Development of Machine Learning Model for Intent Recognition from Air Traffic Radio Communication

Prague University of Economics and Business | Informatics and Statistics

VŠE / FACULTY OF INFORMATICS AND STATISTICS

Ing. Michal Marhan

Supervisor: Ing. Jan Mittner, Ph.D

Motivation and Problem

Radio communication between pilots and ground (ATC) on non-controlled **air traffic** zones (ATZ) is often **single source** of **information** about the current state of planes/pilot's intention [1].

But due to **noise** and large amount of information, it can be sometimes harder for pilots and operators to keep track of current situation in the ATZ.

Problem can be solved using latest state of the art **ML models** for **Speech To Text and Intent recognition** and **fine-tuning** using custom domain specific dataset.[2]

Internet of Things, **object storage, Software defined Radio (SDR)** and **5G mobile networks** allows collection of massive amounts of training data from real radio communication. [3]

Conclusion

This thesis introduces **scalable architecture** for:

- IoT devices doing live recording of ATC communication from remote ATZ locations.
- Backend and storage for data processing from all IoT endpoints. [4]
- **Fined-tuned** machine learning **models** for air traffic communication **transcription** and tags **extraction** in **real-time.** [5]

Architecture has been proven during test deployment. The models are undergoing further tuning and dashboard for pilots and operators in ATZ is being prepared for commercial deployment.

Experiment results

Speech To Text model	WER* (%)
Fine-Tuned Azure SST model	12,10
Generic Azure SST model	53,13
Generic Whisper Large v3	50,62
Fine-Tuned Whisper Large v3	**9,04**

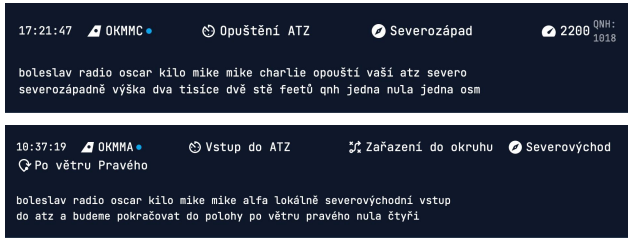
Results based on using own handcrafted dataset of **~750 transcribed** CZ ATC radio conversations. Using custom UX/UI interface for fast and accurate training dataset labelling using experienced humans.

Extreme accuracy for callsign, circuit, altitude, intent transcription.

Work also includes proof of concept for extracting tags above from text using Transformers models with high precision.

*Word error rate. Defines how many words were not detected. Detected wrongly or added.

Examples of live ATC data



The Process

- IoT device listens on ATZ AM radio using Software Defined Radio hardware.
- Detected speech is uploaded to object storage and backend is notified.
- Model 1 does Speech To Text. Model 2 then extracts tags like callsign.
- Valuable info can be then delivered in any form to pilots/operators.

How to deploy nationwide?

Same architecture as in thesis. Using read-only OS, VPN and Containers on Edge devices deployed in Air Traffic Zone.

Model inferencing can be further optimised using libraries like Whisper.cpp and scaled using more hardware for real-time transcription and intention extraction across ~90 ATZ located in the Czech Republic.

This would alllow real-time nationwide data for pilots about current state of Air Traffic Zone.

References

[1] ŘÍZENÍ LETOVÉHO PROVOZU ČESKÉ REPUBLIKY, Š. P. Všeobecná pravidla a postupy [online]. Dostupné z: https://aim.rlp.cz/ais_data/aip/data/valid/e1-1.pdf

[2] openai/whisper [online]. Python. B.m.: OpenAI. 20. duben 2024 [vid. 2024-04-20]. Dostupné z: https://github.com/openai/whisper

[3] FACTOR, M., K. METH, D. NAOR, O. RODEH a J. SATRAN. Object Storage: The Future Building Block for Storage Systems A Position Paper. [online]. Sardinia, Italy: IEEE, 2005, s. 119–123 [vid. 2024-04-03]. ISBN 978-0-7803-9228-1. Dostupné z: doi:10.1109/LGDI.2005.1612479

[4] CHAPMAN, Pete. CRISP-DM 1.0: Step-by-step Data Mining Guide [online]. B.m.: SPSS, 2000. Dostupné z: https://www.ibm.com/docs/it/SS3RA7_18.3.0/pdf/ModelerCRISPDM.pdf

[5] LAKSHMANAN, Valliappa, Sara ROBINSON a Michael MUNN. Machine learning design patterns: solutions to common challenges in data preparation, model building, and MLOps. First edition. Beijing Boston Farnham: O'Reilly, 2020. ISBN 978-1-09-811578-4.