University *of Ljubljana*
Faculty *of Computer and Information Science*

Automatic generation of Slovenian traffic news for RTV Slovenija

Miha Frangež, Matej Medved, and Jan Mrak

Abstract

/

Keywords

traffic news, llm, prompt engineering, slovenian language, news generation, fine tuning

Advisors: Slavko Žitnik

Introduction

Traffic reporting is an essential service for public broadcasters, providing critical real-time information that affects citizens' daily commutes and travel plans. Currently at RTV Slovenija, traffic news is produced manually, with students reviewing and transcribing data from the *promet.si* portal into structured news segments that are broadcast every 30 minutes. This manual process introduces several challenges:

- It is labor-intensive, requiring constant human monitoring.

- It creates potential for human error in transcription and prioritization.

- It limits the speed at which critical traffic information can be disseminated.

- It introduces inconsistencies in reporting style and format.

Our project aims to address these limitations by developing an automated system for generating Slovenian traffic news reports. By leveraging Large Language Models (LLMs), prompt engineering techniques, and fine-tuning approaches, we intend to create a solution that processes structured traffic data from the *promet.si* portal and automatically generates concise, accurate, and standardized traffic reports that adhere to RTV Slovenija's broadcasting guidelines.

Related work

Similar work on this topic includes papers like Fast Hybrid Approach for Thai News Summarization [1]. The paper is solving the problem of news summarization of news in Thai language using a fine-tuned mBART model.

Another paper Enhancing Large Language Model Performance through Prompt Engineering Techniques [2] focuses on prompt engineering techniques to improve the performance of large language models, which is one of the approaches we will use to generate traffic news reports.

In a similar paper Leveraging the Power of LLMs: A Fine-Tuning Approach for High-Quality Aspect-Based Summarization [3] the authors investigate the use of fine-tuning LLMs for purposes of generating summaries. They hypothesize that fine-tuning open-source LLMs like Llama2, Mistral, Gemma, and Aya on a domain-specific dataset will lead to superior aspect-based summaries compared to current state-of-the-art methods.

Initial ideas

As both the input as well as the intended output texts are in Slovene, it is essential that we use a language model that is proficient in that language. As no powerful foundation models trained specifically on Slovene are available, we will most likely employ a model with strong multi-lingual abilities. Some proprietary models such as ChatGPT have been shown to perform well in Slovene language tasks, however, the inability to fine-tune them, as well as the practical issues with using a proprietary model, only available as an API from one provider, make this a suboptimal choice.

Some open-source or open-weights models such as *Llama 3, Gemma 3* and *DeepSeek R1* might be more suitable for the task, as their inferior multilingual ability could be improved by fine-tuning on the example RTV Slovenija texts, perhaps to the point of surpassing proprietary models. Such models

are also more suitable to specifics of this task, as the public broadcaster has very strict rules about the use of Slovene language and it is unlikely this standard could be achieved without fine-tuning.

As both the input and the output text is very short and the task calls for little reasoning, a purely prompt-based approach might be sufficient to achieve semantically correct output. As this leaves significant space in the context window of most models, a significant about of example text (few-shot prompting) as well as written rules about the style can be provided to the model, which may be sufficient to achieve the required writing style. If this is not sufficient, a feedback loop could be constructed, wherein the model is prompted to analyze and revise its writing based on additional examples and/or written rules.

If prompt-only techniques prove insufficient in reproducing the desired style, this could be improved by fine-tuning the model on the example corpus using methods such as *LoRA*. Fine-tuning is also likely to improve the accuracy of task-specific parts of the text, such as place names.

Dataset

Our primary source of data, used as input for traffic news generation, is the *promet.si* website, operated by the *Traffic Information Centre for Public Roads*. The website publishes displays the latest traffic information in text form, separated into sections by both topic and importance: *Important, Accident, Delay, Weather, Obstructions, Road works, Warning, International information, General*.

For training and evaluation of our models, we obtained a three-year archive of both the aforementioned input data, as well as the final texts that were aired on RTV Slovenija.

Methods

Results

Discussion

References

[1] Kietikul Jearanaitanakij, Suratan Boonpong, Kirttiphoom Teainnagrm, Thanakrit Thonglor, Tiwat Kullawan, and Chankit Yongpiyakul. Fast hybrid approach for thai news summarization. *Engineering and Technology Horizons*, 41(3):410307, Sep. 2024.

[2] Arlo Octavia and Meade Cleti. Enhancing large language model performance through prompt engineering techniques, 10 2024.

[3] Ankan Mullick, Sombit Bose, Rounak Saha, Ayan Kumar Bhowmick, Aditya Vempaty, Pawan Goyal, Niloy Ganguly, Prasenjit Dey, and Ravi Kokku. Leveraging the power of llms: A fine-tuning approach for high-quality aspect-based summarization, 2024.