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

Automatic Generation of Slovenian Traffic News for RTV Slovenija

Filip Turk and Tschimy Aliage Obenga

Abstract

The abstract goes here.

Keywords

Generating traffic reports, Large Language Models, NLP, Prompt Engineering, Fine-tuning, Slovenian traffic news, Automated text generation

Advisors: Slavko Žitnik

Introduction

Traffic reporting is a crucial aspect of public broadcasting, especially for real-time updates on road conditions. Currently, RTV Slovenija relies on students to manually check, filter, and type reports from the Promet.si portal every 30 minutes. This process is time-consuming and prone to inconsistencies.

This project aims to automate traffic news generation using a Large Language Model (LLM). The approach includes leveraging prompt engineering techniques, defining evaluation criteria, and fine-tuning an LLM to improve accuracy and relevance. The generated reports must align with RTV Slovenija's guidelines, ensuring clarity, conciseness, and correctness in road naming and event significance.

Existing Solutions and Related Work

Traffic News Automations

Automated traffic reporting systems have been developed worldwide, primarily using Natural Language Processing (NLP) and Machine Learning (ML) techniques. Some existing solutions include:

1. Google Maps Traffic Alerts: Uses real-time data and machine learning to generate concise traffic updates
2. Waze Traffic Reports: Crowdsourced traffic conditions, automatically summarized for users.
3. AI-driven News Generation: Organizations like OpenAI and Google have explored AI-generated news reports in sports, finance, and weather domains.

These solutions utilize structured traffic data and LLMs for automated summarization. However, fine-tuning on local datasets, such as Slovenian road networks and traffic terminology, remains a challenge.

Large Language Models in Text Generation

Recent advancements in LLMs like GPT-4, LLaMA, and T5 have demonstrated strong capabilities in text summarization and structured data interpretation. Studies have shown that fine-tuning domain-specific models improves text coherence and factual accuracy.

Initial Corpus Analysis

The dataset for this project consists of structured traffic data from Promet.si, formatted in Excel. A preliminary analysis reveals the following key attributes:

Road Names and Locations: Entries specify highways, regional roads, and key urban intersections.

Incident Types: Traffic congestion, road closures, weather-related disruptions, and accidents.

Time Stamps: Time of report updates, crucial for real-time relevance.

Severity and Impact: Differentiates minor delays from major traffic disruptions.

Initial observations suggest that filtering out redundant or low-impact reports will be necessary to maintain concise news summaries. To achieve this, categorizing incidents based on importance will help prioritize critical updates. Categories could include:

1. High Priority: Major accidents, full road closures, severe weather disruptions

2. Medium Priority: Traffic congestion, partial road closures, significant delays
3. Low Priority: Minor roadworks, temporary disruptions with minimal impact

This structured approach will enhance the efficiency of traffic news generation by ensuring that significant events receive appropriate attention.

Initial Ideas and Approach

Prompt Engineering

The first step involves experimenting with LLMs through structured prompts to generate effective traffic summaries. Potential techniques include:

Using role-based prompting (e.g., "You are a Slovenian traffic news reporter")

Applying structured templates to ensure consistent output

Fine-tuning prompt phrasing to optimize summarization quality

Evaluation Criteria Definition

To ensure quality traffic reports, evaluation metrics will focus on:

1. Accuracy: Correct road names and locations
2. Conciseness: Keeping reports within an optimal word count
3. Format: Ensuring the report fit the provided format

Parameter-Efficient Fine-Tuning

Once prompt engineering is refined, fine-tuning an LLM (such as GPT-3.5 or T5) with a domain-specific dataset will be explored. LoRA (Low-Rank Adaptation) or Adapter-based methods will be considered to optimize training efficiency.

Interactive Testing Interface

A web-based or command-line interface will be developed to allow interactive testing of generated reports before full automation.

References