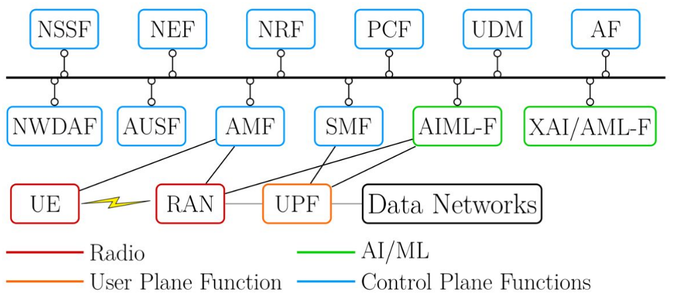
Towards Native Explainable and Robust AI in 6G Networks

G. Attanasio[1,2], S. Moghadas Gholian[1,2], C. Fiandrino[1], M. Fiore[1], J. Widmer[1]
[1]IMDEA Networks Institute [2]University Carlos III of Madrid

Introduction

- Researchers are expecting for 6G networks to face the daunting task of providing services in a mass scale level, that makes the network architecture unprecedentedly complex. This calls for the help of AI-based solutions to automate this process.



NSSF (Network Slice Selection Function)
NEF (Network Exposure Function)
NRF (Network Repository Function)
PCF (Policy Control Function)
UDM (Unified Data Manager)
AF (Application Function)
AUSF (Authentication Server Function)
AMF (Access and Mobility Management Function)
SMF (Session Management Function)
UPF (User Plane Function)
UE (User Equipment)

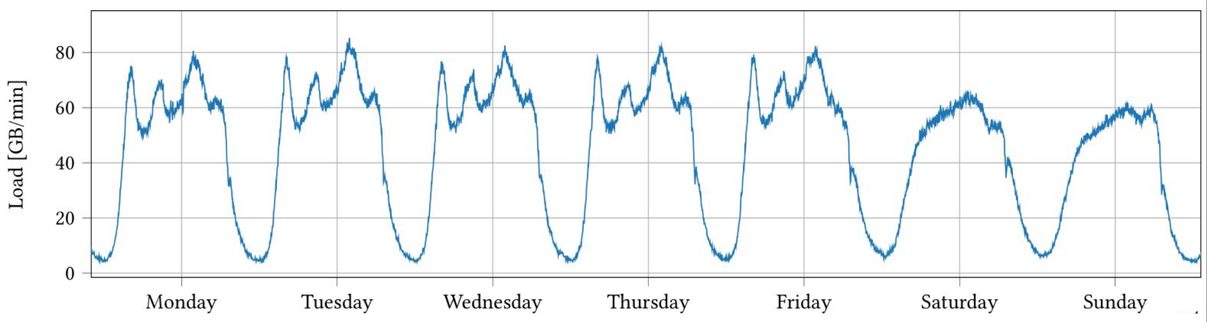
- Utilizing data-based approaches instead of traditional mathematical models, can decrease the trust in our models.
- The lack of transparency raises the issue of vulnerability to attacks and malicious data, i.e., Adversarial ML (AML).

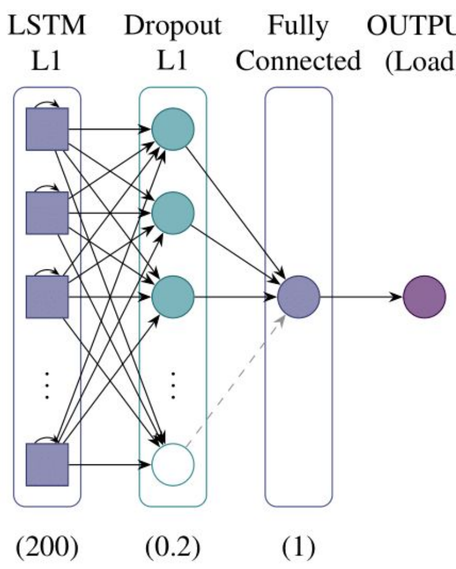
Motivation

- ML models as black-box are difficult to interpret and their decisions are opaque to practitioners.
- Explainable Artificial Intelligence (XAI) allows to detect samples sensitive to perturbations and thus to AML attacks.
- XAI can make AI more transparent and robust.

Method

- First step: investigating XAI with a real-world traffic dataset collected in a 4G network serving a major metropolitan region in Europe.
- Mobile traffic forecasting at 3 minute time scale.

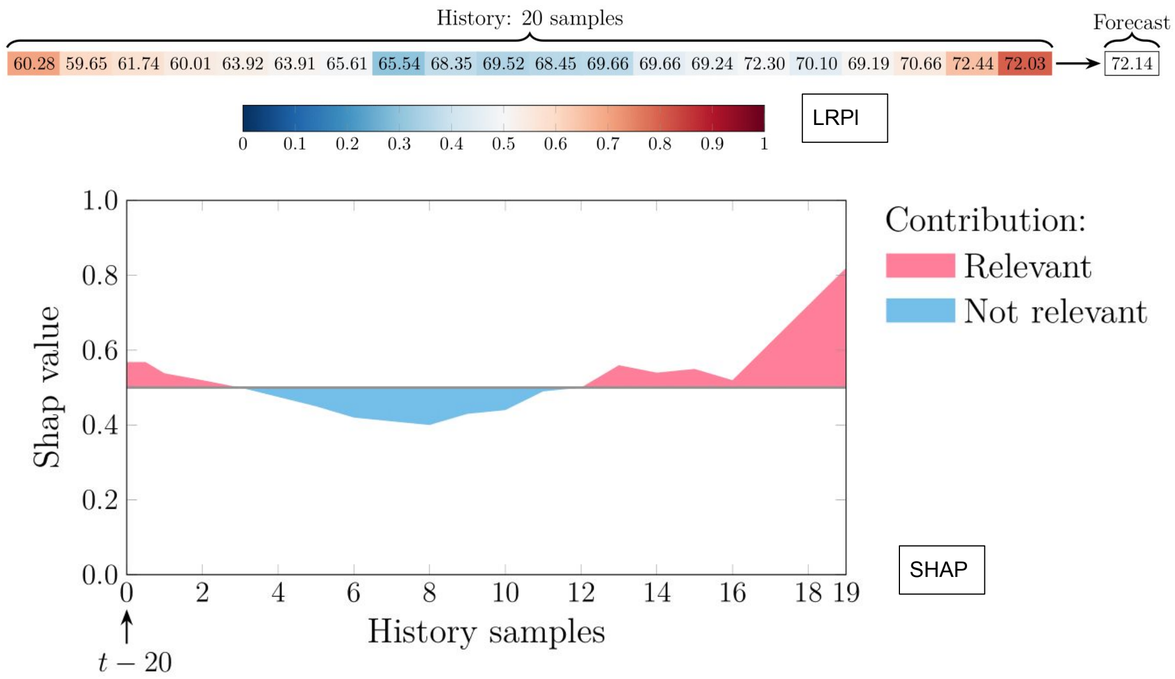




- LSTM recurrent neural networks (RNN) for time series prediction.
- LRP and SHAP methods provide explanations of ML predictions.

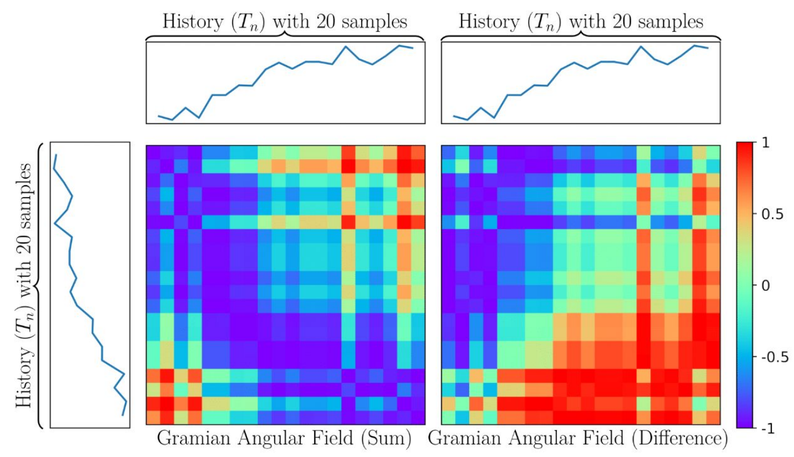
Preliminary Results

- Explain how the model predicts the load in the next 3 minutes from 20 past observations, i.e., 1 hour.



Patterns:
- More recent and old observations have high relevance for the forecast.
- Intermediate observations have low relevance for the forecast.

- Combine interpretability of XAI tools with techniques that mine the input data.



- Gramian Angular Field (GAF) to mine the input data are consistent with explainability methods:
  - GAF shows highly positive values at the extremities.
  - Most relevant observations for low load values and high load ones.

References

[1] Z. Tianhang, B. Li, "Poisoning Attacks on Deep Learning based Wireless Traffic Prediction," in IEEE INFOCOM 2022
[2] S. Lundberg, S. Lee, “A Unified Approach to Interpreting Model Predictions,” in NeurIPS. 2017.
[3] S. Lapuschkin, A. Binder, G. Mantovan, K. Muller, W. Samek in “The LRP Toolbox for Artificial Neural Networks,” in JMLR. 2016
[4] C. Fiandrino, G. Attanasio, M. Fiore, J. Widmer “Toward Native Explainable and Robust AI in 6G Networksç Current State, Challenges and Road Ahead,” under submission in ComCom. 2022.