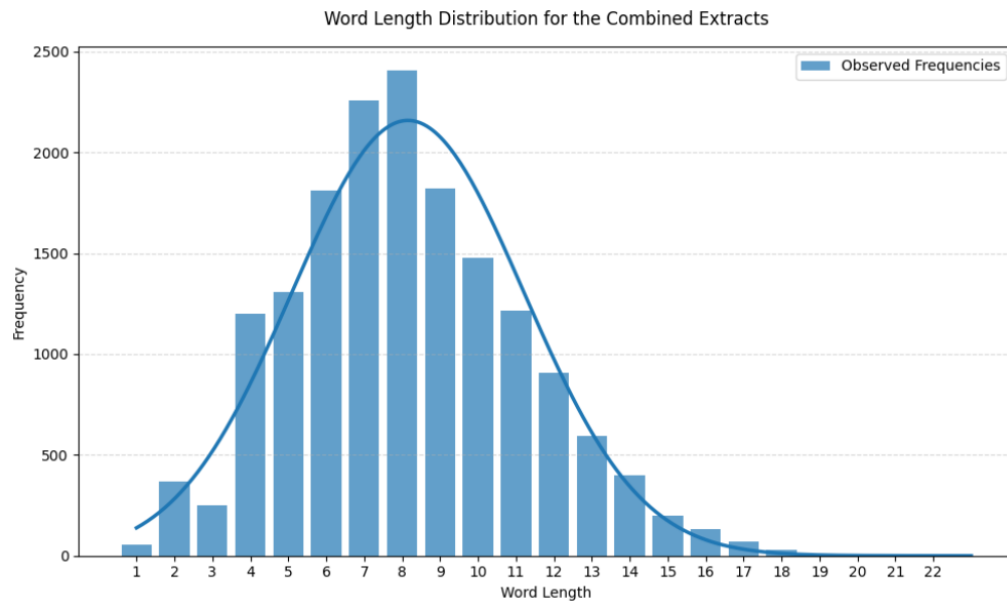


Figure 2. The word length of the QA dataset words, illustrating a normal distribution curve.

Figure 3 shows the top 10 most frequent words by word length, seven, eight and nine. The three-word length groups (seven, eight and nine) have 938, 970 and 950 unique words out of 7478 unique words in the dataset. Due to scarcity of the fine-tuning data, the most frequent word by length, *ngobana* (because) is used 96 times in the isiNdebele QA dataset.

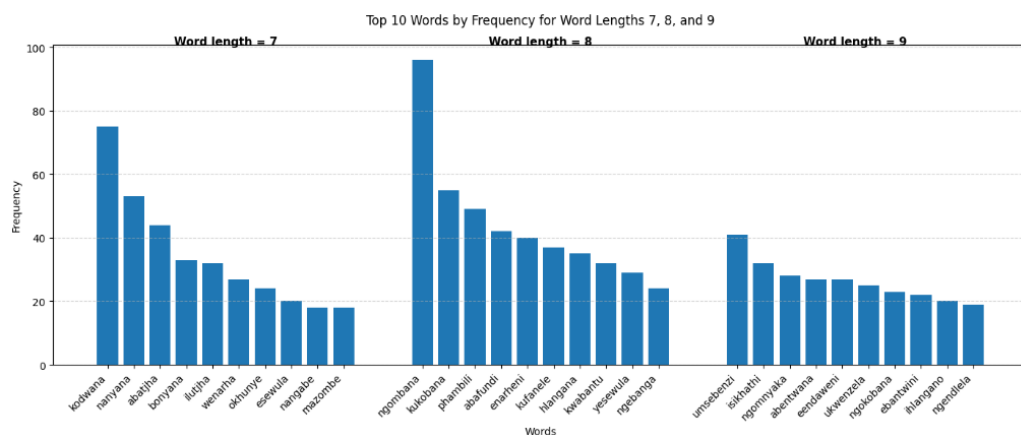


Figure 3. Top 10 most frequent words in the three highest word lengths, 7,8 and 9.

There are 71 unique words of length 3 from 248 words, 208 unique words of length 4 from 1200 words and 368 unique words of length 5 from 1310 words. The words *bona* (to see or to watch) and *khulu* (huge) have the highest frequency of 393 and 97, respectively. The words with length 3, 4 and 5 constitute 2758 out of 16452, and most of these words are verbs, adverbs and adjectives.

There are only six words of length 20, 21 and 22 as depicted in figure 4. These are the longest words in the dataset and they represent the conjunctive form of the language. The word *nangokungayithogomeli* which means “and by not

taking care of it” can be broken down into several morphemes, na-ngoku-ngayi-tlhogomel-i. Each morpheme represents a word or words in English, with the na-morpheme meaning “and”. The morphemes ngoku- and nga- mean “by” and “not” respectively. The object concord “it” is denoted by the yi- morpheme. The verb root in the conjunctive word is -tlhogomel- which refers to “take care of” and to indicate the negative form of the verb the -i is used. This is an indication of how isiNdebele language is agglutinated in its word formation.

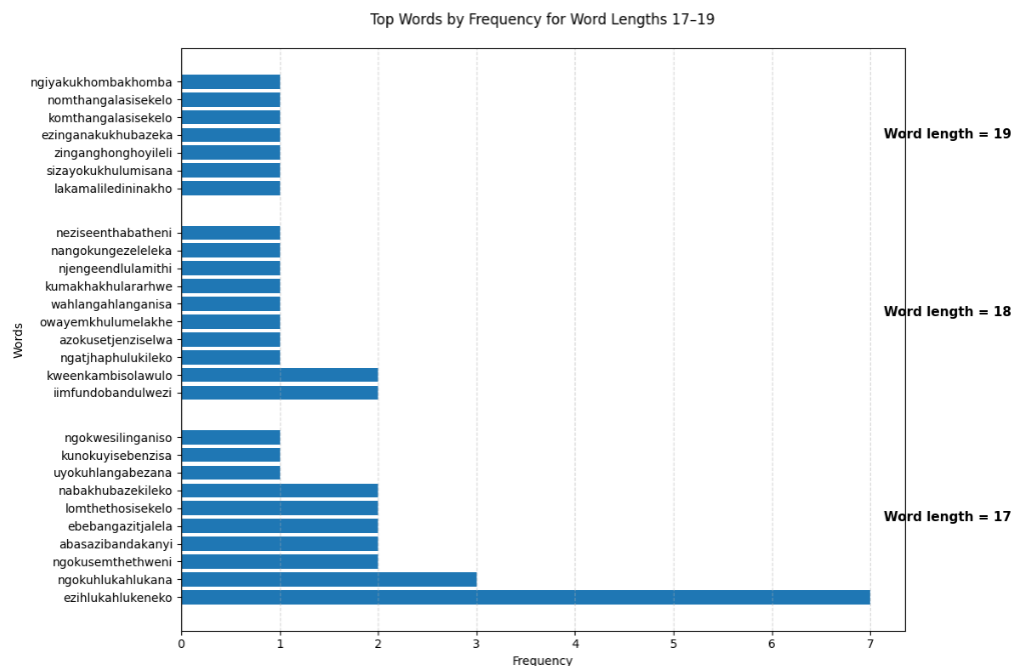


Figure 4. The longest words in the isiNdebele QA datasets derived for word lengths 17, 18 and 19.

Given that some of the most used words are not from three-word lengths with the highest count, we then looked at the words with the highest frequency in the dataset regardless of the word length. Figure 5 illustrates the 15 most frequent words in the isiNdebele QA dataset. The word *bona* is the most used word in the dataset with a frequency of 393, followed by *abantu* with 173. Some of the common words are nouns, *abantu* (people), *afrika* (Africa), *umuntu* (person) and *imali* (money), while others like *phambili* (forward), *khona* (there) are adverbs. There are also conjunction words, *ngombana* (because), *kodwana* (but) and a conjunctive adverb *begodu* (also).

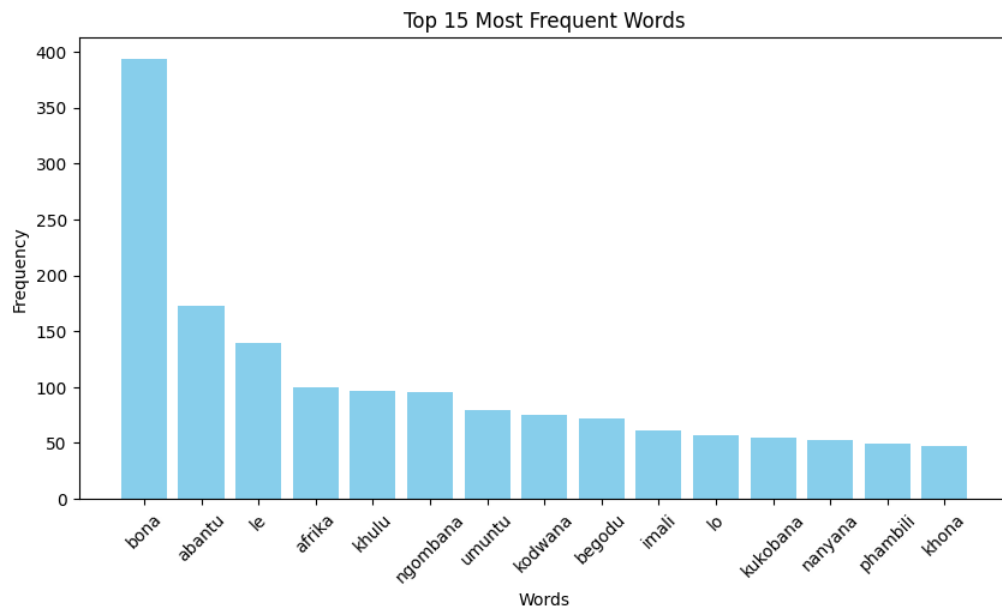


Figure 5. The frequency of the 15 most used words from the isiNdebele QA datasets.

- **Data Format**

The section describes the two question types, the explicit and the implicit isiNdebele QA. For isiNdebele explicit QA dataset, we then used the section A - question 1 of the 45 question papers and their respective memorandums to create the triplets. For every paper, we took the extract as the context and used it for every question-and-answer pair from the same paper, and this was repeated for all the question papers. The developed isiNdebele explicit QA dataset consists of 1276 triplets of context, question and answer.

For the isiNdebele implicit QA dataset, we then concatenated the extracts from all the question papers and split them into sentences. The concatenated extracts were used as the context of the triplets. The resulting file had 3012 sentences, which we randomly masked a word per sentence. The masked word was replaced with “____” in the sentence. And to form the target of each sentence, we used the masked word. The resulting dataset had three columns, the context, question and answer.

In the attempt to increase the isiNdebele implicit QA dataset, we performed an exhaustive technique which involved the masking of each word per sentence. Data augmentation is a widely used technique to deal with data scarcity [20]. We filtered out single-character words to sustain meaningful words only. This process also improves the vocabulary to ensure each word is the target. Unlike in the original implicit QA dataset, the context in this case is the complete sentence not the concatenated extracts. The resulting dataset has 36560 triplets of the context (complete sentence), question (masked sentence) and the answer (the missing word). And the dataset was partitioned into training:validation:testing ratio of 70:10:20 respectively.

- **Data Verification**

The two isiNdebele QA datasets created were evaluated using human-in-the-loop method for their quality to be used in fine-tuning processes. The use of a

human evaluation form has gained momentum in text mining [21], [22], [22], [23]. The number of participants were recruited to investigate and evaluate datasets ranges from 2 to 15 [24], [25], [26].

- Participants

Three native speakers of isiNdebele language were recruited to assist in evaluating the quality of the datasets. All the participants have at least a grade 12 and are currently pursuing their undergraduate qualification.

- Selection of material

A data sampling procedure has been adopted where the humans are only given a percentage ranging from 5% to 10% of the dataset [22]. Samples from both the explicit and implicit QA datasets were randomly generated for the evaluation purpose. From the explicit QA dataset, 5% of the 1276 triplets were randomly sampled for the verification process. The same percentage was also sampled from the 3012 triplets of the implicit QA dataset. For verifying the quality of the two datasets, 64 and 150 triplets were used for the explicit and implicit QA datasets respectively. Table 1 and Appendix A depict implicit and explicit samples that are selected for the verification process. In table 1 both the question and answers are directly extracted from the given context (sentence). And the answers in Appendix A are derived from the given context, they might not be exact extracts from the context as this type of questions tests the comprehension of learners.

Table 1. Three samples of the isiNdebele implicit QA triplets.

Sample	Context	Question	Answer
1	sezizokuphila ipilo ephezulu.	sezizokuphila ____ ephezulu.	Impilo
2	angisakhulumi ngeenthabathaba zeentolo.	angisakhulumi ngeenthabathaba ____.	Zeentolo
3	iimfonyo khulukhulu lokha nabahlangana nabanye abantu.	iimfonyo khulukhulu lokha nabahlangana ____ abantu.	Nabanye

- Test Procedure

There are two protocols developed, for the isiNdebele explicit and implicit QA datasets. The explicit dataset protocol has 3 out of 45 extracts and 64 questions and answers. Questions and answers emanating from the same extracts were grouped together so that the evaluators can evaluate them at once. The task of the evaluator was to read the extract first and check if it is indeed written in isiNdebele language (a YES/NO answer) [27]. Secondly, they were expected to note any errors (semantical or syntactical) and correct them. For each triplet, the context, question and answer, they either accept (1) or reject (0) it based on its correctness.

For the implicit QA dataset protocol, the context is a single sentence. The triplets are in a single row, and the evaluators were also expected to read each sentence and confirm if it is written in isiNdebele language. Since the question and answer are derived from their respective sentences, the evaluators had to confirm that each question has one missing word and that the missing word is the answer. If an error is noted, they indicate and correct it. Lastly, they either accept or reject that triplet by writing either a 1 or 0 respectively [28].

- **Zero-shot Model**

Using the isiNdebele augmented text dataset we then performed a zero-shot on the mT5 model. The mT5 is trained on the largest dataset compared to other

multilingual models [30]. From the 101 languages covered by mT5, there are three South African indigenous languages, isiZulu, isiXhosa and Sotho. The isiZulu and isiXhosa share the same morphological structure with isiNdebele as they belong to the Nguni language group.

From the augmented implicit QA text dataset, 29331 triplets were used for training and validation, and the remaining 7229 triplets were used to test the model. The text dataset created is small compared to datasets used in state-of-the-art QA models. The transfer learning technique with the common morphological structure that the mT5 is trained on will be beneficial to training the isiNdebele language with the minimal dataset created.

To fine-tune the mT5 model for isiNdebele implicit QA we fixed the learning rate and learning rate scheduler to $3e-5$ and cosine respectively. Two batch size experiments were performed in the model development process. The 2x2 and 2x4 were chosen since we have small dataset. Since the answers to this type of questions are exact, to evaluate the zero-shot model we used the ROUGE-L, BERTScore F1, perplexity, accuracy, precision, recall and F1 score (Naaman et al., 2025).

C. Result and Discussion

- **Data Verification**

Three participants responded to both the explicit and implicit data verification. Both the explicit and implicit datasets received exceptional good feedback from all three participants. With all the participants approving all 150 (100%) triplets of the implicit dataset. Participant 1 and participant 3 approved all the 64 (100%) explicit dataset triplets, with participant 2 approving 61 (95.3%) of the triplets. The three triplets were neither accepted or rejected by participant 2. The acceptance level of the two dataset is 100% and 98.33% for implicit and explicit dataset respectively.

- **Zero-shot model**

Figure 6 presents the training and validation losses when training the mT5 for isiNdebele implicit QA. The validation loss of 0.056 from epoch 1 and went up to 0.072 at epoch 2. A gentle decline was then recorded from epoch 3 to epoch 5 with a validation loss of 0.067 at epoch 5 and a convergence between the training and validation loss. These small losses indicate that the model learns fast and performs well on zero-shot. Since the task fill in a missing word is closely related to the span corruption technique used to train the mT5 model, the performance on this task is expected. The model performed exceptionally well with a ROUGE-L and BERTScore F1 of 0,992 and 0,999 respectively. The perplexity and accuracy achieved indicates great performance by the model on the testing dataset with the scores 1,157 and 0,947 respectively. The transfer learning technique proves to be effective as we achieved exceptionally BERTScore F1 of 0.993 and the study of the authors in [15] achieved 0.983.

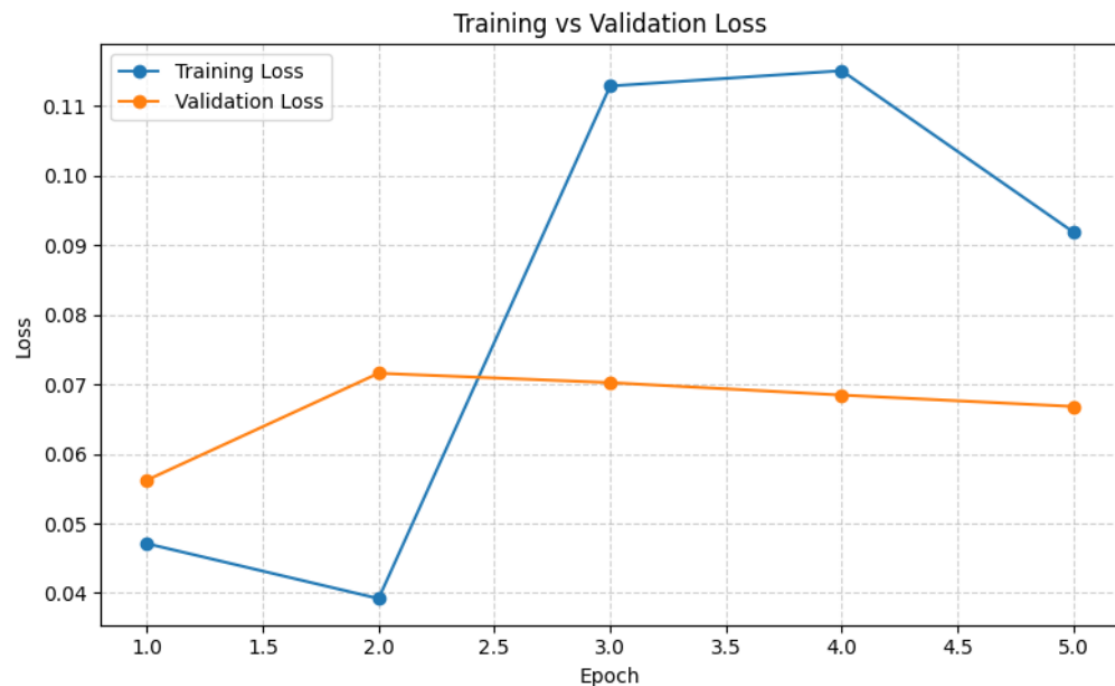


Figure 6. The isiNdebele implicit QA training and validation loss over 5 epochs.

The collected data is very small however the created datasets can serve as a baseline and be increased through more data collection and data augmentation. As the data is downloaded from the government website and was used for the national examinations, there were no errors found in the text. The QA preprocessing helped in achieving a compatible format for mT5. The usage of the created implicit QA dataset in zero-shot proved the capability of multilingual transfers in handling low resourced languages [1]. The isiNdebele implicit QA model developed performed exceptionally well in predicting the missing words. There are 61 incorrectly predicted responses from the 7229 predictions made. Table 2 shows samples of predictions verse the target words and the type of error. From the table we note that the predictions are syntactically related to the target words. For the target word *itjhuguluko* (change) the model predicted *ithuguluko*. For the word *asole* (blame) the prediction is *angasola* (may criticise) which are semantically related.

Table 2. Six samples of predicted missing words and their respective error type.

Example No	Target	Prediction	Error Type
1	2016	/2016	Partial Match
2	Itjhuguluko	Ithuguluko	Exact Mismatch
3	Iindlela	Indlela	Partial Match
4	uqeda	uquqeda	Partial Match
5	ngabazifihlako	ngabaziako	Exact Mismatch
6	asole	angasola	Exact Mismatch

D. Conclusion

This study aimed at creating isiNdebele QA dataset from the minimal data available online. Question papers were downloaded from the South African Basic Education website, and a thorough analysis was performed to understand the dataset. Two QA datasets, explicit and implicit, were created and verified by using human-in-the-loop method. The results showed that the data from the question papers are well structured and proper isiNdebele language is used. To further explore the datasets usage, we performed a zero-shot isiNdebele on the mT5 model using the implicit QA dataset. The developed model performed exceptionally well with a ROUGE-L and BERTScore F1 of 0.992 and 0.999 respectively. The usage of mT5 model for zero-shot in implicit QA task produced good results for a low-resourced language like isiNdebele. Amongst other challenges in the study is the availability of isiNdebele text data especially the explicit QA text data. As part of future work, researcher can look into isiNdebele data collection for different specific tasks to improve the availability

E. References

- [1] A. Alhussain and A. Azmi, "Crosslingual Transfer Learning for Arabic Story Ending Generation," *The Indonesian Journal of Computer Science.*, vol. 13, no. 2, Apr. 2024, doi: 10.33022/ijcs.v13i2.3831.
- [2] Y. Wang *et al.*, "Automated Evaluation of Personalized Text Generation using Large Language Models," Oct. 17, 2023, *arXiv*: arXiv:2310.11593. doi: 10.48550/arXiv.2310.11593.
- [3] M. Kowsher, A. A. Sami, N. J. Prottasha, M. S. Arefin, P. K. Dhar, and T. Koshiba, "Bangla-bert: transformer-based efficient model for transfer learning and language understanding," *IEEE Access*, vol. 10, pp. 91855–91870, 2022.
- [4] N. Wiwatbutsiri, A. Suchato, P. Punyabukkana, and N. Tuaycharoen, "Question generation in the thai language using MT5," presented at the 2022 19th International Joint Conference on Computer Science and Software Engineering (JCSSE), IEEE, 2022, pp. 1–6.
- [5] P. A. Atmajaya, A. Romadhony, and A. F. Ihsan, "Multiple-Choice Question Generation in Low Resource Indonesian Language with T5.," *International Journal of Intelligent Engineering & Systems*, vol. 18, no. 5, 2025.
- [6] A. Moreno-Cediel, J.-A. del-Hoyo-Gabaldon, E. Garcia-Lopez, A. Garcia-Cabot, and D. de-Fitero-Dominguez, "Evaluating the performance of multilingual models in answer extraction and question generation," *Scientific Reports*, vol. 14, no. 1, p. 15477, 2024.
- [7] Salsabila Rabbani, Agustin, Susandri, Rahmiati, and M. Khairul Anam, "Optimalisasi Kinerja Klasifikasi Teks Berdasarkan Analisis Berbasis Aspek Dan Model Hybrid Deep Learning," *The Indonesian Journal of Computer Science*, vol. 13, no. 3, Jun. 2024, doi: 10.33022/ijcs.v13i3.4034.
- [8] M. Szep, D. Rueckert, R. von Eisenhart-Rothe, and F. Hinterwimmer, "A Practical Guide to Fine-tuning Language Models with Limited Data," Nov. 14, 2024, *arXiv*: arXiv:2411.09539. doi: 10.48550/arXiv.2411.09539.
- [9] A. Azmi and T. Alwaneen, "Deep Learning Model for Answering Why-Questions in Arabic," *ijcs*, vol. 12, no. 2, pp. 401–409, Apr. 2023, doi: 10.33022/ijcs.v12i2.3183.

- [10] J. O. Alabi, D. I. Adelani, M. Mosbach, and D. Klakow, "Adapting Pre-trained Language Models to African Languages via Multilingual Adaptive Fine-Tuning," Oct. 18, 2022, *arXiv*: arXiv:2204.06487. doi: 10.48550/arXiv.2204.06487.
- [11] A. Al-Numai and A. Azmi, "LEMMA-ROUGE: An Evaluation Metric for Arabic Abstractive Text Summarization," *ijcs*, vol. 12, no. 2, pp. 470–481, Apr. 2023, doi: 10.33022/ijcs.v12i2.3190.
- [12] M. Korochkina and K. Rastle, "Morphology in children's books, and what it means for learning," *npj Science of Learning*, vol. 10, no. 1, p. 22, 2025.
- [13] S. Maulidia, "Pengaruh Kompensasi, Lingkungan Kerja, dan Gaya Kepemimpinan Terhadap Semangat Kerja Pengajar SMKN 1 Jember," *Jurnal. PERBUP*, no. 3, 2017.
- [14] M. Astuti, Z. Arifin, F. Mutohhari, and M. Nurtanto, "Competency of digital technology: the maturity levels of teachers and students in vocational education in Indonesia," *Journal of Education Technology*, vol. 5, no. 2, pp. 254–262, 2021.
- [15] T. H. Rochadiani, "Pendekatan Transfer Learning Untuk Klasifikasi Tangisan Bayi Dengan Imbalance Dataset," *The Indonesian Journal of Computer Science*, vol. 13, no. 2, Apr. 2024, doi: 10.33022/ijcs.v13i2.3834.
- [16] Fauzan Muhammad, Aniat Murni Arimurthy, and Dina Chahyati, "Transfer learning pada Network VGG16 dan ResNet50," *ijcs*, vol. 12, no. 1, Feb. 2023, doi: 10.33022/ijcs.v12i1.3130.
- [17] A. Shushi and Adnan M. Abdulazeez, "Fashion Design Classification Based on Machine Learning and Deep Learning Algorithms: A Review," *ijcs*, vol. 13, no. 3, Jun. 2024, doi: 10.33022/ijcs.v13i3.3980.
- [18] Hawar Bahzad Ahmad and A. M. Abdulazeez, "Hybrid Transfer Learning Model for Facial Attractiveness Prediction," *ijcs*, vol. 14, no. 6, Dec. 2025, doi: 10.33022/ijcs.v14i6.5018.
- [19] M. Pratiwi and R. Liana Gema, "Machine Learning on Opinion Mining of Netizen's Hate Speech," *The Indonesian Journal of Computer Science*, vol. 13, no. 1, Jan. 2024, doi: 10.33022/ijcs.v13i1.3617.
- [20] J. Turihohabwe, Ssembatya Richard, and Wasswa William, "Exploring Strategies for Optimizing Mobilenetv2 Performance in Classification Tasks Through Transfer Learning and Hyperparameter Tuning with A Local Dataset from Kigezi, Uganda," *The Indonesian Journal of Computer Science*, vol. 14, no. 1, Feb. 2025, doi: 10.33022/ijcs.v14i1.4436.
- [21] S. Schelter, D. Lange, P. Schmidt, M. Celikel, and F. Biessmann, "Automating large-scale data quality verification," 2018.
- [22] N. Polyzotis, M. Zinkevich, S. Roy, E. Breck, and S. Whang, "Data validation for machine learning," *Proceedings of machine learning and systems*, vol. 1, pp. 334–347, 2019.
- [23] C. Federmann, T. Kocmi, and Y. Xin, "NTREX-128-news test references for MT evaluation of 128 languages," presented at the Proceedings of the first workshop on scaling up multilingual evaluation, 2022, pp. 21–24.
- [24] L. Cecchi and P. Babkin, "ReportGPT: Human-in-the-loop verifiable table-to-text generation," presented at the Proceedings of the 2024 Conference on

- Empirical Methods in Natural Language Processing: Industry Track, 2024, pp. 529–537.
- [25] S. Bhatt, R. Jain, S. Dandapat, and S. Sitaram, “A case study of efficacy and challenges in practical human-in-loop evaluation of nlp systems using checklist,” presented at the Proceedings of the Workshop on Human Evaluation of NLP Systems (HumEval), 2021, pp. 120–130.
- [26] H. J. Kang, M. A. Gulzar, N. Peng, and M. Kim, “Human-in-the-loop synthetic text data inspection with provenance tracking,” presented at the Findings of the Association for Computational Linguistics: NAACL 2024, 2024, pp. 3118–3129.
- [27] Y. Wang, J. Pfeiffer, N. Carion, Y. LeCun, and A. Kamath, “Adapting grounded visual question answering models to low resource languages,” presented at the Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2023, pp. 2596–2605.
- [28] S. E. Whang, Y. Roh, H. Song, and J.-G. Lee, “Data collection and quality challenges in deep learning: a data-centric AI perspective,” *The VLDB Journal*, vol. 32, no. 4, pp. 791–813, Jul. 2023, doi: 10.1007/s00778-022-00775-9.
- [29] S. Jumisko-Pyykkö, V. Malamal Vadakital, and M. M. Hannuksela, “Acceptance Threshold: A Bidimensional Research Method for User-Oriented Quality Evaluation Studies,” *International Journal of Digital Multimedia Broadcasting*, vol. 2008, no. 1, p. 712380, 2008.
- [30] D. Waysi Naaman, Berivan Tahir Ahmed, and Ibrahim Mahmood Ibrahim, “Architectural Evolution of Transformer Models in NLP: A Comparative Survey of Recent Developments,” *The Indonesian Journal of Computer Science*, vol. 14, no. 5, Oct. 2025, doi: 10.33022/ijcs.v14i5.4984.

Appendix A. Three samples of the isiNdebele explicit QA triplets.

Context

abantu abaphila nokukhubazeka eendaweni zemisebenzi ngaphasi komthetho wokulingana womnyaka we-2010 umuntu ekuthiwa uphila nokukhubazeka ngonobuthakathaka nofana ongakghoni ukwenza ngefanelo eminye yemisetjenzana yangamalanga. ubuthakathakobu nofana ukungakghonokhu kusuke kubangelwa sisitho nofana zizitho ezithileko zomzimba ezirholopheleko nofana kubangwe kukhubazeka kwengqondo. umuntu angabelethwa nokukhubazeka, angakhubazeka ngebanga lengozi bese uthinteka umgogodlha nanyana alimale iinyawo zingasasebenzi. kesinye isikhathi umuntu uba nokukhubazeka ngebanga lobujamo obuthileko ahlangebezene nabo. ukuphila nokukhubazeka kwabantu kuyahluka. kunokuphila nokukhubazeka okunzinzileko nokuphila nokukhubazeka okurhagalako lokha umuntu nakakhulako, njengokuthi umuntu bekakghona ukwenza izinto ezithileko kodwana ukuya nesikhathi uthole bona akasakghoni ukuzenza. ukuphila nokukhubazeka okubangelwa kusebenzisa khulu iindakamizwa nofana kusela utjwala ngokweqileko akuthathwa njengokukhubazeka ngaphasi komthetho wokulingana womnyaka we-2010. kuyinto ehle nokho bona umuntu ophila nokukhubazeka amukele ukukhubazeka kwakhe. kuliqiniso khona bona ukuphila nokukhubazeka kwabanye abantu kungaba yinto ebudisi kodwana kuneendlela ezinengi umuntu angamukela ngazo ubujamo bokukhubazeka kwakhe abe akghone ukuphila nabo. ukuzihlela komuntu ophila nokukhubazeka ngokuqinisekisa bonyana uhlala ahlwengekile; nendawo ahlala kiyo ihlwengeke, kuphungula igandeleleko angaba nalo bekulethe nokuzithemba. umuntu ophila nokukhubazeka kufanele akhuthazwe bona ahlale azithabulula nofana azibandakanye emidlalweni ayikghonako. lokho kuzokuphungula isikhathi sokuhlala atshwenyekile begodu kuzomenza bona abe nepilo ehle. akangazami ukuphalisana nabantu abangasisebujameni obufana nebakhe ngombana nakungenzeka bamdlule uzokuphela amandla agcine alisile. ufanele abe mumuntu ohlala azithobile nonemikghwa emihle ngaso soke isikhathi. nakungenzeka kube nomuntu olinga ukuhlelekisa ngaye akangaziqaleli phasi, akakhombise ukuzithemba. ngokwenza njalo uzabe ona isithunzi saloyo muntu ohlekisa ngaye kabanye abantu. kuqakathekile bona afune isekelo kibodorhodere bezokwelapha, kubayelelisi bezehlala kuhle nebanganini nofana emalungeni womndeni athembekako. akazifundise ukuhlukanisa phakathi komuntu omsekela ngokweqiniso nalowo olinga ukwehlisa isithunzi sakhe. abantu abaphila nokukhubazeka balisizo elikhulu eendaweni zemisebenzi. isewula afrika, njengezinye iinarha ithlaga iinsebenzi ezinamakghono azokwazi ukuthuthukisa inarha le. abantu abaphila nokukhubazeka banendima eqakathekileko ekufanele bayidlale ekuthuthukiseni umnotho wenarha. ngokwamarhubhululo ahlukahlukeneko kutholakele bona umuntu ophila nokukhubazeka uthuthuka abe sisebenzi esinekgghono elihle tle, esingenisa inzuzo, esikwazi ukusebenzisana nezinye iinsebenzi nesineemfiso ezihle ngomsebenzi esiwenzako. kutholakele bona iinsebenzi eziphila nokukhubazeka zisebenza ngcono khulu kunalezo ezinganakukhubazeka. azitshwili emisebenzini begodu ziyathembeka emakhamphanini nofana emabubulweni eziwasebenzelako. ngalekwalokho abantu abaphila nokukhubazeka bafanele banikelwe amathuba eendaweni zemisebenzi bebathuthukiswe ngombana lokho kumalungelwabo. ukuthuthuka kwethekhnoloji kudlale indima ekulu ekususeni iinqabo ebantwini abaphila nokukhubazeka ukuze bakghone ukungena emabizelweni wemisebenzi abawathandako. abanengi babo abaphila nokukhubazeka bokungaboni, bokungezwa nabakhubazeke ngokomzimba benza kuhle khulu emabizelweni wekhomphyutha. ithekhnoloji yenze bona amalanga la sibe namagqwetha angaboniko emehlweni. kunamalekhtjhara, abotitjhere bemivumo nababonisani bezokuthengisa abangaboniko. abantu abaphila nokukhubazeka kwamehlo bayakghona ukwenza umsebenzi ofana nokwenza iintina kanye nokufaka amathayela. abantu abangezwako nabezwela kude benza ngcono emisebenzini enetjhada efana nokubhodulula iinkoloyi nengaba nemithelela emimbi ebantwini abezwako. amabizelo ahlukahlukeneko afana newezamahlathi, ubukghwari bokudweba, ithekhnoloji yezokwelapha neyamabulungelo enza kuhle ngemisebenzi eyenziwa babantu abaphila nokukhubazeka eendlebeni. abantu abaphila nokukhubazeka kwezitho zomzimba baphumelele emabizelweni afana nokufundisa, wezehlalakuhle, wokulawula amarhwebo nakamanye. iinqabo ezinengi, njengokuthathelwa phasi nokuninana ngokobujamo bepilo kwenze abantu abaphila nokukhubazeka babandlululwe ngokungekho emthethweni emiphakathini abahlala kiyo

nemisebenzini. ngalesi isizathu abantu abaphila nokukhubazeka batlanyelwe umthetho ngaphasi komthetho wokulingana emisebenzini we-1998. umnqopho womthetho lo kususa ukubandlululwa ngokungekho emthethweni nokukhuthaza ukulingana eendaweni zemisebenzi. umthetho lo uvikela abantu abaphila nokukhubazeka ekuninweni ngokungekho emthethweni njengesiqhema esasidinywe amathuba ngaphambilini. abantu abaphila nokukhubazeka bayahlomula ehlelweni lokubonelela labo ababedinywe amathuba ngokobana bakhutjhulelwe eenkhundleni eziphezulu. ungqongqotjhe wezemisebenzi uphasise umthetho wokuphatha kuhle nokuqatjiswa kwabantu abaphila nokukhubazeka ngaphasi komthetho wokulingana kwabasebenzi. umthetho lo uhlahlela abaqatjhi nabasebenzi ukukhuthaza ukulingana kwamathuba nokuphathwa ngendlela efanako nelinganako kweensebenzi eziphila nokukhubazeka. usiza ukwakha nokuyelela inzuzo engalethwa babantu abaphila nokukhubazeka eendaweni zemisebenzi. ngokutjho komthethosisekelo wenarha le omunye nomunye umuntu unelungelo lokuveza nofana lokungavezi ukukhubazeka kwakhe nangabe lokho kukhubazeka akubonakali begodu angeze kwamenza bona angakghoni ukuphumelelisa iindingo zomqatjhi.

Sample	Question	Answer
1	tlola okubili okuvela etheksthini okungaba ziinzathu ezingakwenza bona uqatjhe iinsebenzi eziphila nokukhubazeka nange unebubulo nofana ungsorhwebo.	zingenisa inzuzo. azitshwili emsebenzini.
2	tlola okubili okuvela etheksthini okungaba ziinzathu ezingakwenza bona uqatjhe iinsebenzi eziphila nokukhubazeka nange unebubulo nofana ungsorhwebo.	ziba ziinsebenzi ezihle. zineemfiso ezihle ngemisebenzi eziyenzako.
3	kubayini abantu abaphila nokukhubazeka bayenza kuhle nangokuthembeka imisebenzi abanikelwa yona? ipendulo ayibe liphuzu elilodwa.	kungombana basuke babe nevalo lokobana nakungenzeka baqotjhe ngebanga lokungenzi kuhle imisebenzi angeze basayithola eminye.
4	veza ngomutjho owodwa umehluko phakathi kwabosomarhwebo babomma babantu abanzima bangesikhathi sebandlululo nebanamhlanje.	abomma bangesikhathi sebandlululo bebangakavunyelwa ukutlolisa amarhwebabo embusweni ukuze bababoleke imali kodwana banamhlanje banelungelo lokuboleka imali ngendlela abafuna ngayo begodu namarhwebabo urhulumende uwaqalelela khulu ukudlula wamadoda ukwenzela ukuvala isikhala sangaphambilini.