УДК 004.032.26
DOI: 10.26102/2310-6018/2026.57.6.009

Fake news detection in low-resource languages with LLMs

A.S.M.H. Kabir[1]✉, S.A. Khan[2], A.A. Kharlamov[1], I.M. Voronkov[1]

[1]Moscow Institute of Physics and Technology, Moscow, the Russian Federation
[2]Innopolis University, Innopolis, Republic of Tatarstan, the Russian Federation

Abstract. The proliferation of fake news is a global challenge to tackle in the digital era of information availability. The resourceful languages are tackling this issue through enormous research works whereas the low-resource languages are left behind to address the issue adequately. Bangla is one of the low-resource languages in computation despite being in the top ten most spoken languages in the world. To contribute in the field and address the issue of fake news, this research work focuses on the fake news detection in Bangla language leveraging large recent advancement of language models using cross-lingual prompting techniques for better response from the large language models. We leverage the open source models for resource accessibility and utilize DeepSeek-R1, Llama 3.2 and Qwen 2.5 large language models in our experiments and extensively analyze the fake news detection capacity of each model in Bangla language. We find that Qwen 2.5 outperforms the other models in this specific task achieving a maximum accuracy of 97.5 while it also reports no inconclusive response.

Keywords: fake news, bangla, large language models, low-resource language, cross-lingual prompting.

For citation: Kabir A.S.M.H., Khan S.A., Kharlamov A.A., Voronkov I.M. Fake news detection in low-resource languages with LLMs. *Modeling, Optimization and Information Technology.* 2026;14(6). URL: https://moitvivt.ru/ru/journal/article?id=2368 DOI: 10.26102/2310-6018/2026.57.6.009

Обнаружение фейковых новостей в малоресурсных языках с использованием больших языковых моделей

А.С.М.Х. Кабир[1]✉, С.А. Кхан[2], А.А. Харламов[1], И.М. Воронков[1]

[1]Московский физико-технический институт, Москва, Российская Федерация
[2]Университет Иннополис, Иннополис, Республика Татарстан, Российская Федерация

Резюме. Распространение фейковых новостей представляет собой глобальную проблему в цифровую эпоху доступности информации. Языки с богатыми ресурсами активно решают эту проблему благодаря значительным исследовательским усилиям, тогда как языки с ограниченными ресурсами остаются недостаточно охваченными в этом направлении. Бенгальский язык является одним из таких языков с ограниченными вычислительными ресурсами несмотря на то, что он входит в десятку самых распространённых языков мира. С целью внесения вклада в данную область и решения проблемы фейковых новостей, данное исследование сосредоточено на их обнаружении в бенгальском языке с использованием современных достижений в области языковых моделей, включая методы кросс-лингвистического промтинга для повышения качества ответов больших языковых моделей. В работе используются модели с открытым исходным кодом для обеспечения доступности ресурсов, а именно большие языковые модели DeepSeek-R1, Llama 3.2 и Qwen 2.5. Проводится подробный анализ способности каждой модели обнаруживать фейковые новости на бенгальском языке. Результаты показывают, что модель Qwen 2.5 превосходит другие модели в данной задаче, достигая максимальной точности 97,5 %, при этом не демонстрируя неопределённых ответов.

Ключевые слова: фейковые новости, бенгальский язык, большие языковые модели, языки с ограниченными ресурсами, кросс-языковой промтинг.

© Кабир А.С.М.Х., Кхан С.А., Харламов А.А., Воронков И.М., 2026

Для цитирования: Кабир А.С.М.Х., Кхан С.А., Харламов А.А., Воронков И.М. Обнаружение фейковых новостей в малоресурсных языках с использованием больших языковых моделей. *Моделирование, оптимизация и информационные технологии.* 2026;14(6). (На англ.). URL: https://moitvivt.ru/ru/journal/article?id=2368 DOI: 10.26102/2310-6018/2026.57.6.009

Introduction

In an era of massive information availability, information flows faster across geographical and linguistic boundaries [1]. News is generally produced by newspapers, social media users, independent journalists, online newspapers, media outlets and other authorized sources. Considering the massive content production, fake news is an alarming global concern because the rapid spread of misinformation through fake news causes damage to the flow of healthy information. Fake news can change public opinion, damage financially, create chaos in politics, foster communal violence and so on. Recent developments in the domain of fake news detection in both traditional natural language processing and large language models for fact checking [2] and claim verification [3] are showing promising ways to mitigate this issue. However these research works are mostly focused on English or Chinese language not considering other non-English and non-Chinese languages spoken by billions of people. In a view to address the gap in resources and mitigate the damages caused by fake news, this work focuses on leveraging large language models (LLMs) to detect fake news and serve them as detection tools. Most of the large language models perform well in English or Chinese considering recent developments in OpenAI[1], DeepSeek-AI [4] and multi-lingual support has been provided in Llama 3.2 [5] family models and Qwen 2.5[2] for a few languages. Among these models OpenAI[1] provides proprietary models and DeepSeek-R1 [4], Llama 3.2 [5] and Qwen 2.5[2] are open-source for developers and researchers. Researchers in [6] worked towards a fake news dataset and [7] worked with transformer-based models to tackle the issue of fake news detection in Bangla and have a promising result with BERT based multilingual models. In this work, rather than traditional machine learning or deep learning models we work to leverage large language models for fake news detection. In order to achieve our goal, we use publicly available datasets [6] and extended datasets [8] and preprocess the datasets that are suitable for DeepSeek-R1 [4], Llama 3.2 [5] and Qwen 2.5[2] large language models in fine-tuning and prompting to get expected responses.

Over the years, researchers worked on the issue to mitigate the fabrication and spread of fake news which negatively impacts society and societal narrations. Before the recent advancement of large language models, machine learning and deep learning models were built to identify fake news and filter them out in authentic news. Researchers approached Support Vector Machine (SVM) [9] to identify fake and satirical news, the authors have shown a promising development with considerable accuracy with their curated sample size. Earlier in work [10] authors came up with Convolutional Neural Network (CNN) and BiLSTM to address the issue. In [11] researchers presented a comprehensive review of classification of fake news data collected from social media, newspapers and considered social theories, algorithms aligned with data mining, evaluation metrics and dataset representation. While most works focus on English language and work towards them a very few works have been focused in Bangla language. A scarcity of datasets hinders work on fake news detection in Bangla. Addressing the resource constraint researchers published a highly curated dataset [6] containing around 50000 of unprocessed Bangla newspaper data of authentic and fake news, they have also provided a

[1] OpenAI, Achiam J., Adler S., et al. *GPT-4 Technical Report.* arXiv. URL: https://arxiv.org/abs/2303.08774 (дата обращения: 19.04.2026).

[2] Yang A., Yang B., Zhang B., et al. *Qwen2.5 Technical Report.* arXiv. URL: https://arxiv.org/abs/2412.15115 (дата обращения: 19.04.2026).

portion of labeled dataset annotated by human verifiers. In addition to [6] researchers leveraged LLMs to enrich the dataset and provide a newer version released in [8], this dataset has an independently curated test set of data from credible sources for rigorous evaluation. After the groundbreaking work presented as Transformer [12] the researchers introduced a new simple network architecture based solely on attention mechanisms, dispensing with recurrence and convolutions entirely. Later on a bunch of researchers at Google introduced BERT [13] which stands for Bidirectional Encoder Representations from Transformers, BERT model can be finetuned with just one additional output layer to create state-of-the-art models for a wide range of tasks and provides multilingual support for 104 languages. Researchers [6] and [7] have advanced and shown promising results for fake news detection in Bangla using transformer based BERT [13] models. Studying the prominent advancement in natural language processing and leveraging LLMs we work to address the gap for fake news detection in Bangla utilizing the large language models DeepSeek-R1 [4], Llama 3.2 [5] and Qwen 2.5[2].

Materials and methods

In this work to perform our expected experiments we process the datasets [6, 8], we utilize the LLMs DeepSeek-R1 [4], Llama 3.2 [5] and Qwen 2.5[2], we conduct the experiments through a high-level application programming interface[3] that is compatible with most of the LLMs. We employ Cross-Lingual Prompting (CLP) [14] techniques to ensure better understanding and response against the given fake news in Bangla. In Figure 1 we can see an overview of the procedure of our experiment.

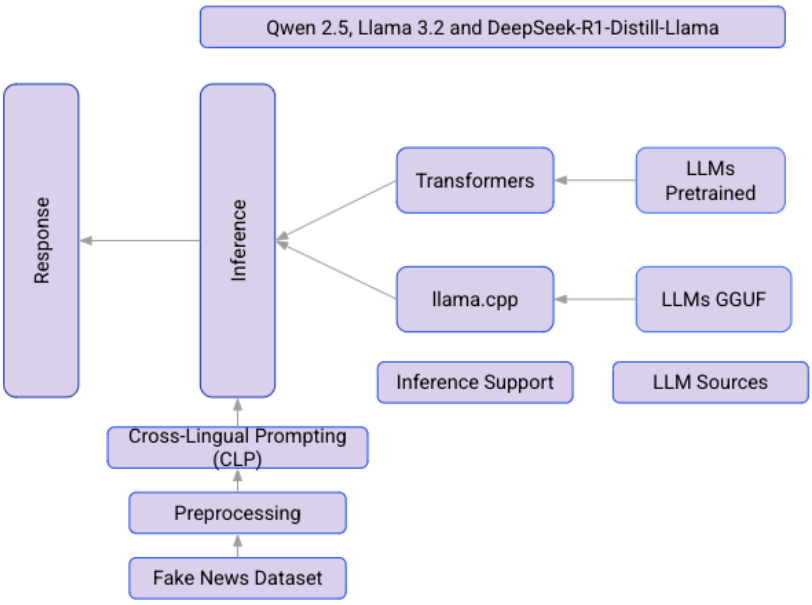

Figure 1 – Detection of fake news in Bangla with LLMs following CLP technique
Рисунок 1 – Выявление фейковых новостей на бенгальском языке с помощью LLM с использованием методики CLP

Cross-Lingual Prompting (CLP). Cross-Lingual Prompting (CLP) [14] is introduced to improve zero-shot chain-of-thought reasoning across languages and significantly increase the efficiency of multilingual tasks for large language models. CLP consists of two parts: (1) cross-lingual alignment prompting and (2) task specific solver prompting. The first part is responsible

[3] abetlen. *llama-cpp-python.* GitHub. URL: https://github.com/abetlen/llama-cpp-python (дата обращения: 19.04.2026).

for alignment of representations for different languages and later on is used to generate the final chain-of-thoughts and give response for the specific task. In our work, we utilize this technique to make the LLM understand the Bangla news piece and respond with task specific solvers to know whether the news is fake or real.

Datasets. In our task, we source 1000 fake news covering various topics in Bangla from publicly available dataset [6] and extended dataset [8]. The news are of diverse categories. Further we clean and preprocess the dataset to use in LLMs for better responses. We cross-check the label and content of the news. The datasets are from credible sources and independent fact-checkers as mentioned in dataset [6] and extended dataset [8] ranked by authenticity of the sources. During preprocessing we have eliminated the unused column and mostly used the "content" and "label" column, where "content" is the news described in Bangla and "label" is either 0 or 1. 0 denotes the fake news and 1 denotes the real news. We feed the content as {news} in our prompt to the large language models and tell it to respond with either 0 or 1. The dataset contains columns "articleID" refers to the article ID, "domain" refers to where it was published electronically, "date" refers to the date of publication, "category" refers to the category of news, "source" refers to the source of news, "relation" refers to the relation of content and headline, "headline" refers to the news headline, "content" refers to the whole content of the news, "label" refers the label of the news as fake or authentic, "F-type" refers to the type of fake news. We extensively worked with the "content" and "label" column in our work with LLMs. A news sample with important metadata with a translation where needed for better understanding is shown in Table 1.

Table 1 – News sample in Bangla with important metadata and translated into English where necessary

Таблица 1 – Подборка новостей на бенгальском языке с важными метаданными и переводом на английский, где это необходимо

Column	In Bangla	In English
Category	–	National
Headline	**চট্টগ্রামে অজগর মেরে হোটেলে মাংস বিক্রির চেষ্টা দুইজনকে ভ্রাম্যমান আদালতের সাজা**	Mobile court sentences two for killing python and trying to sell meat in a hotel in Chittagong.
Content	এবার চট্টগ্রামে অজগরকে পিটিয়ে মেরে চামড়া ছাড়িয়ে নিয়ে পাচাররের চেষ্টা চালিয়েছে একটি অপরাধি চক্র। তবে শেষমেষ আইনের হাত থেকে রেখাই পাননি তারা। তাৎক্ষণিক ভ্রাম্যমান আদালতের অভিযানে হাতেনাতে ধরা পড়ে এ চক্রের দুই সদস্য। পরে দুইজনকে আটক করে বণ্যপ্রানী আইনে ছয় মাসের কারাদন্ড দেন জেলা প্রশাসনের ভ্রাম্যমাণ আদালত। বৃহস্পতিবার সন্ধ্যায় নগরীর নুপুর মার্কেট এলাকা থেকে চামড়া সহ মৃত সাপটি উদ্ধার করা হয়। দন্ডিতরা হলেন রেয়াজউদ্দিন বাজারের নুপুর মার্কেটের 'বন্দর বিতান' নামের একটি দোকানের কর্মচারী রাজীব ও রহমতউল্লাহ। । ভ্রাম্যমাণ আদালত পরিচালনাকারী চট্টগ্রাম জেলা প্রশাসনের নির্বাহী	This time, a criminal gang in Chittagong has tried to smuggle pythons by beating them and skinning them. However, in the end, they did not get a break from the law. Two members of the gang were caught red-handed in an immediate mobile court operation. Later, the mobile court of the district administration arrested the two and sentenced them to six months in prison under the Wildlife Act. On Thursday, the dead snake along with its skin was recovered from the Nupur Market area of Sanbyay Nagari. The convicts are Rajib and Rahmatullah, employees of a shop named 'Bandar Bitan' in Nupur Market in Reazuddin Bazar. . Executive Magistrate Forkan Elahi Anupam of the Chittagong District Administration, who is running the mobile court, said, "They beat up a medium-sized python in the Chittagong

Table 1 (continued)
Таблица 1 (продолжение)

	ম্যাজিস্ট্রেট ফোরকান এলাহী অনুপম বলেন, "বিকালে চট্টগ্রাম রেল স্টেশন এলাকায় মাঝারি আকারের একটি অজগরকে তারা পিটিয়ে মারে। পরে সাপটিকে নিজেদের দোকানের পিছনে নিয়ে চামড়া ছাড়ানো হচ্ছিল। স্থানীয়দের কাছ থেকে খবর পেয়ে ভ্রাম্যমাণ আদালত সেখানে হাজির হয়ে দুইজনকে হাতেনাতে আটক করে। নির্বাহী ম্যাজিস্ট্রেট ফোরকান বলেন, "গ্রেপ্তাররা বলেছে, ওই সাপের চামড়া তারা পশুর চামড়া এবং হাড়গোড় দিয়ে তাবিজ ও ওষুধ তৈরি করে এমন একটি দোকানে বিক্রি করতে চেয়েছিল।" এ অভিযোগে পরে বন্য প্রাণি সংরক্ষণ আইন-২০০৪ এর ৫ ও ১১ ধারা অনুসারে গ্রেপ্তার দুইজনের প্রত্যেককে ছয় মাস করে কারাদন্ডের আদেশ দেয়া হয়। চট্টগ্রাম চিড়িয়াখানার ভারপ্রাপ্ত ডেপুটি কিউরেটর শাহাদাত হোসেন শুভ জানান, মৃত সাপটি চিড়িয়াখানায় পাঠানোর হয়েছে। "সেটি পেলে সংরক্ষণ করব। অজগর সাপের চামড়া খুব বেশি দামে বিক্রি হয় না। অজ্ঞতা থেকেই সাপ মেরে চামড়া তুলে নেওয়ার চেষ্টা করেছিল তারা।" চলতি বর্ষা মৌসুমের শুরু থেকেই নগরী ও মিরসরাই উপজেলার বিভিন্ন স্থানে লোকালয়ে বেশ কয়েকটি বড় ও মাঝারি আকারের অজগর সাপ ধরা পড়েছে। বন বিভাগের কর্মকর্তারা জানান, বর্ষায় পাহাড়ি ঢল ও ভারি বর্ষণের কারণে খাবারের সন্ধানে অজগর লোকালয়ে নেমে আসে। উদ্ধার হওয়া প্রায় সবগুলো অজগরই স্থানীয় বন বিভাগের সহায়তায় সংলগ্ন বনে ছেড়ে দেওয়া এবং চট্টগ্রাম চিড়িয়াখানায় পৌঁছে দেওয়ার উদ্যোগ নেয় স্থানীয়রা।	Railway Station area in the afternoon. Later, the snake was being skinned behind their shop. After receiving information from locals, the mobile court arrived there and arrested the two red-handed. Executive Magistrate Forkan said, "The arrestees said that they wanted to sell the snake's skin to a shop that makes amulets and medicines from animal skins and bones". The two arrested were later sentenced to six months in prison each under Sections 5 and 11 of the Wildlife Conservation Act-2004. Acting Deputy Curator of the Chittagong Zoo, Shahadat Hossain Shuvo, said that the dead snake will be sent to the zoo if it is found. "We will preserve it. Python skins do not sell for very high prices. Out of ignorance, they tried to kill the snake and skin it". Since the beginning of the current monsoon season, several large and medium-sized pythons have been caught in various places in the city and Mirsarai upazila. Forest department officials said that due to hill slopes and heavy rainfall during the monsoon, pythons come down to the locality in search of food. With the help of the local forest department, the locals took the initiative to release almost all the rescued pythons in the adjacent forest and take them to the Chittagong Zoo.
Relation	–	Unrelated
Label	–	0
F-type	–	Clickbaits

Models. We conduct our experiments with instruction-tuned large language models DeepSeek-R1 [4] which is DeepSeek-R1-Distill-Llama-8B (8 billion parameters), Llama 3.2 [5] which is Llama 3.2–3B (3 billion parameters) and Qwen 2.5[2] which is Qwen 2.5–3B (3 billion parameters). We conduct the experiments through a high-level application programming interface[3] that is compatible with most of the LLMs. We have used the quantized

version of their models in GGUF[4] format for efficient inference. In case of multilinguality, DeepSeek-R1 supports 2 languages, Llama 3.2 supports 8 languages and Qwen 2.5 supports 29 languages across the globe. In a single prompt we use the context window size of 8192 tokens. We provide the following system prompt for each LLM: "You are an intelligent assistant who can understand Bangla and detect fake news or real news". And we provide the LLM a prompt following CLP technique: "Please read this news inside the parentheses which is written in Bangla language, this is the news (" + news + "). Now observe the news carefully, understand the meaning, think logically, cross verify information, check the writing style and respond with 0 if you think that the news is fake and respond with 1 if you think that the news is real. In the output please avoid any tokens and return only one digit 1 or 0". Where ("+ news +") is the news content and label will be generated as a response from the LLM.

Evaluation. We evaluate the response from LLMs according to correct response, incorrect response and inconclusive response. Correct response denotes that the LLM flagged the fake news correctly, incorrect response denotes that the LLM returns a wrong assumption and inconclusive response returns neither correct nor incorrect rather it returns random tokens.

$$Correct(\%) = \frac{100 \times Number\ of\ Correct\ Responses}{Total\ Prompts},$$

$$Incorrect(\%) = \frac{100 \times Number\ of\ \text{Incorrect}\ Responses}{Total\ Prompts},$$

$$Inclusive(\%) = \frac{100 \times Number\ of\ \text{Inclusive}\ Responses}{Total\ Prompts}.$$

Results and discussion

After finishing our experiments, we evaluate the results following the evaluation process and present the achieved results for each large language model in Table 2. We focus on accuracy achieved by the models and consider the inconclusiveness while they generate responses against each prompt.

Table 2 – Result of different large language models
Таблица 2 – Результаты различных больших языковых моделей

Large Language Model	Correct (%)	Incorrect (%)	Inconclusive (%)
Qwen 2.5–3B	97.5	2.5	0
Llama 3.2–3B	93	0	7
DeepSeek-R1-Distill-Llama-8B	62	28	10

Table 2 broadly represents the results of different large language models. We find that Qwen 2.5 has outperformed Llama 3.2 and DeepSeekR1. Considering the incorrect response Llama 3.2 reports none while Qwen 2.5 and DeepSeek-R1 reports a few.

Evaluation. We show that in our experiments of fake news detection in Bangla language Qwen 2.5–3B (3 billion parameters) has an accuracy of 97.5 %, Llama 3.2–3B (3 billion parameters) has an accuracy of 93 % and DeepSeek-R1-Distill-Llama-8B (8 billion parameters) has an accuracy of 62 %. After observing the results, Qwen 2.5 outperforms other models in accuracy. In addition, Qwen 2.5 has a higher multilingual support than the other models.

Inconclusive Response. When we observe the inconclusive responses given by the large language models, we see that DeepSeek-R1-Distill-Llama reports a higher number with 10 % response. As of the official information, DeepSeek-R1 primarily targets English and Chinese

[4] ggml. *GGUF (GPT-Generated Unified Format).* GitHub. URL: https://github.com/ggml-org/ggml/blob/master/docs/gguf.md (дата обращения: 19.04.2026).

language. On the other hand, Llama 3.2 reports with 7 % inconclusive responses, Llama 3.2 also lacks multilinguality supporting a lower number. Qwen 2.5 handles inconclusive responses better with 0 % response and that makes it highly qualitative than the other two models. It is to note that Qwen 2.5 has the highest multilingual support among the models used in our experiments.

Comparison with Classification Models. In our research work [7], we explored machine learning, deep learning and neural network based fake news classification models. In this work the overall classification of news articles shows promising results, but the fake news classification as fake minor class achieves lower results. The fake news classification based on Bidirectional Encoder Representations from Transformers (BERT) achieves a maximum of 81 % accuracy, Long short-term memory (LSTM) achieves 65 %, Random Forest (RF) and Support Vector Machine (SVM) achieves 54 %. If we consider these classifiers with our achieved results of LLMs for fake minor class classification with a higher accuracy of maximum 97.5 %, LLMs outperforms classifying fake news in Bangla.

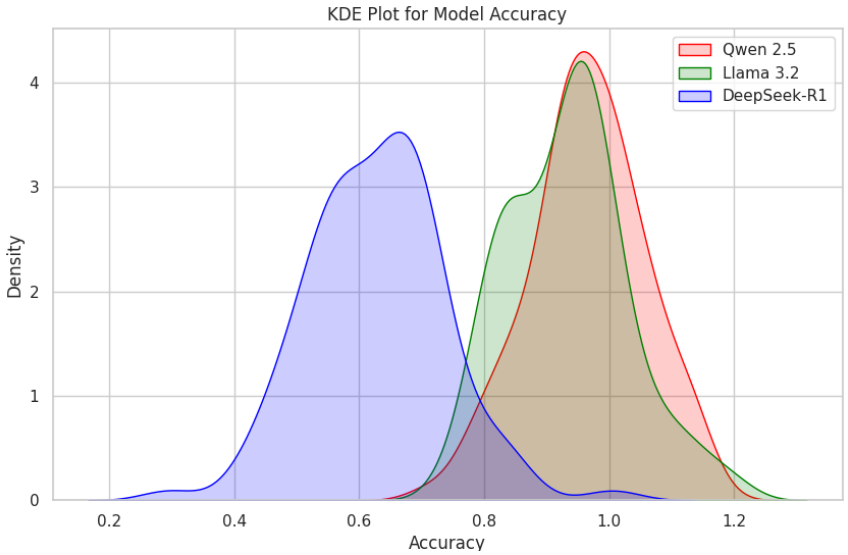

Рисунок 2 – KDE-график точности моделей
Figure 2 – KDE plot for model accuracy

Visualization and Distribution Analysis. We utilize Kernel Density Estimation (KDE) plot to visualize and analyze the model accuracies across various factors. KDE produces the continuous estimation of the probability density function for a given variable. A higher value on the density curve implies a higher probability of achieving that accuracy. And a lower value on the density curve indicates a lower probability of achieving that accuracy. In Figure 2 we observe that the peak of Qwen 2.5 is centered around 0.97 and Llama 3.2 is nearly 0.93 but slightly broader, while the DeepSeek-R1 peaks around 0.65 And we can see that Qwen 2.5 is more consistent in terms of accuracy while DeepSeek-R1 is worse. The overlap between Llama 3.2 and Qwen 2.5 suggests that both models can perform nearly the same.

Conclusion

In this era of rapid production of news and faster proliferation of news, fake news hinders the information among mass people. Fake news is alarming in the digital space and researchers have been working relentlessly to address the issue. Though most of the work to tackle fake news is happening in major languages especially in English, non-English languages lag behind to tackle this issue because of low resources. This work aims to contribute to the low

resource language Bangla and detect fake news in Bangla to mitigate the impact of fake news. We leverage large language models following the cross-lingual prompting technique to detect fake news in Bangla. We experimented with the open-source large language model Qwen 2.5, Llama 3.2 and DeepSeek-R1 which are considered as the latest advancements in large language models. We source publicly available data to utilize in our work and process them to make them suitable for large language models. We consider the multilinguality of the models and observe responses against the fake news from each of the large language models and report the results. After visualizing and analyzing the results we find that Qwen 2.5 outperforms DeepSeek-R1 and Llama 3.2 in this specific task to detect fake news in Bangla. Although our work has a good outcome, it is possible that there is a limitation due to the availability of data in the internet and possibly the large language models are also trained on them as the large language models are very recent advancements. Our work contributes to the domain of natural language processing for social good and advocates for ongoing research on large language models for fake news detection.

REFERENCES / СПИСОК ИСТОЧНИКОВ

1. Darvin R. Language and identity in the digital age. In: *The Routledge Handbook of Language and Identity.* Routledge; 2016. P. 523–540.
2. Lee N., Li B.Z., Wang S., et al. Language Models as Fact Checkers? In: *Proceedings of the Third Workshop on Fact Extraction and VERification (FEVER), 09 July 2020, Seattle, WA, USA.* Association for Computational Linguistics; 2020. P. 36–41. https://doi.org/10.18653/v1/2020.fever-1.5
3. Hoes E., Altay S., Bermeo J. *Leveraging ChatGPT for Efficient Fact-Checking.* OSF. URL: https://doi.org/10.31234/osf.io/qnjkf [Accessed 19th April 2026].
4. DeepSeek-AI, Guo D., Yang D., et al. *DeepSeek-R1: Incentivizing Reasoning Capability in LLMs via Reinforcement Learning.* arXiv. URL: https://arxiv.org/abs/2501.12948 [Accessed 19th April 2026].
5. Grattafiori A., Dubey A., Jauhri A., et al. *The Llama 3 Herd of Models.* arXiv. URL: https://arxiv.org/abs/2407.21783 [Accessed 19th April 2026].
6. Hossain M.Z., Rahman M.A., Islam M.S., et al. BanFakeNews: A Dataset for Detecting Fake News in Bangla. In: *Proceedings of the Twelfth Language Resources and Evaluation Conference (LREC 2020), 11–16 May 2020, Marseille, France.* European Language Resources Association; 2020. P. 2862–2871. URL: https://aclanthology.org/2020.lrec-1.349
7. Kabir A.S.M.H., Kharlamov A.A., Voronkov I.M. Research Methods for Fake News Detection in Bangla Text. In: *Advances in Neural Computation, Machine Learning, and Cognitive Research VII: Selected Papers from the XXV International Conference on Neuroinformatics, 23–27 October 2023, Moscow, Russia.* Cham: Springer; 2023. P. 54–60. https://doi.org/10.1007/978-3-031-44865-2_6
8. Shibu H.M., Datta Sh., Miah M.S., et al. From Scarcity to Capability: Empowering Fake News Detection in Low-Resource Languages with LLMs. In: *Proceedings of the 31st International Conference on Computational Linguistics, COLING 2025 – Workshops, 19–24 January 2025, Abu Dhabi, UAE.* Association for Computational Linguistics; 2025. P. 100–107. URL: https://aclanthology.org/2025.indonlp-1.12
9. Rubin V., Conroy N., Chen Y., et al. Fake News or Truth? Using Satirical Cues to Detect Potentially Misleading News. In: *Proceedings of the Second Workshop on Computational Approaches to Deception Detection, 17 June 2016, San Diego, CA, USA.* Association for Computational Linguistics; 2016. P. 7–17. https://doi.org/10.18653/v1/W16-0802

10. Hossain E., Kaysar M.N., Jalal Uddin Joy A.Z.M., et al. A Study Towards Bangla Fake News Detection Using Machine Learning and Deep Learning. In: *Sentimental Analysis and Deep Learning: Proceedings of ICSADL 2021, 18–19 June 2021, Songkhla, Thailand.* Singapore: Springer; 2021. P. 79–95. https://doi.org/10.1007/978-981-16-5157-1_7
11. Shu K., Sliva A., Wang S., et al. Fake News Detection on Social Media: A Data Mining Perspective. *ACM SIGKDD Explorations Newsletter.* 2017;19(1):22–36. https://doi.org/10.1145/3137597.3137600
12. Vaswani A., Shazeer N., Parmar N., et al. Attention is All You Need. In: *Advances in Neural Information Processing Systems 30: Annual Conference on Neural Information Processing Systems, 04–09 December 2017, Long Beach, CA, USA.* 2017. P. 5998–6008.
13. Devlin J., Chang M.-W., Lee K., et al. BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. In: *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies (NAACL-HLT 2019): Volume 1, 02–07 June 2019, Minneapolis, MN, USA.* Association for Computational Linguistics; 2019. P. 4171–4186. https://doi.org/10.18653/v1/N19-1423
14. Qin L., Chen Q., Wei F., et al. Cross-lingual Prompting: Improving Zero-shot Chain-of-Thought Reasoning across Languages. In: *Proceedings of the 2023 Conference on Empirical Methods in Natural Language Processing (EMNLP 2023), 06–10 December 2023, Singapore.* Association for Computational Linguistics; 2023. P. 2695–2709. https://doi.org/10.18653/v1/2023.emnlp-main.163

ИНФОРМАЦИЯ ОБ АВТОРАХ / INFORMATION ABOUT THE AUTHORS

Кабир А. С. М. Хумаюн, аспирант, кафедра интеллектуальных информационных систем и технологий, Московский физико-технический институт, Москва, Российская Федерация.
e-mail: humaun.kabir@phystech.edu

A. S. M. Humaun Kabir, Postgraduate, Department of Intelligent Information Systems and Technologies, Moscow Institute of Physics and Technology, Moscow, the Russian Federation.

Кхан Самеед Ахмед, аспирант, факультет компьютерных наук и инженерии, Университет Иннополис, Иннополис, Республика Татарстан, Российская Федерация.
e-mail: sameedkhandurrani@gmail.com

Sameed Ahmed Khan, Postgraduate, Faculty of Computer Science and Engineering, Innopolis Univeristy, Innopolis, Republic of Tatarstan, the Russian Federation.

Харламов Александр Александрович, доктор технических наук, профессор, кафедра интеллектуальных информационных систем и технологий, Московский физико-технический институт, Москва, Российская Федерация.
e-mail: kharlamov@analyst.ru

Alexander A. Kharlamov, Doctor of Engineering Sciences, Professor, Department of Intelligent Information Systems and Technologies, Moscow Institute of Physics and Technology, Moscow, the Russian Federation.

Воронков Илья Михайлович, старший преподаватель, кафедра интеллектуальных информационных систем и технологий, Московский физико-технический институт, Москва, Российская Федерация.
e-mail: voronkov.im@phystech.edu

Ilia M. Voronkov, Senior Lecturer, Department of Intelligent Information Systems and Technologies, Moscow Institute of Physics and Technology, Moscow, the Russian Federation.

Статья поступила в редакцию 20.04.2026; одобрена после рецензирования 11.06.2026; принята к публикации 19.06.2026.

The article was submitted 20.04.2026; approved after reviewing 11.06.2026; принята к accepted for publication 19.06.2026.